



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN LINGÜÍSTICA

ANÁLISIS SOCIOCOGNITIVO DE LA PROSODIA DE ENUNCIADOS  
ASEVERATIVOS DE DOS REDES SOCIALES DE LA CIUDAD DE MÉXICO

TESIS

QUE PARA OPTAR AL TÍTULO DE:  
DOCTOR EN LINGÜÍSTICA

PRESENTA:

ALBERTO MANUEL RODRÍGUEZ MÁRQUEZ

TUTOR PRINCIPAL:

DRA. ÉRIKA MENDOZA VÁZQUEZ

COMITÉ TUTOR:

DRA. LEONOR OROZCO VACA

INSTITUTO DE INVESTIGACIONES FILOLÓGICAS

DR. PEDRO MARTÍN BUTRAGUEÑO

EL COLEGIO DE MÉXICO

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO

JUNIO, 2022



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## Índice general

<b>Introducción</b> .....	1
A.1. Aspectos generales .....	1
A.2. Planteamiento del problema.....	2
A.3. Justificación .....	5
A.5. Objetivos .....	7
A.5.1. General .....	7
A.5.2. Específicos .....	7
A.6. Preguntas de investigación.....	8
A.7. Hipótesis .....	8
<b>Capítulo I: Marco Teórico</b> .....	10
1.1. Conceptos teóricos para el análisis prosódico .....	10
1.1.1. Prosodia: definición general.....	11
1.1.2 Elementos prosódicos.....	14
1.1.2.1. Entonación .....	15
1.1.2.2. Intensidad .....	18
1.1.2.3. Duración .....	19
1.1.3. Jerarquía prosódica.....	20
1.1.4. Modelos de análisis prosódico .....	24
1.1.4.1. Modelo Aix-en-Provence .....	24
1.1.4.2. Modelo métrico-autosegmental .....	32
1.1.4.2.1. Sistema de transcripción <i>Tones and Break Indices</i> (ToBI) .....	36
1.1.4.2.2. Consideraciones teóricas sobre el análisis entonativo en dos niveles con base en el proceso de abstracción .....	40
1.3. Conceptos teóricos para el análisis sociolingüístico.....	47

1.3.1. Sociolingüística: definición general .....	47
1.3.2. Las tres olas de la sociolingüística .....	50
1.3.3. Las variables en la sociolingüística .....	53
1.3.4. Prestigio.....	55
1.3.5. Redes sociales .....	57
1.3.6. Modo de vida.....	59
1.3.7. Indexación .....	63
1.4. Lingüística cognitiva.....	67
1.4.1. Lingüística cognitiva: nociones generales.....	67
1.4.2. Esquematización .....	70
1.4.3. Conceptualización .....	72
1.4.4. Categorización.....	72
1.4.5. Prototipo .....	74
1.5. Sociolingüística cognitiva .....	75
1.5.1. Sociolingüística cognitiva: definición general .....	76
1.5.2. La tesis basada en el uso .....	79
1.5.3. Esquemas, prototipos y categorías desde la sociolingüística cognitiva .....	82
1.5.4. Sociofonología cognitiva .....	84
1.5.5. El papel de la percepción .....	90
<b>Capítulo II: Marco metodológico .....</b>	<b>94</b>
2.1. Aspectos generales.....	94
2.1.1. Redes sociales .....	94
2.1.1.1. Características de la red social vecinal .....	95
2.1.1.2. Características de la red laboral.....	97
2.1.2. Modo de vida.....	99

2.1.2.1. Modo de vida 1 .....	99
2.1.2.2. Modo de vida 2 .....	100
2.1.2.3. Modo de vida 3 .....	101
2.2. Fase de producción: elicitación y análisis de datos .....	101
2.2.1. Corpus .....	101
2.2.2. Las entrevistas .....	102
2.2.2.1. Fase 1: Entrevista sociolingüística .....	102
2.2.2.2. Fase 2: Cuestionario sociolingüístico .....	103
2.2.3. Digitalización de los datos .....	104
2.2.4. Selección de enunciados para el análisis .....	104
2.2.5. Unidades de análisis .....	105
2.2.6. Variables .....	107
2.2.7. Método de análisis acústico .....	108
2.2.8. Sistema de transcripción prosódica .....	108
2.2.9. Método de análisis prosódico en dos niveles .....	115
2.2.10. Análisis estadístico .....	118
<b>Capítulo III: Análisis fonético-fonológico .....</b>	<b>121</b>
3.1. Duración, intensidad y frecuencia fundamental (F0) en el enunciado .....	121
3.1.1. Duración en el enunciado .....	123
3.1.2. Intensidad en el enunciado .....	126
3.1.3. Frecuencia fundamental (F0) en el enunciado .....	127
3.2. Primera sílaba tónica .....	128
3.3. Tonemas internos (intertonemas): Configuraciones nucleares de frase entonativa intermedia (ip) .....	134
3.3.1 Tonemas internos (intertonemas) no oxítonos .....	134

3.3.2. Tonemas internos (intertonemas) oxítonos .....	139
3.3.3. Funciones sintáctico-discursivas en los tonemas internos (intertonemas) .....	145
3.4. Tonemas: configuraciones nucleares de frase entonativa mayor (IP) .....	150
3.4.1. Tonemas .....	150
3.4.2. Tonemas no descendentes .....	156
3.4.3. Picos muy tempranos en frases entonativas mayores (IP) .....	159
3.5. Recapitulación del análisis fonético-fonológico .....	170
<b>Capítulo IV: Aspectos sociolingüísticos en la fase de producción</b> .....	<b>174</b>
4.1. Antecedentes de trabajos prosódicos y sociolingüísticos .....	174
4.2. Primera sílaba tónica .....	177
4.3. Tonemas internos (intertonemas): Configuraciones nucleares de frase entonativa intermedia (ip) .....	190
4.3.1. Tonemas internos (intertonemas) no oxítonos .....	190
4.3.2. Tonemas internos (intertonemas) oxítonos .....	203
4.3.3. Funciones sintáctico-discursivas .....	209
4.3.4. Resumen de resultados de los tonemas internos de ip .....	213
4.4. Tonemas: configuraciones nucleares de frase entonativa mayor (IP) .....	213
4.4.1. Tonemas de IP en conjunto con las variables duración, red social y sexo .....	214
4.4.2. Tonemas no descendentes .....	228
4.4.3. Picos muy tempranos en frases entonativas mayores (IP) .....	235
4.4.4. Resumen de resultados de los tonemas de IP .....	239
4.5. Recapitulación del análisis sociolingüístico .....	240
<b>Capítulo V: Análisis de los resultados de la prueba de percepción</b> .....	<b>248</b>
5.1. Antecedentes de trabajos de percepción .....	248
5.2. Metodología de la fase de percepción .....	251

5.2.1. Variables .....	252
5.2.2. Estímulos.....	253
5.2.2.1. Intertonemas oxítonos en ip.....	256
5.2.2.2. Tonemas de IP .....	261
5.2.3. Jueces .....	266
5.2.4. Instrumento .....	268
5.2.5. Validación y prueba piloto .....	274
5.3. Resultados de la primera parte de la prueba: índices sociolingüísticos.....	280
5.3.1. Percepción de los rasgos prosódicos en los intertonemas oxítonos .....	281
5.3.2. Percepción de los rasgos prosódicos en los tonemas .....	289
5.4. Resultados de la segunda parte de la prueba: asociación de patrones prosódicos con el habla de la Ciudad de México.....	297
5.5. Recapitulación de resultados y consideraciones finales de la fase de percepción...	303
<b>Conclusiones</b> .....	316
6.1. Respuestas a preguntas de investigación y evaluación de las hipótesis.....	316
6.2. Conclusiones teóricas y metodológicas .....	319
6.3. Conclusiones sociolingüísticas .....	324
<b>Referencias bibliográficas</b> .....	329
<b>Apéndices</b> .....	344
Apéndice 1. Repertorio de temas y guion de preguntas que sirvieron de guía (no condicional) en la fase 1 de la entrevista (habla espontánea). .....	344
Apéndice 2. Cuestionario sociolingüístico aplicado en la entrevista (adaptación de Villena Ponsoda <i>et al.</i> , 2003; Vida Castro, 2004; Martín Butragueño, Mendoza y Orozco, en preparación). .....	345

Apéndice 3. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de duración de las distintas sílabas pertinentes en la totalidad del corpus. ....	352
Apéndice 4. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de intensidad de las distintas sílabas pertinentes en la totalidad del corpus. ....	357
Apéndice 5. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de la frecuencia fundamental (F0) de las distintas sílabas pertinentes en la totalidad del corpus. ....	364
Apéndice 6. Agradecimiento inicial y breve explicación del proyecto. ....	370
Apéndice 7. Instrucciones generales de la prueba de percepción.....	371
Apéndice 8. Dinámica de la prueba de percepción.....	371
Apéndice 9. Instrucciones de la prueba de audio.....	372
Apéndice 10. Instrucciones para la sección de práctica de la prueba de percepción.....	373
Apéndice 11. Preguntas de la sección de práctica de la prueba de percepción .....	373
Apéndice 12. Mensaje de culminación de la práctica e inicio de la prueba de percepción .....	375
Apéndice 13. Anuncio de culminación del primer ejercicio e instrucciones del segundo .....	376
Apéndice 14. Mensaje de culminación de la prueba de percepción. ....	376
Apéndice 15: Mensaje de bienvenida de la prueba de validación .....	377
Apéndice 16: Instrucciones de la prueba de validación.....	377
Apéndice 17: Mensaje de culminación de la prueba de validación .....	378



## Índice de cuadros

Cuadro 1.1. Ejemplificación de los niveles de la jerarquía prosódica.....	21
Cuadro 1.2. Inventario de acentos tonales monotonaes y bitonaes en el español mexicano y su representación esquemática (De-la-Mota <i>et al.</i> , 2010:320) .....	38
Cuadro 1.3. Inventario de tonos de juntura monotonaes y bitonaes en el español de México y su representación esquemática (De-la-Mota <i>et al.</i> , 2010:321) .....	39
Cuadro 1.4. Representación de la estructura profunda, superficial y esquema acústico de los acentos tonales (tomado de Dorta <i>et al.</i> , 2013b: 72) .....	41
Cuadro 2.1. Distribución de hablantes de acuerdo con las variables pre-estratificadoras .	102
Cuadro 2.2. Descripción del sistema de transcripción prosódica .....	109
Cuadro 2.3. Descripción del tono de juntura !HH% .....	114
Cuadro 3.1. Sílabas con medias de duración (ms) estadísticamente diferentes en la totalidad del corpus.....	124
Cuadro 3.2. Recuento de acentos tonales de la primera sílaba tónica en los niveles concreto y abstracto.....	129
Cuadro 3.3. Medias de duración (ms) de la primera sílaba tónica .....	133
Cuadro 3.4. Recuento de tonemas internos no oxítonos en los niveles concreto y abstracto .....	135
Cuadro 3.5. Medias de duración (ms) de las sílabas nucleares y lindes en los datos generales del corpus.....	138
Cuadro 3.6. Recuento de duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos.....	139
Cuadro 3.7. Recuento de tonemas internos oxítonos en los niveles concreto y abstracto..	140
Cuadro 3.8. Medias de duración (ms) de los intertonemas oxítonos abstractos.....	145
Cuadro 3.9. Recuento de tonemas internos abstractos de ip en cuanto a variables pragmáticas, sintácticas y discursivas.....	146

Cuadro 3.10. Recuento de tonemas internos abstractos de ip en concordancia con la presencia de foco estrecho, expresividad y el conjunto de ambas.....	148
Cuadro 3.11. Recuento de casos de los intertonemas abstractos en elementos extraoracionales.....	150
Cuadro 3.12. Recuento de tonemas en los niveles concreto y abstracto .....	151
Cuadro 3.13. Medias de duración (ms) de las sílabas nucleares y lindes en los datos generales del corpus.....	154
Cuadro 3.14. Recuento de duración relativa de la sílaba nuclear con respecto a la sílaba linde en la totalidad de los tonemas fonológicos del corpus .....	155
Cuadro 3.15. Recuento de tonemas no descendentes de acuerdo con la expresividad en la enunciación.....	158
Cuadro 3.16. Tonemas no descendentes y tipo de foco.....	159
Cuadro 3.17. Tonema de IP con acento nuclear $L+\langle H^*$ en los que la duración de la sílaba nuclear es menor que la del linde (Núcleo < Linde) y parámetros prosódicos .....	161
Cuadro 3.18. Tonema de IP con acento nuclear $L+\langle H^*$ en los que la duración de la sílaba nuclear es mayor que la del linde (Núcleo > Linde) y parámetros prosódicos .....	163
Cuadro 3.19. Tonemas de IP con acento nuclear $L+\langle H^*$ con sostenimiento y posterior descenso de $F_0$ .....	165
Cuadro 3.20. Tonemas de IP con acento nuclear $L+\langle H^*$ en los que la sílaba nuclear inicia con consonante fricativa u oclusiva sorda .....	167
Cuadro 4.1. Acentos tonales en la primera sílaba tónica en cada red social .....	177
Cuadro 4.2. Medias de duración (ms) de la primera sílaba tónica de acuerdo con los acentos tonales producidos por cada red social .....	182
Cuadro 4.3. Recuento de acentos tonales en la primera sílaba tónica en relación con la variable sexo de acuerdo con la red social .....	184
Cuadro 4.4. Medias de duración (ms) de la primera sílaba tónica de acuerdo con los acentos tonales en relación con la variable sexo .....	187
Cuadro 4.5. Tonemas internos no oxítonos en cada red social.....	191

Cuadro 4.6. Tonemas internos abstractos no oxítonos y la variable sexo .....	195
Cuadro 4.7. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos en cada red social .....	197
Cuadro 4.8. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos en cada red social. N=194.....	198
Cuadro 4.9. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de los hombres de ambas redes sociales .....	199
Cuadro 4.10. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de las mujeres de ambas redes sociales.....	200
Cuadro 4.11. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos de los hombres de ambas redes sociales.....	201
Cuadro 4.12. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos de las mujeres de ambas redes sociales .....	201
Cuadro 4.13. Tonemas internos oxítonos en cada red social.....	203
Cuadro 4.14. Tonemas internos oxítonos en variables sexo y red social .....	206
Cuadro 4.15. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos oxítonos en cada red social .....	207
Cuadro 4.16. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de acuerdo con el sexo de los informantes y con la red social .....	209
Cuadro 4.17. Intertonemas abstractos en las distintas funciones sintáctico-discursivas en la red vecinal .....	211
Cuadro 4.18. Intertonemas abstractos en las distintas funciones sintáctico-discursivas en la red laboral.....	211
Cuadro 4.19. Tonemas en cada red social .....	214
Cuadro 4.20. Tonemas de acuerdo con la variable sexo y red social .....	220
Cuadro 4.21. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas en cada red social .....	223

Cuadro 4.22. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por cada red social .....	224
Cuadro 4.23. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas de los hombres de ambas redes sociales .....	226
Cuadro 4.24. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas de las mujeres de ambas redes sociales .....	226
Cuadro 4.25. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por hombres de ambas redes sociales .....	227
Cuadro 4.26. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por mujeres de ambas redes sociales .....	228
Cuadro 4.27. Tonemas no descendentes y red social .....	229
Cuadro 4.28. Tonemas no descendentes y variable sexo .....	231
Cuadro 4.29. Tonemas no descendentes de acuerdo con la expresividad y tipo de foco en cada red social .....	232
Cuadro 4.30. Tonemas no descendentes de acuerdo con la expresividad y tipo de foco en relación con el sexo de los informantes .....	232
Cuadro 4.31. Tonemas con picos muy tempranos en la red vecinal de acuerdo con el tipo de foco y expresividad.....	236
Cuadro 4.32. Tonemas con picos muy tempranos en la red laboral de acuerdo con el tipo de foco y expresividad.....	237
Cuadro 5.1. Distribución de los jueces de acuerdo con sus rasgos sociales.....	267
Cuadro 5.2. Configuraciones entonativas incluidas en el ejercicio de contraste.....	273
Cuadro 5.3. Distribución de los estímulos en la prueba de validación de estímulos .....	275
Cuadro 5.4. Medias de las respuestas de los J1 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	282
Cuadro 5.5. Distribución de las respuestas de los J1 a la cuarta pregunta de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	283

Cuadro 5.6. Medias de las respuestas de los J2 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	285
Cuadro 5.7. Distribución de las respuestas de los J2 a la cuarta pregunta de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	285
Cuadro 5.8. Medias de las respuestas de los J3 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	287
Cuadro 5.9. Distribución de las respuestas de los J3 a la cuarta pregunta de la fase 1 en relación con los patrones en los intertonemas oxítonos .....	287
Cuadro 5.10. Medias de las respuestas de los J1 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas.....	290
Cuadro 5.11. Distribución de las respuestas de los J1 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas.....	291
Cuadro 5.12. Medias de las respuestas de los J2 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas.....	293
Cuadro 5.13. Distribución de las respuestas de los J2 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas.....	294
Cuadro 5.14. Medias de las respuestas de los J3 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas.....	295
Cuadro 5.15. Distribución de las respuestas de los J3 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas.....	296
Cuadro 5.16. Distribución de las respuestas relacionadas con el intertonema más asociado con la Ciudad de México en contraste de pares en cada estímulo.....	299
Cuadro 5.17. Distribución de las respuestas relacionadas con el intertonema más asociado con la Ciudad de México en contraste de pares en cada estímulo.....	301

## Índice de figuras

Figura 1.1. Secciones básicas y esenciales de un grupo melódico (diseño propio con base en Navarro Tomás, 1966).....	16
Figura 1.2. Tonemas planteados por Navarro Tomás (diseño propio con base en Navarro Tomás, 1966).....	17
Figura 1.3. Niveles del módulo prosódico (Tomado de Baqué y Estruch, 2003:125) .....	26
Figura 1.4. Oscilograma y curva estilizada a través de algoritmo MOMEL del enunciado <i>Una vez en el interior, los agentes encontraron a tres niños</i> .....	28
Figura 1.5. Ejemplo de aplicación de las reglas de reajuste y linearización para el enunciado <i>És imprescindible que la tingui demà perquè sóc diabètic</i> (tomado de Baqué y Estruch, 2003:137).....	31
Figura 1.6. Oscilograma, espectrograma y F0 del enunciado <i>¿Entonces sí vienes a cenar?</i> (Tomado de De-la-Mota <i>et al.</i> , 2010: 336).....	34
Figura 1.7. Esquema y prototipo (tomado de Moreno Fernández, 2012:97) .....	83
Figura 1.8. Modelo de categoría radial (tomado de Kristiansen, 2006: 127).....	88
Figura 1.9. Modelo radial de la estructura interna correspondiente a los fonemas /t/ y /d/ del inglés (tomado de Moreno Fernández, 2012: 150).....	88
Figura 1.10. Modelo de red (tomado de Kristiansen, 2006: 129).....	89
Figura 2.1. Mapa de las colonias en las que viven los miembros de la red vecinal (El DeFe, 2021).....	96
Figura 2.2. Oscilograma, espectrograma y curva melódica de dos frases entonativas intermedias del enunciado Rs1M2-29: [( <i>Le puso harta sal</i> ) <sub>ip</sub> ( <i>en la espalda</i> ) <sub>ip</sub> ( <i>y que mi otro hermano</i> ) <sub>ip</sub> ( <i>le lambiera la espalda</i> ) <sub>ip</sub> ] <sub>IP</sub> . .....	115
Figura 2.3. Ejemplo de transcripción en dos niveles (concreto y abstracto) del enunciado Rs1M2-30: [( <i>Mi mamá nos pegaba con lo que ella encontraba</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	116
Figura 2.4. Ejemplo de transcripción en dos niveles (concreto y abstracto) del enunciado Rs2M2-39: [( <i>Mi reto principal y el más grande fue con mi marido</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	117

Figura 3.1. Oscilograma, espectrograma y F0 del enunciado Rs2H1-40: [( <i>Logró salirse</i> ) <sub>ip</sub> (y te puede borrar algo que tu hayas hecho) <sub>ip</sub> ] <sub>IP</sub> con señalamiento de las sílabas analizadas	122
Figura 3.2. Oscilograma, espectrograma y F0 del enunciado Rs2H3-13: [( <i>Esa visión teníamos</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	122
Figura 3.3. Oscilograma, espectrograma y F0 del enunciado Rs2H1-22: [( <i>Esos medios digitales no existían</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	130
Figura 3.4. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente al primer acento tonal.....	132
Figura 3.5. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los intertonemas no oxítonos.....	137
Figura 3.6. Oscilograma, espectrograma y F0 del enunciado Rs2M3-10: [( <i>Se me hace una cultura</i> ) <sub>ip</sub> ( <i>muy</i> ) <sub>ip</sub> ( <i>aislada</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	141
Figura 3.7. Oscilograma, espectrograma y F0 del enunciado Rs1M1-18: [( <i>Primero que ya lo iban a operar</i> ) <sub>ip</sub> (y ya estaba llore y llore) <sub>ip</sub> ] <sub>IP</sub> .....	142
Figura 3.8. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los intertonemas oxítonos.....	143
Figura 3.9. Oscilograma, espectrograma y F0 del enunciado Rs1M2-23: [( <i>Sí entraba</i> ) <sub>ip</sub> ( <i>pero se iba a jugar a las maquinitas</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	147
Figura 3.10. Diagrama representativo de la centralidad (prototipicidad) compartidos por tres categorías cognitivas. ....	149
Figura 3.11. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los tonemas .....	153
Figura 3.12. Oscilograma, espectrograma y F0 del enunciado Rs1H1-18: [( <i>El francés ya no es la México Nuevo</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	157
Figura 3.13. Oscilograma, espectrograma y F0 del enunciado Rs2H3-33: [( <i>Me gusta mucho el mar</i> ) <sub>ip</sub> ( <i>lo respeto mucho</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	157

Figura 3.14. Oscilograma, espectrograma y F0 del enunciado Rs1H3-21: [( <i>Cuando ya estaba</i> ) <sub>ip</sub> ( <i>agarrabas un bote lo llenabas</i> ) <sub>ip</sub> ( <i>y se hacían unas planchas así como la mesa</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	162
Figura 3.15. Ejemplo de acento nuclear L+<H* con descenso tonal en la misma sílaba en la sección final del enunciado Rs1H3-13: [( <i>Cuando yo estaba en la secundaria</i> ) <sub>ip</sub> ( <i>también ya estaba trabajando</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	164
Figura 3.16. Ejemplo de sílaba nuclear con consonante sorda prevocálica en el enunciado Rs1H1-15: [( <i>Había un padre</i> ) <sub>ip</sub> ( <i>que se casó con una Miss Venezuela</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	166
Figura 3.17. Ejemplo de sílaba nuclear con consonante sorda prevocálica en el enunciado Rs2M1-06: [( <i>Toda esa parte está creciendo muchísimo</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	167
Figura 3.18. Oscilograma, espectrograma y F0 del enunciado Rs2H3-35: [( <i>Queríamos hacer un viaje a Cuba</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	168
Figura 3.19. Oscilograma, espectrograma y F0 del enunciado Rs1H2-29: [( <i>Dicen que de tal palo tal astilla</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	169
Figura 3.20. Oscilograma, espectrograma y F0 del enunciado Rs1H2-05: [( <i>Obviamente el sol sale para todos</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	169
Figura 4.1. Representación gráfica del tonema circunflejo L+ <sub>i</sub> H* L% .....	175
Figura 4.2. Oscilograma, espectrograma y F0 del enunciado Rs1M1-15: [( <i>Nosotros no los necesitamos</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	178
Figura 4.3. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M1-24: [( <i>La mayoría de las mamás</i> ) <sub>ip</sub> ( <i>de la escuela de mi hija</i> ) <sub>ip</sub> ( <i>no trabajan</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	178
Figura 4.4. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H3-7: [( <i>Pues mientras no caigan las cabezas</i> ) <sub>ip</sub> ( <i>los peones siguen trabajando</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	180
Figura 4.5. Oscilograma, espectrograma y F0 del enunciado Rs2H1-22: [( <i>Esos medios digitales no existían</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	180



Figura 4.6. Orden de asignación de rasgos prosódicos (acento tonal y duración) en el sistema cognitivo de la red laboral .....	183
Figura 4.7. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1M3-33: [( <i>Conocí a un chavo con el que empecé a salir</i> ) <sub>ip</sub> ( <i>pero me causó muchos problemas porque era muy celoso</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	184
Figura 4.8. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M3-19: [( <i>Si la regresas con la llanta ponchada o raspada o así</i> ) <sub>ip</sub> ( <i>pues te cobro</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	185
Figura 4.9. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2H2-32: [( <i>En ese instante</i> ) <sub>ip</sub> ( <i>para mí desapareció la ventana</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	188
Figura 4.10. Oscilograma, espectrograma y F0 de las tres primeras palabras fonológicas del enunciado Rs2M2-28: [( <i>Según yo todavía en aquel entonces no estaban como ahorita los niños</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	188
Figura 4.11. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M2-1: [( <i>Finalmente estudiamos todo</i> ) <sub>ip</sub> ( <i>y decidimos que</i> ) <sub>ip</sub> ( <i>comunicación era lo mío</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	191
Figura 4.12. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1M2-24: [( <i>Mejor me dieron informes</i> ) <sub>ip</sub> ( <i>de mi sobrino</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	192
Figura 4.13. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente a los intertonemas no oxítonos en ambas redes sociales .....	192
Figura 4.14. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H3-33: [( <i>Ahorita en la actualidad lo he visto</i> ) <sub>ip</sub> ( <i>ya ves a chavos drogándose</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	193
Figura 4.15. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M2-8: [( <i>Ahorita tienes programas bien divertidos</i> ) <sub>ip</sub> ( <i>y la verdad es que sí me hubiera visto ahí perfecto</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	202
Figura 4.16. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H2-26: [( <i>Aquí de marihuana</i> ) <sub>ip</sub> ( <i>saldría esto</i> ) <sub>ip</sub> ( <i>quince pesos</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	202

Figura 4.17. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente los intertonemas oxítonos en ambas redes sociales..	205
Figura 4.18. Oscilograma, espectrograma y F0 de las dos primeras frases entonativas intermedias del enunciado Rs2H3-32: [( <i>La ciudad</i> ) <sub>ip</sub> ( <i>esta de San Cristóbal</i> ) <sub>ip</sub> ( <i>me gustó mucho</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	207
Figura 4.19. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente los tonemas en ambas redes sociales .....	215
Figura 4.20. Oscilograma, espectrograma y F0 del enunciado Rs1H1-18: [( <i>El francés ya no es la México Nuevo</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	217
Figura 4.21. Oscilograma, espectrograma y F0 del enunciado Rs2H1-18: [( <i>Las empresas tienen que ver cómo generar dinero</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	218
Figura 4.22. Oscilograma, espectrograma y F0 del enunciado Rs1M1-6: [( <i>Está tomando el café apenas</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	221
Figura 4.23. Oscilograma, espectrograma y F0 del enunciado Rs2M3-20: [( <i>Con un portuñol espectacular nos comunicábamos perfecto</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	221
Figura 4.24. Oscilograma, espectrograma y F0 del enunciado Rs1H1-6: [( <i>Tendría diez once años</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	222
Figura 4.25. Oscilograma, espectrograma y F0 del enunciado Rs1H3-10: [( <i>La preparatoria te digo nada más la tengo trunca</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	222
Figura 4.26. Oscilograma, espectrograma y F0 del enunciado Rs1M1-1: [( <i>A Óscar le tocó negra</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	230
Figura 4.27. Oscilograma, espectrograma y F0 de la última frase entonativa intermedia del enunciado Rs1M1-14: [( <i>No le salió tampoco tan caro al Carlos</i> ) <sub>ip</sub> ( <i>nada más fue de entrada por salida</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	230
Figura 4.28. Oscilograma, espectrograma y F0 de la última ip del enunciado Rs1M3-19: [( <i>Me acosté</i> ) <sub>ip</sub> ( <i>prendí puse una película</i> ) <sub>ip</sub> ( <i>y</i> ) <sub>ip</sub> ( <i>y me quedé dormida</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	233
Figura 4.29. Oscilograma, espectrograma y F0 del enunciado Rs1H2-6: [( <i>Vivíamos en un cuartito yo y ella</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	233

Figura 4.30. Oscilograma, espectrograma y F0 del enunciado Rs1H1-29: [( <i>La cámara de diputados</i> ) <sub>ip</sub> ( <i>son los dictadores</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	235
Figura 4.31. Oscilograma, espectrograma y F0 de las últimas sílabas del enunciado Rs2H1-6: [( <i>Siempre va a haber algo que trate de detenerte</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	238
Figura 5.1. Oscilograma, espectrograma y F0 del enunciado original [( <i>La mezcla cultural de ese lugar</i> ) <sub>ip</sub> ( <i>lo hacía muy divertido</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	257
Figura 5.2. Oscilograma, espectrograma y F0 del enunciado base [( <i>La mezcla cultural de ese lugar</i> ) <sub>ip</sub> ( <i>lo hacía muy divertido</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la primera batería de estímulos ...	258
Figura 5.3. Oscilograma, espectrograma y F0 del enunciado base [( <i>La mezcla cultural de ese lugar</i> ) <sub>ip</sub> ( <i>lo hacía muy divertido</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la segunda batería de estímulos ..	259
Figura 5.4. Oscilograma, espectrograma y F0 del enunciado original [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> con duración normal en la sílaba final.....	262
Figura 5.5. Oscilograma, espectrograma y F0 del enunciado base [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la primera batería de estímulos con duración normal en la sílaba final .....	262
Figura 5.6. Oscilograma, espectrograma y F0 del enunciado original [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> con duración modificada en la sílaba final .....	263
Figura 5.7. Oscilograma, espectrograma y F0 del enunciado base [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la primera batería de estímulos con duración modificada en la sílaba final.....	263
Figura 5.8. Oscilograma, espectrograma y F0 del enunciado base [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la segunda batería de estímulos con duración normal en la sílaba final .....	264
Figura 5.9. Oscilograma, espectrograma y F0 del enunciado base [( <i>Defendía mucho a mi hermano</i> ) <sub>ip</sub> ] <sub>IP</sub> según criterios de la segunda batería de estímulos con duración modificada en la sílaba final.....	264
Figura 5.10. Oscilograma, espectrograma y F0 del enunciado distractor notoriamente artificial [( <i>La mezcla cultural de ese lugar</i> ) <sub>ip</sub> ( <i>lo hacía muy divertido</i> ) <sub>ip</sub> ] <sub>IP</sub> .....	276

Figura 5.11. Oscilograma, espectrograma y F0 del enunciado distractor notoriamente artificial [(Defendía mucho a mi hermano) <sub>ip</sub> ] <sub>IP</sub> .....	276
Figura 5.12. Oscilograma, espectrograma y F0 del enunciado distractor no comunitario [(La mezcla cultural de ese lugar) <sub>ip</sub> (lo hacía muy divertido) <sub>ip</sub> ] <sub>IP</sub> .....	277
Figura 5.13. Oscilograma, espectrograma y F0 del enunciado distractor no comunitario [(Defendía mucho a mi hermano) <sub>ip</sub> ] <sub>IP</sub> .....	278
Figura 5.14. Oscilograma, espectrograma y F0 del enunciado [(La mezcla cultural de ese lugar) <sub>ip</sub> (lo hacía muy divertido) <sub>ip</sub> ] <sub>IP</sub> grabado por el voluntario venezolano .....	278
Figura 5.15. Oscilograma, espectrograma y F0 del enunciado [(Defendía mucho a mi hermano) <sub>ip</sub> ] <sub>IP</sub> grabado por el voluntario venezolano .....	279
Figura 5.16. Grado de asociación de cada intertonema con la Ciudad de México en cada grupo de jueces .....	299
Figura 5.17. Grado de asociación de cada tonema con la Ciudad de México en cada grupo de jueces .....	301
Figura 5.18. Oscilograma, espectrograma y F0 del enunciado Rs1H2-14: [(Al principio era por relajo) <sub>ip</sub> ] <sub>IP</sub> .....	307
Figura 5.19. Modelo cognitivo básico de la intervención de categorías cognitivas lingüísticas y sociolingüísticas (indexicales) en el proceso de comunicación .....	311
Figura 6.1. Modelo de orden sociocognitivo de introducción de significado social en la prosodia .....	325

## Lista de abreviaturas

*	Símbolo que indica que el tono está anclado a la sílaba tónica
%	Símbolo que señala el tono de frontera de una frase entonativa mayor
-	Símbolo que representa el tono de frontera de una frase entonativa intermedia
dB	Decibelios
F0	Frecuencia Fundamental
H	Tono alto (del inglés <i>high</i> )
Hz	Hercios
ip	Frase entonativa intermedia
IP	Frase entonativa mayor
L	Tono bajo (del inglés <i>low</i> )
ms	Milisegundos
st	Semitonos
SPSS	Statistical Package for the Social Sciences
J-	Jueces participantes en la prueba de percepción cuyo modo de vida corresponderá con el número que reemplace al guion
MdV-	Modo de vida analizado en la prueba de percepción

## Introducción

### A.1. Aspectos generales

La lingüística cognitiva y la sociolingüística, por separado, han aportado información valiosa para la comprensión de la naturaleza y el funcionamiento del lenguaje. Sin embargo, cuando estas disciplinas se combinan, la riqueza de los datos obtenidos se incrementa considerablemente, pues se multiplican las perspectivas de análisis y, por ende, las aristas e implicaciones de los resultados. Esto es posible en el campo de la sociolingüística cognitiva, que “siguiendo las pautas generales de la lingüística cognitiva, se preocupa especialmente por el estudio de los recursos cognitivos implicados en el procesamiento y el uso lingüístico contextualizado” (Moreno Fernández, 2012:17). En otras palabras, esta disciplina toma conceptos fundamentales de la lingüística cognitiva (como las nociones de percepción, esquema, categoría y prototipo) y, a través del estudio de las variedades lingüísticas, las aplica en el contexto social en el que son empleadas. Esta idea se ve reflejada en la tesis basada en el uso, pues se parte de la idea de que se debe estudiar el lenguaje tal y como lo utilizan los hablantes (Kristiansen y Dirven, 2008), lo que permite ver las rutinas mentales que han desarrollado a través del reconocimiento de conductas lingüísticas, durante su interacción social con su entorno (Evans, 2012). Esto implica que la frecuencia con la que un individuo emplea una determinada construcción lingüística es el reflejo del proceso cognitivo de percepción y almacenamiento de la realidad social en la que se encuentra inmerso.

La presente investigación parte de los postulados de la sociolingüística cognitiva, puesto que se propone conjugar elementos propios de las teorías cognitivistas y sociales del lenguaje. De la sociolingüística, se toman las nociones de variación lingüística, redes sociales e indexación de significado social (§1.3) y de la lingüística cognitiva las ideas de esquemas, categorías, prototipos y lengua en uso (§1.4), todas ellas con base en la percepción del mundo por parte de los hablantes, pues los individuos aprenden su lengua materna, construyen su estilo y adoptan los elementos necesarios para la construcción de su identidad social a través de la percepción. Con estas herramientas, se obtendrá información acerca de posibles variaciones sociocognitivas, presentes en el almacenamiento de estructuras lingüísticas portadoras de significado social, por parte de distintos grupos de hablantes de una sociedad.

Esta información permite alcanzar el objetivo del presente trabajo: determinar cuáles son los elementos prosódicos que indexan significado sociolingüístico en los enunciados aseverativos producidos por los hablantes de dos redes sociales de la Ciudad de México. Para ello, se creó un corpus integrado por 12 entrevistas sociolingüísticas a hablantes agrupados de acuerdo con las características sociales correspondientes a dos modos de vida, que podrían considerarse opuestos: uno de ellos propio de trabajadores independientes (comerciantes informales) y trabajadores del escalafón más bajo de la jerarquía laboral (en pequeños establecimientos comerciales o en oficinas gubernamentales), mientras que el otro se corresponde con empleados ejecutivos de una empresa multinacional. De las grabaciones, se extrajeron los datos acústicos necesarios para determinar los patrones prosódicos de manera individual y establecer estadísticamente los patrones prototípicos (los más frecuentes) y periféricos (los menos frecuentes) de cada red social. Posteriormente, se llevó a cabo una prueba de percepción, para determinar la tasa de reconocimiento de los patrones prosódicos e identificar los rasgos prosódicos que indexan significado sociolingüístico. Finalmente, se contrastaron los datos cuantitativos y cualitativos obtenidos, tanto en la etapa de producción como en la de percepción, para someterlos a los análisis que permitieron llegar a las conclusiones de este trabajo.

Las páginas que se presentan a continuación, dentro de este mismo apartado, contienen las siguientes secciones del proyecto: i) planteamiento del problema, en el cual se describe el interés que motivó la investigación; ii) justificación, donde se expresan las razones que respaldan la realización de este trabajo; iii) los objetivos del proyecto; iv) las preguntas de investigación que se pretenden responder; y v) las hipótesis de partida de la investigación.

## A.2. Planteamiento del problema

A pesar de que la lingüística cognitiva dio sus primeros pasos hace más de tres décadas (Lakoff y Johnson, 1980; Langacker, 1986; Lakoff, 1987), los trabajos en el área de la fonología llevados a cabo desde este enfoque no son tan abundantes como podría esperarse; considerando el número de investigadores de la lingüística cognitiva, los que abordan el tema de la cognición fonológica son pocos (entre los que destaca Bybee, 2001), incluyendo los lingüistas cognitivos que estudian activamente la prosodia (Van der Hulst, 2003; Face, 2006;

Neset, 2008, Pinillos, 2016; entre otros).<sup>1</sup> Adicionalmente, la mayoría de los estudios sobre lingüística cognitiva que se han podido consultar hasta este momento se enfocan, principalmente, en la cognición del lenguaje y dejando de lado su componente social, a pesar de que, dentro del marco de la lingüística cognitiva, se reconoce la importancia de los factores sociales que permiten observar de cerca cómo el contexto de un hablante puede condicionar su producción lingüística y, en mayor profundidad, sus estructuras cognitivas (Geeraerts y Cuyckens, 2007; Kristiansen y Dirven, 2008). No obstante, existen algunos lingüistas cognitivos, como Kristiansen (2006), que abogan por una mayor inclusión del componente social basado en el uso, en los trabajos cognitivos.

En la sociolingüística se da la misma situación, aunque en sentido inverso. No son muchos los estudios sociolingüísticos en los que se aborda el componente cognitivo del lenguaje, posiblemente, debido a que las bases de la sociolingüística centran el análisis en el lenguaje empleado en los contextos lingüístico y social sin tomar en cuenta el aspecto psicológico de los hablantes, en términos de procesamiento y almacenamiento del material lingüístico. Sin embargo, en el ámbito hispanico, los sociolingüistas Moreno Fernández (2010) y Martín Butragueño (2015a) ya han hecho avances en el campo de la fonología basada en el uso, que relacionan la sociolingüística con algunos elementos de la lingüística cognitiva. No obstante, no ha sido posible encontrar trabajos que traten el tema de la indexación de significado social desde la perspectiva planteada por estos autores, pues en los trabajos consultados en los que se estudia la indexación de significado social, como los desarrollados por Eckert (1988, 2008), Zhang (2005), Campbell-Kibler (2007), por ejemplo, no se plantea el análisis de la organización de las variantes estudiadas dentro de la categoría cognitiva a la que pertenecen.

De igual manera, a pesar de que las investigaciones sociofonéticas han probado que en algunos grupos de hablantes existen patrones prosódicos más frecuentes que en otros dentro de una misma comunidad de habla (cfr. Martín Butragueño, 2004; Enbe y Tobin, 2008; Orozco, 2016; Rodríguez, 2016; Martín Butragueño y Mendoza, 2018, entre otros); y que los trabajos en fonología cognitiva han demostrado la existencia de categorías fonológicas y determinado sus ejemplares prototípicos y periféricos (cfr. Samuel, 1982;

---

<sup>1</sup> A pesar de que se incluyen la referencias Face (2006) y Pinillos (2016) como muestras de un trabajo que aborda el estudio de la prosodia desde una perspectiva cognitiva, es necesario señalar que estos autores se especializan en estudios fonológicos, no en cognitivos.



Lively y Pisoni, 1997; Lotto, 2000), en ninguno de los trabajos consultados, se plantea analizar la manera en que las construcciones prosódicas prototípicas y periféricas de una categoría cognitiva pueden transmitir significado sociolingüístico. Así mismo, aunque existen trabajos centrados en el estudio perceptivo de rasgos fonéticos de la prosodia del español (como Face, 2006; Face, 2007; Dorta *et al.*, 2013b; Díaz y Dorta, 2015, entre otros), no se han podido encontrar trabajos en los que se estudie, en concreto, la indexación de significado sociolingüístico en los elementos prosódicos desde un enfoque perceptivo.

Aparte del trabajo de Pinillos (2016), enfocado en el plano segmental, no fue posible obtener datos relacionados con investigaciones sociofonéticas, planteadas desde los principios de la sociolingüística cognitiva señalados por Moreno Fernández (2012). En el resto de los trabajos consultados, no se observa una correlación entre el plano social y el cognitivo en el planteamiento metodológico. Dichas investigaciones se centran en la descripción de las distintas variantes de una variable, mas no abordan un tema relevante en el análisis del lenguaje humano como un fenómeno cognitivo y social: la correspondencia entre el análisis sociolingüístico de los datos de producción de habla con la realidad perceptiva y, por ende, cognitiva de los hablantes. Desde esta perspectiva, se considera necesario contrastar las variantes lingüísticas estadísticamente significativas con la percepción de los hablantes de una comunidad de habla, para determinar si, en efecto, el análisis estadístico revela rasgos verdaderamente vinculables a grupos sociales o si, por el contrario, los elementos estadísticamente significativos no son portadores de significado social.

El aspecto social de la comunidad de habla de la Ciudad de México también representa un caso de interés, debido a la complejidad de su estructura social. De acuerdo con Martín Butragueño y Lastra (2011), la ciudad tiene características geográficas particularmente llamativas, debido a su gran extensión territorial y al crecimiento de las zonas conurbadas que aumentan las dimensiones del área metropolitana. De igual forma, comentan los autores, la ciudad cuenta con un gran número de migrantes provenientes de distintas regiones del país y del extranjero (Martín Butragueño y Lastra, 2011: x), lo que aumenta su variedad lingüística. Adicionalmente, dentro del territorio de la ciudad, existe un desequilibrio en la distribución de la población económicamente activa, pues aquellos con menor poder adquisitivo se localizan principalmente al sur y al este de la ciudad; mientras

que aquellos con ingresos más altos se encuentran al oeste de la ciudad (Lastra y Martín Butragueño, 2000; Martín Butragueño y Lastra, 2011). Este fuerte desnivel entre la mayoría con menos ingresos (que puede llegar a recibir menos de dos salarios mínimos, de acuerdo con los autores) y la minoría con mayores ingresos puede verse en la alcaldía de Miguel Hidalgo; alcaldía que Lastra y Martín Butragueño (2000) ubican como la cuarta con mayor población económicamente activa y que ingresa más de cinco salarios mínimos. Lo observado por ambos autores en esta alcaldía resulta de gran interés para este trabajo, pues en la referida zona coexisten, en bastante cercanía, las dos redes sociales que forman parte de este estudio.

Por todo lo expuesto en este apartado, se considera que existe un vacío informativo en cuanto a la posible existencia de variación sociolingüística relacionada con la constitución de las categorías cognitivas que contienen información prosódica. En concreto, no se cuenta con suficiente conocimiento acerca de cuáles pueden ser los patrones prosódicos prototípicos y periféricos de estas categorías, ni de cuáles son las pistas prosódicas (índices sociolingüísticos) empleadas por los individuos de una comunidad de habla, para identificar determinadas características sociales de sus interlocutores.

### A.3. Justificación

Con esta investigación, se pretende hacer un aporte al vacío informativo existente, mencionado en el apartado anterior; pues se manejará una metodología mixta que permita definir patrones prosódicos, delimitar su estatus en una categoría cognitiva (prototípicos o periféricos) e identificar los elementos prosódicos que indexan significado sociolingüístico (en términos de Eckert 2008). Para ello, se empleará el enfoque lingüístico basado en el uso (Bybee, 2001; Kristiansen, 2006; Moreno Fernández, 2010; y Martín Butragueño, 2015a), aplicado al aspecto prosódico del lenguaje. Por lo tanto, este trabajo se enmarca en el ámbito de la prosodia basada en el uso (Martín Butragueño y Velásquez, 2014), un área poco explorada desde el enfoque basado en el uso.

Esta investigación busca complementar la información disponible respecto a los rasgos prosódicos del español de Ciudad de México y creará una propuesta para definir el estatus cognitivo (prototípico o periférico) de construcciones prosódicas pertenecientes a una categoría cognitiva concreta (en este caso, enunciados aseverativos). Los datos recolectados

aportarán una nueva perspectiva, sobre la forma en la que la comunidad de habla de la Ciudad de México almacena la información lingüística en su sistema cognitivo. Esto se hará mediante un análisis fundamentado en la teoría basada en el uso (Bybee, 2011), el cual permitirá observar la disposición de distintos esquemas prosódicos, dentro de las estructuras cognitivas de dos redes sociales que forman parte de esta comunidad. En síntesis, se plantea un trabajo con un enfoque sociolingüístico, en un ámbito de categorías cognitivas lingüísticas, para aportar nuevos datos de acuerdo con la metodología de la sociolingüística cognitiva.

Debido a que ya se cuenta con trabajos que documentan las características prosódicas de los hablantes de la Ciudad de México (Matluck, 1952; Martín Butragueño, 2001, 2004, 2006a, 2006b, 2014, 2015a, 2015b, 2021; Sagastuy y Fernández Planas, 2014; entre otros) esta investigación se propone ir más allá, al ahondar en el campo poco explorado de la percepción. Posterior al análisis sociolingüístico de los datos obtenidos de dos redes sociales, que establecerá cuáles son los rasgos prosódicos con potencial carga de significado social, los índices sociolingüísticos potenciales se someterán a una prueba de percepción que determinará su valor indexical. Con esta prueba, se pueden sentar las bases metodológicas para proponer un modelo de estudio perceptivo que permita validar el significado social de los elementos lingüísticos relevantes, según el análisis sociolingüístico; y que posibilite determinar la medida en la que dicho significado está asociado con esquemas cognitivos prototípicos y periféricos de la categoría cognitiva correspondiente. Todo el proceso se plantea dentro del marco de la lingüística basada en el uso, un campo poco explorado en las sociedades de habla hispana (Dąbrowska, 2008; Dugua *et al.*, 2009; Chevrot *et al.*, 2009; Martín Butragueño, 2015a; Zeschel y Proske, 2015, entre otros). La falta de exploración hace necesario un estudio que contraste los resultados del análisis sociolingüístico (obtenidos mediante herramientas estadísticas) con la percepción de los hablantes de una comunidad de habla, como es el caso de este trabajo, puesto que, como explica Irvine (2001), los índices funcionan como signos relacionados con el mundo social de los hablantes; y estos requieren intérpretes que establezcan su significado dentro de ese mundo social, es decir, deben ser entendidos por personas que identifiquen las implicaciones sociales que conllevan los signos (índices).

Como se mencionó en §A.2, en la Ciudad de México existe una complejidad social que no solo favorece la contribución de los trabajos sociolingüísticos, sino que los vuelve necesarios para tener una mejor comprensión de las formas en las que las dinámicas sociales se ven reflejadas en el habla. En la presente investigación, se observarán en detalle las diferencias lingüísticas, concretamente prosódicas, de dos grupos con modos de vida considerablemente diferentes que coexisten en espacios bastante cercanos. Con esto, la gama de contrastes observables permitirá un mejor entendimiento de cómo la realidad social de estos hablantes de la Ciudad de México se manifiesta en términos sociolingüísticos y, a través del enfoque cognitivo basado en el uso, observar el grado en el que la exposición a esta realidad social configura la organización de los esquemas prosódicos pertenecientes a la categoría cognitiva de enunciados aseverativos.

#### A.5. Objetivos

Todos los procesos metodológicos están orientados al cumplimiento de los objetivos que se presentan a continuación.

##### A.5.1. General

Determinar los elementos prosódicos que indexan significado sociolingüístico en enunciados aseverativos, en dos redes sociales pertenecientes a la Ciudad de México.

##### A.5.2. Específicos

- a. Establecer los parámetros prosódicos (duración, intensidad, frecuencia fundamental, acentos tonales y tonos de juntura) de los enunciados aseverativos, en dos redes sociales pertenecientes a la ciudad de México.
- b. Definir los rasgos prosódicos prototípicos y periféricos de los enunciados aseverativos en cada red social, con base en el enfoque de Prosodia basada en el uso (Martín Butragueño y Velásquez Upegui, 2014; Martín Butragueño, 2015a).
- c. Identificar cuáles son los rasgos prosódicos (prototípicos o periféricos) que pueden ser índices sociolingüísticos.
- d. Determinar si los índices sociolingüísticos están constituidos por un solo rasgo prosódico o por varios.

- e. Confirmar, a través de un estudio perceptivo, la condición de índices sociolingüísticos de aquellos rasgos prosódicos identificados como índices potenciales.
- f. Establecer el grado de centralidad (prototipicidad) cognitiva de los rasgos prosódicos portadores de significado sociolingüístico.

#### A.6. Preguntas de investigación

- a. ¿Cuáles son los parámetros prosódicos de los enunciados aseverativos producidos por los hablantes de cada red social?
- b. ¿Se puede aplicar en el estudio prosódico el mismo enfoque basado en el uso que se ha empleado para la delimitación de elementos categoriales prototípicos y periféricos en el plano segmental?
- c. ¿Cuáles son los rasgos prosódicos prototípicos y periféricos de cada una de las redes sociales estudiadas?
- d. ¿Los índices sociolingüísticos están conformados por un solo rasgo prosódico o por la combinación de varios?
- e. ¿Los rasgos prosódicos identificados como índices sociolingüísticos en el análisis estadístico también son considerados como tales por los jueces en la prueba de percepción?
- f. ¿Son los rasgos prosódicos prototípicos o los periféricos de cada red social los que indexan significado sociolingüístico?

#### A.7. Hipótesis

Esta investigación parte de siete hipótesis concatenadas.

- a. Los parámetros prosódicos de los enunciados aseverativos de cada red corresponden con los previamente identificados por otros autores como propios de la comunidad de habla de la Ciudad de México, aunque hay diferencias sustanciales entre los parámetros de ambas redes (Matluck, 1952; Kvavik, 1979; Sosa, 1999; Martín Butragueño, 2004, 2006a, 2006b, 2011, 2021).
- b. Es posible aplicar el enfoque basado en el uso en un análisis sociofonético, para determinar elementos prototípicos y periféricos en el habla de dos redes sociales pertenecientes a una misma comunidad de habla.

- c. Los rasgos prosódicos prototípicos corresponden con los identificados previamente como rasgos propios de la Ciudad de México (Matluck, 1952; Kvavik, 1979; Sosa, 1999; Martín Butragueño, 2004, 2006a, 2006b, 2011, 2021), mientras que los periféricos son aquellos que no se alinean con las tendencias determinadas por las investigaciones previas.
- d. Los índices sociolingüísticos están constituidos por las características conjuntas de más de un rasgo prosódico.
- e. Los jueces serán capaces de reconocer el significado sociolingüístico en los rasgos identificados como índices durante el análisis sociolingüístico.
- f. Los rasgos prosódicos prototípicos de cada red son los que indexan el significado social que les permiten a los miembros de la comunidad de habla (en este caso, de la Ciudad de México) identificar a un hablante como perteneciente a un grupo social o a otro.

Para responder las preguntas de investigación y validar las hipótesis, este trabajo se estructura en seis capítulos. El capítulo I contiene el marco teórico, donde se exponen las bases en el ámbito de la fonología, sociolingüística, lingüística cognitiva y sociolingüística cognitiva. En el capítulo II, se detallan los aspectos metodológicos de la fase de producción. El capítulo III consiste en la presentación del análisis de los datos de producción desde una perspectiva fonética y fonológica. En el capítulo IV, se expone el análisis sociolingüístico de los datos obtenidos a través del análisis de los enunciados extraídos del corpus. En el capítulo V, se muestra el procedimiento seguido para diseñar la prueba de percepción, junto con el análisis de los datos obtenidos durante la fase de percepción. Finalmente, se exponen las conclusiones de la investigación.

# Capítulo I

## Marco teórico

En este capítulo se presentan las bases teóricas sobre las que se llevó a cabo la investigación. En primer lugar, se explica qué es la prosodia, cuáles son sus componentes y los modelos de análisis prosódicos relevantes para este estudio. Luego, se presentan aspectos teóricos relacionados con la sociolingüística, como su definición general; las tres olas (o enfoques principales) que se han dado dentro de esta disciplina; lo que se entiende por “variable sociolingüística”, y los conceptos de “prestigio”, “redes sociales”, “modos de vida”, “estilo” e “indexación”. Posteriormente, se aborda el tema de la lingüística cognitiva; su noción general y los términos “esquemización”, “conceptualización”, “categorización” y “prototipo”. Finalmente, se plantea una definición general de la sociolingüística cognitiva; de la teoría basada en el uso (cuya base se encuentra en la lingüística cognitiva); de los conceptos de “esquema”, “prototipo” y “categoría”, planteados desde este enfoque; de la sociofonología cognitiva y se expone brevemente el papel de la percepción dentro de esta área de la sociolingüística.

### 1.1. Conceptos teóricos para el análisis prosódico

La prosodia ha sido estudiada como parte del sistema lingüístico desde el tiempo de la antigua cultura griega (Mora y Asuaje, 2009), aunque el concepto que se tenía de ella en ese entonces ha variado considerablemente hasta llegar al punto en el que se encuentra hoy. La prosodia ha logrado asentarse dentro de la lingüística como un elemento fonológico con carácter autónomo, pues, a través de ella, se pueden expresar significados semánticos y una amplia variedad de sentidos pragmáticos que no están condicionados por el material segmental del enunciado (por ejemplo, la prosodia es el elemento que otorga significado aseverativo o interrogativo a dos enunciados segmentalmente idénticos y, de igual modo, puede plantear una afirmación o una duda sin necesidad de articular cualquier fono consonántico o vocálico). En este apartado, se presenta un concepto general de la prosodia junto con la descripción de algunos de sus elementos constitutivos (los más relevantes para esta investigación), la jerarquía de los distintos niveles prosódicos que han sido postulados y una descripción de los

dos modelos de análisis prosódico que sirven como referencia para esta investigación (Aix-en-Provence y el modelo Métrico-Autosegmental).

### 1.1.1. Prosodia: definición general

Mora y Asuaje (2009) explican que los estudios sobre la prosodia se originaron en la antigua Grecia, donde el término se acuñó a partir de la preposición *pros* (προς), que significa ‘hacia’, y *ôidê* (ὀδῆ), que significa ‘canto’, por lo que “prosodia” significa ‘hacia el canto’ o ‘acorde al canto’. Estas autoras también señalan que la prosodia, a su vez, constituía una parte de la gramática del lenguaje y consistía en el estudio de los acentos, la ortografía y la duración silábica.

El interés en el estudio de la prosodia no es un fenómeno reciente, pues ha gozado de atención a lo largo de la historia. En el siglo XVIII, Álvarez (1785:3) señaló lo siguiente:

Prosodia, según su origen griego, es lo mismo que *cantus*, ò *accentus notitia*; por lo que puede definirse: *Faculas* [sic] *agens de accentu*: una Facultad que trata universalmente de los acentos, ò según el P. Ricciolo: *Ratio dimètiendi syllabarum quantitatem*. Arte, ò Facultad, que enseña la cantidad de las sílabas. Pero uniendo las dos explicaciones, podrase decir mejor, que la Prosodia es: *Facultas agens de accentu*, y *quantitate syllabarum*.<sup>2</sup>La cantidad de la silaba, es la longitud, ò brevedad con que se pronuncia: de suerte que, quando la sílaba es larga, se dice, que tiene cantidad larga, quando es breve, que tiene cantidad breve [...] El accento es el tono, con que se levanta, ò baxa la voz en la pronunciación de las sílabas.<sup>3</sup>

Según se puede ver en el texto de Álvarez (1785), desde finales del siglo XVIII ya se tenía la idea de que el tono y la duración silábica formaban parte de la prosodia. Con una perspectiva más actual, Mora y Asuaje (2009) afirman que la prosodia engloba todo lo relativo a la musicalidad y métrica del lenguaje, lo que permite la organización del habla en una serie de modulaciones musicales en tiempos específicos. La prosodia, indican las autoras, está constituida por la entonación, la acentuación y el ritmo, que en términos físicos se manifiestan como la variación en la frecuencia fundamental (F0), la intensidad y la duración de los segmentos. Ellas definen la prosodia como “la música de la expresión oral, [que] contribuye en la organización de las palabras que conforman enunciados, determina el sentido de las palabras enmarcadas en diferentes contextos lingüísticos y situacionales” (p. 23).

---

<sup>2</sup> Facultad que trata del acento y la cantidad de sílabas.

<sup>3</sup> El uso de cursivas en la cita de Álvarez (1785) corresponden con el texto original del autor.



Navarro Tomás (1957), quien utiliza el término *acento* para referirse a la prosodia, afirma que este es la combinación de tono, timbre, cantidad (duración silábica) e intensidad, el cual difiere de una lengua a otra. Para este autor, el tono es la altura musical de un sonido, el timbre corresponde a la vibración que generan los sonidos y consta de una vibración principal y de otra o más vibraciones secundarias (amplificadas por los órganos resonadores), mientras que la cantidad consiste en la duración del sonido y la intensidad responde al grado de fuerza espiratoria con que se pronuncia un sonido (manifestada acústicamente por la amplitud de las vibraciones). Adicionalmente, Navarro Tomás (1957) señala que la diferencia entre los valores de estos elementos no solo es perceptible en cada lengua, sino también en el habla de individuos de distintas localidades de un país que emplean un mismo idioma.

Matluck (1965), por su parte, define prosodia como el conjunto de rasgos suprasegmentales que acompañan a las vocales y a las consonantes; y que se sobreponen a ellas. Este autor reconoce siete rasgos suprasegmentales: i) tono, definido como la altura musical de la vocal que se puede medir en ciclos por segundos;<sup>4</sup> ii) intensidad, entendida como la fuerza de emisión de la vocal medida en decibeles (dB); iii) transición final,<sup>5</sup> referente al ascenso o descenso brusco del tono inmediatamente anterior a la pausa; iv) transición abierta, definida como la pausa pequeña (que puede ser real o aparente) entre palabras dentro del grupo fónico cuya función es separar las palabras; v) cantidad silábica, conceptualizada como la duración de la sílaba en términos de tiempo desde su intención hasta su distensión;<sup>6</sup> vi) ritmo, establecido como la relación existente entre la cantidad silábica y el número de sílabas emitidas en una unidad de tiempo; y vii) entonación, que Matluck (1965), retomando las palabras de Navarro Tomás (1957), define como la línea de altura musical determinada por la serie de sonidos sucesivos que componen una palabra, frase o discurso.

---

<sup>4</sup> La medida de ciclos por segundo también es conocida como Hercios o Hertzios. 1 ciclo por segundo = 1 Hz. (Quilis 1981).

<sup>5</sup> A lo que Matluck (1965) llama “transición final”, Quilis (1981) lo llama “tono de juntura final”, y Sosa (1999) lo llama “entonema” o “tonema”. Sosa (1999) define este término desde una perspectiva ajustada a los hallazgos de los estudios acústicos, como “el conjunto de tonos que marcan el final de un enunciado y que coincide con las sílabas finales a partir de la que lleva el último acento” (p. 31).

<sup>6</sup> Esta unidad se mide generalmente en milisegundos (ms).

Matluck (1965) afirma que prevalece la tendencia de considerar el tono, la intensidad y la transición como únicos elementos de la entonación. Este autor comenta que, si bien estos tres rasgos son los más importantes de la entonación, no se pueden dejar de lado el ritmo y la cantidad, ya que “la entonación incluye más específicamente los tres rasgos mayores, pero, en términos generales, son inseparables de ella todos los rasgos suprasegmentales” (p. 9). Esta visión concuerda con la de Mora (1996, 1998) y Mora *et al.* (1997), quienes afirman que, si bien es cierto que el movimiento melódico es el rasgo principal de la entonación, no se puede obviar el rol que desempeñan el acento y el ritmo. Por tal motivo, en estos tres trabajos se establece una diferencia conceptual entre entonación y melodía; y se plantea que la segunda es parte constituyente de la primera, junto con el ritmo, la intensidad y las pausas. Tal distinción implica que los autores, al menos en estos trabajos, emplean el término “entonación” de forma análoga a lo que se entiende por “prosodia” en este trabajo, pues su definición no es muy distinta de la de Ladd (2008), que se presenta unos párrafos más adelante.

Según Quilis (1981), la entonación (como elemento constituyente de la prosodia) cuenta con una sustancia y una forma. La sustancia, explica el autor, es un *continuum* sobre el que se delimita la entonación en unidades discretas, para poder establecer sus patrones melódicos y la naturaleza de sus elementos. La forma, por otra parte, es la descripción estructural de la entonación, que viene dada por la composición lingüística del número de elementos que la integran, sus relaciones y sus funciones. Quilis (1981) señala que los diferentes niveles de análisis de la prosodia son difíciles de abordar, debido a que, debajo del nivel estructural donde se da una serie de oposiciones limitadas, hay un valor simbólico susceptible de ser dicotómico. Dicho valor simbólico, según este autor, provoca reacciones diferentes en el oyente, por lo que posee carga significativa, y por debajo de este nivel está toda la sustancia (los datos) que solo se puede elaborar en función del nivel estructural. Adicionalmente, de acuerdo con D’Introno *et al.* (2010), dicho nivel posee una naturaleza contrastiva, pues la diferenciación entre estos rasgos prosódicos no radica en la distinción entre los elementos alternantes (como sucede en el caso de los segmentos), sino que se puede percibir entre hechos concurrentes y sucesivos en el tiempo. D’Introno *et al.* (2010) explican que la cualidad acústica de un segmento que posee un determinado rasgo prosódico “consiste en el valor peculiar que alcanza un cierto parámetro con respecto a los valores con que se

presenta el mismo parámetro en otros segmentos de la misma secuencia” (p. 124); por ejemplo, el carácter tónico de una sílaba no se determina por los rasgos acústicos de dicha sílaba aislada, sino por el contraste de la tonicidad de esa sílaba con aquellas de su entorno.

De todos los conceptos presentados hasta este punto, la definición de Ladd (2008) es la que se alinea más con la perspectiva manejada como base en la presente investigación, pues este autor emplea el término “prosodia” para referirse al uso de los rasgos fonéticos suprasegmentales. Dichos rasgos son la frecuencia fundamental (F0), intensidad y duración. Debido a las diferentes definiciones del término “prosodia”, Ladd (2008) establece una distinción entre las propiedades físicas, psicológicas y fonéticas de los rasgos suprasegmentales. Para este autor, el acento es una cualidad fonética (una amalgama perceptual indirectamente relacionable con las dimensiones psicológica y física), el volumen es una cualidad psicológica y la intensidad una física. De igual forma, la F0 es una propiedad física del habla, cuyo correlato psicológico es el tono. Sin embargo, Ladd (2008) reconoce que el uso indistinto de estos términos genera poca ambigüedad fuera del contexto de la psicofísica.

Una vez entendido el concepto de prosodia, no se puede pasar por alto que las cualidades de sus rasgos, junto con sus funciones, varían de una lengua a otra. Hirst (2013) explica tal fenómeno como la “paradoja prosódica”; que, de acuerdo con el autor, consiste en entender la prosodia como un elemento lingüístico universal, debido a que todos los lenguajes lo poseen y, simultáneamente, como un elemento específico de cada lengua y cada dialecto, ya que cada lengua tiene características prosódicas específicas, en un sentido general; y cada dialecto de la misma lengua posee rasgos prosódicos que lo distinguen de los demás. Navarro Tomás (1935) provee ejemplos de este hecho al contrastar las diferencias entre la prosodia, o “acentos” (según la terminología empleada por el autor), de varias lenguas (español, francés, italiano, alemán e inglés) y de varios dialectos del español (andaluz, gallego, catalán, caraqueño, bonaerense, entre otros).

### 1.1.2 Elementos prosódicos

De acuerdo con lo presentado en el apartado anterior, no existe una única visión en relación con los elementos que integran la prosodia. Como se pudo ver, diferentes autores pueden ofrecer perspectivas distintas y emplear diferentes términos para hacer referencia a un mismo

elemento. No obstante, la mayoría de los autores mencionados hasta este punto consideran que la prosodia está integrada por, al menos, la entonación, la intensidad y la duración. En este apartado se describen cada una de ellas con especial atención a su función y características en la lengua española.

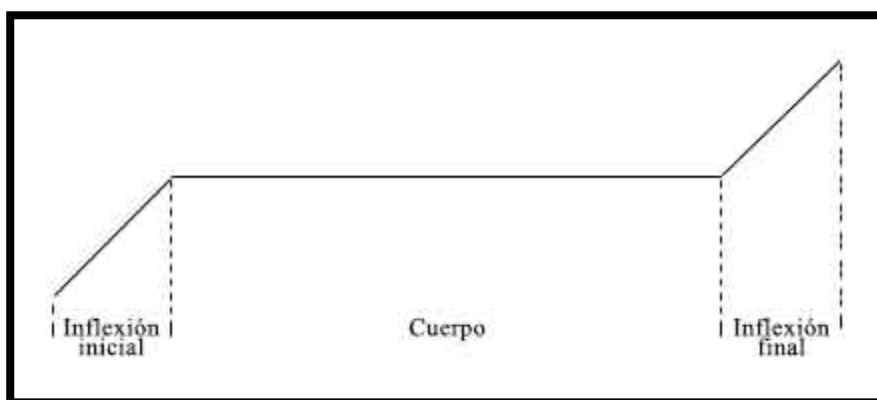
#### 1.1.2.1. Entonación

Si bien se observa una falta de consenso en cuanto a lo que se entiende por el término “entonación”, como se observó en §1.1.1, es necesario aclarar que la noción de “entonación” manejada en este trabajo difiere de aquellas definiciones que comparten la perspectiva retomada por la Real Academia Española (2011: 435), según la cual la entonación es “el movimiento melódico con el que se pronuncian los enunciados y [que] fonéticamente constituye la suma de un conjunto de variaciones en el tono, duración e intensidad”. En esta investigación, se entiende dicho término de acuerdo con los planteamientos de D’Introno *et al.* (2010), quienes no consideran que la entonación está constituida, en parte, por la duración y la intensidad, sino que la definen como la secuencia de tonos de las sílabas de una oración, los cuales corresponden con la frecuencia fundamental (también llamada altura musical) de la vocal que constituye el núcleo silábico. Esta frecuencia fundamental (o F0) corresponde con la vibración de las cuerdas vocales al momento de la fonación, la cual se ve posteriormente multiplicada por los distintos resonadores hasta generar la onda acústica compleja que percibimos en el habla (Quilis, 1981).

Navarro Tomás (1957 y 1966) divide la entonación de un enunciado en español en tres secciones, las cuales (de una u otra forma, según el enfoque de cada estudio) se siguen considerando como las secciones básicas y esenciales de un grupo melódico: inflexión inicial, cuerpo e inflexión final de la unidad (ejemplificadas en la figura 1.1). De acuerdo con Navarro Tomás (1966), las sílabas átonas al inicio del enunciado son producidas con tonos más graves que el de la primera sílaba tónica, a menos que la primera sílaba del enunciado sea tónica, en cuyo caso la entonación inicial se encontraría en un nivel medio. La sección que va desde la sílaba que sigue a la primera tónica y hasta la inmediatamente anterior a la última tónica del enunciado constituye el cuerpo de la unidad melódica. En dicha sección, la altura de la voz aumenta en la primera sílaba postónica para luego ajustarse en un nivel relativamente uniforme. En caso de que el enunciado sea producido de manera enfática, los

valores tonales de ambas secciones del enunciado son propensos a ser considerablemente superiores que en los casos no enfáticos. Por último, el final de la unidad, llamado “tonema”,<sup>7</sup> comprende el material lingüístico desde la última sílaba acentuada hasta el final del enunciado.

Figura 1.1. Secciones básicas y esenciales de un grupo melódico (diseño propio con base en Navarro Tomás, 1966)

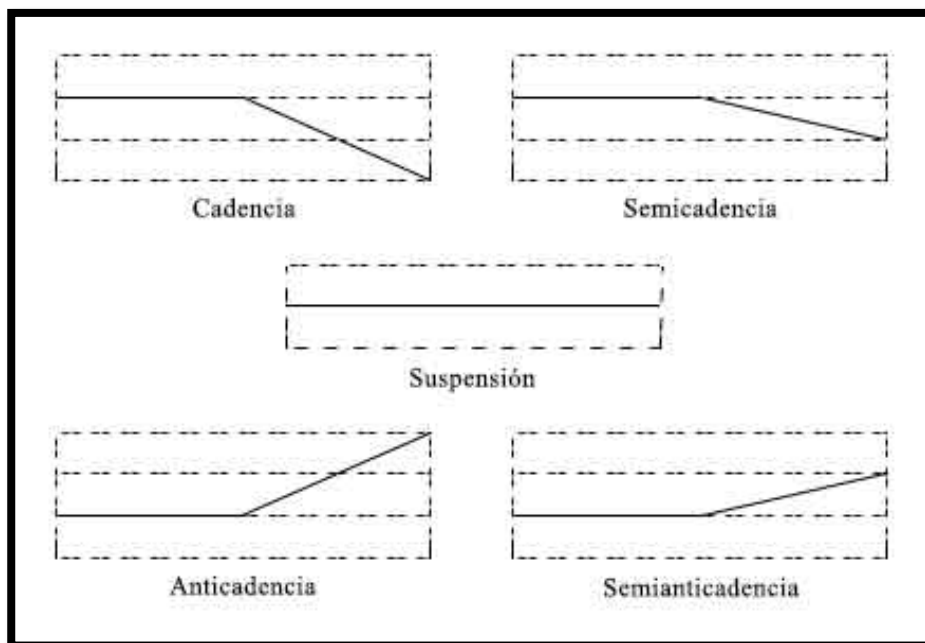


Navarro Tomás (1966) comenta que el desarrollo del tono en el principio y en el cuerpo del enunciado ofrece menos valor significativo, mas no ocurre lo mismo con el tonema, que es la parte más significativa de la unidad enunciativa, ya que sella el carácter fonológico de la unidad en que se halla. El autor propone la existencia de cinco tipos de tonemas en español (ejemplificados en la figura 1.2): i) cadencia, que implica una terminación grave que expresa la terminación absoluta de la frase enunciativa; ii) anticadencia, constituida por una terminación alta, por encima del cuerpo, que constituye el fin de la rama tensiva y es esencial en el relieve predicativo, en el contraste y en la oposición de conceptos; iii) semicadencia, reconocible como una terminación grave, aunque no tanto como la cadencia; y expresa la proposición en serie semántica, la aseveración de forma insegura y una idea definida parcialmente; iv) semianticadencia, compuesta por una terminación menos alta que la anticadencia, se da en unidades interiores, pues posee un sentido continuativo y señala oposiciones y contrastes secundarios; y v) suspensión, que es

<sup>7</sup> El concepto de tonema se abordará nuevamente en §1.1.4.2., donde será explicado con mayor profundidad debido al papel que desempeña en el modelo métrico-autosegmental.

una culminación en el mismo nivel que el cuerpo del enunciado y expresa el corte de una idea pendiente de continuación.

Figura 1.2. Tonemas planteados por Navarro Tomás (diseño propio con base en Navarro Tomás, 1966)



Por otra parte, Quilis (1981) afirma que la entonación en español cumple funciones lingüísticas y expresivas. En el plano lingüístico, el autor señala que la entonación desempeña dos funciones: contrastiva y demarcativa. La función contrastiva consiste en oponer, en primer lugar, los enunciados declarativos (caracterizados por un movimiento descendente) a los interrogativos (caracterizados tanto por un movimiento ascendente en los interrogativos absolutos, como por uno descendente en los pronominales). La función demarcativa o delimitadora puede poseer función distintiva, establecida por la combinación de oraciones explicativas y especificativas, la oposición entre estilo directo e indirecto y preguntas en las que se ofrecen dos alternativas; y una función no distintiva, que no deja de ser imprescindible para la interpretación del mensaje, como en el caso de la pregunta disyuntiva, la enumeración completa a final de frase, la enumeración incompleta a final de frase, la enumeración interna de la frase, el complemento, y en la coordinación. En el nivel expresivo, Quilis (1981) señala que la entonación se comporta de la siguiente manera: tanto en la afirmación enfática como

en la pregunta enfática, reiterativa y en la relativa se observa un ascenso inicial hasta el tono más alto y un posterior descenso hasta el final de la frase; en la pregunta pronominal con matiz de cortesía se evidencia un ascenso inicial seguido de un descenso y un posterior ascenso hasta el final de la frase; en la pregunta confirmativa se documenta un continuo descenso hasta el tono más bajo seguido de un ascenso hasta el final de la frase; en la pregunta imperativa se maneja únicamente un nivel alto; y en la exclamación se registra un descenso acusado desde la primera sílaba acentuada.

Antes de culminar esta sección, se considera necesario hacer un paréntesis para explicar la forma en que la frecuencia fundamental (F0) puede verse afectada por ciertos fenómenos fisiológicos. Debido a que la F0 se genera como producto de la vibración de las cuerdas vocales, esta puede sufrir alteraciones debido a la influencia de dos tipos de fonación no modal que alteran la onda acústica: voz laringizada (o *creaky*) y voz susurrada (o *breathy*). De acuerdo con Stevens (2000), el efecto de laringización se produce cuando la totalidad de los pliegues vocales, excepto en la sección de los bordes superiores, son presionados fuertemente el uno contra el otro durante la fase estática de la vibración, lo que reduce el flujo de aire entre ellas. En la voz susurrada, explica el autor, ocurre el fenómeno contrario, pues durante este tipo de fonación se mantiene una leve apertura por la cual pasa el aire entre los pliegues vocálicos durante la fase de relajación, lo que causa transiciones más suaves entre la apertura y el cierre de las cuerdas vocales. Según Stevens (2000), ambos tipos de fonación causan que la frecuencia descienda durante su ejecución, lo cual se puede percibir como descensos tonales.

#### 1.1.2.2. Intensidad

El parámetro de intensidad hace referencia a la energía con la que se ponen en movimiento las partículas de aire al momento en que se produce un sonido; y está directamente relacionada con la amplitud de la onda de dicho sonido (Quilis, 1981). Esta cualidad acústica del habla ha sido relacionada con los acentos léxicos en el discurso. Navarro Tomás (1957) afirma que el acento de intensidad (o dinámico) es el rasgo que influye más que ningún otro en la estructura prosódica, ya que el cambio de la intensidad en la única sílaba tónica de la palabra puede modificar su acentuación y alterar su significado. Él señala que las sílabas tónicas poseen un mayor grado de intensidad, mientras que las átonicas tienen uno mucho

menor. Navarro Tomás (1966) establece una relación entre el tono y la intensidad, ya que, al pronunciar una palabra aislada, el tono se mueve en forma similar a la intensidad. Esto aplicaría, según este autor, incluso cuando ambos elementos se estudian en un contexto sintáctico, pues ambos suben en las sílabas tónicas y descienden en las átonas.

Quilis (1981), quien también relaciona la acentuación con incrementos en la intensidad, aunque no de una forma tan directa como Navarro Tomás (1966), incluye la presencia de un acento enfático o de insistencia. Este acento se puede dar por un énfasis especial o por afectación del hablante y consiste en la introducción de un segundo acento de intensidad en la palabra que modifica una sílaba átona en tónica. Este autor pone como ejemplo la frase *bajo mi responsabilidad*, que transcribe “/baxo mi rēspoNsabilidadD/” (p. 319), en la cual se enfatiza “responsabilidad” al introducir un acento nuevo en la primera sílaba.

Desde otras perspectivas, se ha considerado que la intensidad no está directamente relacionada con la acentuación. D'Introno *et al.* (2010) afirman que la creencia de que la tonicidad silábica implica un aumento en la intensidad no es cierta: “los análisis electroacústicos demuestran claramente que los picos de intensidad no coinciden necesariamente con las sílabas que se perciben como tónicas.” (p.127). Estos autores explican que las sílabas tónicas suelen implicar la producción de más energía, pero dicho aumento también puede ocurrir en sílabas átonas, del mismo modo que también puede ocurrir que una sílaba tónica se emita con menos energía si tiene una sílaba átona contigua.

### 1.1.2.3. Duración

La duración hace referencia al tiempo que transcurre durante la producción de un sonido de habla. De acuerdo con Malmberg (1977) y Obediente (2005), la duración silábica no desempeña una función diferenciadora de significado en el español, sino que cumple una función expresiva. Esto se debe a que las vocales tienen duraciones similares en el habla corriente, es decir, en el habla no afectada o enfática (Navarro Tomás, 1957). Sin embargo, como afirmó Andrés Bello (1847) hace poco menos de dos siglos, no todas las vocales tienen la misma duración en español, pues las vocales tónicas son más largas que las átonas. Cabe señalar que, tal como aclaran Navarro Tomás (1957) y Malmberg (1977), en nuestra lengua no hay vocales que se puedan denominar propiamente largas, en comparación con lo que



ocurre en otras lenguas. Navarro Tomás (1957) explica: “la denominación de largas que se aplica a estas vocales en determinados casos, sólo representa cierto grado de superioridad relativa con respecto a los diversos matices de duración que el uso distingue dentro de la brevedad general de todas ellas” (p. 200).

La diferencia relativa entre las vocales largas y cortas en español, presentada por Navarro Tomás (1957), se ha visto confirmada en trabajos más recientes, como Mora (1996), Mora *et al.* (2009) y D’Introno *et al.* (2010). Todos estos autores afirman que los estudios acústicos han evidenciado la existencia de una diferencia entre la duración de las sílabas tónicas y las átonas, siendo las primeras más largas que las segundas. Sin embargo, la duración de las sílabas finales de un enunciado no necesariamente se adecua a esto, ya que, en el núcleo de un enunciado (desde la última sílaba tónica hasta la final), la duración tiende a extenderse tanto en la última sílaba tónica como en la átona (en caso de que la última palabra no sea aguda), por lo que la mayor duración en esta parte puede ser una pista cuya función es delimitar el dominio prosódico (Dorta, *et al.*, 2013a; Dorta *et al.*, 2013c).

### 1.1.3. Jerarquía prosódica

Al momento de llevar a cabo un estudio prosódico, es necesario tener en cuenta cuáles son los niveles que se han propuesto para el análisis de la prosodia. La jerarquía prosódica plantea distintos elementos, clasificados de acuerdo con su rango en el enunciado. En el caso específico de esta investigación, son especialmente relevantes los dos niveles superiores que se presentan algunos párrafos más adelante.

Las representaciones fonológicas de los elementos suprasegmentales se organizan de acuerdo con una jerarquía de constituyentes, o categorías prosódicas. De acuerdo con Selkirk (1986), los distintos elementos de la jerarquía funcionan como dominios característicos para la aplicación de reglas fonológicas y, aunque no son isomórficos con la estructura sintáctica, se definen en relación con ella, al menos parcialmente. La estructura prosódica participante en la representación fonológica (según Selkirk, 1986) tiene las siguientes propiedades:

- a. Consta de las categorías prosódicas (fonológicas): sílaba, pie, palabra prosódica, frase fonológica, frase entonativa y enunciado fonológico.
- b. Para cualquier categoría prosódica, la oración es analizada exhaustivamente en una secuencia de estas categorías.

- c. Las categorías prosódicas están ordenadas en una jerarquía y las representaciones fonológicas se organizan estrictamente en capas de acuerdo con la jerarquía.
- d. La organización jerárquica de categorías prosódicas genera una agrupación bien formada.

La autora explica que las propiedades antes mencionadas, organizadas jerárquicamente, crearían la estructura mostrada en el cuadro 1.1, que ejemplifica la buena formación en términos prosódicos. Este modelo de jerarquía prosódica ha sido tomado por otros autores, quienes han propuesto algunas modificaciones de acuerdo con sus perspectivas. Prieto (2006), Nespor y Vogel (2007) y D’Introno *et al.* (2010), por ejemplo, plantean un modelo de jerarquía prosódica que consiste en siete unidades: enunciado fonológico ( $U$ ), frase entonativa ( $I$ ), frase fonológica ( $\phi$ ), grupo clítico ( $C$ ), palabra fonológica ( $\omega$ ), pie silábico ( $\Sigma$ ) y sílaba ( $\sigma$ ). Esta jerarquía difiere de la de Selkirk (1986) únicamente en la inclusión del dominio prosódico correspondiente al grupo clítico entre la frase fonológica y la palabra fonológica. Por otra parte, se encuentra la propuesta de Toledo (2008) quien plantea los dominios: frase entonativa mayor (IP), frase entonativa intermedia (ip), frase fonológica ( $\phi$ ), palabra prosódica ( $\omega$ ) y sílabas ( $\sigma$ ). A diferencia de los demás modelos mencionados, hasta el momento, este autor no toma en cuenta el grupo clítico ni el pie silábico y, adicionalmente, propone un cambio en la nomenclatura del enunciado fonológico y frase entonativa a frase entonativa mayor (IP) y frase entonativa intermedia (ip), respectivamente.

Cuadro 1.1. Ejemplificación de los niveles de la jerarquía prosódica  
(diseño propio con base en Nespor y Vogel, 2007)

Enunciado	(_____)													
Frase entonativa	(_____)							(_____)						
Frase fonológica	(_____)				(_____)				(_____)					
Palabra prosódica	(_____)			(_____)		(_____)			(_____)			(_____)		
Pie	(_)	(_____)	(_____)	(_)	(_____)	(_____)	(_)	(_____)	(_____)	(_)	(_____)	(_____)	(_)	(_____)
Sílaba	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)	(_)

Cabe señalar que, en el caso concreto del español, no todos los autores están de acuerdo con el planteamiento de una jerarquía prosódica que contemple la existencia de frases entonativas

intermedias (ip) dentro de una frase entonativa mayor (IP). Sosa (1999:55-56) afirma que, a pesar de que el planteamiento de distintos tipos de grupos melódicos ha sido identificado productivamente en ciertas lenguas, en el caso del español: “no encontramos justificación alguna para postular más de un tipo de grupo melódico, por lo que entendemos que hay un solo nivel o dominio, definido por el contorno entonacional y delimitado por el tonema”. No obstante, autores como Prieto y Roseano (2010b) afirman que las frases entonativas pueden ser subdivididas en español, al igual que ocurre en el inglés (Pierrehumbert, 1980). De igual forma, Hualde (2002) aboga por el reconocimiento de frases entonativas intermedias en español; y sugiere que lo que permite dar diferentes interpretaciones a textos ambiguos es precisamente la posibilidad de modificar las fronteras de estas frases intermedias.

Como se ve en los párrafos anteriores, las propuestas relacionadas con los niveles de la jerarquía prosódica son diversas, lo que hace que el número de dichos niveles y sus características difieran entre los distintos modelos. Es por ello que se considera necesario delimitar los niveles que resultan pertinentes para este trabajo. El presente estudio se enfoca en el análisis de los dos niveles superiores de las propuestas de Prieto (2006), Nespor y Vogel (2007) y D’Introno *et al.* (2010): enunciado fonológico (*U*) y frase entonativa (*I*). Sin embargo, en esta investigación se usará la nomenclatura empleada por Toledo (2008) para referirse a estos mismos niveles: frases entonativas mayores (IP) e intermedias (ip), respectivamente. La preferencia de estos términos no se relaciona con una diferencia conceptual entre las propuestas de los autores mencionados previamente y la de Toledo (2008), pues ambas son equivalentes en lo que respecta a estos dos niveles. Esta decisión se tomó debido a que se considera que así se hace un mayor énfasis en el papel de la entonación como elemento delimitador de los dos niveles prosódicos pertinentes, basándose en las funciones demarcativas de la entonación de Quilis (1981), mencionadas previamente (§1.1.2.1), y la propuesta del código biológico de producción postulada por Gussenhoven (2004), explicada tres párrafos *infra*.

Se maneja la noción de que una frase entonativa mayor (IP) puede estar constituida por una o varias frases intermedias (ip), como ocurre en los siguientes enunciados:

- i) [(*María va a venir luego*)<sub>ip</sub>]<sub>IP</sub>;
- ii) [(*Si tocan la puerta*)<sub>ip</sub> (*aunque te llamen*)<sub>ip</sub> (*no abras*)<sub>ip</sub>]<sub>IP</sub>.

Nótese que una frase intermedia (ip) únicamente puede corresponder con una mayor (IP) cuando ambas tienen la misma extensión (como en el primer ejemplo). Si en un enunciado hay más de una frase intermedia (ip), la extensión de la frase mayor corresponderá con la del conjunto de todas las intermedias (ip) que constituyen el enunciado, como se ve en el segundo ejemplo.

Von Heusinger (2007) indica que existen ciertos criterios fonéticos, sintácticos y semánticos que ayudan a identificar la extensión de una frase entonativa mayor, los cuales coinciden en cierta medida con los planteamientos de Nespor y Vogel (2007) para el enunciado fonológico. Los parámetros delimitadores de una IP mencionados por Von Heusinger (2007) son los siguientes: i) puede estar precedida o seguida por una pausa; ii) contiene la presencia del acento más prominente; iii) el linde está marcado por un tono de juntura y el ajuste tonal también juega un papel importante; iv) el linde puede bloquear algunos fenómenos de juntura; v) los lindes corresponden con los de algunos constituyentes sintácticos; vi) el material lingüístico debe constituir una unidad de sentido o información.

En lo que respecta a las frases entonativas intermedias (ip), Hualde (2003) señala que estas indican un grado de separación menor que el que se da en las frases entonativas mayores (IP). A su vez, Prieto (2006) comenta que estas poseen un acento prominente en la última sílaba tónica y, opcionalmente, un ascenso indicador de continuación en el linde derecho de la frase. Este ascenso tonal corresponde con la elevación empleada por los hablantes para hacer referencia a la no conclusión del enunciado que están formulando, el cual puede ser explicado por el código biológico de producción de Gussenhoven (2004). De acuerdo con este autor, dicho ascenso se opone al descenso tonal al final de una frase entonativa mayor, originado por la pérdida de energía que se da desde el inicio de la producción oral, que denota la finalización de la enunciación. No obstante, Prieto (2006) explica que también es posible que se perciba una interrupción al final de la frase entonativa intermedia (ip) sin que se produzca el mencionado ascenso de continuidad. Por ende, afirma la autora, en español, la marcación tonal del límite derecho de la frase intermedia es opcional. Por su parte, Toledo (2007) señala que la frase intermedia suele producirse entre pausas y que, además, tiene un tono de frontera que puede ser alto (H-) o bajo (L-). Las frases con tonos H- presentan un incremento tonal en la última sílaba de la palabra fonológica final, cuyo alargamiento produce una prolongación de este tono. De igual forma, el autor menciona que las frases

intermedias que culminan con tonos L- indican la separación de esa frase de la siguiente, mientras que aquellas que culminan con tonos H- señalan que la frase tiene cohesión semántica con la siguiente frase.

#### 1.1.4. Modelos de análisis prosódico

A lo largo del estudio de la prosodia, se han creado varios modelos de análisis con el fin de proporcionar una descripción que optimice nuestro conocimiento acerca de este elemento de la fonología. Entre estos modelos se encuentran el análisis por configuraciones (escuela británica), el análisis por niveles (escuela americana), el modelo IPO (escuela holandesa), el modelo Aix-en-Provence y el métrico-autosegmental (Pierrehumbert 1980). A continuación, se presenta una explicación de los dos últimos modelos, pues son los que tienen relación directa con este trabajo.

##### 1.1.4.1. Modelo Aix-en-Provence

El modelo de Aix-en-Provence, desarrollado en el laboratorio *Parole et Language* de la universidad de *Aix-en-Provence*, se enmarca en la línea del generativismo, pues distingue los ámbitos lingüísticos en dos niveles básicos: un nivel abstracto (cognitivo) que está relacionado con la competencia lingüística del hablante nativo y un nivel de representación analítico que refleja la enunciación del material lingüístico (Hirst y Di Cristo, 1998; Hirst *et al.*, 2000; Baqué y Estruch, 2003). Al respecto, Baqué y Estruch (2003) señalan que en el modelo se percibe la prosodia como un elemento autosuficiente dentro del componente fonológico de la lengua, debido a que cumple con el principio de autonomía, es decir, constituye un módulo independiente del sentido lingüístico que se transmite. Por su parte, Hirst *et al.* (2000) afirman que la prosodia es una representación del componente fonológico, debido a que proporciona información necesaria para la pronunciación de un enunciado y para su interpretación sintáctica y semántica; es decir, cumple con el principio de interpretabilidad, según el cual, las representaciones de los niveles intermedios deben ser interpretables desde los dos niveles adyacentes, tanto el más abstracto como el más concreto.

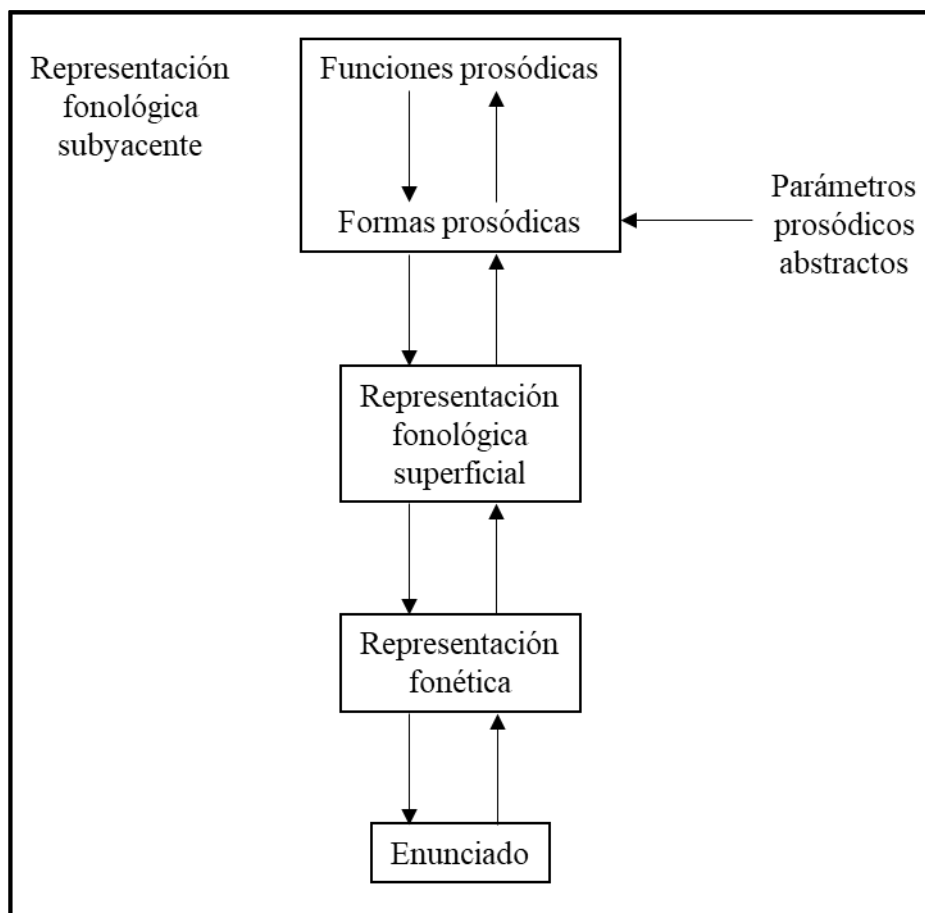
En su explicación del modelo Aix-en-Provence, Baqué y Estruch (2003) resaltan que no se limita a proveer una descripción de los movimientos de la curva melódica, sino que presenta un análisis pluriparamétrico que engloba las variaciones de la frecuencia

fundamental, la intensidad y la duración de los enunciados, para establecer los constructos prosódicos empleados en la codificación de cada una de las funciones lingüísticas presentes, con forma variable, en cada lengua. Por lo tanto, explican las autoras, se consideró necesario ampliar la perspectiva enfocada en dos niveles (abstracto y fonético) a cuatro niveles:

- a. Nivel fonológico profundo o subyacente: representaciones funcionales que proporcionan la información pertinente para la interpretación semántica y sintáctica de la prosodia de un enunciado.
- b. Nivel fonológico superficial: categorías discretas que permiten describir los fenómenos lingüísticos superficiales asociados con aquellos producidos en la oralidad.
- c. Nivel fonético: fenómenos variables a partir de los cuales se manifiestan las condiciones universales de producción y percepción prosódicas.
- d. Nivel físico: características acústicas de un enunciado.

Baqué y Estruch (2003) exponen la representación de la figura 1.3 (tomada de Hirst *et al.*, 2000) para ejemplificar la dinámica existente entre los cuatro niveles, durante la producción y percepción de un enunciado. En este esquema, se muestra cómo las formas prosódicas subyacentes, consistentes en parámetros prosódicos abstractos, se relacionan con las funciones que tienen los distintos patrones prosódicos en una lengua. Estas formas prosódicas subyacentes, a su vez, se asocian con categorías prosódicas discretas en la representación fonológica superficial, que se ponen de manifiesto a través de las representaciones fonéticas que generan las cualidades acústicas de un enunciado. Cabe señalar que el proceso descrito es el que se realiza en la producción de un enunciado, el cual ocurre de forma inversa durante la percepción de un acto de habla.

Figura 1.3. Niveles del módulo prosódico (Tomado de Baqué y Estruch, 2003:125)



Según Baqué y Estruch (2003), el modelo Aix-en-Provence considera la entonación como una serie de segmentos discretos yuxtapuestos a lo largo de un enunciado, en donde la altura tonal de cada uno posee especial relevancia en el análisis. Las autoras explican que se considera que los segmentos tonales discretos son más importantes que un contraste entre distintos tonos de la misma secuencia melódica; y que, de esta manera, se presta atención tanto a los segmentos tonales particulares como a aquellos incluidos en una rejilla métrica que contiene la distribución de acentos. Baqué y Estruch (2003) señalan que dicha metodología de análisis, influida por el estructuralismo europeo, posee un razonamiento deductivo que lleva a definir reglas de reajuste y linearización (que permitan una secuencia de tonos abstractos en un fenómeno pronunciable) y una estructura lineal jerarquizada de unidades prosódicas: i) unidad entonativa (UT), entendida como el segmento tonal comprendido entre dos pausas; ii) unidad rítmica (UR), conceptualizada como una unidad

intermedia consistente en un único acento léxico y una o más unidades tonales y iii) unidad tonal (UT), definida como la agrupación de una sílaba tónica con las átonas que la rodean.

Por otro lado, Hirst *et al.* (2000) establecen que la superposición de dos componentes suprasegmentales es lo que genera la prosodia de un enunciado como tal: el componente microprosódico, relacionado con las variaciones prosódicas originadas por las características propias de los elementos segmentales (alófonos) articulados, y el componente macroprosódico, vinculado con las características de la curva melódica empleada por un hablante durante la pronunciación de un enunciado.

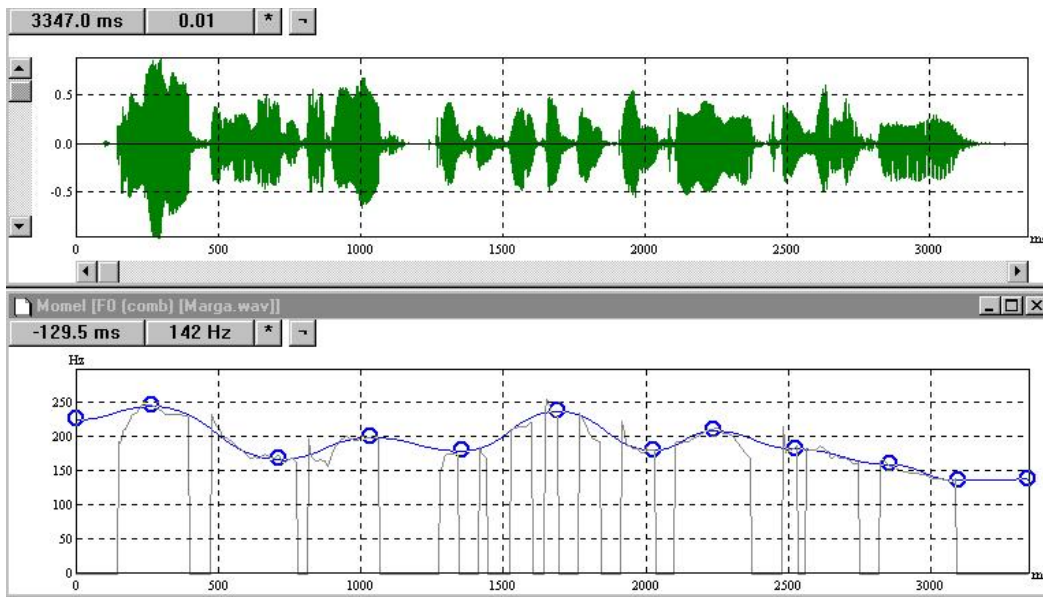
A continuación, se detallan los procedimientos que se siguen en el análisis de cada uno de los niveles contemplados entre los componentes de este modelo:

- a. Nivel físico: este nivel es universal, debido a que se sustenta en la hipótesis de que todas las lenguas presentan las mismas restricciones físicas en la producción y percepción de las variaciones de la frecuencia fundamental (F0), la intensidad y la duración, por lo que en esta fase se extraen los valores acústicos de los elementos suprasegmentales de una señal de audio digitalizada (Baqué y Estruch, 2003; Hirst *et al.*, 2000).
- b. Nivel fonético: se abordan los fenómenos constantemente variables, cuyas restricciones universales en la producción y percepción del sonido han sido calculadas (Hirst *et al.*, 2000). En esta fase se extrae la información macroprosódica en dos etapas: i) la estilización o modelización de la curva melódica a través de la sustitución de los valores de la frecuencia fundamental (F0) por una función numérica que conserva la información macroprosódica original y descarta la información no relevante (en otras palabras, se substituye la curva melódica original por una curva estilizada y simplificada que conserva los valores perceptivamente equivalentes); ii) la codificación simbólica, que convierte la curva estilizada en una secuencia de categorías discretas, que pueden ser utilizadas para regenerar una configuración melódica percibida de forma lingüísticamente equivalente a la original (Baqué y Estruch, 2003). El proceso de estilización se lleva a cabo de forma automática por medio del uso del algoritmo MOMEL (*Modelling Melody*), que une mediante parábolas los puntos de inflexión (*target points*) de la curva de F0, a través de una función cuadrática (*quadratic spline function*) (Hirst *et al.*, 2000). En la sección



inferior de la figura 1.4, se puede observar cómo se empleó el algoritmo MOMEL para llevar a cabo la estilización de una curva melódica a partir de los puntos de inflexión.

Figura 1.4. Oscilograma y curva estilizada a través de algoritmo MOMEL del enunciado *Una vez en el interior, los agentes encontraron a tres niños* (Llisterri, 2018)



- c. Nivel fonológico superficial: es el nivel de las categorías distintivas discretas con las que se pueden describir fenómenos propios de las representaciones fonológicas superficiales de forma interlingüística (Hirst et al., 2000). En este nivel se maneja la hipótesis de que la prosodia de diferentes lenguas se puede describir mediante una misma serie de símbolos (Baqué y Estruch, 2003). Hirst *et al.* (2000) explican que los puntos de inflexión obtenidos con el algoritmo MOMEL se codifican a través del sistema de anotación INTSINT (*International Transcription System for Intonation*), donde cada uno de los puntos es codificado como un tono absoluto (definido en función del rango frecuencial del hablante) o como un tono relativo (determinado como un punto de inflexión intermedio en secuencias ascendentes o descendentes). Estos tonos, a su vez, se distinguen entre iterativos (aquellos intermedios en curvas ascendentes o descendentes) y no iterativos (que no corresponden con tonos absolutos)

y que no constituyen puntos intermedios en los movimientos tonales). Los tonos se codifican de la siguiente manera (Hirst y Di Cristo, 1998): los tonos absolutos se registran como T (*Top*) para la altura tonal máxima del hablante ( $\uparrow$ ), B (*Bottom*) para su altura mínima ( $\downarrow$ ) y M (*Mid*) para los valores medios ( $\Rightarrow$ ); los relativos no iterativos se señalan como H (*Higher*) para los puntos más altos que los adyacentes ( $\uparrow$ ), L (*Lower*) para los más bajos ( $\downarrow$ ) y S (*Same*) para los que tiene la misma altura que el tono anterior ( $\rightarrow$ ). Los tonos relativos iterativos se codifican como U (*Up – Upstepped*) cuando es un punto en una secuencia ascendente ( $<$ ) y como D (*Down – Downstepped*) cuando está en una secuencia descendente ( $>$ ) (Hirst y Di Cristo, 1998).

d. Nivel fonológico profundo: corresponde a las decisiones lingüísticamente significativas que interactúan con un número de parámetros prosódicos específicos para cada lengua con el fin de generar patrones entonativos. A continuación, se presentan las reglas que determinan el patrón entonativo de un enunciado según este modelo (Hirst et al., 2000; Baqué y Estruch, 2003):

- División del enunciado en unidades entonativas.
- División de las unidades entonativas en unidades tonales.
- Asignación de los patrones entonativos a las unidades entonativas.
- Asignación de los patrones a las unidades tonales.
- Aplicación, si es preciso, de la regla de reajuste.

Baqué y Estruch (2003) dividen los procesos que se llevan a cabo en el nivel fonológico profundo en tres etapas. En la primera, se establece el árbol fonológico y la rejilla métrica asociada al enunciado sin dejar de lado el hecho de que puede ser necesario tener en cuenta los aspectos relacionados con una posible acentuación enfática o distintiva, la desacentuación cuando hay presencia de dos sílabas tónicas consecutivas, entre otros factores. En esta fase, se divide el enunciado en unidades entonativa (UE), unidades rítmicas (UR) y unidades tonales (UT). A continuación, se presenta el árbol fonológico, junto con la rejilla métrica correspondiente, de la frase en catalán *És imprescindible que la tingui demà perquè sóc diabètic* (*Es imprescindible que la tenga mañana porque soy diabético*), donde se divide el enunciado en unidades entonativas (UE), unidades rítmicas (UR) y unidades

tonales (UT); además, se señalan las sílabas prominentes (+) y las no prominentes (.) (Baqué y Estruch, 2003: 135).

[És imprescindible que la tingui demà]<sub>UE</sub> [perquè sóc diabètic]<sub>UE</sub>  
 [{És imprescindible}<sub>UR</sub> {que la tingui demà}<sub>UR</sub>]{UE} [{perquè sóc}<sub>UR</sub> {diabètic}<sub>UR</sub>]{UE}  
 [{(És)<sub>UT</sub> (im pres cin di)<sub>UT</sub> ble}<sub>UR</sub> {(que la tin)<sub>UT</sub> (gui de mà)<sub>UT</sub>}{UR}]{UE} [{(per què)<sub>UT</sub>  
 (sóc)<sub>UT</sub>}{UR} {(di a bè)<sub>UT</sub> tic}<sub>UR</sub>]{UE}

3 (+ . . . . + . . . + . . +) . . + . . + .)

2 (+ . . . . + .) (. . + . . +) (. + +) (. . + .)

1 (+) (+ + + +) + (+ + +) (+ + +) (+ +) (+) (+ + +) +

S És im pres cin di ble que la tin gui de mà per què sóc di a bè tic

En una segunda etapa, se introducen tonos predeterminados a las unidades tonales y entonativas según la lengua que se está estudiando y la estructura del enunciado. En el ejemplo del enunciado catalán (*És imprescindible que la tingui demà perquè sóc diabètic*), las unidades tonales emplean los tonos LH y las entonativas LL en los extremos, como se ve en el patrón entonativo resultante (Baqué y Estruch, 2003:136):

L

(LH) (LH) (LH) (LH) L] [L (LH) (LH) (LH) L].

Finalmente, se aplican las siguientes reglas de reajuste y linearización (Baqué y Estruch, 2003:136) esquematizadas en la figura 1.5:

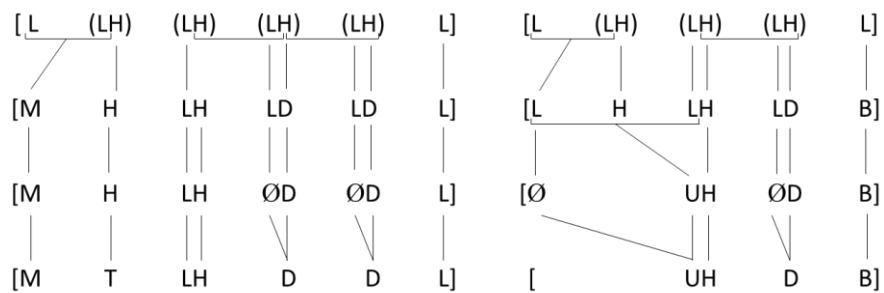
- a. De acuerdo con los principios de simplificación, los tonos L iniciales de la representación se fusionan en un tono medio M, propio del registro de referencia del locutor.
- b. El tono final L se realiza como un tono bajo absoluto B, correspondiente al valor límite inferior del registro del locutor.
- c. En cumplimiento de la regla de reajuste de escalonamiento descendente (*downstep*), un tono H precedido de una secuencia de tonos H y L en la misma unidad entonativa es bajado y realizado como un tono D, debido a la presencia del tono L, que desaparece posteriormente.
- d. Debido a la aplicación de la regla de escalonamiento ascendente (*upstep*), el tono L precedido de una secuencia de tonos L y H en la misma unidad entonativa se realiza

como un tono relativo iterativo U, pues el tono H anterior, no realizado fonéticamente, crea un realce del tono L en cuestión.

- e. Los segmentos tonales de las diferentes unidades prosódicas se organizan en diferentes líneas, lo que implica la ubicación de los tonos correspondientes a las unidades de rango prosódicamente superior en las unidades tonales adecuadas de acuerdo con las reglas de linearización, para obtener así una secuencia fonológica pronunciable.
- f. Se reducen las secuencias de segmentos tonales idénticos en una misma unidad tonal a un único segmento tonal, para evitar la colisión de dos tonos idénticos en la misma unidad. Esto es para dar cumplimiento a la regla de simplificación tonal.
- g. El tono inicial H se realiza como un tono absoluto T, correspondiente al límite superior del registro tonal del locutor, pues representa la máxima frecuencia de la oración.

Todo este proceso permite obtener la secuencia [MTLHDDL UHDB] derivada del patrón prosódico presentado anteriormente. La aplicación de las distintas reglas y las fases de este procedimiento se pueden observar en la figura 1.5.

Figura 1.5. Ejemplo de aplicación de las reglas de reajuste y linearización para el enunciado *És imprescindible que la tingui demà perquè sóc diabètic* (tomado de Baqué y Estruch, 2003:137)



Como se puede ver en los párrafos anteriores, el modelo Aix-en-Provence representa un sistema complejo de análisis prosódico. Lo que se retoma de dicho modelo, en esta investigación, es la perspectiva generativista que contempla el estudio en un nivel superficial y en uno subyacente, así como el enfoque pluriparamétrico del análisis prosódico. Con respecto al primer punto, dado el aspecto cognitivo de este trabajo, se comparte la idea de

que es necesario estudiar el componente suprasegmental, teniendo en cuenta tanto las realizaciones superficiales (fonéticas) como las subyacentes (fonológicas). Sin embargo, a lo no se realizará un análisis consistente en cuatro niveles, como propone el modelo, sino que se abordarán solo dos de ellos: el “concreto” y el “abstracto” (ambos términos equiparables a “superficial” y “subyacente”, respectivamente, como se explicará posteriormente en §1.1.4.2). Además de esto, este trabajo también parte de la premisa de que la prosodia es un elemento pluriparamétrico y autónomo del lenguaje, por lo que, durante el análisis, se deben tener en cuenta todos los constituyentes que la integran. Desde esta perspectiva, se considera que un análisis detallado y pormenorizado de cada uno de los rasgos suprasegmentales en el nivel concreto (acústico) puede brindar los datos necesarios para identificar las formas abstractas que operan en la mente de los hablantes de una lengua.

#### 1.1.4.2. Modelo métrico-autosegmental

El modelo métrico-autosegmental es el modelo de referencia en el análisis entonativo desde la década de 1990; y tiene sus orígenes en la tesis doctoral de Pierrehumbert (1980), quien llevó a cabo un análisis entonativo del inglés. El objetivo de este modelo, explica Hualde (2003:155), es “la identificación de los elementos contrastivos del sistema entonativo cuya combinación produce los contornos melódicos que encontramos en los enunciados posibles de una lengua”. Como una metodología de análisis enraizada en la teoría fonológica autosegmental (Goldsmith, 1976), Pierrehumbert (1980) plantea una perspectiva desde la cual la entonación de los enunciados constituye un segmento autónomo e independiente de los demás rasgos fonológicos. Por lo tanto, los tonos presentes en la pronunciación de un enunciado constituyen autosegmentos asociados a un nivel segmental por medio de reglas universales y, a la vez, específicas para cada lengua (Hualde, 2003).

De acuerdo con Sosa (1999), el modelo métrico-autosegmental posee dos características cruciales en contraste con sus predecesores: la descripción de los movimientos de la F0 con base en tonos altos (H) y bajos (L) y la expresión cuantitativa de las realizaciones fonéticas basándose en contorno de la F0. En relación con la primera de estas características, las secuencias de tonos altos (H) y bajos (L) que se alinean con las sílabas tónicas de un enunciado, marcados con un asterisco (\*) junto a dos tonos adicionales, el de frase, señalado con un guión a modo de superíndice (´), y el de juntura, identificado con el símbolo porcentual

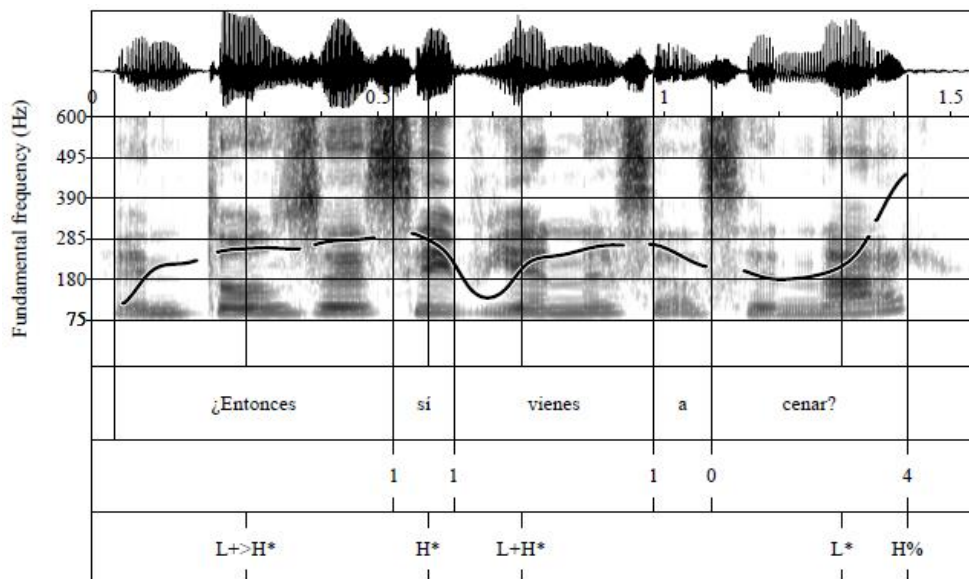
(%), que describen los movimientos melódicos al final de la frase. Pierrehumbert (1980) explica que los acentos tonales se alinean con las sílabas métricamente fuertes; y que pueden estar constituidos por tonos de diferentes alturas (H y L), de los cuales H corresponderá con el pico del contorno de F0, mientras que L corresponderá con un valle anterior o posterior al pico.

De igual forma, la autora explica que el tono de frontera se encuentra al final de la frase sin importar la estructura métrica del texto; y que el acento de frase<sup>8</sup> se localiza inmediatamente después del acento principal del enunciado (el último acento tonal) y antes del tono de frontera. Por otra parte, Pierrehumbert (1980) aclara que los acentos nucleares pueden ser monotonaes, constituidos por un solo tono H o L, o bitonaes, conformados por secuencias de dos tonos (H+L o L+H), pero tanto los tonos de frase como los de frontera solo pueden ser monotonaes (se puede ver un ejemplo tanto de los acentos tonales como los tonos de juntura en la figura 1.6). Adicionalmente, existe una asociación directa entre el tono de juntura y el texto, pues dicho tono se localiza en el linde derecho de la frase, independientemente de la estructura métrica de esta; es decir, el tono de juntura se alinea con el linde derecho de la frase entonativa mayor. Estos tonos de juntura, de acuerdo con Sosa (1999), son los movimientos tonales con función delimitadora que aparecen al final de las secuencias donde, en el habla corrida, no siempre se dan pausas evidentes y son los que demarcan las fronteras entre los grupos melódicos.

---

<sup>8</sup> Como se explica unos párrafos más adelante, el acento de frase no es un elemento relevante en la prosodia del español.

Figura 1.6. Oscilograma, espectrograma y F0 del enunciado *¿Entonces sí vienes a cenar?* (Tomado de De-la-Mota *et al.*, 2010: 336)



Según explica Sosa (1999:31), “el conjunto de tonos que marcan el final de un enunciado y que coincide con las sílabas finales a partir de la que lleva el último acento” se denomina “tonema” o “núcleo”. Este concepto, como se mencionó en §1.1.2.1, también fue empleado por Navarro Tomás (1966) y coincide, igualmente, con el término empleado por D’Introno *et al.* (2010) para hacer referencia al mismo conjunto de tonos. Este mismo elemento se maneja, aunque con otros términos, en el modelo por configuraciones de la escuela británica y en el de niveles de la escuela americana. En el primero de ellos, explica García-Lecumberri (2003), la última sílaba tónica del grupo entonativo se denomina “núcleo”, después del cual se encuentra la cola, constituida por la sílaba o el grupo de ellas que se da después del núcleo y que tiene una función significativa. En el modelo por niveles, de acuerdo con Martínez Celdrán (2003), se emplea el término “macrosegmento” para hacer referencia a las sílabas que se emiten en una frase entonativa, el cual incluye un centro y una inflexión terminal. El centro, según este autor, está constituido por la sílaba más prominente de la entonación, que suele ser la última acentuada; y a la que le sigue la inflexión final que marca el límite entre macrosegmentos sucesivos en una emisión. Esta sección, en términos de este modelo, es el equivalente al tonema.

Por otra parte, en referencia al acento de frase, Pierrehumbert (1980) señala que es importante describir lo que ocurre al final de la frase, luego del último acento tonal, pues los tonos en ese segmento se relacionan con la interpretación de la fuerza expresiva de la frase como tal. Sin embargo, esta afirmación se llevó a cabo en un estudio entonativo del inglés, lo que quiere decir que no necesariamente ocurre lo mismo en otras lenguas. De hecho, Prieto y Roseano (2010b) comentan que el acento de frase ha sido causa de debate, debido a que varios autores, entre los que mencionan a Sosa (1999), no consideran que el acento de frase tenga lugar en la descripción entonativa del español, puesto que, a diferencia del inglés, el español tiende a localizar el acento principal (nuclear) al final de los enunciados, por lo que no hay material tonal entre el último acento tonal y el tono de frontera del que se necesite rendir cuenta bajo el concepto de acento de frase. En caso de que se dé un movimiento complejo de F0, se emplea un tono de frontera bitonal para hacer referencia a este.

En relación con la segunda característica importante del modelo métrico-autosegmental que menciona Sosa (1999), la expresión cuantitativa de las realizaciones fonéticas de acuerdo con el contorno de la F0 (y no en forma de transcripción fonética estrecha). Hualde (2003) señala que el método de transcripción basado en el modelo autosegmental no es equivalente a una transcripción fonética (como es el caso del alfabeto fonético internacional). Esto se debe, explica el autor, a que, con los sistemas de transcripción fonética, podemos representar el aspecto (supra)segmental de una lengua que no comprendemos y cuyos contrastes fonológicos no conocemos, mientras que el sistema de notación prosódica fundamentado en el modelo métrico autosegmental requiere el previo conocimiento de las sílabas con acento léxico y el del sistema de contrastes tonales posibles en la lengua que se estudia. Pierrehumbert (1980) aclara que el modelo métrico autosegmental permite obtener una transcripción fonológica de la entonación, de acuerdo con el análisis de elementos tonalmente contrastivos de la lengua que se transcribe. Por lo tanto, la autora estableció reglas locales sensibles al contexto, que derivan la secuencia tonal en valores cuantitativos que determinan el contorno de la frecuencia fundamental (F0). Estas reglas, explica Pierrehumbert (1980), se procesan de izquierda a derecha e incluyen reglas de escalonamiento ascendente y descendente junto con otras que permiten develar características entonativas no locales, como la configuración general del contorno.



#### 1.1.4.2.1. Sistema de transcripción *Tones and Break Indices* (ToBI)

El sistema de transcripción desarrollado con base en el modelo métrico-autosegmental, una de cuyas variantes es empleada en esta investigación, es el llamado *Tones and Break Indices* (ToBI). Este modelo “fue concebido como un sistema de referencia para la transcripción de la entonación y de las estructuras prosódicas de los enunciados de una variedad lingüística, el inglés norteamericano” (Sosa, 2003:187). Sosa (2003) comenta que la intención de este sistema es ofrecer a la comunidad de lingüistas un vocabulario y convenciones comunes que facilitarían el análisis de corpus orales, a través de un sistema estandarizado de etiquetaje prosódico diseñado para transcribir bases de datos digitalizadas. Esto se lleva a cabo a través del uso de un repertorio de acentos tonales finito, que pueden variar en forma y número dependiendo de la lengua que se estudie (Hualde, 2003).

La variante de ToBI empleada para el español (y por lo tanto en este trabajo) es el Sp\_ToBI, el cual, según afirman Prieto y Roseano (2010b), se diferenció, desde el inicio del modelo estándar, por la existencia de un tono de frontera medio (M%), opuesto a la tendencia de emplear únicamente tonos H y L. Los autores comentan que este tono se ha documentado en español desde los trabajos de Navarro Tomás (1944) y que se registra actualmente en distintas variedades del español, como la castellana, la canaria, la andina venezolana, la chilena y la argentina. Según explican Sosa (2003) y Beckman *et al.* (2002), el sistema de transcripción Sp\_ToBI consiste en los siguientes elementos:

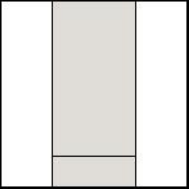
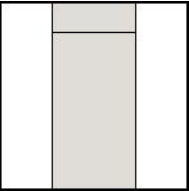
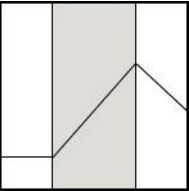
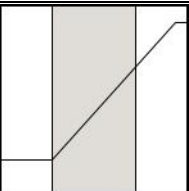
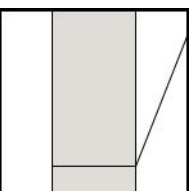
- a. Estrato de palabra: transcripciones ortográficas con segmentación del enunciado en palabras ortográficas, como en los sistemas ToBI para otras lenguas con escritura alfabética.
- b. Estrato de sílabas: transcripción de los segmentos, sílaba por sílaba.
- c. Estrato de índices de disyunción: señalización de disyunción entre pares de palabras y antes de cada pausa, como en otros sistemas ToBI. Se distinguen los siguientes niveles:
  - 0: cualquier ejemplo claro de reducción silábica por contacto de vocales entre palabras, como en casos de sinalefa.
  - 1: cualquier otra juntura ordinaria entre palabras.
  - 4: frase entonativa (grupo melódico)

Se deben reservar los niveles 2 y 3 para unidades con marcadores fonológicos claros y un sentido de disyunción intermedio entre 4 y 1; posibles candidatos para estos niveles podrían ser la frase intermedia, el grupo tónico (acentual) y el grupo clítico.

- d. Estrato de tonos: distinción entre tipos de acentos tonales.
- e. Estrato misceláneo: etiquetado de fenómenos como pausas distributivas, disfluencias, risas o cualquier otra incidencia que complique el análisis del patrón tonal y/o del fraseo melódico, al igual que análisis alternativos propuestos para el enunciado.
- f. Estrato de código: identificación del dialecto (y sociolecto) del hablante, de conocerse, y eventualmente, para casos de alternancia de códigos entonacionales.

Los estudios llevados a cabo con el sistema Sp\_ToBI han permitido develar una amplia gama de acentos tonales en distintas variedades de español, como se puede observar en los diferentes capítulos de Prieto y Roseano (2010a). Con el fin de mostrar un ejemplo de los resultados del sistema Sp\_ToBI, a continuación, se muestra el repertorio de acentos tonales (cuadro 1.2), de frontera y configuraciones tonemáticas (cuadro 1.3) registrados por De-la-mota *et al.* (2010) en el español mexicano, por ser la variedad de estudio en este trabajo:

Cuadro 1.2. Inventario de acentos tonales monotonaes y bitonaes en el español mexicano y su representación esquemática (De-la-Mota *et al.*, 2010:320)

Acentos tonales monotonaes		
	L*	Este acento se realiza fonéticamente como una meseta baja en la parte más baja del rango del hablante. Se encuentra en posición nuclear de enunciados de foco amplio, enunciados de contradicción, interrogativas absolutas que solicitan información desconocida, interrogativas absolutas imperativas, interrogativas absolutas de eco, interrogativas como invitaciones o solicitudes corteses, interrogativas parciales de eco y vocativos.
	H*	Este acento se realiza fonéticamente como una meseta alta sin un valle previo de F0. Se registra en posición prenuclear en enunciados con foco amplio, enunciados de contradicción, interrogativas absolutas de confirmación, interrogativas parciales con función de invitación y en interrogativas parciales.
Acentos tonales bitonaes		
	L+H*	Este acento se realiza fonéticamente como un movimiento tonal ascendente durante la sílaba acentuada, con el pico de F0 situado al final de dicha sílaba. Se encuentra comúnmente en la posición nuclear de enunciados con foco estrecho y amplio, enunciados exclamativos, declaración de lo obvio, interrogativas parciales, interrogativas parciales de eco, interrogativas parciales exclamativas, interrogativas parciales imperativas, órdenes, solicitudes corteses y vocativos.
	L+>H*	Este acento se realiza fonéticamente como un ascenso tonal en la sílaba acentuada con el pico de F0 alineado en la sílaba postónica. Se encuentra en posición prenuclear en enunciados con foco amplio, enunciados exclamativos e interrogativas absolutas imperativas.
	L*+H	Este acento se realiza fonéticamente como un valle de F0 en la sílaba acentuada con un ascenso subsecuente en la sílaba postónica. Se registra en posición prenuclear en interrogativas absolutas contraexpectantes e interrogativas parciales de eco.

	<p>H+L*</p>	<p>Este acento se realiza fonéticamente como un descenso de F0 dentro de la sílaba acentuada. Se encuentra en posición prenuclear en interrogativas parciales imperativas.</p>
--	-------------	--

Cuadro 1.3. Inventario de tonos de juntura monotonaes y bitonaes en el español de México y su representación esquemática (De-la-Mota *et al.*, 2010:321)

Tonos de juntura monotonaes		
	<p>L%</p>	<p>L% es realizado fonéticamente como un tono bajo sostenido o descendente en la línea base del rango del hablante. Se encuentra al final de enunciados de foco amplio y estrecho, enunciados exclamativos, interrogativas parciales exclamativas, interrogativas parciales imperativas, órdenes y vocativos.</p>
	<p>M%</p>	<p>M% se realiza fonéticamente como un movimiento tonal ascendente o descendente hasta un punto medio del rango del hablante. Se encuentra en interrogativas parciales exhortativas, enunciados de incertidumbre y vocativos.</p>
	<p>H%</p>	<p>H% se realiza fonéticamente como un movimiento tonal ascendente que procede de un acento bajo o alto. Se registra en interrogativas absolutas que buscan información nueva.</p>
Tonos de juntura bitonaes		
	<p>HH%</p>	<p>HH% se realiza fonéticamente como un ascenso agudo al final de la frase y que usualmente alcanza el nivel más alto del rango del hablante. Es típico de invitaciones corteses e interrogativas absolutas con función de petición.</p>
	<p>LH%</p>	<p>LH% se realiza fonéticamente como un valle de F0 seguido por un ascenso. Se registra en interrogativas absolutas que buscan información nueva, interrogativas absolutas de eco e interrogativas absolutas imperativas.</p>













		HL%	HL% se realiza fonéticamente como un pico de F0 seguido por un descenso. Se encuentra en enunciados de contradicción, interrogativas parciales, solicitudes y vocativos.
		LM%	LM% se realiza fonéticamente como un valle de F0 seguido por un movimiento hasta un punto medio en el rango del hablante. Se registra en enunciados de obviedad.

Lo presentado en los cuadros 1.2 y 1.3 corresponde con una propuesta de sistema de notación para los fenómenos entonativos documentados en el español. Sin embargo, a lo largo del tiempo se han presentado distintas modificaciones en las etiquetas de transcripción, cuyo empleo ha dependido de las perspectivas teóricas y planteamientos metodológicos de los autores al momento de realizar sus investigaciones. En el caso concreto de este trabajo, el sistema de notación empleado se especifica en el capítulo siguiente (§2.2.8), donde se presenta el conjunto de etiquetas utilizadas, se describe el movimiento tonal que representan y se hace referencia a los trabajos que sirvieron como guía para su diseño.

1.1.4.2.2. Consideraciones teóricas sobre el análisis entonativo en dos niveles con base en el proceso de abstracción

Debido a que el modelo métrico-autosegmental propone la notación tonal desde una perspectiva fonológica, como se mencionó anteriormente, es posible que ciertas secuencias tonales en el plano acústico sean variantes superficiales de un tono o secuencia tonal subyacente, tal como ocurre entre los alófonos y los fonemas. Este caso es planteado por Dorta *et al.* (2013b), cuando presentan el inventario de acentos tonales en las variedades de español canaria, cubana y venezolana mostrado en el cuadro 1.4.

Cuadro 1.4. Representación de la estructura profunda, superficial y esquema acústico de los acentos tonales (tomado de Dorta *et al.*, 2013b: 72)

ESTRUCTURA PROFUNDA	ESTRUCTURA SUPERFICIAL	NIVEL ACÚSTICO
Invariante fonológica	Variantes	Esquema del contorno tonal
/L*+H/	[L*+H]	
	[L*+!H]	
/L+H*/	[L+H*]	
	[L+>H*]	
/L*/	[L*]	
	[!H+L*]	
/H*/	[H*]	
	[¡H*]	
	[!H*]	
	[L+H*]	
/H*/	[L+¡H*]	
/!H*/	[L+!H*]	

Esta noción acerca de que determinados movimientos tonales en el plano acústico pueden ser variantes de una secuencia tonal subyacente es un aspecto que se considera en la metodología de este trabajo. Si bien es cierto que uno de los planteamientos principales del sistema de notación ToBI es su carácter fonológico (Sosa, 2003; Jun, 2005), autores como Hualde y Prieto (2016) ya han hablado acerca de la necesidad de sistematizar la transcripción entonativa con base en un análisis tanto fonético como fonológico de la entonación. De esta

manera, se podrían determinar las realizaciones alofónicas de categorías tonales subyacentes. El planteamiento de una transcripción que tome en cuenta el aspecto fonético y fonológico cobra sentido al tener en cuenta la cualidad cognitiva del lenguaje (en cuanto a que la lengua es un producto de la mente humana). Partiendo de la premisa de que los rasgos entonativos subyacentes son abstracciones de los hechos físicos concretos producidos en la oralidad, es posible inferir que uno o más tonos fonéticos pueden estar asociados con un solo elemento fonológico. Sin embargo, antes de proseguir con otros argumentos y ejemplos que soporten el análisis que contempla la transcripción en los niveles fonético y fonológico, es necesario revisar previamente el concepto de abstracción.

Al momento de definir “abstracción”, Burgoon *et al.* (2013) explican que existe un consenso claro en que la abstracción tiene influencia en la forma cómo se aprende información nueva, se generan juicios y se regulan comportamientos. Estos autores definen la abstracción como un proceso de reducción de información que permite almacenar y recuperar conocimiento central eficientemente, lo que se logra a través de la identificación de características centrales invariables constituidas por aquellas propiedades que aumentan la probabilidad de identificar, de forma precisa, un elemento determinado cuando se encuentra en varios contextos. Pease *et al.* (2009) establecen una analogía entre la abstracción cognitiva y la informática, donde la abstracción es el resultado del razonamiento de la estructura común compartida por un grupo de datos o entidades matemáticas sin tomar en cuenta aquellas propiedades que difieren de un elemento a otro. El resultado de este proceso son las representaciones abstractas (o subyacentes) que, de acuerdo con Watkins *et al.* (2008), transmiten el sentido y significado esenciales de eventos y acciones. Las representaciones abstractas también se pueden considerar esquemas cognitivos, de los cuales se hablará posteriormente (§1.4.2)

En el sistema de notación prosódica Sp\_ToBI empleado en esta investigación (véase cuadro 2.2 en §2.2.8), se puede observar que existen diacríticos empleados para representar fenómenos que se producen en la entonación de un enunciado; por ejemplo, del escalonamiento descendente (*downstep*), señalado a través del diacrítico del signo de admiración de cierre junto con un tono alto (!H), o el de apertura junto a un tono alto (;H), para indicar un escalonamiento ascendente (*upstep*). Debido a que ambos diacríticos hacen referencia a un tono alto (con altura variable, aunque en cualquier caso opuesto a un tono

bajo) en relación con la totalidad del rango tonal de un hablante al momento de producir un enunciado, se puede decir que estos constituyen características secundarias del acento tonal. De acuerdo con Shapira *et al.* (2012), al pasar a un constructo de nivel superior (abstracto), se realiza una distinción entre características primarias, es decir, rasgos definitorios relativamente estables e invariables, y características secundarias (más concretas), que pueden variar en la medida en que se presentan cambios en el contexto y que son, por lo tanto, omitidas en la representación de nivel superior.

De acuerdo con todo lo expuesto, por ejemplo, si en la primera sílaba tónica del enunciado hipotético *Comí mucho helado*, cinco hablantes distintos producen los acentos tonales  $L+H^*$ ,  $L+;H^*$ ,  $L+!H^*$ ,  $L+<!H^*$  y  $L+>;H^*$ , sería difícil afirmar que estos acentos corresponden a cinco representaciones subyacentes (o abstracciones) diferentes, tal como concluyen Díaz-Campos y Ronquest (2007), en su investigación relacionada con los acentos tonales  $L+H^*$  y  $L^*+H$ . La variación entre un acento tonal y otro puede deberse a diversos factores, como el choque tonal que ocurre entre la última sílaba de la primera palabra y la primera de la segunda (ambas tónicas), foco contrastivo en alguna de estas dos palabras, expresividad en la enunciación, entre otros. Debido a que los diacríticos empleados en el SP\_ToBI señalan características particulares de los acentos tonales, para especificar la altura alcanzada y el lugar en el que se ubica el pico tonal con respecto al linde de la sílaba tónica, su función es la de agregar mayor grado de detalle a la descripción entonativa y, por ende, hacerla más concreta (más fonética). Este grado de especificación (aunque no absoluto en términos acústicos) resta abstracción al fenómeno entonativo, pues de algún modo se enfoca en los detalles del ascenso tonal y no en el ascenso como tal, que en este caso constituye la abstracción detrás de la producción acústica y, en consecuencia, el verdadero hecho fonológico. Cabe señalar que con esto no se quiere decir que los diacríticos tengan siempre una naturaleza fonética, puesto que también pueden representar hechos fonológicos. Por ejemplo, Hualde y Prieto (2015) transcriben los tonos de juntura medios como  $!H\%$ , donde el diacrítico es parte del elemento fonológico abstracto y no una especificación de una cualidad acústica inherente a la juntura. De igual forma, no se descarta la posibilidad de que los diacríticos en la transcripción formen parte de la descripción entonativa de un hecho fonológico en alguna otra lengua distinta al español.



Es posible establecer una analogía entre las cualidades concretas y abstractas de una secuencia tonal y las de otro fenómeno fonológico, aunque este último se da en el plano segmental en español.<sup>9</sup> Es sabido que uno de los fonemas que presenta más variación en español es /s/ en posición posnuclear. En este caso, el fonema /s/ tiene cuatro realizaciones posibles en el dialecto andaluz y en el español de América, de acuerdo con D'Introno *et al.* (2010): relajada [s̺], aspirada [h], asimilada a la consonante siguiente (/mis.mo/ se realiza como ['mim.mo], /as.no/ como ['an.no], /is.la/ como ['il.la], /pes.ca/ como ['pek.ka], /des.pa.sjo/ como [dep.'pa.sjo], etc.) y elisión [∅]. A esta categorización le podemos sumar la realización plena [s]. Si bien es cierto que cada uno de estos alófonos tiene características articulatorias suficientemente distintas como para pensar que corresponden a representaciones subyacentes diferentes, el hecho de que la variación no represente una alteración semántica evidencia que todos ellos son realizaciones concretas de un mismo elemento abstracto subyacente: /s/. Esta misma afirmación se podría hacer acerca de la relación que guardan las realizaciones concretas de los acentos tonales L+H\*, L+<sub>i</sub>H\*, L+!H\*, L+<!H\* y L+><sub>j</sub>H\* con el acento abstracto L+H\*. A pesar de que estos acentos tonales pueden considerarse fonológicamente distintos desde la perspectiva teórica del modelo métrico-autosegmental, puede que correspondan a distintas formas concretas de un mismo elemento abstracto.

Sin importar la productividad de un acento tonal, es decir, la frecuencia con la que sea usado por hablantes de una zona en particular o incluso del español general, este siempre corresponderá con una abstracción subyacente que no necesariamente debe ser idéntica a la realización concreta. Por ejemplo, el acento prenuclear más común en la primera sílaba tónica de enunciados con foco amplio en el español de la Ciudad de México es L+>H\* (De-la-Mota *et al.* 2010), pero el hecho de que el pico entonativo sea principalmente situado en la sílaba postónica no quiere decir que la representación subyacente necesariamente implique esta característica. La abstracción que se encuentra en el constructo cognitivo de nivel superior en la mente del hablante es una elevación tonal (que puede corresponder con el acento tonal L+H\*) en un nivel fonológico. Las características concretas (acústicas) relacionadas con la

---

<sup>9</sup> Con el ejemplo que se plantea en este párrafo no se pretende sugerir que la naturaleza o las dinámicas lingüísticas o cognitivas de los elementos segmentales y los suprasegmentales son iguales, sino que se busca ilustrar el hecho de que diversos elementos fonéticos con cualidades acústicas distintas pueden estar asociados con una misma forma fonológica subyacente.

melodía de un enunciado no necesariamente deben formar parte de la abstracción subyacente de la prosodia.

De vuelta al ejemplo segmental de /s/ posnuclear, el hecho de que la realización aspirada ([h]) del fonema /s/ en posición final de sílaba sea la más común en el habla de Caracas (Carrasquero, 2010) no quiere decir que [h] sea la abstracción subyacente. Al igual que en el caso del acento tonal L+>H\* en la primera sílaba tónica, [h] en posición posnuclear es la forma concreta más frecuente de un constructo de nivel superior en esta posición específica. Con estos ejemplos no se propone plantear un paralelismo absoluto entre los fenómenos en el plano segmental y en el suprasegmental, sino ilustrar la idea de que las abstracciones fonológicas no están necesariamente condicionadas por la frecuencia de uso de una realización fonética en concreto. Las emisiones orales están directamente relacionadas con un constructo abstracto en la mente de los hablantes, el cual también se puede manifestar acústicamente de otras formas menos frecuentes (como las otras realizaciones del acento tonal L+H\* y del fonema /s/ presentadas previamente).

Los casos de focalización contrastiva (o estrecha) no representan una excepción a la variación de producciones concretas. Retomando la analogía entre el acento tonal ascendente (L+H\*) y el fonema fricativo alveolar sordo (/s/), aunque ahora relacionados en términos del énfasis que aportan al discurso, de la misma forma en la que se ha establecido el acento tonal L+;H\* como una tendencia principal para marcar el foco contrastivo en el español del centro de México (Martín Butragueño y Mendoza, 2018). se ha documentado que la producción de la sibilante plena [s] en posición posnuclear es un marcador de énfasis en aquellas variedades del español donde la norma en esta posición es la aspirada [h] o la elisión [ø] (Terrell, 1978; Torreblanca, 1978; Broś, 2013). En una situación hipotética en que una madre regaña a su hijo por haber desorganizado algo, ella podría indicarle que arreglara el desorden que hizo empleando el enunciado enfático *¡Ordena esto!*, el cual, de acuerdo con los trabajos mencionados previamente, podría ser producido con un acento nuclear L+;H\* si fuese emitido por una hablante del centro de México y con una sibilante plena [s] si fuese caraqueña (venezolana). En ambos casos, lo que se observa es la implementación de una variante concreta de un elemento abstracto que aporta significado pragmático al discurso, sin modificar el contenido semántico primario, con la finalidad de enfatizar un aspecto del discurso sobre los demás o agregar un matiz expresivo. Esto responde a que, en los dos

ejemplos, la hablante hace uso del código de esfuerzo (Gussenhoven, 2004) a través de un empleo de mayor energía articuladora con la finalidad de transmitir un sentido específico en el enunciado. En cualquier caso, dado que la utilización de una variante ( $L+_{i}H^*$  o /s/) por otra ( $L+>H^*$  o /h/) en esta posición española no altera el significado semántico en la lengua, sino que agrega matices pragmáticos, el uso de uno u otro no implica un contraste entre distintos elementos abstractos (fonológicos), sino una diferencia del valor pragmático entre dos realizaciones concretas (fonéticas) que incluso pueden aportar información de índole sociolingüística.

Si bien con esta propuesta de abstracción de acentos tonales y tonos de juntura se afirma que distintas etiquetas del sistema de transcripción pueden hacer referencia a diferentes realizaciones concretas (fonéticas) de un mismo elemento abstracto (fonológico), no se sugiere que este planteamiento implique que distintas configuraciones tonales que comparten una misma función lingüística también corresponden con un mismo tono subyacente. Como se explicó anteriormente, la abstracción consiste en omitir rasgos específicos variables que acompañan a los elementos centrales invariables, los cuales conforman la forma abstracta subyacente. Es posible que distintos elementos subyacentes compartan la misma función discursiva, es decir, que exista variación entre ellos, como puede ser el foco contrastivo o estrecho. Martín Butragueño y Mendoza (2018) comentan que los acentos tonales  $L^*$  y  $!H^*$  son usados junto con  $L+H^*$  como marcadores de foco estrecho por hablantes de español del centro de México. En este caso, tres acentos tonales distintos son empleados con la misma función pragmática. En ningún caso se plantea que estos tres acentos puedan estar relacionados de forma subyacente en ningún otro plano lingüístico además del funcional, pues sus características esenciales son suficientemente distintas como para no compartir características centrales invariables y, por ende, se puede descartar que compartan la misma forma abstracta subyacente.

De igual modo, se considera necesario hacer hincapié en el hecho de que, hasta este punto, los planteamientos se han formulado con base únicamente en la información recolectada y estudiada en investigaciones acerca de la entonación, con especial atención en aquellas relacionadas con la lengua española. Es importante tener en cuenta esta aclaración, ya que es posible que, por ejemplo, la excursión tonal en un ascenso melódico pueda ser un rasgo fonológico en otra lengua, en la cual, los acentos tonales  $L+H^*$  y  $L+_{i}H^*$  tengan una

oposición en términos de significado semántico (no pragmático). Por lo tanto, este planteamiento se debe entender solo como una aproximación teórica formulada para la entonación del español de acuerdo con el fenómeno cognitivo de la abstracción. Adicionalmente, se quiere resaltar, una vez más, que la analogía establecida entre la operación del sistema suprasegmental y el segmental presentada previamente (con el acento tonal L+H\* y el fonema /s/) no se debe interpretar como el planteamiento de una relación directa entre el funcionamiento de uno y otro, pues las dinámicas propias de uno de ellos no se dan de forma idéntica en el otro. Tal analogía se expone únicamente como ejemplo de una propuesta teórica sobre la cual es necesario profundizar y ahondar en términos lingüístico-cognitivos (ya que se entiende que un enfoque como este es el más eficiente para comprender la organización de los elementos lingüísticos en la mente humana a través de métodos empíricos).

### 1.3. Conceptos teóricos para el análisis sociolingüístico

Debido a que este trabajo se encuentra directamente relacionado con la sociolingüística variacionista, es necesario explicar algunos aspectos relevantes que se retoman para el análisis que se presenta en el capítulo IV. A grandes rasgos, el principal objetivo de la sociolingüística, desde sus orígenes en la década de 1960 con los trabajos de William Labov, consiste en comprender la forma en la que se producen cambios lingüísticos y cómo se extienden entre los hablantes por medio del estudio empírico de la relación entre variables lingüísticas y variables sociales. Esta noción general se desglosa en los distintos incisos de este apartado, en donde se muestra, en primer lugar, una definición general de lo que se entiende como sociolingüística, seguida de una explicación de las tres perspectivas de estudios en este ámbito, una exposición de lo que se entiende como variable y prestigio en la sociolingüística y la exposición de tres conceptos claves en el desarrollo de esta investigación: redes sociales, modos de vida e indexación.

#### 1.3.1. Sociolingüística: definición general

Meyerhoff (2006) afirma que crear una definición de sociolingüística en una sola frase es una tarea ardua, pues esta comprende un campo bastante amplio y el término puede ser empleado para describir múltiples formas de estudio lingüístico. Tanto Hernández Campoy

y Almeida (2005) como Moreno Fernández (2009) afirman que las definiciones que se han ofrecido de la sociolingüística no carecen de problemas y críticas por parte de lingüistas que se dedican al estudio de otras disciplinas, pues se ha cuestionado desde la naturaleza de su metodología hasta su verdadero objeto de estudio. En general, muchos autores coinciden al momento de definirla como una disciplina que estudia el lenguaje en su contexto social o sociocultural (Labov, 1972; Hudson, 1980; Lavandera, 1988; Hernández Campoy y Almeida, 2005).

La mayoría de los autores se orientan hacia una explicación del objeto de estudio y características metodológicas de la sociolingüística para definirla. Tal es el caso de Silva-Corvalán (2001), quien señala que la sociolingüística es una disciplina que se interesa ampliamente en el estudio de la lengua en su entorno social. Esto se lleva a cabo a través del análisis de fenómenos lingüísticos relacionados con factores sociales, como los diferentes sistemas de organización política, económica, social y geográfica de una sociedad, los factores individuales que afectan la organización social en general (edad, raza, sexo y nivel de instrucción), los factores históricos y étnico-culturales, y la situación inmediata que enmarca la interacción. Según Moreno Fernández (2009), el objeto de estudio de la sociolingüística es la lengua como sistema de signos, pero vista desde un contexto social, por lo que al sociolingüista le interesa la relación entre los estratos sociales y la estructura lingüística. De acuerdo con Tagliamonte (2006), la sociolingüística parte de la premisa de que el lenguaje existe en un contexto dependiente tanto de quien lo emplea como del lugar y la razón por la que se usa. Esto se debe a la idea de que el lenguaje se relaciona con el uso que los hablantes hacen de él y a que el discurso tiene una función social, como medio de comunicación y de identificación de grupos sociales. Meyerhoff (2006) señala que la sociolingüística emplea una serie de métodos empíricos para analizar el lenguaje en uso y las actitudes sobre la forma en la que se usa, por lo que algunos patrones solo pueden ser observados sistemáticamente a través del examen de un gran volumen de datos y un buen entendimiento acerca del trasfondo o lugar de un hablante en una comunidad.

Desde la visión de Hernández Campoy y Almeida (2005), la sociolingüística es una rama de la lingüística que plantea una perspectiva distinta, pues contempla el lenguaje como un fenómeno social y cultural. Por lo tanto, esta estudia el lenguaje en su contexto social, en situaciones de la vida real, por medio de la investigación empírica. En esta disciplina se

emplea una metodología relacionada con los contenidos y procesos de las ciencias sociales, principalmente la antropología social y la sociología.

Según Moreno Fernández (2009), existen dos tipos de sociolingüística. La primera de ellas es propiamente lingüística, pues estudia la estratificación interna de los diasistemas. Dichos diasistemas se definen como un conjunto de sistemas interrelacionados que constituyen una lengua variable (en términos sociales, funcionales y geográficos) y en constante cambio (Areiza Londoño *et al.* 2012). La segunda estudia la relación entre el diasistema y la sociedad. Estos dos tipos de sociolingüística pueden complementar las nociones de micro-sociolingüística y macro-sociolingüística expuestas por Hernández Campoy y Almeida (2005). La primera de estas nociones, explican los autores, describe y analiza tanto las lenguas como las características de sus usuarios en estudios de interacción directa, análisis del discurso, análisis de la conversación y otras áreas que requieran el estudio de grupos relativamente pequeños. Por otra parte, la macro-sociolingüística, de acuerdo con los precitados autores, estudia el uso de la lengua dentro de la interacción grupal, lo que en ocasiones podría incluir ramas como la sociología del lenguaje, la lingüística secular (sociolingüística laboviana), la dialectología y la geolingüística.

Los estudios sociolingüísticos muestran especial interés por la interacción entre el lenguaje y los diasistemas en la descripción del sistema lingüístico. Hernández Campoy y Almeida (2005) afirman que existen cuatro ejes fundamentales en el estudio sociolingüístico (diatópico, diastrático, diafásico y diacrónico). El eje diatópico está relacionado con factores geográficos (por ejemplo, el estudio del habla de una comunidad rural y una urbana, espacialmente distantes), el eje diastrático tiene que ver con elementos sociales o culturales (por ejemplo, un análisis del habla de personas con nivel socioeconómico o sociocultural bajo, en contraste con aquellos de nivel alto); el eje diafásico maneja elementos relacionados con el registro -o estilo- dado por un contexto situacional (por ejemplo, una comparación entre el habla en un contexto formal y la de un contexto familiar) y el eje diacrónico, que tiene que ver con factores temporales (por ejemplo, un estudio del habla de una comunidad en 1950 y en 2010).

La idea de emplear estos diasistemas es identificar elementos que llevan al cambio lingüístico, para lo cual se toma como componente de estudio la variación en la lengua. Esta variación es, según Moreno Fernández (2009), la alternancia de dos o más formas de un

mismo elemento lingüístico, condicionada por factores sociales, sin que se dé un cambio de naturaleza semántica. El variacionismo se basa en el principio de uso lingüístico, según el cual “el uso lingüístico (natural) solo puede tener lugar en contextos sociales y situacionales concretos” (Moreno Fernández, 2009: 39), lo que supone que el lenguaje debe ser observado en contextos y situaciones reales. Gracias a esto, señala el autor, los trabajos sociolingüísticos han dado a conocer que las variables que influyen en el proceso de variación lo hacen de un modo específico, en cada comunidad de habla, e involucran fenómenos lingüísticos concretos, lo que implica que la variación no ocurre de la misma manera en todas las comunidades; ya que las características sociales de sus miembros difieren considerablemente. Estos cambios en las comunidades de habla, de acuerdo con Moreno Fernández (2009), hacen que sea necesario un análisis sociológico previo de la comunidad, junto con estudios exploratorios que permitan conocer cuáles son las variables importantes en la estructura social y cuáles son las que, presumiblemente, influyen más en el uso social de la lengua.

La definición de sociolingüística que se considera en esta investigación consiste en una amalgama de varias de las presentadas previamente. Así, se concibe la sociolingüística como una disciplina que estudia el cambio lingüístico de acuerdo con el entorno social en el que se produce; y cuyo objetivo es el estudio empírico de la influencia que ejercen distintas variables sociales (edad, sexo, grado de instrucción, origen, entre otros) en el uso de la lengua, de acuerdo con el contexto social. De igual manera, se contempla la variación como la alternancia de dos o más formas de un mismo elemento lingüístico, cuando esta no genera un cambio en el significado semántico, de tal manera que la sustitución de un elemento por el otro no conlleva cambio en el significado semántico de lo enunciado.

### 1.3.2. Las tres olas de la sociolingüística

De acuerdo con Eckert (2012), desde sus inicios, la sociolingüística ha atravesado por tres olas, es decir, tres enfoques o perspectivas de análisis. Cada una de estas olas representa un enfoque distinto, una forma diferente de concebir y estudiar los fenómenos sociolingüísticos que, incluso, pueden llegar a complementarse mutuamente, si el objetivo de la investigación lo requiere. Las características de cada ola se exponen en los siguientes incisos (basado en Eckert 2012):

- a. La primera ola del variacionismo inició en 1966, con el trabajo de William Labov en los grandes almacenes de Nueva York, donde estudió las diferentes realizaciones de /r/ en coda silábica. Esta primera ola estableció una amplia correlación entre las variables lingüísticas y las categorías macrosociales de clase socioeconómica, sexo, etnia y edad. Los estudios dentro de este marco identificaron un patrón regular en la estratificación sociolingüística de las formas lingüísticas, pues se observó una gran diferenciación regional y étnica en el nivel más bajo de la jerarquía socioeconómica, al igual que un gran uso de formas no estándares. Las formas estigmatizadas en una comunidad de habla decrecen en frecuencia a medida que se asciende en la jerarquía de clases. Estos trabajos aportaron un nuevo empirismo cuantitativo a la lingüística, ya que, al agrupar a los hablantes de acuerdo con características macrosociológicas, cada uno de ellos pasó a ser definido a través de las características demográficas que representaba, lo que permitió el empleo de una metodología empírica. Una noción central de este primer enfoque era la noción de lengua vernácula, definida entonces como la primera producción lingüística que se realiza de forma automática y que, por lo tanto, es altamente sistemática. La lengua vernácula se posicionó como el objeto principal de la investigación, pues se entendía que no era afectada por correcciones socialmente motivadas ni por las reflexiones lingüísticas de los individuos. El estilo, por su parte, se entendía como una producción resultante de la atención variable que se prestaba al discurso. La primera ola percibía el cambio lingüístico como el resultado de las presiones en el sistema lingüístico que, primero, afectaban el habla de aquellos menos influenciados por la lengua estándar y que se esparcía hacia afuera a través de la población con mayor resistencia al cambio. Las variables lingüísticas indicaban estatus socioeconómico, mientras que las dinámicas de estilo y género eran vistas como el resultado de los efectos de estas categorías en la orientación de los hablantes hacia sus posiciones asignadas en la jerarquía.
- b. La segunda ola inició en la década de 1980 con el estudio de Lesley Milroy acerca de la variación fónica en redes sociales en Belfast; y se destaca por el empleo de un enfoque etnográfico. Esta ola aportó una perspectiva local a los estudios, lo que estableció la conexión entre las categorías macrosociológicas de la primera ola y categorías locales más concretas, junto con una configuración que dio a las segundas,



un significado. Los estudios bajo este enfoque se centran en categorías de hablantes aparentemente estáticas; y establecen una relación de igualdad entre la identidad de los hablantes y la afiliación a una determinada categoría. La segunda ola le atribuye un papel social al uso vernáculo, asociado con una expresión de identidad local o de clase. Desde esta perspectiva, se concibe la variación como el resultado de un complejo estilístico más amplio, que incluye el territorio y el rango total de consumo. Por otra parte, se entiende que las variables lingüísticas no indexan categorías sociales (subgrupos de una sociedad), sino características sociales (como el grado de instrucción, nivel de ingresos, grupo etario, entre otros).

- c. La tercera ola, que cuenta con algunas características que ya se habían observado en Labov (1963), inicia formalmente en la primera década del siglo XXI. Su atributo principal fue el cambio en la perspectiva etnográfica de variación, como reflejo de identidad social de la práctica lingüística, en la que los hablantes se sitúan en el panorama social a través de la práctica estilística. Este modelo posee tres características: i) sostiene que la variación constituye un sistema semiótico social robusto, que tiene el potencial de expresar el rango total de asuntos sociales en una comunidad dada; ii) plantea variables más específicas (en vez de características demográficas generales) de acuerdo con los estilos empleados en distintas situaciones sociales; iii) percibe la variación no solo como un reflejo, sino también como un elemento constructor del significado social, por lo que es una fuerza en el cambio social. Mientras que en las otras se concibió el significado de la variación como un resultado incidental del espacio social, la tercera ola la ve como un rasgo esencial del lenguaje, debido a la multiplicidad de recursos disponibles para generar un mismo enunciado. Por lo tanto, la variación constituye un sistema social semiótico capaz de expresar el rango total de las necesidades comunicativas. Debido a que estas necesidades cambian constantemente, las variables no pueden ser marcadores consensuales de significado estático, sino que el propósito de los hablantes, manifestado a través de sus prácticas estilísticas, produce cambios semiótico-sociales, reinterpretando variables y combinándolas en un continuo proceso de construcción. El orden indexical es determinante en la mutabilidad de los signos indexicales. En un punto inicial, un grupo social (o una comunidad entera) puede resaltar entre los demás

hablantes de una lengua por su uso de determinados elementos lingüísticos; y sus rasgos distintivos pueden atraer la atención. Una vez reconocidos los rasgos distintivos, estos pueden ser transformados en índices de membresía en esa población. Además, un índice puede ser evocado en movimientos ideológicos relacionados con la población, identificado con formas de pertenencia o características asociadas a ella. El índice también puede ser empleado por personas ajenas al grupo, para señalar estereotipos asociados con este, al igual que puede ser utilizado por miembros del grupo para hacer una distinción entre ellos. Por lo tanto, el orden indexical no es lineal, sino que progresa simultánea y gradualmente en múltiples direcciones, permaneciendo subyacente a un repertorio de significados relacionados que pueden constituir, en cualquier momento particular, un campo indexical, es decir, una constelación de significados ideológicamente conectados que pueden ser invocados en contexto, de forma individual. El enfoque que esta ola hace en el estilo lleva más allá de las variables regionales no estándares, que han sido el punto focal de las otras dos olas. Los estudios de la tercera ola a menudo comienzan investigando la forma en la que los estilos son distintivos, en un intento de identificar los recursos y significados que le dan vida social al lenguaje.

### 1.3.3. Las variables en la sociolingüística

Uno de los elementos característicos de la sociolingüística es la variable sociolingüística. Hernández Campoy y Almeida (2005: 46) definen esta variable como “aquellos rasgos lingüísticos que presenten variabilidad de formas en su uso y cuyas variantes denoten un significado social y/o contextual”. Asimismo, los autores afirman que la selección de las variables de estudio se puede justificar de acuerdo con el grado de significación social reflejada en el uso del elemento estudiado y el grado de diferenciación lingüística implicada. Por su parte, Silva-Corvalán (2001: 95) indica que la variable sociolingüística es “la covariación entre fenómenos lingüísticos y sociales” y presenta una discusión relacionada con la pertinencia de incluir en este tipo de variables aquellas que no presentan estratificación social. Esta autora señala que los patrones característicos de la covariación permiten diferenciar entre tres tipos de variables sociolingüísticas (previamente definidos por Labov, 1972). La primera de estas tres variables mencionadas por la autora son los indicadores, que

covarían de forma regular con el grupo étnico, generacional, socioeconómico, etc., pero no presentan variaciones con respecto a la situación ni el estilo, por lo que varían directamente con el hablante. La segunda son los marcadores, que consisten en variables con posibilidad de estar estratificadas estilística y socialmente, por lo que se considera que varían de acuerdo al hablante y al uso que este hace de la lengua. Por último, el tercer tipo de variable son los estereotipos que la comunidad reconoce conscientemente como tales, pero no corresponden necesariamente al uso lingüístico de los hablantes, además de que se considera un rasgo definitorio de un grupo social.

Otra variable empleada en esta disciplina, que bien puede ser la piedra angular del variacionismo, es la variable lingüística, consistente en un elemento lingüístico abstracto que puede manifestarse de distintas maneras, al punto que podría considerarse como formas alternas de decir lo mismo (Moreno Fernández, 2009; Silva-Corvalán, 2001; Meyerhoff, 2006). Moreno Fernández (2009:21); explica que las formas que puede tomar una variable en una emisión de habla reciben el nombre de variantes y las define como “cada una de las manifestaciones o expresiones de una variable”. Por su parte, Meyerhoff (2006) conceptualiza las variantes como las realizaciones reales de una variable, análoga a la realización fonética de un fonema. A su vez, Silva-Corvalán (2001:86) agrega que las variantes “son idénticas en cuanto a su valor referencial o de verdad, pero se oponen en cuanto a su significado social y/o estilístico”. Finalmente, Hernández Campoy y Almeida (2005) indican que la variable lingüística debe tener al menos cinco propiedades principales para poder ser estudiada en una comunidad: i) relevancia en la comunidad desde la perspectiva de la variación y del cambio lingüístico; ii) susceptibilidad de ser estudiada en términos de detección, cuantificación y correlación; iii) suficiente frecuencia del fenómeno que permita contar con los casos necesarios para la realización de un análisis estadístico; iv) estructura, como para formar parte de un sistema lingüístico mayor; v) alto grado de estratificación, al correlacionarse con los factores sociales u otras variables lingüísticas.

El estudio de la variable lingüística debe ser exhaustivo, para que el investigador pueda dar cuenta fiel del comportamiento del fenómeno que está estudiando. Para ello, los sociolingüistas se rigen por el principio de responsabilidad (*accountability principle*). Este principio sostiene que todas y cada una de las variantes que forman parte del contexto variable deben ser tomadas en cuenta en todos los contextos de ocurrencia posible

(Tagliamonte, 2006; Silva-Corvalán, 2001), aunque el elemento estudiado no esté presente en la totalidad de los casos, ni tenga la interpretación más sencilla en todos ellos (Meyerhoff, 2006). Dicho de otra forma, los datos de aparición de una variante deben ir acompañados de los datos de ausencia de dicha variable. Esto implica, según explican Tagliamonte (2001) y Moreno Fernández (2009), que no se deben estudiar únicamente las variantes nuevas, interesantes, inusuales o no estándares, sino que, también, se deben estudiar las variantes en todos los contextos en los cuales estas serían posibles, aunque estén ausentes, por lo que es insuficiente enfocar la atención únicamente en lo que se ve o se oye.

Las variables extralingüísticas, tal como las definen Hernández Campoy y Almeida (2005), son los elementos constantes que pueden determinar significación sociolingüística y la posible distribución y condiciones de variabilidad de distintas variantes en una comunidad de habla. El rango de variables extralingüísticas que se pueden emplear en estos estudios es bastante amplio; y la selección de una u otra depende de los objetivos e intereses mismos de la investigación que se está llevando a cabo. Algunas de las variables extralingüísticas (entre las planteadas por Silva Corvalán, 2001; Hernández Campoy, 2005; y Moreno Fernández, 2009) son las siguientes: edad, sexo, nivel de instrucción, clase social, profesión, procedencia, raza, etnia, religión y estilo de vida. No obstante, como se mencionó unas líneas atrás, las variables empleadas en un estudio dependerán de sus características y de la información que se desee obtener.

#### 1.3.4. Prestigio

El prestigio que se atribuye a determinadas variedades de una lengua por parte de sus usuarios en general puede variar de acuerdo con los valores sociales de estos usuarios. De acuerdo con Trudgill (2000), debido a que el lenguaje como fenómeno social está vinculado con la estructura y los sistemas de valores de una sociedad, distintos dialectos pueden ser evaluados de forma distinta por parte de sus integrantes. El autor comenta que el inglés estándar, por ejemplo, tiene mucho más estatus y prestigio que otras variedades, ya que, el primero es altamente valorado por muchas personas y quienes lo hablan pueden hacerse acreedores de ciertos beneficios económicos, sociales y políticos. Otras variedades no estándar, comenta el autor, suelen ser consideradas como una desviación de la norma, originada por causa de la pereza, ignorancia o falta de inteligencia. Debido a que existen variables lingüísticas en una

comunidad de habla que se encuentran en covariación con la clase social (Labov, 1966; Trudgill, 2000), se puede considerar que existe cierta presión social sobre los hablantes para que adquieran las variantes con mayor prestigio, aunque esto no ocurre en todos los casos.

Tanto Labov (1966) como Trudgill (1972) encontraron evidencia de que el habla de la clase trabajadora en las comunidades que estudiaron (Nueva York y Norwich, respectivamente) tiene connotaciones positivas para los hablantes masculinos. Ambos concluyeron que las variantes no estándar de la clase baja podían tener un tipo de prestigio para algunas personas, concretamente hombres, no pertenecientes a la clase trabajadora. Labov (1966) denominó la cualidad de estas variantes como prestigio encubierto (*covert prestige*), debido a que las actitudes favorables hacia estas variantes no estándar no suelen expresarse abiertamente y se desvían considerablemente de los valores sociales más predominantes, de los que todos los hablantes están conscientes (Trudgill, 2000). Esta noción se contrapone directamente a la de prestigio abierto (*overt prestige*), que corresponde con el prestigio otorgado a una variedad por parte de la comunidad y que se asocia a lo correcto, adecuado o normativo (Moreno Fernández, 2009).

En el trabajo de Trudgill (1972) en Norwich, se documentó que las mujeres utilizaban un mayor número de variantes con prestigio social, consideradas correctas y educadas; mientras que los hombres hacían un mayor uso de variantes con menor grado de valoración comunitaria. En su investigación, Trudgill (1972) determinó el grado de prestigio de las variantes estudiadas, tras analizar los resultados de una serie de autoevaluaciones realizadas por sus entrevistados, la cual estaba relacionada con las formas lingüísticas que ellos mismos emplearon durante la recolección de muestras de habla. En estas autoevaluaciones, los hombres reportaron un mayor uso de las variantes menos prestigiosas y un menor número de formas mejor valoradas socialmente, por lo que el autor concluye que las variantes con menor prestigio social tienen un valor oculto (prestigio encubierto) para los hombres de Norwich. El hecho de que los hombres se atribuyeran un mayor empleo de formas menos valoradas, afirma el autor, indica la presencia de un prestigio encubierto de estas variantes entre los hombres de Norwich. Por otra parte, las mujeres reportaron un mayor uso de las formas lingüísticas mejor valoradas y un menor uso de las peor valoradas, lo que resalta notoriamente el prestigio abierto de dichas variantes.

Si bien los trabajos de Labov (1966) y Trudgill (1972) se enmarcan en la primera ola de la sociolingüística, por su perspectiva metodológica y de análisis, la noción de prestigio puede relacionarse con la de indexación (explicada en §1.3.7) presente en la tercera ola. Aunque un elemento lingüístico con prestigio abierto o encubierto no necesariamente aporta significado social, pues es utilizado por personas de distintos grupos sociales (como demostraron ambos autores), sí puede poseer significado relacionado con las características de un grupo social. Por ejemplo, el hecho de que los hombres de las comunidades estudiadas por Labov (1966) y Trudgill (1972) otorgaran cierto valor positivo a rasgos del habla de la clase trabajadora puede deberse a que los hombres no pertenecientes a dicha clase consideraran que estos rasgos estaban relacionados con cualidades como masculinidad, fuerza, agresividad, rudeza o autoridad, entre otros posibles atributos. Este tipo de asociaciones entre elementos lingüísticos y características de grupos sociales puede ser de utilidad para explicar la presencia de variantes más frecuentes en un grupo social en miembros de otro grupo, como es el caso en el análisis sociolingüístico llevado a cabo en la presente investigación.

### 1.3.5. Redes sociales

El concepto de “red social” se originó en propuestas antropológicas a mediados de la década de 1950, aunque su verdadero desarrollo (en antropología social) ocurrió a partir de la década de 1970 (Lastra y Martín Butragueño, 2000). Hasta la década de 1980, los estudios de redes sociales se manejaban bajo dos corrientes principales; por una parte, se encuentran los trabajos centrados en los atributos de las relaciones entre los integrantes de la red (amistosa, familiar, laboral, etc.) y, por la otra, se observan aquellos que analizan las características de los vínculos que constituyen la red y el comportamiento de los miembros que la integran (Blanco Canales, 2000).

Milroy y Milroy (1992) explican que una red social puede visualizarse como un tejido de vínculos que se extienden a través de una sociedad, conectando unas personas a otras, aunque en ocasiones sea de forma remota. De acuerdo con Blas Arroyo (2004), las redes sociales representan modelos de interacción de los individuos que hacen vida en una comunidad, en los que se toman como punto de partida las relaciones reales y potenciales con otros miembros de la sociedad, con base en los vínculos de una persona central. Una red

social, según Lastra y Martín Butragueño (2000), actúa como un mecanismo de intercambio de bienes y servicios que impone obligaciones y otorga derechos a sus miembros, por lo que la emisión de un mensaje debe representar un costo menor al beneficio obtenido a través de él.

Asimismo, señalan que las redes sociales poseen una densidad y multiplicidad determinadas; y explican que la densidad consiste en el resultado de la división del número de relaciones reales entre el de relaciones posibles, lo que arroja una aproximación cualitativa; mientras que la multiplicidad es la razón entre el número de vínculos múltiples y el número de vínculos reales entre individuos. En redes con mayor densidad y multiplicidad, afirman Milroy y Milroy (1992), todos los individuos conocerían a todos los demás (densidad) y los actores se conocerían unos a otros en un rango de capacidades (multiplicidad). Adicionalmente, las redes con vínculos estrechos, que varían en el grado en el que se aproximan a una red idealizada de máxima densidad y multiplicidad, tienen la capacidad de mantener y reforzar las convenciones y normas locales, incluyendo las normas lingüísticas; incluso, pueden proveer los medios de oposición en contra de valores institucionales dominantes y la norma lingüística estandarizada.

Además de las cualidades de densidad y multiplicidad, Milroy y Milroy (1992) agregan que, para complementar el análisis de la relación entre redes sociales, se necesita una teoría social como la del modo de vida (véase §1.3.6), que ayude a conectar un análisis de red basado en subgrupos dentro de la sociedad, con un análisis de estructura social a nivel político, institucional y económico. La información obtenida a través de la incorporación de dicha teoría es relevante en cuanto a que puede ayudar a establecer la relación de poder entre las redes analizadas. Esto se debe a que, exponen los autores, las redes poderosas tienen la capacidad de imponer sus normas lingüísticas y culturales sobre las otras; mientras que aquellas menos poderosas solo pueden usar los recursos de la red para mantener y renovar, de la mejor forma posible dentro de sus facultades, sus propias normas lingüísticas y culturales.

Un estudio sociolingüístico basado en redes sociales es provechoso debido a que permite explicar el sentido de los intercambios y comportamientos lingüísticos cotidianos (Lastra y Martín Butragueño, 2000). De igual modo, señalan Milroy y Milroy (1992), el acercamiento a grupos sociales de hablantes desde una perspectiva de redes, en vez de una

de clases sociales, puede determinar la relación entre variabilidad lingüística y estructura social, al tomar en cuenta el papel que cada miembro tiene dentro de la red, su conexión con otros integrantes de la red (y con personas externas) y la fuerza de los vínculos que los unen.

### 1.3.6. Modo de vida

De acuerdo con Milroy y Milroy (1992), el diseño del concepto de modo de vida es atribuido al antropólogo marxista danés Thomas Højrup, quien tenía una perspectiva de clase social más consistente con el análisis de redes. En su trabajo, Højrup (1983) presenta las redes sociales como un proceso de gran escala, con raíces económicas, que divide a la población en subgrupos que comparten ciertas características económicas y sociales; al igual que un modo de vida que emerge de la división que vagamente se podría considerar una diferencia de clases sociales. El análisis de Højrup (1983) se basa en estudios etnográficos llevados a cabo en Dinamarca y a un análisis extenso de la estructura social y económica de otros países de Europa occidental. Este autor propone una división poblacional basada en subgrupos, los cuales son descritos en términos de tres modos de vida. El origen de estos subgrupos no es social o culturalmente arbitrario, ya que son el resultado de estructuras sociales fundamentales, que dividen a la población en modos de vida esencialmente diferentes (Milroy y Milroy, 1992). El análisis de Højrup (1983) propone una integración de los conceptos de redes y clases sociales, ya que los diferentes tipos de estructuras de redes pueden percibirse como el resultado de las diferencias entre los distintos modos de vida.

Los tres modos de vida propuestos por Højrup (1983) se pueden resumir de la siguiente manera:

- a. Modo de vida 1: el modo de vida del autoempleado, con vínculos cercanos centrados en la familia, poca distinción entre el trabajo y el ocio y una fuerte ideología de solidaridad. El negocio familiar puede ser la agricultura, pesca, tienda pequeña o restaurante. Las relaciones sociales en forma de vínculos familiares o relaciones cooperativas entre colegas unen a los productores en una unidad de producción cohesiva, cuya principal preocupación es mantener la producción andando y que la familia y otros productores afiliados formen parte del proceso. El propósito del negocio es mantener el estatus de autoempleo. El concepto de tiempo libre tiene poco valor, debido a que el productor trabaja por sí mismo, para ganar independencia; por



lo tanto, los conceptos de ocio y trabajo tienen un significado diferente del que tendrían para un asalariado.

- b. Modo de vida 2: asalariados incorporados en largos y complejos procesos de producción que no poseen ni controlan. El propósito del trabajo es proveer un ingreso que permita tener una vida satisfactoria durante el tiempo libre del trabajador. La familia está separada de las actividades laborales del asalariado y es el marco en el que se desarrollan las actividades fuera del trabajo. El trabajador carece de compromiso con el puesto y lugar de trabajo y está preparado para vender su labor a quien provea un estímulo adecuado. La solidaridad que surge entre los asalariados que ganan poco suele reflejarse bajo la forma de sindicatos. En su comunidad de residencia, la solidaridad se ve representada a través de redes estrechas, como la de la sociedad de la clase trabajadora tradicional. No obstante, si las condiciones políticas y económicas les permiten a los trabajadores sentirse seguros acerca de sus planificaciones para el futuro, si ganan lo suficiente para volverse móviles y ascender socialmente, la solidaridad puede verse afectada y los vínculos debilitarse.
- c. Modo de vida 3: asalariados profesionales o con labores gerenciales que tienen un alto nivel de habilidad en áreas específicas. La habilidad es una mercancía en venta y al asalariado se le paga para organizar, monitorear y controlar el proceso de producción. Las metas de estos asalariados son ascender en la jerarquía y obtener control a través de papeles profesionales y gerenciales para tener, progresivamente, más poder y escapar del control de los demás y, así, manejar recursos y ejercer el poder por cuenta propia. Esto demanda una inmersión en el trabajo, actitud competitiva hacia los colegas y un desvanecimiento del límite entre el trabajo y el ocio. La familia y su modo de vida cumplen un papel de apoyo con relación a la carrera. El trabajo es la actividad principal en la vida, con la importancia más alta; y el concepto de libertad no corresponde con el tiempo libre, sino con la situación laboral y la perspectiva de carrera. Los asalariados serán social y geográficamente móviles, ya que siguen su carrera, por lo que forman múltiples vínculos débiles, particularmente en el campo profesional, a través de los que transmiten innovación e influencia. Sin embargo, también forman vínculos estrechos y coaliciones en sus

redes personales, a través de las que pueden controlar un considerable número de recursos.

Milroy y Milroy (1992) resaltan que el concepto de modo de vida, al igual que el de red, es estructural, en el sentido de que las características ideológicas y culturales de un modo de vida en particular son determinadas por su contraste con otros modos de vida en la formación social, lo cual implica que la interrelación entre los tres modos de vida y las prácticas culturales asociadas con cada uno de ellos tomarán diferentes formas de acuerdo con el país en el que se observen. En este sentido, Lastra y Martín Butragueño (2000) proponen una modificación de los tres modos de vida originales planteados por Højrup (1983), para que puedan ser aplicables en sociedades latinoamericanas. La propuesta de los autores considera que el modelo original no corresponde con las características de las sociedades latinoamericanas, lo que representa un problema práctico al momento de aplicar la concepción original del concepto a los estudios sociolingüísticos en América Latina. Según los autores, el modelo de Højrup (1983) es uno de consenso, en el que se contempla un continuo social, a través del cual, los individuos pueden desplazarse, lo que quiere decir que los hablantes pueden ascender en la escala sociolingüística. El inconveniente que se presenta en Latinoamérica, explican Lastra y Martín Butragueño (2000), es que en esta zona existe un grado de marginalidad, lo que crea una polémica cultural que se debate entre si los marginados forman una comunidad con valores que difieren de la sociedad general o no, pues “se ha observado que viven en zonas determinadas, que su reivindicación fundamental es el espacio antes que la producción, que tienden a solicitar asistencia antes que derechos, que están fuertemente identificados con los barrios, [y] que desconfían de las instituciones” (pp. 25).

Por lo tanto, Lastra y Martín Butragueño (2000) explican que, bajo estas condiciones, lo más apropiado es plantear el modo de vida bajo un modelo de conflicto (al igual que Milroy y Milroy, 1992), en el que se plantea la existencia de varios grupos sociolingüísticos que se oponen entre sí y que cuentan con fuertes barreras entre ellos, lo que elimina la necesidad de trasladarse de unos grupos a otros. Esto quiere decir que un modelo basado en tres modos de vida puede resultar muy general en el caso concreto de las sociedades latinoamericanas, por lo que Lastra y Martín Butragueño (2000) plantean un modelo con ocho modos de vida,

propuestos especialmente para el caso concreto de la Ciudad de México (comunidad que conforma el eje central del problema social planteado en su trabajo y en la presente investigación):

- a. Modo de vida A, la tradición semirural: se conservan parcialmente estructuras de producción agrarias, mundos culturales agrícolas y redes sociales más típicas de los pueblos que de las ciudades.
- b. Modo de vida B, el empleo informal: es la cara más productiva de la marginalidad, presenta despegue de las instituciones sociales, escapa de las normas municipales, hacendarias e incluso policiales. Son autoempleados (por ejemplo, los vendedores ambulantes).
- c. Modo de vida C, los pequeños comerciantes: pequeñas unidades de producción básicamente familiares, que atienden negocios de barrio (tiendas de abarrotes, talleres mecánicos, tiendas de autopartes, estéticas, panaderías, etc.).
- d. Modo de vida D, los subempleados: empleados por otras personas que separan el trabajo de otras esferas vitales y no tienden a agruparse profesional o gremialmente ni a hacer reclamos ante las instituciones.
- e. Modo de vida E, los obreros: empleados formales del sector secundario (fábricas).
- f. Modo de vida F, los burócratas: empleados formales del sector terciario (oficinas).
- g. Modo de vida G, los cuellos blancos: empleados que hacen carrera dentro de empresas o instituciones que les permiten tomar decisiones, pero en las que se mueven a través de alguna clase de escalafón, como aquellos con cargos de gerencia en una empresa, por ejemplo.
- h. Modo de vida H, los profesionales liberales: trabajadores que son un poco más independientes. Toman decisiones acerca de su trabajo, pero no forman parte de empresas. Si tienen subordinados, estos son personal de apoyo (con labores de asistencia o secretariales, por ejemplo), pero no existe propiamente una estructura de mando.

Lastra y Martín Butragueño (2000) aclaran que no pretenden afirmar categóricamente que los modos de vida que proponen sean todos los existentes en la Ciudad de México, que los límites entre ellos sean claramente reconocibles ni que sea imposible reagruparlos de alguna otra manera. No obstante, consideran que estos modos de vida representan figuras

sociales, redes e individuos fácilmente reconocibles para muchos de los habitantes de la zona metropolitana de la Ciudad de México. De igual manera, estos autores afirman que el análisis de los modos de vida no debe ser considerado como una alternativa al trabajo con redes sociales, sino que los modos de vida son una compensación para reflejar los aspectos macrosociológicos que no se pueden percibir en los estudios de redes sociales, que se enfocan en los aspectos microsociológicos del fenómeno sociolingüístico.

### 1.3.7. Indexación

Antes de entrar en la definición de indexación, es pertinente presentar brevemente lo que se entiende por estilo en sociolingüística. De acuerdo con Eckert (2008), los estilos son niveles de práctica social que corresponden con las distinciones en el terreno en el que se estudia la variación, por lo que se pueden encontrar distintos estilos en diversos grupos sociales, lo que implica que las variables ocurren solo como componentes de los estilos. Según Irvine (2001), los estilos son parte de un sistema de distinción en el que un determinado estilo, junto con su significado social, contrasta con otros estilos posibles. Por tal motivo, las características de un estilo particular no pueden ser explicadas independientemente de otros, sino que se debe prestar atención a las relaciones entre ellos (contrastes y similitudes), para identificar los principios y procesos de las diferencias estilísticas dentro de un sistema sociolingüístico en constante cambio. Irvine (2001) resalta el hecho de que los estilos en el habla involucran las formas en las que los hablantes, como agentes del espacio social (y sociolingüístico), negocian sus posiciones y objetivos dentro del sistema de distinciones y posibilidades. Los participantes en una comunidad no son observadores completamente objetivos del comportamiento que muestran cada uno de ellos, sino que sus propios actos están profundamente influenciados por sus percepciones e interpretaciones de estos comportamientos. Para Eckert (2008), la práctica estilística, relacionada con la interpretación y producción de estilos, inicia cuando una persona percibe un estilo e identifica al individuo o grupo que lo emplea, lo que establece una correlación entre el significado del estilo y sus usuarios.

Los índices son elementos resultantes de formaciones sociales que forman parte de los distintos estilos presentes en una sociedad; y solo pueden informar acerca de la acción social si funcionan como un signo interpretable (Irvine, 2001). Por lo tanto, señala Irvine

(2001), los índices deben estar incluidos en el entendimiento del mundo social de los participantes y de los recursos semióticos disponibles en él. En palabras de Eckert (2008), los índices se asignan en consonancia con el trasfondo de experiencias previas de contacto estilístico. El significado indexical, explica Silverstein (2003), está conformado por dos componentes: la adecuación indexical de acuerdo con los parámetros contextuales (lo que ya se ha establecido entre los usuarios de los signos, al menos implícitamente, como el contexto al que alude la propiedad de una forma lingüística) y la efectividad indexical en contexto (relativa a cómo los parámetros contextuales parecen emerger por el hecho del uso del signo o elemento indexical en sí mismo). Para Silverstein (2003), la relación de presuposición/vinculación no es lineal o unidimensional, como la línea temporal de un “antes” y un “después” de un evento indexical, sino una relación compleja y mediada que resulta en un objeto llamado “texto en contexto”.

Eckert (2018) señala que el significado de las variables no es preciso o fijo, sino que constituye un campo de significados potenciales, al que llama “campo indexical”. Este término hace referencia a una constelación de significados ideológicamente relacionados, cualquiera de los cuales puede ser activado en el momento en el que se emplea una variable. Dicho campo, explica la autora, es fluido y cada activación nueva tiene el potencial de modificar conexiones ideológicas. Por ende, afirma Eckert (2018), la variación constituye un sistema indexical que involucra la ideología dentro de la lengua y que, a la vez, es parte integrante de la construcción de ideologías. Esto conduce a una nueva empresa sociolingüística, comenta la autora, en la que se estudia la variación como un sistema indexical, tomando como punto de partida el significado, en lugar de los cambios concretos que generalmente han gobernado las variables estudiadas.

Por otra parte, es común en la sociolingüística, afirman Irvine y Gal (2000), el hecho de que una forma lingüística puede indexar grupos sociales, ya que, como parte del comportamiento cotidiano, el uso de una forma lingüística puede convertirse en un indicativo o indexador de identidades sociales y de las actividades típicas de los hablantes. De acuerdo con las autoras, se establece una relación entre características lingüísticas e imágenes culturales más amplias de personas y actividades, debido a que las diferenciaciones lingüísticas se construyen sobre aspectos ideacionales, conformados por nociones que los participantes y observadores usan para enmarcar su comprensión de las variedades

lingüísticas. Las ideologías lingüísticas, explican Irvine y Gal (2000), son sostenidas por todos los participantes en un sistema sociolingüístico, quienes ubican los fenómenos lingüísticos dentro de lo que ellos creen que es un contraste sistemático entre grupos sociales indexados en términos conductuales, estéticos, afectivos y morales.

Desde la perspectiva de Eckert (2008), el trabajo local de indexación que los hablantes llevan a cabo a través de la variación es empequeñecido en importancia por el poder de los movimientos internos del sistema lingüístico macrosocial. Por ello, la autora explica que buscar explicaciones para los cambios en cadena que ocurren en las construcciones de significado cotidianas sería un sinsentido, pero resalta que ignorar lo que las personas hacen con los elementos de los cambios en cadena para construir significado social sería cegarse ante un aspecto de la competencia humana. Por ello, afirma Eckert (2008), la exploración del significado de la variación requiere que se examinen los elementos subyacentes de las generalizaciones que se pueden establecer a simple vista, pues el hecho de que las mismas variables se puedan estratificar con múltiples categorías (género, etnicidad y clase, por ejemplo) indica que sus significados no están directamente relacionados con esas variables, sino con algún elemento común a todas ellas. Dicho de otra forma, las variables indexan, indirectamente, categorías demográficas a través de sus asociaciones con cualidades pertenecientes al constructo de dichas categorías.

La relación entre los marcos de análisis microsociales y macrosociales de cualquier fenómeno lingüístico se puede comprender con mayor detalle al tomar en cuenta el orden indexical. Silverstein (2003), el principal exponente de este concepto, explica que este orden corresponde con los varios niveles de abstracción que pueden surgir de las relaciones entre formas lingüísticas y significados sociales. Silverstein (2003) indica que cualquier índice del orden (independientemente de su número o posición) presupone que el contexto en el que es normativo su uso tiene una esquematización de algún tipo en particular. Adicionalmente, señala, existe una tendencia a que se presente también una vinculación indexical regularmente producida por el uso de un elemento de cualquier orden, como una consecuencia directa del grado de compromiso ideológico que los usuarios manifiestan con respecto a la significatividad del orden indexical.

Silverstein (2003) afirma que hay dos tipos de contextos en los que se puede observar la indexación: microsociales y macrosociales. En el primero de ellos, el significado indexical está

conformado por dos componentes: la adecuación (de acuerdo con lo que ya se ha establecido entre los usuarios de los signos, como el contexto al que alude la propiedad de una forma lingüística) y la efectividad en contexto (emergencia de los parámetros contextuales, por el uso del signo en sí mismo). Los contextos microsociológicos están compuestos por distinciones categóricas que constituyen los elementos de una estructura dinámica en desarrollo. El autor considera que estas dinámicas en proceso constituyen eventos interaccionales, definidos como eventos de acción social desarrollados basándose en significados culturales interpretables en la medida en que evocan indexicalmente particiones macrosociológicas del espacio social.

El segundo contexto al que Silverstein (2003) hace referencia, y el verdaderamente relevante para los propósitos de esta investigación, es el macrosocial. El autor define este contexto como los marcos de diferenciación social en términos cualitativos o cuantitativos, convencionalizados o institucionalizados, que están presupuestos por o vinculados a las especificaciones presentes en los microcontextos que se desarrollan durante la interacción. Esto quiere decir que los individuos son identificados socioculturalmente (partiendo de una presuposición) o situados (mediante un vínculo indexical en el uso del lenguaje). En este contexto, las estructuras duraderas de diferenciación categorial están constituidas por los rasgos sociológicos de edad, género, clase social y socioeconómica, profesión y otros aspectos de lo que se considera identidad institucional o posicional.

Johnstone *et al.* (2006), basándose en los planteamientos de Silverstein (2003), proponen un orden indexical de tres niveles. La indexicalidad de primer orden, explican los autores, está relacionada con el tipo de correlaciones entre una forma lingüística y una identidad sociodemográfica o función pragmática observable desde la perspectiva de una persona ajena a un grupo social. Johnstone *et al.* (2006) aclaran que este orden es potencialmente indexical, ya que requiere que alguien se percate de su existencia para ser significativo. El segundo orden indexical tiene lugar cuando los usuarios de una lengua empiezan a utilizar correlaciones del primer orden para realizar acciones sociales (sean interpretativas o performativas). Finalmente, el tercer orden indexical de la propuesta de Johnstone *et al.* (2006) se da cuando una forma es introducida explícitamente en el repertorio lingüístico local, adquiere sus significados y refleja identidades locales. En este tercer orden,

los significados de las formas lingüísticas están cada vez más vinculados a la localidad geográfica.

#### 1.4. Lingüística cognitiva

La lingüística cognitiva, una de las disciplinas lingüísticas en las que se enmarca esta investigación, es una escuela de la lingüística y de la ciencia cognitiva iniciada al principio de los años 80, que da especial importancia al papel del significado, los procesos conceptuales y a la experiencia corporal en el estudio del lenguaje, de la mente y de cómo ambos se entrelazan (Evans 2007). En este apartado se desarrollan las nociones generales de la lingüística cognitiva y las definiciones de los procesos de esquematización, conceptualización y categorización, al igual que la presentación del concepto de prototipo.

##### 1.4.1. Lingüística cognitiva: nociones generales

Tanto Langacker (2008) como Evans (2012) afirman que la lingüística cognitiva es parte de la tradición funcional, ya que, como permite observar Nuyts (2007), algunos lingüistas que trabajan con el enfoque funcional como Chafe, Givón, Slobin, Wierzbicka habían incluido la cognición en sus trabajos desde la década de 1970. Además, agrega el autor, la cognición ha llamado la atención de algunos lingüistas del ámbito funcional en las últimas décadas, como es el caso de la Gramática Funcional (Simon C. Dik) y la Gramática de Rol y Referencia (Robert Van Valin). Nuyts (2007) afirma que la lingüística funcional y la cognitiva tienen algunas perspectivas en común en cuanto a su orientación: ambas emplean un concepto de gramática basada en el uso, comparten la noción de que el conocimiento lingüístico es la competencia para su uso, ven el sistema lingüístico como un subcomponente de la cognición humana y rechazan la visión modular de la facultad lingüística postulada por la tradición generativista. Por lo tanto, asegura Nuyts (2007), se asume que los principios inherentes en el lenguaje son (potencialmente) instantáneos y, más generalmente, cognitivos, puesto que se percibe la gramática como completamente interconectada y sintonizada con otras dimensiones de funcionamiento cognitivo, incluyendo el sistema conceptual. Por su parte, Kristiansen (2006) señala que tanto la lingüística cognitiva como la funcional comparten la idea de que el significado lingüístico no puede ser estudiado de forma aislada, separado de la



naturaleza y el propósito de lo que se comunica y de la dinámica de la comunicación como tal.

En este punto, es pertinente hablar del concepto de *cognición*. Para Evans (2007), la cognición tiene que ver con todos los aspectos de la función mental, conscientes e inconscientes, ya que está relacionada con los fenómenos mentales (mecanismos y procesos) y el conocimiento involucrado en una gran cantidad de tareas, que van desde las percepciones de objeto de bajo nivel hasta las tareas de toma de decisiones. Según Evans (2007), este concepto es el eje central de la lingüística cognitiva, la cual se enfoca en el estudio del lenguaje y la mente (cognición lingüística). Evans (2012) explica que la preocupación por el estudio combinado de la cognición y la experiencia física-social es lo que distingue a la lingüística cognitiva de los demás enfoques lingüísticos contemporáneos.

Langacker (2008) señala que la lingüística cognitiva reconoce completamente las bases del lenguaje en la interacción social, pero insiste en que incluso esa interacción social es críticamente dependiente de la conceptualización de la información obtenida a través de los sentidos. De igual forma, el autor establece que la lingüística cognitiva se resiste a la imposición de barreras entre el lenguaje y cualquier otro fenómeno psicológico, debido a que, en la medida de lo posible, la estructura lingüística recurre a otros sistemas básicos de habilidades (memorización, percepción y categorización, por ejemplo), de los que no se puede ser separada, pues el lenguaje es un aspecto integral de la cognición y no una entidad, módulo o facultad mental distinta e independiente.

Geeraerts y Cuyckens (2007) indican que la lingüística cognitiva percibe el lenguaje como un instrumento de organización, procesamiento y transmisión de información; y se enfoca en el análisis de las bases conceptuales y experimentales de la categorización lingüística. Como explican Cuenca y Hilferty (1999), la lingüística cognitiva parte de una rama filosófica denominada “experiencialismo” o “realismo experiencial”, según la cual, el pensamiento va más allá de la abstracción simbólica, ya que el procesamiento cognitivo depende de la estructura global del sistema conceptual, basado en la experiencia corpórea (la percepción), el movimiento corporal y la experiencia física y social. Esto es el resultado, señalan Cuenca y Hilferty (1999), de las propiedades gestálticas del pensamiento, pues los conceptos tienen una estructura global que trasciende la mera suma de bloques conceptuales creados a partir de reglas generales. Esta base teórica es la piedra angular del componente

cognitivo del presente trabajo, pues se entiende que la complejidad del sistema lingüístico en la estructura cognitiva está construida sobre la base de las experiencias de vida que tienen los usuarios de una lengua durante sus interacciones sociales.

Con base en la perspectiva del lenguaje como instrumento de conceptualización, se pueden mencionar los siguientes principios fundamentales (Cuenca y Hilferty, 1999:18):

- a. El estudio del lenguaje no puede separarse de su función cognitiva y comunicativa, lo cual impone un enfoque basado en el uso.
- b. La categorización, como proceso mental de organización del pensamiento, no se realiza a partir de condiciones necesarias y suficientes que determinan fronteras infranqueables entre las categorías cognitivas, sino a partir de estructuras conceptuales, relaciones prototípicas y de semejanza de familia que determinan límites difusos entre categorías.
- c. El lenguaje tiene un carácter inherentemente simbólico. Por lo tanto, su función primaria es significar, es decir, aportar significado a través de un mensaje. De ello se deduce que el componente gramatical está estrechamente relacionado con el semántico: la gramática no constituye un nivel formal y autónomo de representación, sino que también es simbólica y significativa.
- d. La gramática consiste en la estructuración y simbolización del contenido semántico a partir de una forma fonológica. Así pues, el significado es un concepto fundamental y no derivado en el análisis gramatical. En otras palabras, la información semántica contenida en una forma fonológica es introducida en una estructura lingüística que compone un símbolo, lo que implica que el significado semántico es el elemento primario y no el resultante de la gramática.
- e. Se impone una caracterización dinámica del lenguaje que difumina las fronteras entre los diferentes niveles del lenguaje (la semántica y la pragmática, la semántica y la gramática, la gramática y el léxico) y muestra las dificultades e inadecuaciones que resultan de la aplicación rígida de ciertas dicotomías, como la que opone la diacronía y la sincronía, competencia y actuación, denotación y connotación. La gramática es una entidad en evolución continua, “un conjunto de rutinas cognitivas, que se constituyen, mantienen y modifican por el uso lingüístico” (Langacker, 1987: 57).

La metodología de la lingüística cognitiva se maneja bajo dos principios fundamentales que reciben el nombre de compromisos: el compromiso cognitivo y el compromiso de generalización (Evans, 2012). Según Evans (2007), el compromiso cognitivo representa la idea de que los principios de las estructuras lingüísticas deben reflejar lo que se conoce sobre la cognición humana desde otras disciplinas, particularmente, desde las ciencias cognitivas (filosofía, psicología, inteligencia artificial y neurociencia). La autora señala que de este compromiso se desprende la idea de que el lenguaje y la organización lingüística deben reflejar los principios cognitivos generales, en vez de los principios cognitivos específicos de cada lengua. Evans (2012) explica que parte de este compromiso es que las teorías lingüísticas no pueden incluir estructuras o procesos que violen lo que se conoce sobre la cognición humana.

Con respecto al compromiso de generalización, Evans (2007) dice que este se relaciona con la obligación de caracterizar los principios generales responsables de todos los aspectos del lenguaje humano, pues se basa en la premisa de que el lenguaje refleja mecanismos y procesos cognitivos generales. Por lo tanto, afirma la autora, este principio guía a los cognitivistas en la búsqueda de principios comunes de organización en diferentes sistemas lingüísticos, como la fonología, sintaxis, semántica y demás. Evans (2012) aclara que, debido a que algunos enfoques del estudio del lenguaje separan lo que llaman “facultad del lenguaje” en las diferentes áreas mencionadas, suele haber poca base para hacer generalizaciones con respecto a dichos aspectos del lenguaje y a su interrelación. Evans (2012) comenta que, a pesar de que los lingüistas cognitivos reconocen que la separación de ciertos elementos del lenguaje puede ser beneficiosa en ocasiones, no se parte de la suposición de que los subsistemas del lenguaje están organizados de forma significativamente divergente, por lo que el compromiso de generalización implica investigar abiertamente cómo los distintos aspectos del conocimiento lingüístico emergen de un grupo común de habilidades cognitivas humanas y enfocarse en los aspectos lingüísticos comunes.

#### 1.4.2. Esquematización

La esquematización está estrechamente relacionada con el proceso de abstracción mencionado previamente (§1.1.4.2.2). La creación de esquemas consiste en la extracción de los elementos comunes inherentes en múltiples experiencias, para llegar a una concepción

que representa un nivel más elevado de abstracción (Langacker, 2008). Este proceso puede ser llevado a diferentes grados, dependiendo de la diversidad de elementos en los que se base. Por ejemplo, el sentido básico de “estrella” (‘cuerpo celeste que brilla en el cielo nocturno’) está asociado a un elemento específico en un contexto determinado. Sin embargo, “estrella” también se puede usar para hacer referencia a una persona famosa, una celebridad, por lo que se le puede atribuir un valor esquemático diferente con respecto al básico. Este nuevo valor esquemático constituye una instancia, es decir, una variante más específica de un esquema determinado dentro de una red estructurada de esquemas (Evans, 2007; Langacker, 2008). Para Evans (2007), la esquematización es un tipo especial de abstracción del que provienen las representaciones menos detalladas que las producciones lingüísticas reales de las que surgen. Este proceso es una de las habilidades cognitivas más relevantes, debido a que se refiere al grado de detalle de la especificación cualitativa de un concepto en la mente con relación a otros conceptos (Tuggy, 2007).

El resultado del proceso de esquematización son los esquemas, definidos por Tuggy (2007) como una abstracción conceptual, en cuya construcción se emplea un grado de información variable. Grady (2008) los define como patrones mentales asociados con clases más amplias de patrones recurrentes conceptuales o de experiencias corpóreas. Evans (2007) afirma que los esquemas son alcanzados al dejar de lado las diferencias entre las estructuras reales y tomar solo los aspectos comunes, moldeados en forma jerárquica en términos más abstractos. Por ejemplo, en español, al pensar en el concepto de acento tonal ascendente L+H\* (que representa un ascenso tonal que inicia en la sílaba tónica), se incluye en él la información necesaria para agrupar varios tipos de ascensos melódicos que conforman el repertorio entonativo de una lengua y, al mismo tiempo, se excluye información para diferenciarlos de otros acentos que contienen la característica ascendente entre sus rasgos distintivos, por ejemplo, el acento tonal L\*+H (que hace referencia a un tono bajo en la sílaba tónica seguido por un ascenso que inicia en la sílaba postónica). Tuggy (2007) también afirma que los esquemas son inherentemente imprecisos y flexibles, por lo que son composicionalmente variables, lo que le permite al lenguaje ser una herramienta efectiva de pensamiento y comunicación.

### 1.4.3. Conceptualización

La conceptualización está relacionada con los procesos con base experimental que sigue la mente al momento de establecer el sentido de los elementos lingüísticos y la imposición del significado (Geeraerts y Cuyckens, 2007). Según explica Evans (2007), el resultado de la conceptualización son los conceptos, o representaciones, que son la unidad de conocimiento central de la categorización. La conexión entre la experiencia de la realidad cultural es la que resulta de especial interés en este trabajo, pues se considera que, a través de la conceptualización, los rasgos prosódicos no solo adquieren significado lingüístico, sino que también obtienen el significado social que los convierte en índices sociolingüísticos.

Para Langacker (2008), la conceptualización es un proceso inherentemente dinámico, debido a que se basa en el procesamiento mental que toma cierto tiempo para llevarse a cabo. De hecho, Evans (2007) afirma que los conceptos son inherentes al sistema cognitivo, desde etapas tempranas en la infancia, punto desde el cual son redescritos de acuerdo con la experiencia y mediante procesos de análisis de significado perceptual, lo que da paso a los conceptos más básicos, también llamados esquemas de imágenes. Langacker (2008) también comenta que la conceptualización es multifacética y altamente flexible, ya que en el significado de cualquier elemento lingüístico existe inherentemente una forma particular de construir el contenido conceptual, de tal forma que algunas facetas son focalizadas mientras que otras no. Por ejemplo, al momento de conceptualizar un grupo de sonidos del habla o de acentos tonales, dicho grupo se puede considerar homogéneo o heterogéneo, dependiendo de cuál aspecto de la complejidad del grupo se observe.

### 1.4.4. Categorización

El proceso de categorización es fundamental para comprender cómo se agrupan los diferentes esquemas y conceptos prosódicos en el sistema cognitivo de los hablantes. Este proceso consiste en la interpretación de experiencias con respecto a estructuras cognitivas previamente existentes, a través de la comparación de un elemento nuevo con el recuerdo de cada una de las instancias que ya existen en la categoría en la que se va a incluir, por lo que el proceso está basado en la suma de las similitudes entre la nueva instancia y las ya almacenadas (Langacker, 2008; Johnson, 1997). Según Cuenca y Hilferty (1999), la categorización es un mecanismo de organización de la información recibida a través de la

percepción de la realidad, que permite simplificar la complejidad de la realidad a partir de dos procesos complementarios: la generalización y la abstracción. La generalización, explican los autores, consiste en ignorar las diferencias existentes entre dos o más elementos y agruparlos según sus similitudes. Por ejemplo, la clasificación de H+L\* y HL% como movimientos tonales descendentes. Por el contrario, la abstracción o discriminación se fija en los rasgos que distinguen a distintos elementos para evitar confundirlos entre sí (Cuenca y Hilferty, 1999); por ejemplo, la clasificación de H+L\* como acento tonal y HL% como tono de juntura, de acuerdo con la sílaba del enunciado en la que tengan lugar los descensos melódicos. Cuenca y Hilferty (1999) también señalan que la categorización fundamenta los procesos de comprensión y producción lingüística al clasificar la información en categorías cognitivas, que son conceptos mentales guardados en nuestro sistema cognitivo. Por su parte, Langacker (2008) define el concepto de categoría como un grupo de elementos juzgados como equivalentes para algún propósito, como los sentidos alternantes de elementos léxicos.

Kristiansen (2006) explica que la asimilación de un elemento lingüístico a una categoría de prototipo suele verse en términos de relativa similitud con respecto a otro miembro de la categoría en cuestión, sea este periférico o central. Sin embargo, las categorías de prototipo son entidades flexibles en el sentido de que sus fronteras son extensibles, por lo que nuevos elementos pueden ser incluidos en ellas, para que la cognición se adapte al cambio, innovación y descubrimientos en un mundo físico y social complejo. La autora afirma que, en teoría, un nuevo miembro de una categoría es suficientemente similar a otro miembro como para ser aceptado e incluido en ella, pero también debe ser percibido suficientemente diferente, para merecer una nueva subcategorización o extensión y no solaparse con un miembro ya existente. Tanto Langacker (2008) como Tuggy (2007) afirman que la categorización de una unidad de significado puede ser directa, por esquematización completa (Tuggy 2007), si un elemento (A) perteneciente a una categoría es empleado para categorizar otra unidad (B) dentro de la misma categoría, lo que implica que A es esquemática para B y que este segundo representa una instancia (un ejemplo particular) de A. Así, Langacker (2008) y Tuggy (2007) representan esta relación de la siguiente manera:  $A \rightarrow B$ ; lo que indica que B es completamente compatible con las especificaciones de A, pero que es caracterizada con mayor precisión y detalle. Sin embargo, también puede ocurrir que B entre en conflicto con las especificaciones de A, pero se asimila a la misma categoría con

base en una asociación o similitud percibida, por lo que A es entonces un prototipo y B es una extensión de ella, lo que Tuggy (2007) define como una esquematización parcial. La representación que ambos autores proveen para esta relación es la siguiente:  $A \rightarrow B$ .

#### 1.4.5. Prototipo

La noción de prototipo es fundamental dentro del modelo de análisis basado en el uso y, por ende, de esta investigación, puesto que dicho modelo constituye uno de los ejes centrales del análisis de este trabajo y se explicará detalladamente más adelante (§1.5.2). De acuerdo con Geeraerts (1989), la teoría de los prototipos se originó a mediados de la década de 1970 cuando Eleanor Rosch investigó la estructura interna de las categorías. Geeraerts (1989) afirma que hay dos puntos fundamentales acerca de esta teoría. El primero de ellos es que no es necesario contar con un solo conjunto de atributos definitorios que conforman los requerimientos de necesidad y suficiencia. El segundo es que esta teoría se rehúsa a aceptar la idea de que existe una estructura semántica autónoma en los lenguajes naturales, que pueda ser estudiada de forma independiente con respecto a las otras habilidades cognitivas del ser humano.

La teoría de los prototipos sostiene que hay dos principios básicos que guían la formación de categorías en la mente humana: el principio de economía cognitiva y el principio de estructura del mundo percibido (Evans, 2007). El primer principio establece que el ser humano intenta adquirir tanta información como sea posible acerca de su ambiente, empleando un mínimo de esfuerzo y de recursos cognitivos, generando la formación de categorías (formadas por agrupaciones de estímulos similares) en vez de almacenar información separada sobre cada estímulo individual. Por lo tanto, aclara Evans (2007), el ser humano privilegia categorías formadas en ciertos niveles de inclusión y complejidad informativa. Con respecto al principio de estructura del mundo percibido, la autora explica que el mundo alrededor de nosotros tiene correlatos estructurales. Por ejemplo, las ruedas se asocian con una forma circular y con la facilitación del desplazamiento terrestre en vez de con una forma cuadrada y la facilitación del desplazamiento acuático. Evans (2007) señala que este principio establece que los seres humanos confían en estructuras correlacionadas para formar y organizar categorías, lo que da pie a los prototipos.

El prototipo consiste en el ejemplar (representación esquemática relativamente abstracta) más distintivo y representativo de las instancias (o ejemplares específicos) de una categoría, puesto que es el que contiene las características más sobresalientes o centrales asociadas con miembros de la categoría en cuestión (Cuenca y Hilferty, 1999; Evans, 2007). Cuenca y Hilferty (1999) afirman que las categorías son entidades difusas, debido a la gradualidad que existe entre sus límites y la marcadez de los elementos periféricos. Las categorías se organizan en torno a una imagen central (prototípica) de su miembro más representativo, por lo que la categorización de un elemento se realiza a partir del grado de similitud con el prototipo, aunque esta similitud no condiciona necesariamente a todos los miembros de la categoría. Al igual que Cuenca y Hilferty (1999), Evans (2007) señala que el prototipo sirve para organizar una categoría determinada, lo que se conoce como estructura de prototipo, y esto crea el efecto de typicalidad, constituido por el juicio que se le da a una instancia por ser más o menos representativa de una categoría específica.

No obstante, el carácter central de un elemento puede ser susceptible al contexto. Bybee (2001) indica que una categoría lingüística determinada no tiene solo un prototipo, sino varios, uno para cada contexto frecuente. Este almacenamiento redundante se origina debido a que los oyentes emplean múltiples pistas acústicas para asignar el estímulo a una categoría. Según Bybee (2001), el proceso de categorización produce categorías complejas, centradas alrededor de diversos prototipos, jerarquizados de acuerdo con su frecuencia de uso, los cuales están representados en un grado de detalle considerable.

### 1.5. Sociolingüística cognitiva

En líneas generales, la sociolingüística cognitiva involucra la implementación de una perspectiva cognitiva en el estudio sociolingüístico, con la intención de explicar la forma en la que los procesos cognitivos se ejecutan en el contexto social, a través de métodos empíricos. En esta sección, se presentan las nociones de esquemas, categorías y prototipos, desde esta perspectiva; se expone brevemente cómo se aborda el estudio de elementos fonológicos en la sociolingüística cognitiva y se explica el papel que desempeña la percepción en los fenómenos sociocognitivos.



### 1.5.1. Sociolingüística cognitiva: definición general

Al partir de la premisa de que el objeto de la lingüística es el estudio del lenguaje como un fenómeno humano, es natural pensar que, eventualmente, algunas de las disciplinas lingüísticas concebidas desde enfoques esencialmente diferentes puedan llegar a converger en determinados aspectos. Esto se debe a que, a fin de cuentas, comparten el mismo objeto de estudio complejo y multidimensional, como es el caso de la sociolingüística y la lingüística cognitiva. Ambas disciplinas comparten, además del lenguaje humano como objeto de estudio, el interés por los esquemas de nivel bajo (menos abstractos) relacionados con la producción real (Kristiansen, 2006). De hecho, según Wolf y Polzenhagen (2009), el giro de la lingüística cognitiva hacia la sociolingüística es parte de una evolución inevitable, debido al principio fundamental de la primera como un enfoque basado en el uso. Precisamente, por tener este enfoque, comentan Kristiansen y Dirven (2008), la lingüística cognitiva está sujeta a estudiar problemas sociolingüísticos como la variación regional y social del lenguaje.

El resultado de la combinación de elementos teóricos y metodológicos de la lingüística cognitiva y de la sociolingüística recibe el nombre de sociolingüística cognitiva, que “siguiendo las pautas generales de la lingüística cognitiva, se preocupa especialmente por el estudio de los recursos cognitivos implicados en el procesamiento y el uso lingüístico contextualizado” (Moreno Fernández, 2012: 17). Al respecto, Püts *et al.* (2012) afirman que la sociolingüística cognitiva explora el uso del lenguaje a través de la convergencia de métodos y marcos teóricos típicamente asociados con la lingüística cognitiva y la sociolingüística, lo que provee numerosas posibilidades que benefician la investigación de la variación a nivel teórico y metodológico, pues la integración de ambas perspectivas representa un importante paso en el aporte de un enfoque genuino. Según estos autores, la sociolingüística y la lingüística cognitiva comparten aspectos en común en cuanto a metodología de estudio de la variación, ya que en ambas disciplinas se presta una importancia especial al uso real del sistema lingüístico y al conocimiento que tiene el hablante sobre dicho uso.

Con base en esta idea, se pueden incluir en el marco de la sociolingüística cognitiva todas aquellas investigaciones que se propongan examinar dimensiones sociales y cognitivas, pues como afirman Kristiansen y Dirven (2008), esta disciplina pretende extender el paradigma cognitivo a los patrones lingüísticos regionales y sociales involucrados en la

simbolización lingüística, para ser estudiados como fenómenos concretos o como un elemento paralelo a la estructura conceptual. Esta extensión del paradigma se debe, según Kristiansen y Dirven (2008), a que así se añade una dimensión social, necesaria para la lingüística cognitiva, al prestar especial atención a la forma en que las diferentes conceptualizaciones, preferencias gramaticales y léxicas y diferencias en la prominencia de connotaciones caracterizan el uso del lenguaje en diferentes grupos regionales y sociales. Esta disciplina propone que un entendimiento más completo de la naturaleza del lenguaje, de acuerdo con su uso, solo es posible si se consideran sistemáticamente, junto con los elementos cognitivos, un rango de factores sociales y culturales que moldean los eventos de uso a través del análisis de la manera en que las normas culturales y comunitarias salen a flote en los eventos de uso individuales; y considera la extensión en la que los factores sociales, culturales y cognitivos se entrelazan (Püts *et al.*, 2012).

De acuerdo con Moreno Fernández (2012), el análisis al que aspira la sociolingüística desde esta perspectiva es relacional, en vez de formalista o funcionalista, pues se busca explicar la dinámica de los componentes de la lengua y la lengua en sí, en relación con su entorno natural, cultural, social y situacional. El objetivo de una sociolingüística cognitiva es “entender y explicar la lengua desde una posición cognitivista, lo que significa atender a lo lingüístico y lo social como realidades analizadas desde la cognición humana.” (Moreno Fernández, 2012: 10). El autor explica que, según este enfoque, la sociolingüística presta especial atención al conocimiento y percepción que los hablantes tienen acerca de la lengua y de su uso en contextos sociales e incorpora información de los entornos comunicativos y los procesos de interacción a las nociones de variación, cambio lingüístico y el modo en que son percibidos. Esto está en consonancia con los planteamientos de Eckert (2008), para quien el acto de hablar en el mundo social involucra un continuo proceso de análisis e interpretación de categorías, grupos, tipos y personas (y de sus diferencias en la producción lingüística), lo que, en términos de cognición social, significa el desarrollo de esquemas que surgen a medida que se identifican diferencias, se establecen distinciones y se les otorga un significado. Por lo tanto, afirma la autora, construimos un panorama social mediante la segmentación del terreno social y construimos un panorama lingüístico, a través de la segmentación de las prácticas lingüísticas en dicho terreno.

Por todo esto, indica Moreno Fernández (2012), la sociolingüística cognitiva se presenta como una metateoría que agrupa propuestas de diversas teorías sociolingüísticas, desde un enfoque dinámico del uso de la lengua (superando la división epistemológica de la sociolingüística tradicional), que establece una adaptación a cada objeto de estudio para proponer un denominador común a todos los análisis sociolingüísticos. Esta agrupación de propuestas teóricas se ve favorecida por postulados de la lingüística cognitiva pues, en este marco, se asume que las conceptualizaciones basadas en la cultura subyacente al lenguaje o a la variación lingüística son sistemáticas y pueden ser estudiadas a través de métodos y herramientas teóricas desarrolladas desde el enfoque cognitivo (Wolf y Pozenhagen, 2009). Desde esta perspectiva, señala Moreno Fernández (2012), la sociolingüística cognitiva se preocupa por estudiar los entornos socioculturales y comunicativos en los que se produce el lenguaje, al igual que la forma en la que dichos entornos influyen en las manifestaciones lingüísticas y en la percepción subjetiva que tienen los propios hablantes, tanto de los contextos en los que usan el lenguaje como de las lenguas en sí.

Moreno Fernández (2012) comenta que la sociolingüística cognitiva también tiene el propósito de explicar cómo se originan los procesos de categorización de los objetos lingüísticos y sociales, por lo que el desarrollo de una sociolingüística cognitiva exige la aplicación de conceptos como los de prototipo, categoría cognitiva, centralidad, ejemplar o construcción, por ejemplo; a partir de los cuales sería posible afrontar los conceptos teóricos de comunidad de habla, clase, agrupación social, variable y variante sociolingüística, entre otros, de un modo diferente al de la sociolingüística más tradicional. Aunado a esto, Kristiansen (2008) establece que existen al menos tres razones para aplicar un marco lingüístico cognitivo a la variación lectal, es decir, a la variación de cualquier estilo de discurso dado, incluyendo las categorías tradicionales referidas como variedades estándar, dialectos regionales, sociolectos, basilectos, acrolectos, registros y estilos. La primera es porque la lingüística cognitiva es un modelo basado en el uso. La segunda se debe a que el estudio de la variación estructurada al nivel de las categorías lectales es altamente relevante para otras áreas de investigación, incluyendo la fonología, morfología, sintaxis y semántica. Finalmente, la tercera es que, de no hacerlo, difícilmente seríamos capaces de responder a un número de preguntas concernientes al vínculo cognitivo existente entre el lenguaje y la sociedad.

### 1.5.2. La tesis basada en el uso

Además de los compromisos primarios, la lingüística cognitiva también presenta un número de hipótesis que surgen de los compromisos esenciales: la tesis de la cognición corporizada, la tesis de la semántica enciclopédica, la tesis simbólica, la tesis de que el significado es conceptualización y la tesis basada en el uso (Evans, 2012). Para fines de la presente investigación, únicamente se desarrollará esta última, pues se toma como base el uso de la lengua en el contexto social en el que se produce.

Bybee (2001), la principal exponente de este modelo de estudio, resume sus principales cualidades teóricas en seis puntos clave. El primero de ellos es que las experiencias afectan las representaciones cognitivas, lo que quiere decir que el uso de formas y patrones, tanto en la producción como en la percepción, afectan su representación en la memoria, razón por la cual las palabras y frases con alta frecuencia de uso tienen una mayor representación en cuanto a que su acceso es más sencillo y son menos propensas a sufrir un cambio analógico. El segundo punto es que las representaciones mentales de los objetos lingüísticos tienen las mismas propiedades que las representaciones de otros objetos, por lo que no tienen propiedades dentro de sí mismas, sino que están firmemente basadas en la categorización de elementos reales. La tercera característica es que las categorías se basan en la similitud entre los esquemas almacenados. El siguiente aspecto consiste en que las generalizaciones de las formas lingüísticas no están separadas de las demás representaciones almacenadas, sino que emergen directamente de las relaciones entre ellas, con base en sus similitudes fonéticas y/o semánticas, por lo que las nuevas formas se pueden producir basándose en formas preexistentes. El quinto punto es que la organización léxica proporciona generalizaciones y segmentación en varios grados de abstracción y generalidad; y debido a que el almacenamiento en este modelo es altamente redundante, los esquemas pueden describir el mismo patrón en diferentes niveles de generalidad. El sexto y último punto consiste en que el conocimiento gramatical es procedimental (pues un hablante nativo puede crear una oración aceptable automáticamente, pero no necesariamente puede explicar cómo la construyó o mencionar cuáles son las propiedades de una oración aceptable), por lo cual, la fonología se convierte en una parte fundamental del procedimiento de producción y

decodificación de construcciones, en vez de permanecer como un sistema psicológico completamente abstracto.

La tesis basada en el uso sostiene que la gramática mental del usuario de una lengua está conformada por la abstracción de unidades simbólicas de instancias, o ejemplares, de usos lingüísticos. Una consecuencia importante de la adopción de una tesis basada en el uso es que no hay distinción rígida entre el conocimiento de un lenguaje y su uso, por lo que el conocimiento del lenguaje es el conocimiento de cómo se usa el lenguaje (Evans, 2012). Según Ellis (2019), este enfoque propone que la competencia lingüística creativa de un individuo emerge del conjunto de los recuerdos de todas las experiencias de interacción significativa en su historia como usuario de la lengua, pues la percepción reiterada de una estructura lingüística da como resultado su emergencia. El autor afirma que esto empieza a ocurrir desde que se comienza a aprender la lengua, pues el proceso de aprendizaje de una lengua involucra el aprendizaje de asociaciones entre construcciones que, a su vez, implican el análisis distribucional de la lengua y el análisis paralelo de actividad perceptual, por lo que las construcciones abstractas son aprendidas desde la percepción de ejemplares concretos de uso, siguiendo mecanismos de aprendizaje que relacionan el *input* con la cognición del sujeto.

Kristiansen y Dirven (2008) señalan que una lingüística basada en el uso toma como base de su estudio el lenguaje tal como es empleado por hablantes reales, en situaciones reales, en un momento histórico específico. Este modelo, de acuerdo con Croft y Cruse (2004), contrasta con el generativismo, debido a que en dicho modelo únicamente la estructura de las formas gramaticales determina su representación en la mente del hablante. En cambio, en el modelo basado en el uso, las propiedades del uso lingüístico en la comunicación determinan la representación mental de unidades gramaticales en la mente del hablante. Esta representación mental se genera en forma de redes, cuya redundancia, cuando la hay, se percibe en términos de organización jerárquica de relaciones ejemplar-esquema sostenidas entre unidades simbólicas más esquemáticas y más específicas (Evans, 2012).

Püts *et al.* (2012) explican que este enfoque basado en el uso no solo constituye un repositorio de significados asociados con formas concretas que se emplean cotidianamente, sino que el lenguaje también es el producto del uso diario de la lengua y que se sustenta en las experiencias de los usuarios, debido a que el significado tiene una naturaleza experiencial y la variación en el significado (incluyendo el cambio de este) emerge a través de inferencias

en los eventos de uso lingüístico. Por ende, afirman los autores, el análisis de corpus desde esta perspectiva implicaría la investigación sistemática de datos lingüísticos producidos de forma natural, puesto que los métodos empíricos (especialmente los estadísticos) son herramientas necesarias para comprobar la veracidad de los constructos psicolingüísticos de ciertos fenómenos con base en su uso. En este enfoque, Harder (2012) *dixit*, se considera que el uso es un agente impulsor del cambio lingüístico dentro del ciclo evolutivo de la lengua, en el que el correlato sincrónico es el eje primario en el entendimiento del lenguaje, por lo que se le asigna un papel focal al uso individual y local en oposición a constructos generales.

Según Evans (2012), las unidades simbólicas que se almacenan en la mente del hablante, denominadas rutinas mentales, surgen a través de procesos de abstracción y esquematización, basados en patrones de reconocimiento; y consisten en emparejamientos convencionales de forma y significado. Estas unidades simbólicas presentan grados de afianzamiento variables para diferentes usuarios, cuya frecuencia relativa de uso afectará la naturaleza de la gramática, lo que implica que las unidades simbólicas con mayor frecuencia de uso (consideradas prototípicas en este enfoque) se afianzan más; y estas moldean el sistema lingüístico en términos de patrones de uso (Bybee, 2001). Por otra parte, Langacker (2008) define este proceso como el dominio de una estructura compleja, a través de la repetición y el ensayo, al punto de que su uso se vuelve virtualmente automático y requiere poco monitoreo consciente. El autor explica que una estructura pasa a través de un afianzamiento progresivo y, eventualmente, se convierte en una unidad establecida.

La importancia del método basado en el uso, afirman Püts *et al.* (2012), yace en el hecho de que puede explicar fenómenos que son problemáticos para los enfoques basados en reglas; pues los modelos basados en el uso (o modelos basados en ejemplares) conllevan la idea de que la producción y comprensión del lenguaje humano operan sobre la base del almacenamiento de experiencias lingüísticas concretas, en lugar de reglas lingüísticas abstractas. De este modo, las vivencias lingüísticas de un individuo son las que determinan la forma en que se almacena la información lingüística en el sistema cognitivo. Harder (2012) comenta, desde un punto de vista fundamentado en la evolución, que la lingüística basada en el uso provee una base teórica sólida para la idea clave de que el nivel de (re)producción individual de enunciados es el fenómeno básico sin el cual ninguna otra parte del sistema lingüístico sería posible, pero también demuestra por qué los actos individuales de

(re)producción de expresiones lingüísticas no pueden ser entendidas como lo único existente. Adicionalmente, comenta este autor, el que la existencia del lenguaje cambie diacrónicamente requiere necesariamente tanto un nivel de orden colectivo como social para ser posible, pues sin la participación de los miembros de una población sería inverosímil postular que las lenguas cambian, ya que solo habría actos de habla individuales.

### 1.5.3. Esquemas, prototipos y categorías desde la sociolingüística cognitiva

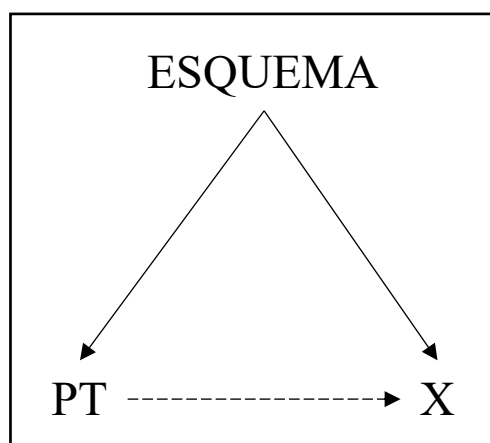
Püts *et al.* (2012) explican que, en la sociolingüística, la categorización de personas, sus relaciones y actividades sociales constituyen la estructura social, con la cual se trata de establecer una relación en términos de variación lingüística, donde el uso del lenguaje de un individuo es determinado por su propia cognición (su percepción de los grupos sociales y las redes alrededor de él). Los autores indican que los individuos perciben la estructura social a su alrededor en términos de categorías (o modelos cognitivos) que muestran todas las propiedades que la lingüística cognitiva predice en términos de miembros centrales (prototípicos), menos centrales y marginales (periféricos). En términos lingüísticos, la identificación de variación alofónica, por ejemplo, se vuelve parte de los estereotipos lingüísticos, en cuanto a que los llamados acentos pueden funcionar como puntos de referencia en la cognición social. Dicho de otra forma, el empleo notoriamente predominante de una variante por parte de un determinado grupo social (en contraste con un reducido o nulo uso de esta por otros grupos) puede crear un estereotipo lingüístico dentro de una sociedad, pues dicha variante pasará a ser asociada directamente con el grupo en cuestión.

Desde la perspectiva basada en el uso, Moreno Fernández (2012: 96) establece que los usos que las personas hacen del lenguaje dan lugar a los esquemas de percepción y los define como “un conjunto estático de elementos almacenados y categorizados a partir de un proceso de generalización de unidades lingüísticas”. Este autor indica que el almacenamiento se lleva a cabo mediante un sistema de conexiones léxicas, originado a partir de rasgos fónicos y semánticos y, dependiendo de la frecuencia de uso de dichas conexiones, estas se asientan como fonológicas o semánticas. De igual forma, Moreno Fernández (2012) explica que un esquema proporciona menos información que los elementos concretos que lo constituyen; y es compatible con una amplia serie de elaboraciones, lo que hace posible que se organicen jerárquicamente de acuerdo a su grado de esquematicidad y admitan un rango

variable de variaciones. En la prosodia, por ejemplo, se puede entender el acento tonal ascendente (L+H\*) como un esquema que posee menos información que sus realizaciones y que puede estar constituido por varios elementos concretos, como un ascenso mayor a 3 st (L+;H\*) o el pico tonal desplazado a la sílaba postónica (L+>H\*), entre otras posibles opciones que cuentan con mayor o menor grado de detalle (L+>;H\*, L+<H\*, L+<;H\*).

Moreno Fernández (2012: 97) define los prototipos como “ejemplos típicos de una categoría determinada, de modo que otros elementos podrían asimilarse a la categoría general dependiendo de la similitud percibida respecto al del prototipo”. La conexión establecida entre esquemas y prototipos a través de la percepción de la realidad se establece de acuerdo con el grado en que un elemento X puede asimilarse a una categoría definida por un prototipo PT, de modo que la completa adecuación de X a las especificaciones de PT permitiría reconocerlo como un elemento prototípico de la categoría. De no existir conflictos en la asociación resultante de la conexión perceptiva de variables y variantes lingüísticas de una variedad lingüística (o lecto), el prototipo (PT) se interpretaría como una proyección de un esquema compatible con el elemento X. La relación descrita previamente entre esquema y prototipo puede ejemplificarse en el diagrama presentado en la figura 1.7.

Figura 1.7. Esquema y prototipo (tomado de Moreno Fernández, 2012:97)



En la figura superior, se puede ver ejemplificado que, para validar la pertenencia de un elemento X a un esquema, se contrastan sus cualidades con las del prototipo (PT) para determinar si existe una adecuación completa entre ambos. En caso de que se pueda establecer exitosamente una relación, el elemento en cuestión puede incluirse en el esquema



definido por PT. Esta inclusión se llevaría a cabo con base en la similitud de rasgos compartidos por ambos elementos.

#### 1.5.4. Sociofonología cognitiva

Con el inicio de la sociolingüística se empezó a estudiar sistemáticamente la correlación entre las variables sociales y las variables fónicas (véase Labov, 1963, 1966; Wolfram, 1969; Cedergren, 1973; Trudgill, 1974; y Macaulay, 1977, entre otros). La lingüística cognitiva, en cambio, no ha tomado en cuenta el hecho de que las expresiones fónicas pueden variar de acuerdo con la influencia de distintas variables sociales, pero ha aportado al conocimiento lingüístico importantes nociones teóricas como las de esquema, prototipo y categoría cognitiva. Kristiansen (2006) explica que, en una categoría de prototipo fonémico, las variantes se agrupan alrededor de un elemento central, que contrasta totalmente con respecto a los miembros centrales de categorías circundantes. Por lo tanto, de acuerdo con este planteamiento, se podría pensar que las categorías lingüísticas de fonemas estructuradas alrededor de elementos centrales productiva y perceptivamente contrastantes conforman el repertorio fonémico de las lenguas humanas. Sin embargo, Kristiansen (2006) afirma que este planteamiento no explica la posibilidad de variación fónica en la lengua.

El enfoque de la sociolingüística cognitiva ayuda, entre otras cosas, a llenar ese vacío en la lingüística cognitiva. Moreno Fernández (2012) explica que la sociolingüística cognitiva propone la interpretación de los fonemas bajo parámetros cognitivistas de esquema y prototipo. Según el autor, los usuarios de la lengua organizan los rasgos percibidos como acústicamente relevantes en categorías fonéticas discretas, privilegiando la información relativa a la identidad de la categoría y diluyendo la información perteneciente a la forma acústica concreta. En el establecimiento de categorías y prototipos, afirma Moreno Fernández (2012), desempeñan un papel fundamental tanto la frecuencia de las producciones fónicas como la similitud de los rasgos percibidos, pues de acuerdo con ello se generan las posibilidades de variación.

En términos amplios, la variación fónica presente en la forma de hablar de un individuo conforma lo que comúnmente las personas en general (fuera del ámbito de los estudios lingüísticos) denominan “acento”. Los acentos, en palabras de Kristiansen (2006), son socialmente diagnósticos, al grado de que los niños, gradualmente, adquieren

conocimiento acerca de las variedades lectales y aprenden a relacionarlas sistemáticamente con dominios sociales, valores sociales y percepciones estereotípicas, por lo que un hablante puede ser caracterizado socialmente únicamente por su forma de hablar. Esto se debe, explica la autora, a que los patrones de habla adquiridos durante la infancia temprana no se alteran fácilmente, por lo que los estereotipos lingüísticos constituyen un marcador especialmente confiable de la identidad social, ya que las emisiones lingüísticas registran significado social en un nivel subconsciente.

Moreno Fernández (2012) señala que los elementos tradicionalmente denominados “fonemas” constituyen esquemas cognitivos (realidades cognitivas emergentes constituidas a partir de los ejemplares concretos en el habla contextualizada), mientras que los prototipos son categorías fónicas perceptivas (conjuntos de variantes cercanas que se agrupan por su similitud), cuya manifestación relevante depende del contexto (comunicativo, social, lingüístico, etc.). Esta relación entre las categorías fónicas y el contexto de manifestación ofrece una explicación teórica sobre los mecanismos cognitivos que dan lugar a la variación lingüística, puesto que se plantea que distintas circunstancias sociales o situacionales pueden generar la emisión de una u otra forma del prototipo de una categoría. Para Kristiansen (2006) existen dos procesos universales que explican la variación: “forticiones” y “leniciones”. Las forticiones (*fortitions*) seleccionan, de entre todos los sonidos posibles en el lenguaje humano, aquellos que constituyen fonemas de un lenguaje particular y que definen los efectos de prototipo para los sonidos del lenguaje, lo que ayuda a definir los repertorios fónicos como categorías cuyos miembros prototípicos demuestran un grado absoluto de diferencia perceptual. Por otra parte, las leniciones (*lenitions*) crean variantes alofónicas, es decir, las extensiones de un miembro prototípico dentro de una categoría, lo que explicaría los numerosos procesos de asimilación.

Este cambio de perspectiva con respecto a la naturaleza de los fonemas es positivo para la sociolingüística cognitiva. Moreno Fernández (2012) explica que la ventaja de este enfoque radica en la naturaleza abstracta del fonema que vincula directamente cada manifestación ejemplar con el conjunto al que pertenece, de modo que no es posible manejar la idea de un esquema de forma independiente del conjunto. Esto supone que todo ejemplar es plenamente compatible con un esquema, mientras que la compatibilidad de cualquier

elemento con el prototipo es relativa. Al aplicar esta relación al área de la fonología, ha de aceptarse la siguiente serie de premisas teóricas (Moreno Fernández, 2012: 149):

- a. Los sonidos de una lengua no se limitan a clasificarse en un conjunto cerrado de fonemas, sino que afectan los niveles suprasegmentales, como se viene preconizando desde la fonología autosegmental, prosódica o de dependencias.
- b. Los sonidos de una lengua no son percibidos del mismo modo por todos los hablantes; algunos de ellos, incluso, pasan desapercibidos.
- c. Las variantes fónicas pueden darse tanto en un nivel de consciencia como en un nivel inconsciente.
- d. Los sonidos de una lengua pueden ser considerados como ejemplares que reflejan sus posibilidades de variación alofónica.

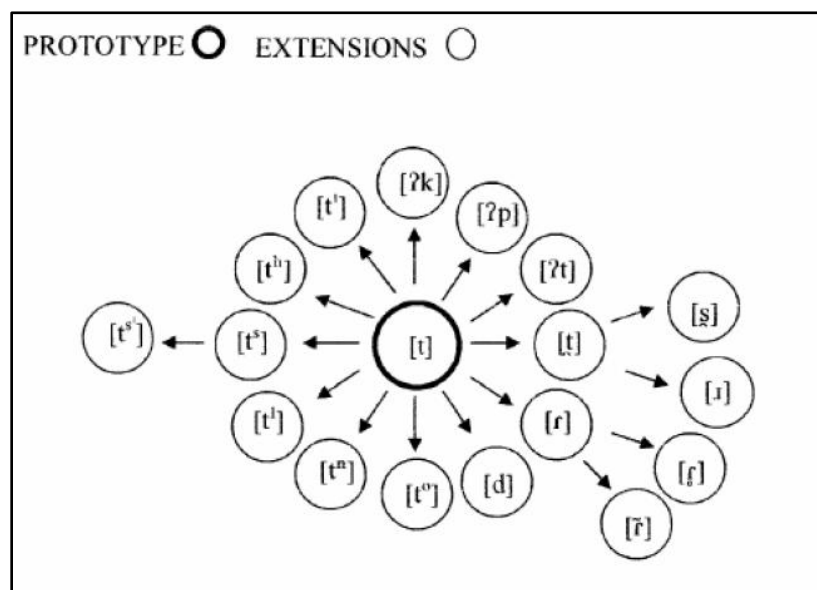
En la sociolingüística cognitiva, afirma Moreno Fernández (2012), se podrían distinguir tres niveles de abstracción en la configuración de la variación. El primer nivel, el más abstracto, es el del esquema, que correspondería con lo que el hablante cree que pronuncia, es decir, lo que tradicionalmente se define como fonema. El almacenamiento de los ejemplares de un esquema estaría asociado con el almacenamiento de palabras, realizado sobre bases semánticas y fónicas, además de la información cuantitativa referente a la frecuencia relativa de uso (lo que no implica la presencia de un aparato probabilístico como tal). El segundo nivel sería el del prototipo, que correspondería con lo que el hablante cree que debe pronunciar, por lo que las producciones fónicas (propias de y ajenas al hablante) podrían caracterizarse de acuerdo con el grado de centralidad o similitud con respecto al prototipo. Moreno Fernández (2012) acota que entre lo que el hablante cree que debe pronunciar se encuentran sonidos con cualidades que difieren de las percibidas, pues la información se almacena en la mente del usuario y se asocia a los esquemas correspondientes, pero con elementos relativos a los usos en contexto. El tercer nivel corresponde al de los sonidos que realmente se pronuncian, que pueden presentar características inhabituales o no percibidas por el hablante, incluyendo los casos de elisión de sonidos (en los cuales no se percibe la ausencia de un fonema, sino que la elisión corresponde con un alófono de dicho fonema).

Estos tres niveles se pueden ejemplificar en el plano prosódico partiendo del planteamiento realizado acerca de la abstracción de los elementos tonales de la prosodia

(§1.1.4.2.2) y del concepto de esquematización (§1.4.2). En un hablante de la Ciudad de México, por ejemplo, en el nivel más abstracto, lo que el hablante cree que pronuncia, se puede encontrar el acento tonal ascendente  $L+H^*$  en la primera sílaba tónica de un enunciado con foco amplio (entendido en este nivel como un ascenso tonal sin especificar el punto de culminación del acento, pues se trata de la noción más abstracta en el sistema cognitivo de los hablantes). En el segundo nivel, lo que el hablante cree que debe pronunciar, se puede encontrar el acento  $L+>H^*$  (con pico tonal desplazado a la sílaba postónica), que es el más común en la primera sílaba tónica de este tipo de enunciados en el centro de México (De la mota *et al.*, 2010). Por último, en el tercer nivel, lo que el hablante realmente pronuncia, se podrían registrar los acentos  $L+<H^*$  (es menos habitual que el prototipo),  $L+>|H$  (con una mayor elevación a la habitual que pudo haber pasado desapercibida para el hablante) o  $L+H^*$  (la representación esquemática más abstracta).

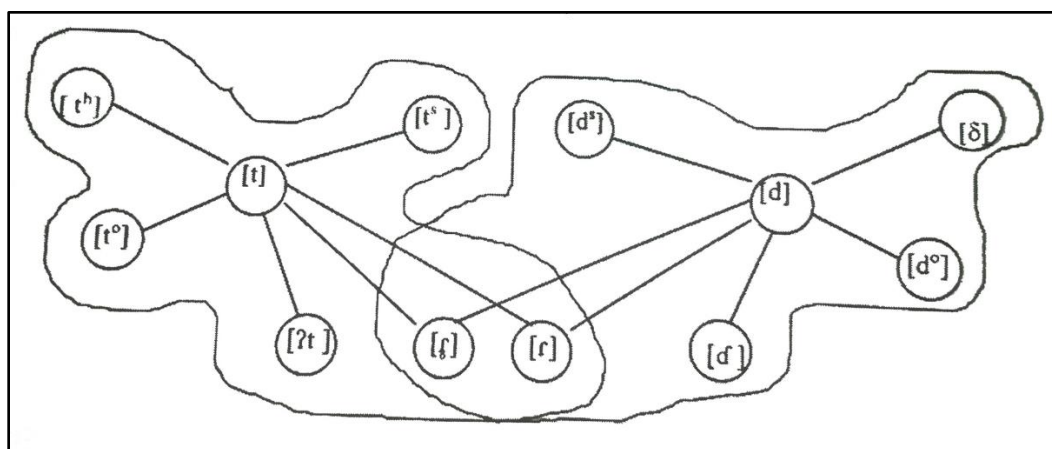
La forma en que todos estos elementos se ubican en una categoría cognitiva se puede explicar a través de modelos de organización. Kristiansen (2006) señala que existen dos modelos prevalentes sobre la estructura de categorización fonémica en la lingüística cognitiva: el modelo de categoría radial y el modelo de red. El modelo de categoría radial asume que los miembros menos prototípicos se organizan alrededor de un miembro prototípico, en términos de extensiones asimiladas a la categoría bajo el principio de la similitud relativa (Kristiansen, 2006 y Moreno Fernández, 2012). La cadena de relaciones resultantes, que va desde el centro a la periferia, se presenta como una rueda (ver figura 1.8). Los miembros organizados en un mismo radio de la rueda no necesariamente tienen alguna relación o similitud con los miembros de los radios adyacentes, salvo su conexión con el elemento central (Kristiansen, 2006).

Figura 1.8. Modelo de categoría radial (tomado de Kristiansen, 2006: 127)



Algunos de estos componentes se pueden conectar a otros elementos pertenecientes a otras redes, lo que puede dar lugar a una serie de solapamientos que genere redes descentralizadas que hacen más comprensibles fenómenos como la neutralización fonética o la equivalencia acústica, como se puede observar en la figura 1.9.

Figura 1.9. Modelo radial de la estructura interna correspondiente a los fonemas /t/ y /d/ del inglés (tomado de Moreno Fernández, 2012: 150)

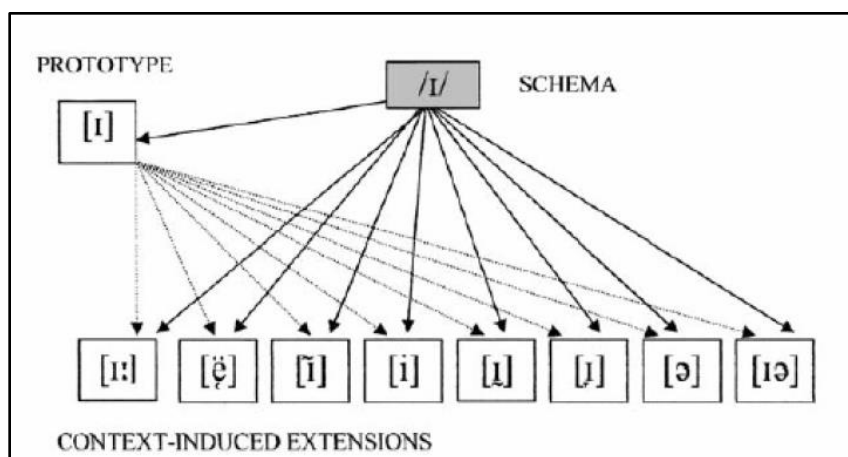


No obstante, Kristiansen (2006) afirma que este tipo de representación resulta insuficiente si se quieren describir las formas en la diferencia perceptual y los contrastes intrafonémicos existentes entre los miembros de una categoría, que permiten la transmisión de significado

social a los usuarios de una lengua. Para esta autora, la representación gráfica debería permitir expresar la posibilidad de mostrar cómo emergen las relaciones contrastivas en una manera no lineal, a través de una representación más apropiada del dominio fuente.

Por otra parte, el modelo de red involucra una categoría de prototipo, extensiones inducidas por el contexto que parten de este prototipo y un esquema que captura las similitudes percibidas en varias extensiones (Kristiansen, 2006). Un ejemplo gráfico de este modelo se puede apreciar en la figura 1.10.

Figura 1.10. Modelo de red (tomado de Kristiansen, 2006: 129)



Según explica Kristiansen (2006), la mejoría de este modelo con respecto al radial radica en el hecho de que los modelos de red operan en una estructura de dos niveles: el nivel de ocurrencia real, generalmente con variantes disímiles, y una unidad esquemática, que es una abstracción sobre los eventos basados en el uso.

La sociolingüística cognitiva reconoce la presencia de tres niveles conceptuales traídos desde una perspectiva funcional: la zona funcional, el espacio de variabilidad y la variación (Moreno Fernández, 2012). La zona funcional hace referencia al espacio de producción fonética en un *continuum*; puede estar delimitada por la capacidad de discriminar el significado de las unidades léxicas; por ejemplo, las variantes de /s/ en español pertenecen a la misma zona. El espacio de variabilidad consiste en un área de variación a menudo compartida con la zona funcional, pero que puede moverse a diferentes áreas y abarcar más de una a la vez, a causa de alguna circunstancia lingüística; por ejemplo, el posible uso de /r, l/ en coda silábica, que sería el área de distensión, ocurre debido a que las dos unidades operan

en ese espacio de variabilidad. Finalmente, la variación corresponde a la expresión diferenciada y estructurada de una variable en entidades perceptibles en un análisis fónico; si la variación no produce cambio en el significado, se le llama variación no funcional (como la alternancia de [n] y [ɲ] en posición final de palabra en español), pero si implica un cambio se le llama variación funcional (como el cambio de [s] por [θ] en el español peninsular).

Un planteamiento desarrollado bajo las premisas de los estudios fónicos desde la sociolingüística cognitiva (en concreto, la tesis basada en el uso) es el de la prosodia basada en el uso (PBU), propuesto por Martín Butragueño y Velásquez Upegui (2014). Martín Butragueño (2015a) señala que la PBU pretende brindar una perspectiva realista acerca de los procesos lingüísticos, mediante su empleo como hilo conductor entre los datos y un planteamiento lingüístico general. Martín Butragueño (2015a) explica que el término “realista” debe entenderse desde al menos tres principios metodológicos: la representatividad social y discursiva, la ponderación del error y la exhaustividad. El autor expone que la PBU parte de los mismos propósitos y supuestos que la gramática basada en el uso propuesta por Bybee (2001), por lo que permite generar ciertas hipótesis: “a) el carácter fonético-fonológico de una estructura prosódica depende de la cantidad de contexto necesario para describirla [...]; b) la frecuencia de una estructura permite el asentamiento social, la negociación de identidades, el aprovechamiento conversacional y el aprendizaje apropiado [...]; c) el sentido pragmático de la aportación prosódica a los enunciados se construye a partir del uso en un contexto dado” (Martín Butragueño, 2015a: 98). La PBU proviene de un marco teórico y metodológico próximo al de los estudios de variación y cambio, por lo que toma las comunidades de práctica como marco de referencia para la indexación de significados sociales, al mismo tiempo que considera la estratificación, el contacto, el avance y retracción de fenómenos específicos y las diferencias entre estilos contextuales.

#### 1.5.5. El papel de la percepción

Según explica Bybee (2001), las primeras investigaciones sociocognitivas enfatizaron la importancia de los límites entre las categorías fonémicas, lo que llevó a la exploración de fenómenos de percepción categorial. Dicha exploración permitió determinar que los oyentes pueden discriminar más fácilmente los estímulos que trascienden los márgenes de una categoría que aquellos dentro de ella. La autora indica que esto origina la teoría de que los

oyentes construyen categorías fonémicas discretas basándose en las características acústicamente relevantes, lo cual ocurre a través de la retención de la información determinante para la identidad de la categoría y el descarte de los detalles acústicos. El alto nivel de similitud en el detalle articulatorio alcanzado por los hablantes de un mismo dialecto atestigua el afinamiento del sistema perceptivo, por lo que es poco probable que los oyentes que ya han adquirido las cualidades articulatorias de su dialecto perciban erróneamente palabras ya adquiridas al grado de causar un cambio lingüístico (Bybee, 2001). Sin embargo, explica esta autora, la percepción desempeña dos papeles en el cambio lingüístico. El primero de ellos consiste en un problema de percepción y adquisición de una configuración fonética relativamente difícil por parte de los niños, quienes no lograrían asimilar una articulación específica y la sustituirían por otra, lo que daría lugar a un cambio lingüístico. El segundo papel ocurre en aquellos casos en los que un cambio contextual se haya dado por razones articulatorias y que sea extendido por un reanálisis perceptivo. Algunos casos de elisión, explica Bybee (2001) se pueden deber a que un solapamiento produce un efecto acústico de la articulación que queda oculto y que, por ende, se perciba como elidido.

Según Moreno Fernández (2012), la percepción subjetiva del lenguaje desempeña una función vital en la dinámica de las variedades lingüísticas, pues cada una de ellas debe su forma (en cierto grado) a la manera en que es percibida por sus hablantes y por quienes emplean otras variedades, ya que responde a un proceso de categorización basado en un aprendizaje discriminatorio que permite percibir la variación y las variedades lingüísticas. Este autor comenta que la percepción es un proceso cognitivo de aplicación compleja debido a que “puede dirigirse a entidades individuales o colectivas, de modo modular u holístico, incorporando informaciones, creencias y prejuicios particulares o grupales, afectando a la dimensión afectiva y con la posibilidad de inducir a acciones lingüísticas y conductas extralingüísticas” (Moreno Fernández, 2012: 213), por lo que su estudio puede tener como objeto tanto las variedades lingüísticas como las comunidades de habla, las agrupaciones de hablantes o los hablantes individuales

Los conceptos perceptuales tienen un estatus especial en el pensamiento humano, cuyo papel como elementos organizadores de la cognición se aclara al definir los esquemas de imagen como representaciones esquemáticas de la experiencia física (Grady, 2008). Moreno Fernández (2012) señala que el proceso de categorización se ha explicado desde tres



propuestas teóricas diferentes: la teoría del rasgo, la del ejemplar y la del prototipo. La primera plantea que la presencia o ausencia de una determinada característica origina la adscripción de un elemento a una variedad lingüística determinada (por ejemplo, la obstrucción parcial o total del flujo de aire durante la articulación de un fono lleva a incluirlo en la categoría de consonantes y a excluirlo de la de vocales). La teoría del ejemplar afirma que la categorización depende de la asociación de un elemento lingüístico a un ejemplar o caso concreto y a la categoría a la que pertenece, es decir, a una variedad específica (tal sería el caso de [ð], que en español se puede incluir junto con [d] en de la categoría del fonema /d/, debido a su alto grado de asociación en esta lengua). Por último, la teoría del prototipo plantea la categorización de la realidad lingüística de acuerdo con prototipos dialectales, por lo que el proceso consistiría en exponer a los posibles miembros de una categoría ante los prototipos, para comprobar su grado de asociación con dicha categoría (un ejemplo de este proceso consiste en las diferentes realizaciones de /s/ en posición posnuclear en el español, ya que el almacenamiento de los distintos alófonos dentro de la categoría del fonema /s/ depende del prototipo que se emplee en el dialecto correspondiente).

Esta última afirmación, tal como la presenta Moreno Fernández (2012), puede dar pie a estudios empíricos para comprobar el grado de prototipicidad de un conjunto de rasgos prosódicos, pues se consideraría lógico pensar que las similitudes y diferencias entre distintos fenómenos prosódicos estadísticamente relevantes pueden ser percibidas con relativa facilidad (Sagastuy y Fernández Planas, 2014). Sin embargo, existe la posibilidad de que aquello estadísticamente significativo no sea perceptible por los hablantes, lo que hace que las pruebas de percepción cobren mayor importancia como herramientas empíricas, para comprobar si existe una correlación entre un análisis cuantitativo de datos de producción de habla y lo que perciben los hablantes de una comunidad. Asimismo, el desarrollo de este tipo de pruebas es necesario debido a que, según indica Moreno Fernández (2012), no suele prestarse mucha atención a la percepción de los no lingüistas, quienes representan la mayoría de la sociedad y tienen mayor incidencia sobre los usos lingüísticos y su valoración colectiva.

A lo largo del capítulo se ha presentado una síntesis de todos los aspectos teóricos de la prosodia, la sociolingüística, la lingüística cognitiva y el enfoque sociocognitivo que enmarcan este trabajo. Toda esta información permite tener una idea acerca del estado de la

cuestión, hasta el momento en el que se desarrolla esta investigación. Adicionalmente, todos los puntos abordados constituyen la base teórica de los complejos fenómenos abordados y analizados en este estudio. En el próximo capítulo se exponen las bases metodológicas y se describen los procesos que se siguieron en el estudio de la prosodia de dos redes sociales de la Ciudad de México.

## Capítulo II

### Marco metodológico

Los aspectos metodológicos del presente trabajo se dividen en dos secciones. La primera, denominada “aspectos generales”, está relacionada con aquellos elementos esenciales de la investigación: la descripción de las redes sociales estudiadas y las definiciones de los tres modos de vida tomados en cuenta en la investigación. Finalmente, la segunda parte, la “fase de producción”, consiste en la información y procedimientos vinculados con la extracción y el análisis (fonético, fonológico y sociolingüístico) de los datos de habla que constituyen el corpus. Es importante señalar que, para facilitar la lectura y comprensión de todo lo relacionado con la fase de percepción (diseño y ejecución de la prueba de percepción), los aspectos metodológicos relacionados con esta fase se presentan posteriormente, en el capítulo V.

#### 2.1. Aspectos generales

La metodología utilizada para construir el corpus de habla está basada en el concepto sociolingüístico de redes sociales (§1.3.5), es decir, los tejidos de vínculos que se extienden a través de una sociedad, conectando unas personas a otras (Milroy y Milroy, 1992). Como un refuerzo teórico y utilitario para la noción de redes sociales, también se maneja el concepto antropológico (empleado en la sociolingüística) de modos de vida, el cual se concibe como un tipo de división poblacional resultante del efecto de estructuras sociales fundamentales (Højrup, 1983).

##### 2.1.1. Redes sociales

Los colaboradores que proporcionaron los datos lingüísticos para llevar a cabo esta investigación se distribuyen en las dos redes sociales descritas en §2.1.1.1 y §2.1.1.2. Vale la pena señalar que se seleccionaron estas redes sociales debido a dos razones fundamentales. La primera es que representan dos grupos sociales muy diferentes, al punto que podrían considerarse opuestos, por lo que pueden emplear variantes lingüísticas que permitan comprender el panorama sociolingüístico de la Ciudad de México. La segunda es que no se ha llevado a cabo en la Ciudad de México, hasta donde se ha podido revisar, un trabajo que ponga en contraste a dos redes sociales con las diferencias aquí planteadas, lo que permite

ahondar en la investigación de este terreno sociolingüístico poco explorado y aportar información complementaria a las observaciones de Martín Butragueño (2021) en su estudio sobre cinco redes sociales. La distribución de hablantes pertenecientes a cada red que conforman la muestra se presenta en el cuadro 2.1 (§2.2.1). A continuación, se describen ambas redes sociales a partir de los datos recolectados por medio de la aplicación del cuestionario sociolingüístico explicado en §2.2.2.2.

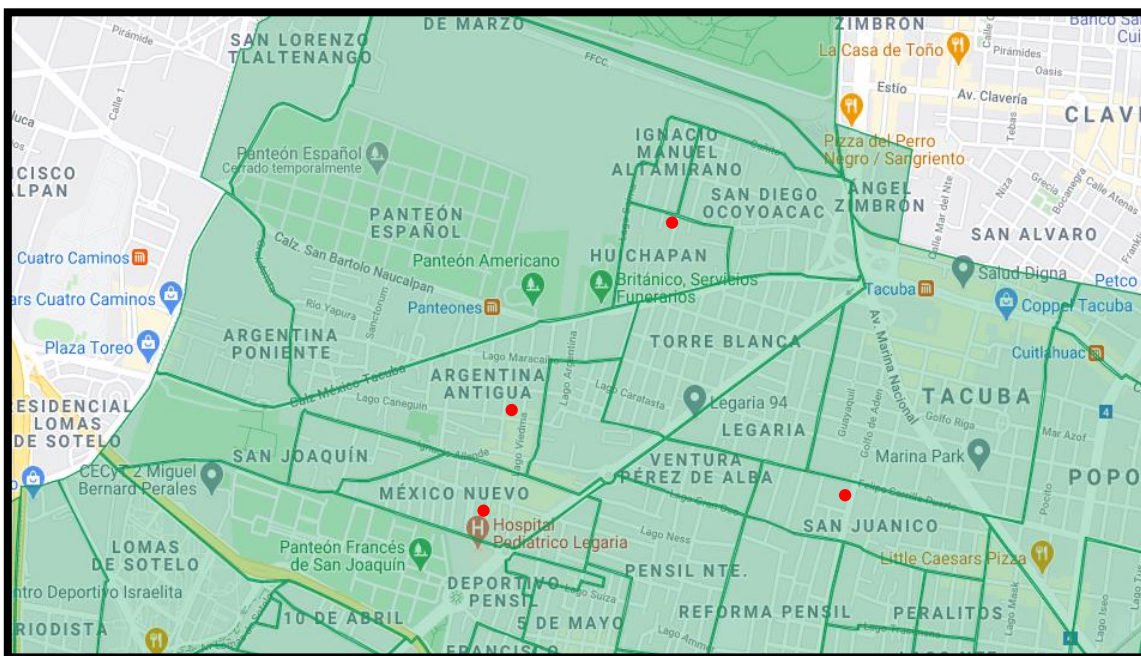
#### 2.1.1.1. Características de la red social vecinal

Esta red está constituida por 6 personas de colonias adyacentes de la zona noroeste de la alcaldía Miguel Hidalgo de la Ciudad de México. Los integrantes de esta red comparten vínculos vecinales, familiares y de amistad; y cuenta con una intensidad reticular promedio de 0.68.<sup>10</sup> Aunque no todos los miembros se conocen entre sí, hay vínculos de uno u otro tipo que los conectan. A pesar de que no todos los integrantes de la red viven en la misma colonia, todos ellos viven en colonias cercanas (todas ellas resaltadas con un punto rojo en la figura 2.1), puesto que cinco de ellos residen en colonias adyacentes (México Nuevo, Argentina Antigua, Huichapán) y uno en una colonia bastante cercana (San Juanico), que se encuentra a aproximadamente cinco cuadras de la colonia Argentina Antigua.

---

<sup>10</sup> El procedimiento para determinar el valor dentro de la escala de intensidad reticular se explica posteriormente en el §2.2.2.2.

Figura 2.1. Mapa de las colonias en las que viven los miembros de la red vecinal  
(El DeFe, 2021)



En cuanto al nivel de estudios, ninguno de los miembros de la red tiene estudios universitarios; uno solo terminó la primaria, cuatro terminaron la secundaria y uno finalizó la preparatoria. En el ámbito laboral, de los seis informantes, cuatro son autoempleados (tres de ellos en el sector de la venta ambulante y el restante se emplea como taxista), uno trabaja como personal de atención al público en una cafetería y el último coordina actividades culturales en su comunidad (labor que desempeña en un centro comunitario). En términos financieros, el nivel de ingresos general se encuentra entre 2,700 y 5,400 pesos, lo que va entre uno y dos sueldos mínimos mensuales (para la fecha en la que se recolectaron los datos, primer cuatrimestre del 2019); y solo un informante, el que trabaja en la cafetería, afirmó tener ingresos mensuales entre los 6,000 y 7,000 pesos, por lo que se puede decir que todos se encuentran en un nivel socioeconómico bajo.

La mayoría de los informantes se autopercibe como perteneciente a la clase socioeconómica media-baja, aunque dos de ellos consideran que pertenecen a la baja (ambas mujeres) y uno a la media (un hombre). La mayoría de sus amistades viven en la misma zona donde habitan o en sus alrededores y frecuentan, principalmente, a sus vecinos y familia en sus domicilios, en la calle y en el mercado. Casi todos los miembros de esta red, quienes en

su totalidad han viajado a distintas ciudades dentro del país, participan en actividades de la comunidad en la que habitan debido a que les agrada, aunque solo algunos se sienten parte de ella (el hombre Rs1H2 y la mujer Rs1M2). Los vínculos más importantes de estos informantes son los familiares, seguido por amistades fuera del trabajo, amistades dentro del trabajo y compañeros de trabajo. Cabe señalar que esta importancia en los vínculos corresponde con los datos masculinos, pues las mujeres dan más importancia a las amistades dentro del trabajo que fuera de este. Adicionalmente, los lazos de confianza que comparten con familiares y amistades son bastante fuertes y afianzados.

El medio de comunicación al que están más expuestos los integrantes de esta red es el internet (páginas con contenido audiovisual), seguido por la televisión (especialmente entre las mujeres) y por la radio. Casi todos los miembros de la red, con excepción de una informante femenina (Rs1M1), afirman dar mucha o bastante importancia al uso normativo de la lengua, pues consideran que este uso les puede resultar conveniente para obtener mejores resultados en sus interacciones sociales. De hecho, todos los miembros de esta red afirman hablar de forma diferente cuando se dirigen a clientes y otras personas que, a su juicio, merezcan un tipo de respeto especial (como pueden ser autoridades de las escuelas de sus hijos o médicos).

#### 2.1.1.2. Características de la red laboral

La red laboral está conformada por empleados con cargos ejecutivos en una empresa multinacional, cuya sede central en México se encuentra ubicada en la colonia Polanco 1ª Sección (aproximadamente a dos kilómetros en línea recta de la colonia México Nuevo). Los integrantes de esta red, cuyo valor medio de intensidad reticular es 0.25, habitan en distintas zonas de la ciudad (alcaldías Gustavo A. Madero, Azcapotzalco, Benito Juárez y Coyoacán), pero comparten vínculos laborales y de amistad dentro del entorno laboral. Esto se debe a que todos ellos se conocen entre sí, a pesar de que la empresa cuenta con más de ciento cincuenta empleados en esta sede en México.<sup>11</sup> En el aspecto educativo, todos los integrantes de esta red tienen títulos de licenciatura (dos de ellos hicieron estudios de posgrado) y alcanzaron un nivel de competencia alto en al menos un idioma (el inglés). El uso de esta lengua es un requisito en su trabajo, ya que deben mantener comunicación constante con

---

<sup>11</sup> Este dato fue mencionado por el hablante Rs2H3 durante su entrevista.

trabajadores de la misma multinacional en otros países. Adicionalmente, viajan con relativa frecuencia al extranjero por motivos laborales. A pesar de que la mayoría de los integrantes de la red ostentan cargos con responsabilidades de gerencia y dirección, todos ellos comparten en común el hecho de que tienen superiores inmediatos que se encuentran en la misma oficina y tienen fuertes motivaciones de superación laboral. Por otra parte, todos los miembros de esta red tienen un nivel de ingresos mucho más elevado que el de los miembros de la red vecinal, pues sus ingresos mensuales superan los 30,000 pesos mensuales (más de cuatro veces el ingreso mensual del miembro con mayor ganancia económica de la red vecinal).

La mayoría de los informantes de esta red considera que pertenece a la clase social media-alta, aunque solo uno se percibe como clase media (el hablante Rs2H1). La mayoría de sus amistades viven en zonas diferentes a aquellas en las que habitan o trabajan los colaboradores, por lo que frecuentan principalmente a sus compañeros de trabajo en las instalaciones de la empresa. Cuando se relacionan con amistades, lo hacen principalmente en lugares de encuentros sociales (restaurantes, clubs, salones de fiesta). Todos ellos participan en actividades de la comunidad en la que habitan, pues les agrada y se sienten parte de ella. En esta red, el vínculo más importante es la familia en todos los casos, seguido por amistades fuera del trabajo, amistades dentro del trabajo y compañeros de trabajo. Al igual que en la red vecinal, las mujeres priorizan los vínculos dentro del trabajo a los existentes fuera del área laboral. Los miembros de esta red comparten fuertes lazos de confianza con sus familiares y con sus amigos (dentro y fuera del trabajo), mas no con compañeros de trabajo con los que no tienen una relación cercana.

En relación con su exposición a los medios de comunicación, todos los integrantes afirman dedicar menos de 10 horas a ver televisión, escuchar radio y navegar en internet. Los miembros masculinos consumen un mayor contenido de programas de radio que de cualquier otro medio y también lo hacen con más frecuencia que las mujeres. Los miembros de esta red afirman prestar bastante importancia al uso normativo de la lengua, debido a que consideran que de esa forma pueden obtener un mejor perfil laboral y mejores oportunidades dentro de los proyectos que ejecuta la empresa. Casi todos ellos consideran que se expresan de la misma manera en todo momento, independientemente de donde se encuentren o a quién

se dirijan, aunque dos (los hablantes Rs2H2 y Rs2M3) afirman hablar de manera distinta cuando se comunican con clientes y con personas con cargos superiores dentro de la empresa.

### 2.1.2. Modo de vida

Los tres modos de vida que se manejan en esta investigación, dos de los cuales están estrechamente relacionados con las redes sociales estudiadas, se construyeron de acuerdo con el modelo propuesto por Lastra y Martín Butragueño (2000) para las sociedades latinoamericanas, presentado previamente en §1.3.6. Para los fines prácticos de este trabajo, los 8 modos de vida planteados por los autores se reagruparon en tres categorías. Esto se hizo con el objetivo de poder generar un modo de vida que fuese compatible con las características sociales de la red vecinal, otro compatible con la red laboral y otro que sirviera para delimitar los rasgos macrosociales de un tercer grupo que, si bien no está incluido en el corpus de habla, forma parte importante de la fase de percepción de este estudio (ver §5.2.3, §5.2.4 y §5.2.5). Cabe señalar que esta nueva clasificación fue posible debido a que los rasgos de los distintos modos de vida planteados por Lastra y Martín Butragueño (2000) no son radicalmente distintos ni incompatibles entre sí y, al ubicarlos en estos nuevos conjuntos, representan los rasgos macrosociales de las dos redes sociales estudiadas.

#### 2.1.2.1. Modo de vida 1

El primer modo de vida corresponde con los medios de producción presentes en la red vecinal; y engloba, en principio, los modos de vida B, D, E y F de la clasificación de Lastra y Martín Butragueño (2000). Se incluyen en esta categoría, por ende, los empleados informales (vendedores ambulantes), subempleados (trabajadores de negocios pequeños que se encuentran en el nivel más bajo de la línea de mando), obreros (empleados de fábricas que desarrollan labores manuales) y burócratas (empleados formales de oficina). No obstante, vale la pena hacer dos precisiones en este sentido. La primera de ellas es que, a pesar de que ninguno de los miembros de la red vecinal realizaba labores manuales en fábricas al momento de hacer la entrevista, se incluye en esta categoría el modo de vida D propuesto por Lastra y Martín Butragueño (2000), porque cuatro de los miembros de la red habían desempeñado trabajos manuales en fábricas anteriormente. Además, se incorpora en este grupo el modo de vida F, establecido por Lastra y Martín Butragueño (2000), debido a que, teóricamente, una



informante forma parte del personal adscrito a un centro cultural comunitario, aunque, en realidad, la mayor parte de sus labores se desarrollan fuera de la oficina y consisten, en menor medida, en labores de asistencia a su superior.

#### 2.1.2.2. Modo de vida 2

El propósito metodológico de este modo de vida es el de categorizar a aquellos individuos que se encuentran en un nivel intermedio entre los modos de vida 1 y 3 (personas cuyas características sociales no corresponden con ninguno de los otros dos modos de vida planteados en este estudio), ya que se considera un grupo con estas características para el desarrollo de una serie de pruebas de percepción que se incluye en el capítulo V (§5.2). Por lo tanto, se puede incorporar a esta categoría el modo de vida F, planteado por Lastra y Martín Butragueño (2000), es decir, empleados que forman parte del sistema burocrático y que desempeñan labores de oficina. Sin embargo, debido a que existe una informante en la red vecinal que se clasifica dentro del modo de vida F, es necesario tomar en cuenta ciertos criterios que establezcan una diferencia precisa entre las características de los burócratas que se pueden incluir en este modo de vida y aquellas de la informante ubicada en el modo de vida 1. Por ende, se pueden incluir en el modo de vida 2 de esta investigación trabajadores de oficina que cuenten con algún tipo de estudio técnico o universitario (en cualquier caso, de un nivel más alto que el de la preparatoria) y que trabajen en empresas o instituciones privadas. Adicionalmente, dado que Lastra y Martín Butragueño (2000) definen su modo de vida F como propio de empleados del sector terciario (servicios), se pueden incorporar en el modo de vida 2 de este trabajo a aquellos que desarrollen algunas de las siguientes actividades productivas: periodistas, personal de sanidad del sector público (enfermeras, médicos y demás) y maestros. Además, aunque no se incluyen en la clasificación de Lastra y Martín Butragueño (2000), se pueden incorporar a este modo de vida personas que no estén integradas formalmente al sistema de producción, pero que cuenten con estudios universitarios recién concluidos o cercanos a concluir, pues se considera que, igualmente, forman parte de la comunidad de habla de la Ciudad de México y que se encuentran suficientemente alejados de las características sociales de la red vecinal (que no cuenta con estudios universitarios) y de la laboral (que ya han hecho carrera profesional por suficiente tiempo como para tener cargos ejecutivos).

### 2.1.2.3. Modo de vida 3

Este modo de vida corresponde con la red laboral y con el modo de vida G de la clasificación de Lastra y Martín Butragueño (2000). Incluye empleados con un cargo ejecutivo que hacen carrera dentro de una empresa privada, que tienen el poder de tomar decisiones importantes y relevantes en el funcionamiento de la empresa, que tienen aspiraciones de ascenso en el plano profesional y que dedican gran parte de su tiempo al trabajo y lo separan notablemente de las demás facetas de su vida.

## 2.2. Fase de producción: elicitación y análisis de datos

Durante esta fase se constituyó el corpus, mediante la realización de entrevistas sociolingüísticas y se extrajeron y analizaron datos acústicos. Todo esto se llevó a cabo con la finalidad de poder identificar, mediante un estudio cuantitativo (estadístico) y cualitativo, aquellos rasgos prosódicos que son potenciales portadores de significado social, es decir, índices sociolingüísticos en potencia. En este apartado, se explica el procedimiento que se siguió en todas las etapas de esta fase.

### 2.2.1. Corpus

Para este estudio se emplearon datos de 12 colaboradores, distribuidos equitativamente entre hombres y mujeres, procedentes de las dos redes sociales mencionadas en §2.1.1. Los criterios de selección de la muestra se presentan a continuación:

- a. Red social: vecinal y laboral.
- b. Sexo: masculino y femenino.
- c. Modo de vida: 1 (empleados informales, subempleados y obreros) y 3 (empleados ejecutivos).
- d. Origen: nacidos en la Ciudad de México o llegados a la ciudad antes de los 5 años.

En el Cuadro 2.1, se presenta la distribución de los hablantes de acuerdo con las variables sociales pre-estratificadoras. En el código asignado a cada hablante, las dos primeras letras (*Rs*) hacen referencia a la red social identificada por el número que le sigue (1 para la red vecinal y 2 para la laboral), después se encuentra la codificación del sexo (*H*

para hombre y *M* para mujer) y un número final (1, 2 o 3) cuya función es diferenciar a los hablantes de forma tal que cada código sea único.

Cuadro 2.1. Distribución de hablantes de acuerdo con las variables pre-estratificadoras

	Sexo		Total
	Hombres	Mujeres	
Red social vecinal	3	3	6
	Rs1H1	Rs1M1	
	Rs1H2	Rs1M2	
	Rs1H3	Rs1M3	
Red social laboral	3	3	6
	Rs2H1	Rs2M1	
	Rs2H2	Rs2M2	
	Rs2H3	Rs2M3	
Total	6	6	12
	12		

Con respecto a los datos mencionados en el cuadro anterior, cabe acotar que la hablante Rs1M1, quien nació en Guanajuato, llegó a la Ciudad de México cuando tenía un año de edad.

### 2.2.2. Las entrevistas

Las entrevistas con las que se construyó el corpus para esta investigación están constituidas por dos fases. La primera consiste en una entrevista semidirigida, de la cual se extrajeron los datos para realizar el análisis sociolingüístico, mientras que la segunda se trató de la aplicación de un cuestionario con el que se obtuvieron datos con valor sociolingüístico.

#### 2.2.2.1. Fase 1: Entrevista sociolingüística

Se entabló una conversación guiada con cada hablante, con preguntas relativas a experiencias de vida que buscaron la producción de enunciados aseverativos. Se contempló un repertorio de temas y preguntas como guía para la realización de una entrevista semidirigida (ver apéndice 1). Sin embargo, no se consideró que los elementos de la guía debían condicionar la espontaneidad de los entrevistados, debido a que lo primordial fue mantener un ambiente tan informal como fuese posible (dentro de las limitaciones contextuales del entorno de la entrevista), para que los informantes se expresaran naturalmente. Durante la entrevista, el

entrevistador mantuvo un papel de observador participante, es decir, pudo participar en el diálogo para favorecer la dinámica conversacional, la creación de un ambiente de familiaridad (relativa) y la espontaneidad del entrevistado. En cualquier caso, el entrevistador buscó mantener la “actitud del aprendiz” (Labov, 1981) durante las entrevistas con los integrantes de ambas redes, tanto cuando la entrevista se apejó a los módulos conversacionales (no condicionales) como cuando la conversación se alejó de ellos.

Se consideró oportuno implementar la técnica de entrevista sociolingüística debido a que, según Moreno Fernández (2012), permite obtener muestras de lengua hablada en la cantidad y con la calidad necesarias para su análisis cuantitativo, aunque no se debe perder de vista que la percepción cognitiva del contexto en el que se lleva a cabo puede reducir el grado de espontaneidad del entrevistado al momento de generar su discurso. Esto responde al hecho de que la situación comunicativa de la entrevista (con un entrevistador haciendo preguntas y la grabadora a la vista) dificulta la obtención de muestras de habla completamente naturales, lo que se relaciona directamente con la paradoja del observador (Labov, 1966). Adicionalmente, este tipo de entrevistas es compatible con la metodología del enfoque de la prosodia basada en el uso, ya que este enfoque parte de la premisa de que el objeto de estudio debe ser el habla cotidiana y promueve el uso de datos obtenidos de la forma más espontánea posible (Martín Butragueño y Velásquez Upegui, 2014; Martín Butragueño, 2019).

#### 2.2.2.2. Fase 2: Cuestionario sociolingüístico

El cuestionario está constituido por catorce secciones, cada una busca recolectar datos específicos con respecto a las características sociolingüísticas de los hablantes (ver apéndice 2). Las partes del cuestionario son las siguientes: datos generales del informante (del inciso 1 al 8), datos generales de los padres del informante (del inciso 9 al 18), ubicación de contactos (del inciso 19 al 24), contacto con otras comunidades (del inciso 25 al 28), actitud e integración en la cultura local (incisos 29 y 30), actitud e integración a la red (inciso 31), densidad de la red (incisos del 32 al 38), fuerza de los vínculos (inciso 39), uso de la red (del inciso 40 al 42), contacto con medios de comunicación (del inciso 43 al 45), contacto con la norma (inciso 46), creencias lingüísticas (del inciso 47 al 52), escala de intensidad reticular (del inciso 53 al 66). Cabe señalar que la mayor parte del cuestionario fue una adaptación del

empleado por Vida Castro (2004), con excepción de los apartados “densidad de la red”, “fuerza de los vínculos” y “uso de la red”, basados en Villena Ponsoda *et al.* (2003) y del apartado “contacto con la norma”, que está fundamentado en Martín Butragueño, Mendoza y Orozco (en preparación).

Para determinar el valor de intensidad reticular de la red, a cada una de las 14 preguntas de la última sección del cuestionario se le asignó un valor de uno (1) a las respuestas afirmativas y cero (0) a las negativas. Posteriormente, se sumaron los valores de las respuestas afirmativas de cada hablante y se dividió entre el número de preguntas (14), para obtener el valor correspondiente a la intensidad reticular de cada hablante en un rango que va de 0 (menor intensidad posible) a 1 (mayor intensidad posible). Luego, para determinar el grado de intensidad de cada red de acuerdo con la escala, se calculó la media resultante de los valores individuales de los miembros de cada red, lo que arrojó un valor que se maneja, igualmente, del 0 al 1.

### 2.2.3. Digitalización de los datos

Con la finalidad de registrar los datos de habla de la entrevista, se empleó una grabadora digital portátil que almacena los datos directamente en formato .wav (formato digital sin compresión), con 16 bits de profundidad y velocidad de muestreo de 48.000 Hz, en formato estéreo, lo que garantizó la alta definición de la muestra de audio. Posteriormente, a través del programa Adobe Audition CS6 (Adobe Systems Incorporated, 2013), se convirtió la señal de audio a formato mono para hacer el posterior tratamiento de los datos acústicos de manera óptima.

### 2.2.4. Selección de enunciados para el análisis

De cada entrevista, se extrajeron 30 enunciados aseverativos, lo que suma 180 enunciados de cada red y 360 enunciados en total. Todos los enunciados están constituidos por una sola frase entonativa mayor (IP). No obstante, la mitad de ellas están conformadas por más de una frase entonativa intermedia (ip), como en el ejemplo [*Si tú amenazas con romper ese tipo de*]<sub>ip</sub> [*ganancia que te genera*]<sub>ip</sub> [*siempre va a haber algo que trate de detenerte*]<sub>ip</sub> ]<sub>IP</sub> (hablante Rs2H1). La otra mitad, en cambio, consisten en una frase entonativa mayor (IP) dentro de la cual hay solo una frase intermedia (isomórfica con la mayor), tal como se ve en

el enunciado [(*Nosotros no lo necesitamos*)<sub>ip</sub>]<sub>IP</sub> (hablante Rs1MM1). Esto responde a que se considera que pueden existir diferencias entonativas relevantes en las configuraciones prosódicas de cada tipo de frase, especialmente en las secciones nucleares situadas al final de las frases entonativas intermedias (ip) no finales (cuyo final no coincide con el de la frase mayor) y de las frases entonativas mayores (IP).

Los criterios para la selección de los enunciados fueron los siguientes: i) ser emotivamente neutros, es decir, se evitó la selección de enunciados producidos con algún tipo de emoción por parte del hablante (tristeza, rabia, etc.); ii) puede haber o no presencia de algún matiz expresivo o exclamativo relacionado con énfasis, pues no se considera que la focalización estrecha (sintáctica o prosódica) afecte la configuración prosódica del enunciado al mismo nivel al que lo pueden hacer las emociones; iii) la última sílaba del enunciado (es decir, de la frase entonativa mayor) debía ser átona, para poder observar la variación prosódica entre las sílabas que constituyen el tonema; iv) los enunciados no debían contener interrupciones provocadas por risas, muletillas o reformulaciones del enunciado, aunque se tomaron en cuenta enunciados con pausas entre dos frases entonativas intermedias (ip).

#### 2.2.5. Unidades de análisis

Las unidades lingüísticas que se tomaron en cuenta para llevar a cabo este estudio son las siguientes:

- a. Fonológica: sílaba y frase entonativa. La sílaba, localizada en la base de la jerarquía prosódica (Nespor y Vogel, 2007) y sobre la que se establecen las demás categorías prosódicas (Selkirk, 1986), es la unidad mínima de análisis, debido a que de ella se obtuvieron los datos acústicos necesarios para construir los patrones prosódicos y para la asignación de los acentos tonales, que son fundamentales para poder configurar los patrones prosódicos. Por su parte, la frase entonativa se toma como unidad de análisis más amplia, debido a que un enunciado puede estar constituido por distintas frases entonativas intermedias (ip), cuyo conjunto conforma una sola frase entonativa mayor (IP) (Toledo, 2008) o por una sola frase intermedia (ip), cuya extensión coincide exactamente con una frase mayor (IP). De igual modo, tanto la ip

como la IP poseen un tono de frontera (Hualde, 2003).<sup>12</sup> En este punto, se considera necesario hacer la siguiente aclaración terminológica que deberá ser tomada en cuenta en lo sucesivo: el término “frase entonativa intermedia” (ip) es utilizado en todo momento para hacer referencia a las frases intermedias cuyas secciones nucleares no coinciden con la de la frase mayor, mientras que “frase entonativa mayor” (IP) se emplea para señalar la frase entonativa cuya extensión coincide con la del enunciado, la cual engloba las frases intermedias existentes en su interior (sean una o varias).

- b. Discursiva-pragmática: enunciado aseverativo. Debido a que el enunciado es una unidad semánticamente autosuficiente y sintácticamente independiente (Eguren y Fernández Soriano, 2006), se puede establecer una asociación entre esta unidad y la frase entonativa, que también está sujeta a criterios semánticos de buena formación (Selkirk, 1986).

Debido a que los movimientos tonales en los núcleos de las frases entonativas se han considerado relevantes para delimitar su extensión (Prieto, 2006; Von Heusinger, 2007; Nespor y Vogel, 2007; Toledo, 2007), se considera que es necesario tenerlas en cuenta como elementos de análisis. Ahora bien, aun cuando el núcleo de frase entonativa mayor (IP) ha sido denominado “tonema” (Sosa, 1999), no se pudo encontrar una palabra empleada previamente para denominar el núcleo de la frase entonativa intermedia (ip). Por lo tanto, se consideró oportuno crear el término operativo “tonema interno” o “intertonema”, para hacer referencia al conjunto de tonos que van desde la última sílaba acentuada hasta el final de la frase entonativa intermedia (ip) no final, es decir, de aquellas ip cuya sección nuclear no coincidiera con el de la frase entonativa mayor (IP).

Este término no solo busca especificar el hecho de que el segmento entonativo en cuestión se encuentra dentro del enunciado y no en su parte final, sino también hacer notar el hecho de que el valor de la configuración nuclear de una ip es diferente al de aquel de una IP. Mientras que el tonema conlleva en sí una carga semántica que marca el contraste entre un enunciado aseverativo y uno interrogativo (Quilis, 1981), el “tonema interno” no tiene esta función, sino la de establecer una separación fonológica, con ayuda de pausas o sin ellas,

---

<sup>12</sup> La relación entre una frase entonativa intermedia (ip) y una mayor (IP), se puede ilustrar a través del enunciado interrogativo *¿quieres agua, café o té?*, donde, dada la naturaleza del enunciado, se esperaría una breve pausa entre cada elemento listado precedida por un tono de frontera alto (H-):  $[[\text{¿quieres agua}]_{ip} [\text{café}]_{ip} [\text{o té?}]_{ip}]_{IP}$ .

entre las frases entonativas intermedias que constituyen una frase entonativa mayor. Un ejemplo de esto se puede encontrar en el enunciado *si no va nadie, yo tampoco voy*, cuyos rasgos entonativos al final de ambas ip (prótasis y apódosis) corresponderían con lo descrito por Navarro Tomás (1957 y 1966) para enunciados aseverativos con subordinación interna. En el enunciado con la estructura fonológica [(*si no va nadie*)<sub>ip</sub> (*yo tampoco voy*)<sub>ip</sub>]<sub>IP</sub> se daría un intertonema que abarcaría tanto la sílaba nuclear (*na*) como la final (*die*) de la primera ip, cuya función sería marcar una separación entre esta frase intermedia (que corresponde sintácticamente con la prótasis, en este caso) y la que le sigue (que comprende la misma extensión que la apódosis). La configuración entonativa de la última palabra del enunciado (*voy*) no se considera un intertonema debido a que, si bien es el núcleo de una frase intermedia, esta sílaba es la última de la frase entonativa mayor, por lo cual, los movimientos tonales en esta sección se clasifican como tonema. En vista de todo lo anterior, se puede plantear que el intertonema tiene como función expresar la no conclusión de la enunciación, por lo que opera como elemento delimitador de la extensión de una frase entonativa intermedia (isomórfica con una unidad sintáctica o no), carente de autonomía semántica (pues el mensaje está inconcluso) y, por ende, establece una dependencia de significado entre los elementos lingüísticos contenidos dentro de sus fronteras y aquellos incluidos en las demás ip de una frase entonativa mayor. El tonema, por su parte, demarca el rango de una unidad discursiva con significado completo e independiente, que puede estar constituida por una o varias frases entonativas intermedias y aporta información relacionada con el tipo enunciativo, ya que puede indicar, por ejemplo, que la IP consiste en una aseveración y no en una pregunta.

#### 2.2.6. Variables

Las variables dependientes son de dos tipos: fonético y fonológico. Las variables fonéticas, constituidas por la duración silábica (milisegundos), intensidad (decibelios) y frecuencia fundamental o F0 (hercios y semitonos). Por su parte, las variables fonológicas están conformadas por los acentos tonales y tonos de juntura.

Se tomaron en cuenta dos tipos de variables independientes: extralingüísticas y lingüísticas. Las variables extralingüísticas consisten en la red social a la que pertenecen los hablantes y su sexo. Las variables lingüísticas, por otra parte, están constituidas por la



posición de la sílaba dentro del enunciado, que puede ser alguna de las siguientes sílabas o la convergencia de más de una de ellas (por ejemplo, cuando la primera sílaba del enunciado es tónica): primera sílaba, primera sílaba tónica, núcleo de frase entonativa intermedia (ip), linde de frase entonativa intermedia (ip), núcleo de frase entonativa mayor (IP), linde de frase entonativa mayor (IP) y sílabas del cuerpo del enunciado.

#### 2.2.7. Método de análisis acústico

A través del programa Praat 6.1.06 (Boersma y Weenink, 2019), se dividieron los 360 enunciados en sílabas, lo que generó un total de 5,899 sílabas. De cada una de ellas, se extrajeron los datos acústicos: duración silábica (milisegundos), intensidad (decibelios) y frecuencia fundamental (estandarizada en semitonos). De igual forma, este programa fue utilizado para obtener los espectrogramas y las curvas de F0 que se usaron como referencia para la transcripción prosódica con el sistema de notación Sp\_ToBI.

El material prosódico fue analizado en dos etapas complementarias. En primer lugar, se registraron los datos acústicos relativos a la cantidad, intensidad y frecuencia fundamental de cada una de las sílabas de los enunciados analizados. Con estos datos, se llevó a cabo un análisis pluriparamétrico de acuerdo al modelo Aix-en-Provence (cf. Baqué y Estruch, 2003), para determinar los patrones presentes en los correlatos acústicos de cada hablante y de los grupos de hablantes.

En una segunda etapa, de acuerdo con los espectrogramas correspondientes a cada uno de los enunciados y en la configuración de la curva melódica observada, se asignaron los acentos tonales anclados en cada una de las sílabas tónicas y los tonos de juntura alineados en los lindes de las frases entonativas, de acuerdo con el sistema de notación prosódica Sp\_ToBI que se desarrolla en el siguiente apartado.

#### 2.2.8. Sistema de transcripción prosódica

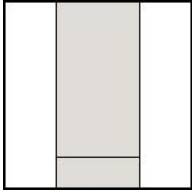
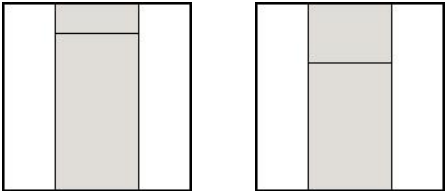
La versión del sistema de transcripción empleado en esta investigación es el propuesto por Estebas Vilaplana y Prieto (2008), junto con algunas adaptaciones planteadas por Escudero *et al.* (2014), Hualde y Prieto (2015) y Martín Butragueño y Mendoza (2018),<sup>13</sup> todo esto

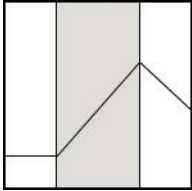
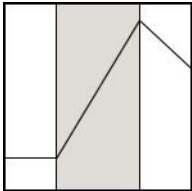
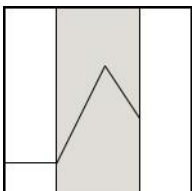
---

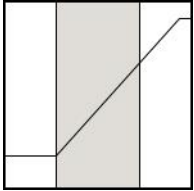
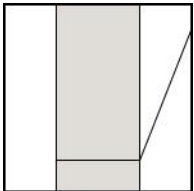
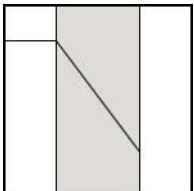
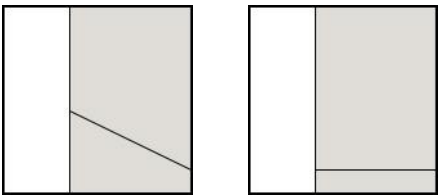
<sup>13</sup> Las adaptaciones realizadas por Martín Butragueño y Mendoza (2018) se hicieron con base en un estudio sobre la prosodia del español de la Ciudad de México.

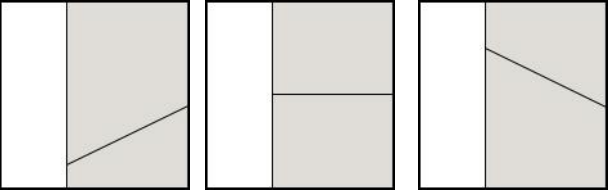
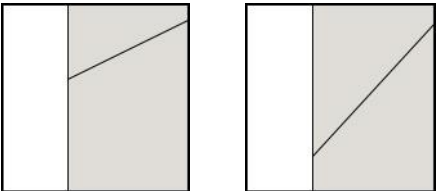
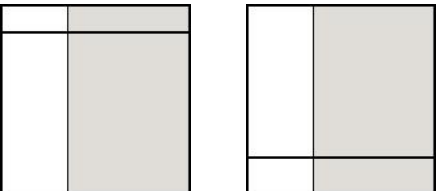
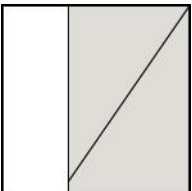
enmarcado dentro de los principios del modelo métrico-autosegmental (Pierrehumbert, 1980). Es necesario señalar que, para la asignación de acentos tonales, se marcaron como bitonos los movimientos entonativos mayores a 1.5 semitonos, pues es la diferencia tonal que se considera significativa en el nivel psicoacústico (Pamies Bertrán *et al.*, 2002; Murrieta, 2016), y se usó el diacrítico “¡” cuando los ascensos tonales fueron mayores a los 3 semitonos (Martín Butragueño, 2015b). La descripción de la transcripción prosódica de los acentos tonales y tonos de juntura empleados en este trabajo, junto con la fuente de la que se tomó cada elemento, se puede apreciar en el cuadro 2.2.

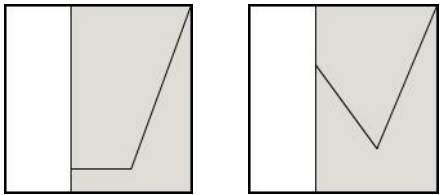
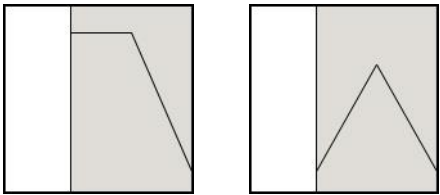
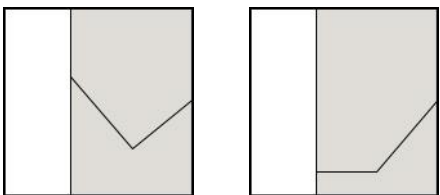
Cuadro 2.2. Descripción del sistema de transcripción prosódica

Acentos tonales monotonaes		
	L*	<p>Acento fonéticamente realizado como una meseta baja situada en el límite inferior del rango tonal del hablante.</p> <p>(Estebas Vilaplana y Prieto, 2008)</p>
	H*, ¡H*	<p>Este acento se visualiza como una meseta alta que no es precedida por un valle de F0. Cuando la meseta no es tan elevada, H* se representa con la forma ¡H*.</p> <p>(Martín Butragueño y Mendoza, 2018)</p>

Acentos tonales bitonales		
	<p>L+H*</p>	<p>Este acento se realiza fonéticamente como un ascenso tonal en la sílaba acentuada, en cuyo final se encuentra el pico de F0. (Estebas Vilaplana y Prieto, 2008)</p>
	<p>L+;H*</p>	<p>Este acento tiene la misma realización fonética que el acento anterior, con la diferencia de que, en este caso, el ascenso tonal supera los 3 semitonos. (Martín Butragueño y Mendoza, 2018)</p>
	<p>L+&lt;H*</p>	<p>Este acento, conocido como “pico muy temprano” (<i>very early peak</i>), consiste en un ascenso de F0, cuyo pico está situado cerca del punto medio de la sílaba acentuada. (Martín Butragueño y Mendoza, 2018)</p>

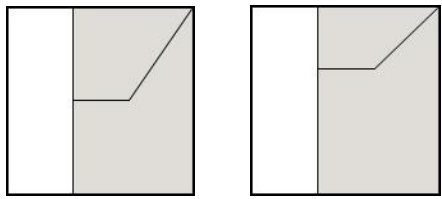
	<p><math>L \rightarrow H^*</math></p>	<p>La realización fonética de este acento consiste en un movimiento melódico ascendente en la sílaba acentuada que culmina en la sílaba postónica, donde se encuentra el pico de F0. (Estebas Vilaplana y Prieto, 2008)</p>
	<p><math>L^* + H</math></p>	<p>En este acento consiste en un valle de F0 en la sílaba acentuada con un posterior ascenso en la sílaba postónica. (Estebas Vilaplana y Prieto, 2008)</p>
	<p><math>H + L^*</math></p>	<p>Este acento se realiza fonéticamente como un descenso de F0 dentro de la sílaba acentuada. (Estebas Vilaplana y Prieto, 2008)</p>
<p>Tonos de juntura monotonaes</p>		
	<p><math>L\%</math></p>	<p>Esta juntura consiste en un tono bajo sostenido o descendente en la base del rango tonal del hablante. (Estebas Vilaplana y Prieto, 2008)</p>

	!H%	Este tono de juntura se realiza fonéticamente como un ascenso o descenso tonal hasta un punto medio del rango tonal o un tono medio sostenido (Hualde y Prieto, 2015).
	H%	Esta juntura se realiza como un movimiento ascendente de F0 desde un acento tonal bajo o alto (Estebas Vilaplana y Prieto, 2008). En este trabajo también se emplea cuando el tono final se encuentra a la misma altura que un acento H*.
	=%	Este tono de juntura hace referencia a un sostenimiento tonal desde el tono anterior (Escudero <i>et al.</i> , 2014). En este trabajo, esta juntura se emplea únicamente en el nivel fonético.
Tonos de juntura bitonales		
	HH%	Esta juntura consiste fonéticamente en un ascenso agudo al final de la frase, generalmente alcanzando el punto más alto del rango tonal del hablante (Estebas Vilaplana y Prieto, 2008).

	<p>LH%</p>	<p>Esta juntura se realiza fonéticamente como un valle de F0 seguido de un ascenso. (Estebas Vilaplana y Prieto, 2008)</p>
	<p>HL%</p>	<p>Esta juntura consiste en un pico de F0 seguido de un descenso. (Estebas Vilaplana y Prieto, 2008)</p>
	<p>L!H%</p>	<p>Esta juntura se realiza fonéticamente como un valle de F0 seguido por un movimiento hasta un punto medio del rango tonal del hablante. También se manifiesta como un descenso melódico seguido de un ascenso hasta un punto medio. (Hualde y Prieto, 2015)</p>

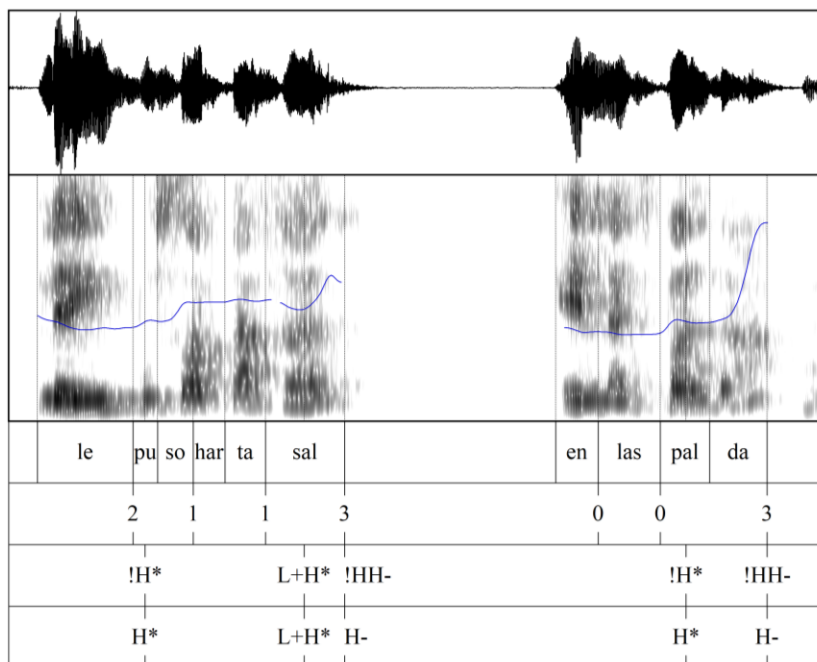
Adicionalmente, se tomó la decisión de incorporar el tono de juntura bitonal !HH%, cuya descripción se presenta en el cuadro 2.3.

Cuadro 2.3. Descripción del tono de juntura !HH%

	<p>!HH%</p>	<p>Este tono de juntura se realiza fonéticamente como un tono sostenido alrededor del punto medio del rango tonal seguido por un ascenso extendido hasta el final de la sílaba.</p>
---	-------------	---

La decisión de incluir el tono de juntura !HH% atiende a que se considera que el tono LH% implica que la sílaba inicia con un tono ubicado en la parte inferior del rango tonal del hablante, por lo que dicho tono no es representativo de un ascenso iniciado desde el punto medio o superior del rango tonal. Por lo tanto, se considera que la juntura LH% no podría representar el movimiento que ocurre al final de las dos frases entonativas intermedias (ip) como las que se muestran en la figura 2.2. En la primera de ellas, la ip termina en sílaba tónica, por lo que el acento nuclear y la juntura ocurren en la misma sílaba. La segunda ip, en cambio, culmina con una palabra paroxítona, por lo que se puede ver que el tono del acento nuclear alto se sostiene hasta poco más allá del inicio de la última sílaba, desde donde la F0 se eleva hasta alcanzar el pico en la parte alta del campo tonal.

Figura 2.2. Oscilograma, espectrograma y curva melódica de dos frases entonativas intermedias del enunciado Rs1M2-29: [(Le puso harta sal)<sub>ip</sub> (en la espalda)<sub>ip</sub> (y que mi otro hermano)<sub>ip</sub> (le lambiera la espalda)<sub>ip</sub>]<sub>IP</sub>.



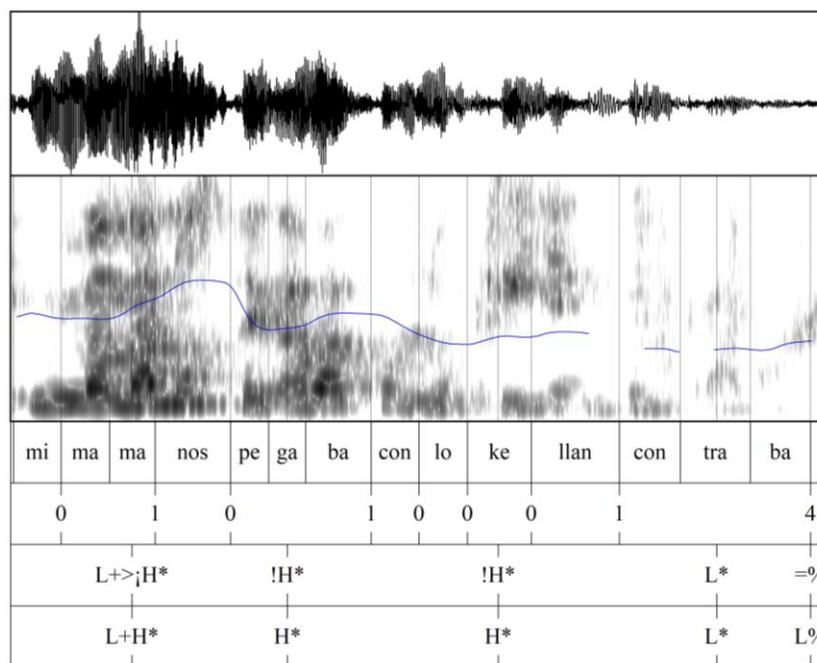
### 2.2.9. Método de análisis prosódico en dos niveles

En esta investigación, se propone una transcripción prosódica en dos niveles: uno concreto y otro abstracto (esto según el planteamiento sobre la abstracción entonativa expuesto previamente en §1.1.4.2.2). Esta noción se ejemplifica en la figura 2.3, donde la tercera grada corresponde al nivel concreto y la cuarta, al abstracto. Aunque, en cierta forma, el nivel concreto se puede asociar con el ámbito fonético y el abstracto con el fonológico, no se emplean estos términos debido a que, al menos en el caso del español, no se encontraron estudios perceptivos que permitieran concluir hasta qué punto distintas realizaciones concretas pueden representar diferentes realizaciones fonéticas de un mismo elemento entonativo fonológico o si cada realización representa, en sí misma, un fenómeno fonológico (todo esto con base en las investigaciones que se han podido revisar en el desarrollo de este trabajo). Sin embargo, sí es posible postular este método de análisis desde las bases teóricas presentadas en §1.1.4.2.2, pues se parte de nociones cognitivas que permiten agrupar distintas emisiones concretas suficientemente similares dentro de un esquema cognitivo abstracto, de acuerdo con sus características centrales invariables, al menos en el caso del español.



Con esta convención de notación se puede obtener información más detallada acerca de la relación entre los elementos concretos y abstractos de la entonación, sin comprometer el nivel de detalle del análisis ni su cualidad fonológica esencial propuesta por Pierrehumbert (1980) en el planteamiento del modelo métrico-autosegmental. En la figura 2.3 se puede observar que el primer acento nuclear en el nivel concreto consiste en un ascenso melódico con pico tonal desplazado a la sílaba postónica ( $L+>_jH^*$ ), mientras que los otros dos acentos tonales son moderadamente altos ( $!H^*$ ) y el tonema es  $L^* =\%$ , que implica un tono bajo en la sílaba nuclear que se sostiene hasta el final del enunciado. Ahora bien, en el nivel abstracto, se puede observar que el primer acento nuclear corresponde con el constructo de nivel superior (ascenso melódico,  $L+H^*$ ), los otros dos acentos nucleares se relacionan con el acento abstracto  $H^*$  y que el tonema se transcribe como  $L^* L\%$ , pues corresponde con un acento nuclear bajo y una juntura igualmente baja.

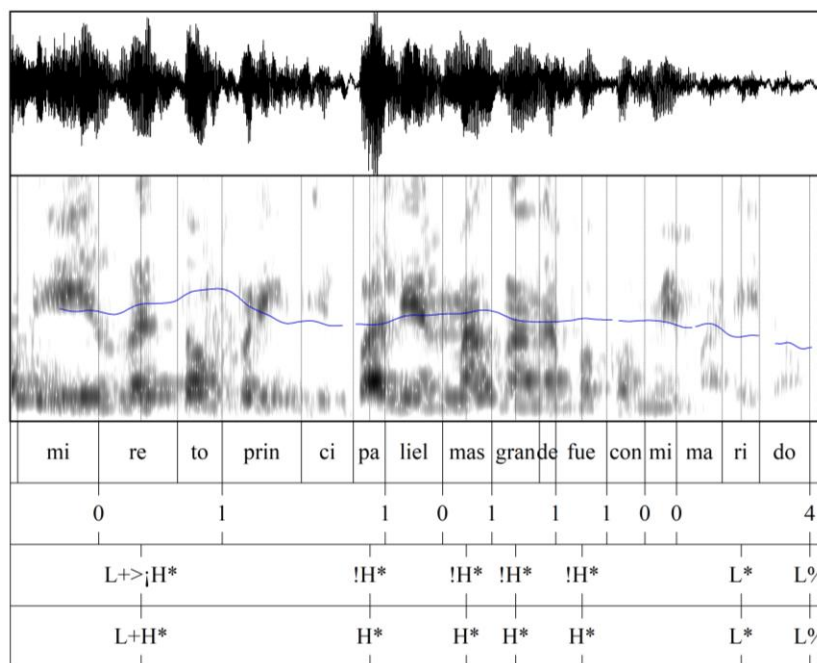
Figura 2.3. Ejemplo de transcripción en dos niveles (concreto y abstracto) del enunciado Rs1M2-30: [(*Mi mamá nos pegaba con lo que ella encontraba*)<sub>ip</sub>]<sub>IP</sub>



La principal ventaja metodológica de este tipo de análisis es que se puede llevar a cabo un estudio detallado de fenómenos entonativos concretos, en los que es posible observar el grado de variabilidad presente en los distintos acentos tonales y tonos de juntura sin perder de vista

la abstracción subyacente de la que proceden (como en el trabajo de Dorta *et al.*, 2013b). De igual forma, permite observar el detalle de la realización concreta de una forma subyacente que puede ser fonéticamente ambigua. Por ejemplo, el tonema abstracto  $L^* L\%$  puede tener dos realizaciones, una de ellas es la suspensión tonal que se presenta en la figura anterior ( $L^* =\%$ ); y la otra sería la realización que se presenta en la figura 2.4, en la que el tono del acento nuclear no se sostiene hasta la juntura, sino que sigue descendiendo hasta un tono aún más bajo. La distinción entre estas dos juntas, *a priori*, puede que no parezca relevante, pero no se puede desechar la idea de que un tono de juntura bajo sostenido desde la sílaba nuclear ( $=\%$ ) y una juntura más baja que el acento nuclear ( $L\%$ ), probablemente, respondan a una variación del significado pragmático del enunciado, como un sentido de mayor certeza acerca de lo enunciado, marcado con un descenso final más pronunciado (por ejemplo) o a algún otro factor lingüístico.

Figura 2.4. Ejemplo de transcripción en dos niveles (concreto y abstracto) del enunciado Rs2M2-39: [(*Mi reto principal y el más grande fue con mi marido*)<sub>ip</sub>]<sub>IP</sub>



Otra ventaja que aporta la presencia del nivel abstracto es que permite incluir múltiples realizaciones concretas con rasgos centrales invariables bajo una sola representación abstracta, lo que permite tener una idea más clara del elemento entonativo (subyacente)

estudiado. Como se observa en las figuras 2.3 y 2.4, los dos primeros acentos tonales de cada figura (distintos en el nivel concreto) corresponden con una misma forma en el nivel abstracto en cada enunciado, mientras que ambos tonemas (igualmente diferentes en el nivel concreto) convergen en una misma configuración en el nivel abstracto. De esta manera y como se verá más adelante en este trabajo (en los capítulos III y IV), al hacer un estudio de los tonemas, en los que se maneja la combinación de acento nuclear y tono de juntura en una misma variable, se puede pasar de un conjunto elevado de elementos a uno más reducido sin comprometer la validez del proceso de agrupación, pues se realiza con el soporte de una teoría cognitiva sólida.

#### 2.2.10. Análisis estadístico

Los valores acústicos y fonológicos registrados en la base de datos se procesaron a través del programa estadístico SPSS (IBM Corp., 2011). Este programa resultó de gran utilidad, debido a que cuenta con una gran batería de pruebas estadísticas que se pueden realizar en estudios humanísticos y sociales, lo que funciona bien en el ámbito sociolingüístico. Adicionalmente, permite manejar las variables de forma eficiente durante la ejecución de más de una prueba a la vez, lo que agiliza el proceso de obtención de resultados y, por ende, el descarte o validación de distintos cruces de variables de forma simultánea.

El primer análisis se llevó a cabo con los valores acústicos de la duración (milisegundos), intensidad (decibelios) y frecuencia fundamental (hercios) en determinadas sílabas de los enunciados, situadas al inicio, medio y final de las frases entonativas intermedias y mayores. Esto se hizo porque se han documentado fenómenos entonativos relevantes tanto al inicio como al final de los enunciados, especialmente relacionados con inflexiones iniciales, acentos tonales en la primera sílaba tónica y rasgos de duración y entonación en las sílabas que conforman el tonema (Navarro Tomás, 1957 y 1966; Sosa, 1999; De-la-Mota *et al.*, 2010). Por lo tanto, se consideró oportuno analizar estas secciones de los enunciados para observar si en estos puntos se pueden encontrar índices sociolingüísticos. Las sílabas que se tomaron en cuenta fueron las siguientes: primera sílaba, primera sílaba tónica, sílabas del cuerpo del enunciado,<sup>14</sup> núcleo de frase entonativa

---

<sup>14</sup> Con base en los planteamientos de Navarro Tomás (1966), se considera cuerpo del enunciado aquellas sílabas que se encuentran entre la primera y la última sílaba tónica (exceptuando la última sílaba tónica y la final de las frases entonativas intermedias).

intermedia (ip), linde de frase entonativa intermedia (ip),<sup>15</sup> núcleo de frase entonativa mayor (IP), linde de frase entonativa mayor (IP).

En relación con las sílabas analizadas, es necesario hacer dos precisiones. La primera es que se toman como sílabas pertenecientes al cuerpo del enunciado todas aquellas que no son la primera sílaba, la primera sílaba tónica, núcleo o linde de ip o IP. La segunda aclaración es que se tomaron en cuenta, adicionalmente, todas aquellas sílabas que constituyeran una combinación de dos o más de las categorías mencionadas, excluyendo el cuerpo, por lo que, por ejemplo, se tomó como una categoría aparte aquellas sílabas que son, simultáneamente, la primera del enunciado y la primera tónica (como ocurriría en un enunciado que comenzara con una sílaba tónica). En estos análisis, por tratarse de variables continuas, se aplicaron las pruebas estadísticas ANOVA y HSD de Tukey, en caso de que la primera prueba descartara la hipótesis nula,<sup>16</sup> las cuales permiten conocer si las diferencias de medias son estadísticamente significativas y cuáles niveles de una variable son los que presentan medias significativamente diferentes, respectivamente (LeBlanc, 2004).

Los datos de la primera sílaba tónica, intertonemas (tonemas internos) y tonemas fueron procesados en contraste con las variables extralingüísticas (red social y sexo), mediante tablas de contingencia, sobre las que se aplicó la prueba no paramétrica del estadístico exacto de Fisher, la cual permite establecer la relación entre variables cuando se dan pocos casos (5 o menos) en alguno de sus niveles (Pedersen, 1996). De igual modo, se empleó la prueba ANOVA, para identificar diferencias significativas entre las duraciones absolutas de las medias de duración de la primera sílaba tónica, producida por los distintos grupos de hablantes (agrupados por red social y sexo). En el caso de los tonemas (tanto de ip como de IP), se analizó, adicionalmente, la duración relativa de las sílabas que constituyen

---

<sup>15</sup> Cabe recordar que cuando se emplea el término “frase entonativa intermedia” (ip) se hace referencia a las ip no finales, es decir, a aquellas cuya sección nuclear (al extremo derecho o final) no coincide con la de la frase entonativa mayor (IP). Por otra parte, el término “frase entonativa mayor” (IP) está vinculado con la unidad prosódica cuya extensión coincide con la del enunciado. Aunque el núcleo y linde derecho de una IP correspondan con los de una ip, se consideran únicamente como núcleo y linde de IP, debido a que esta unidad tiene un mayor rango en la jerarquía, al tener no solo una mayor extensión, sino significado semántico completo.

<sup>16</sup> En estadística, la hipótesis nula es la afirmación contraria a la hipótesis de partida de un análisis (llamada hipótesis alternativa). Por ejemplo, si se plantea la hipótesis alternativa de que un fenómeno ocurre de forma diferente en dos grupos poblacionales, la hipótesis nula corresponde a la afirmación de que tal diferencia no existe. En este trabajo, se usó como referencia una significación ( $p$ ) con valor menor a 0.05 para validar la hipótesis alterativa (y descartar la hipótesis nula).

sus configuraciones nucleares, es decir, la duración más prolongada de una de ellas en detrimento de la otra sílaba.

Adicionalmente, se analizaron algunas funciones sintáctico-discursivas que pudieran estar incidiendo en el empleo de un tipo de configuración nuclear. En el caso de las frases entonativas intermedias (ip), se tomaron en cuenta las funciones: foco informativo estrecho y expresividad como elementos pragmáticos; condicionalidad (que la ip fuese la prótasis de una oración condicional) y yuxtaposición (que la ip analizada se realizara antes de una cláusula yuxtapuesta) como rasgos sintácticos; explicación (que la ip estudiada tuviera algún sentido explicativo relacionado con otro elemento cotextual o contextual), y extraoracionalidad (subdividida en tema, iniciador, orientador de escenario y orientador de tiempo)<sup>17</sup> como características discursivas. Cabe señalar que únicamente las variables pragmáticas (foco estrecho y expresividad) presentan solapamiento con respecto a otras variables. Este solapamiento se debe a que se registraron casos en los que se producía, por ejemplo, una cláusula condicional (prótasis) con foco estrecho en su sección final, un elemento extraoracional producido de forma enfática (con matiz expresivo) o, incluso, una cláusula con función explicativa pronunciada con expresividad y foco estrecho. Se tomó tal decisión debido a que no se considera que las variables pragmáticas sean mutuamente excluyentes ni que su presencia excluya alguna de las otras funciones clasificadas como sintácticas o discursivas. Estas dos funciones, en cambio, sí se consideraron mutuamente excluyentes, por lo que los casos registrados en cada uno de los niveles de estas variables no coinciden con alguno de las demás categorías. Por ejemplo, una cláusula clasificada como prótasis de una oración condicional no se clasificó como preyuxtapuesta, explicativa o extraoracional. En el caso de las frases entonativas mayores (IP), por otra parte, solo se tomaron en cuenta los focos informativos estrechos y los matices expresivos, dejando de lado las demás funciones sintáctico-discursivas (ya sea porque no eran pertinentes, como la condición de prótasis, o porque no se registraron casos).

---

<sup>17</sup> Las categorías consideradas como elementos extraoracionales se retomaron del trabajo de Martín Butragueño (2019)

## Capítulo III

### Análisis fonético-fonológico

En este capítulo, se expone un análisis de las características prosódicas de las muestras de habla recolectadas en el corpus, en términos fonéticos y fonológicos. De igual forma, se presentan algunos datos sociolingüísticos complementarios, que proporcionan información acerca de los rasgos prosódicos de cada red social, los cuales permiten observar el grado de heterogeneidad prosódica en el habla de ambas redes. La presentación de resultados consta de los siguientes aspectos: características de la duración, intensidad y frecuencia fundamental en determinadas sílabas del enunciado, distribución de acentos tonales en la primera sílaba tónica, en el tonema interno (o intertonema) y en el tonema. Cabe señalar que el análisis presentado en este capítulo será la base para la interpretación sociolingüística expuesta en el capítulo V.

#### 3.1. Duración, intensidad y frecuencia fundamental (F0) en el enunciado

En esta sección, se analizan los valores de duración, intensidad y F0 de siete categorías silábicas del enunciado, considerando que, por su naturaleza de linde demarcativo, pueden existir diferencias significativas en las emisiones de los individuos que conforman el corpus. Dichas categorías, que se pueden ver ejemplificadas en la figura 3.1 de acuerdo con el número asignado, son las siguientes: (1) primera sílaba, (2) primera sílaba tónica, (3) núcleo de frase entonativa intermedia (ip), (4) linde de frase entonativa intermedia (ip), (5) núcleo de frase entonativa mayor (IP), (6) linde de frase entonativa mayor (IP) y (7) sílabas del cuerpo del enunciado. Vale la pena recordar que se entienden por sílabas del cuerpo del enunciado todas aquellas que no constituyen una de las sílabas mencionadas anteriormente. Asimismo, cuando se mencionan las frases entonativas intermedias (ip) se hace referencia únicamente a aquellas no finales, es decir, a las ip cuyo núcleo no coincide con el de la frase mayor (IP). Además, se toman en cuenta aquellos casos en los que una sílaba pertenece a más de una de las categorías mencionadas, por ejemplo, cuando una sílaba es la primera del enunciado y la primera tónica al mismo tiempo, tal como ocurre en el enunciado de la figura 3.2. Casos como este se incorporaron al análisis general para determinar si la convergencia de categorías incide en los valores de los rasgos prosódicos o no.

Figura 3.1. Oscilograma, espectrograma y F0 del enunciado Rs2H1-40: [(*Logró salirse*)<sub>ip</sub> (y te puede borrar algo que tu hayas hecho)<sub>ip</sub>]<sub>IP</sub> con señalamiento de las sílabas analizadas

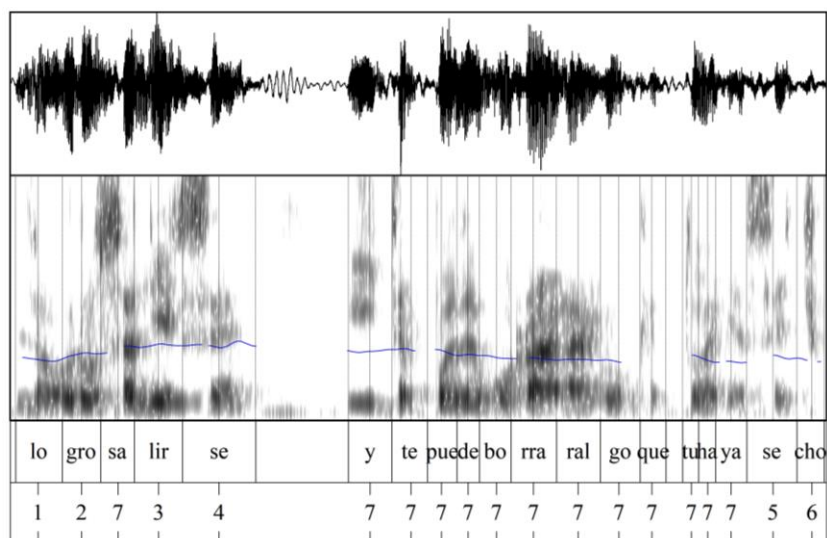
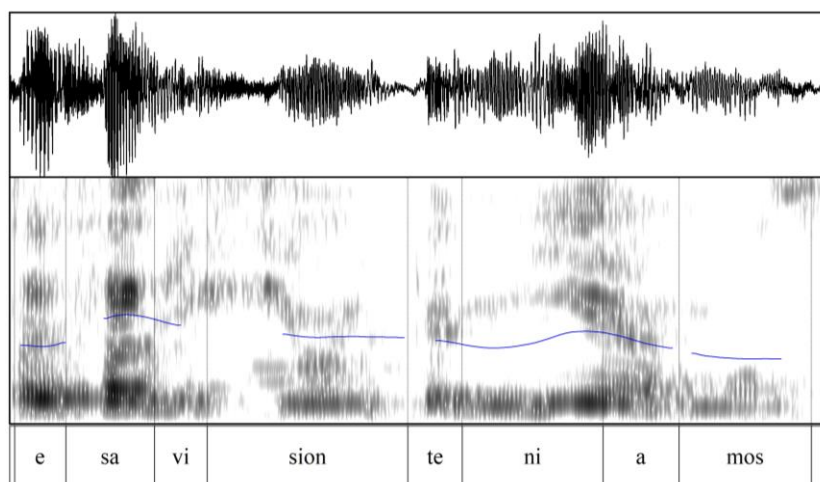


Figura 3.2. Oscilograma, espectrograma y F0 del enunciado Rs2H3-13: [(*Esa visión teníamos*)<sub>ip</sub>]<sub>IP</sub>



Debido a que las tres variables lingüísticas estudiadas son continuas, se aplicó en todos los casos la prueba ANOVA de un factor, para identificar diferencias significativas en los valores medios presentes en cada sílaba analizada. En aquellos casos en los que la diferencia fue significativa, se aplicó la prueba HSD de Tukey para identificar cuáles son las sílabas con valores medios contrastantes. Es importante mencionar que los patrones fonéticos

identificados en las secciones siguientes (§3.1.1, §3.1.2 y §3.1.3) no permiten un análisis cognitivo basado en el uso, debido a que se manejaron valores absolutos y no de frecuencia de producción. La razón de esto es que, al estudiar datos de habla no controlada, el número de producciones de cada tipo silábico por enunciado resulta impredecible, lo que imposibilita la creación de algún sistema de normalización de las variables analizadas que permitiese cuantificar la frecuencia con la que una sílaba es más larga o tiene mayor valor de intensidad o F0 que otra.

### 3.1.1. Duración en el enunciado

En general, la duración de las sílabas estudiadas se manifiesta de forma menos sistemática que la intensidad y la frecuencia fundamental. No obstante, emergen ciertos patrones de comportamiento relevantes, pues la prueba ANOVA resultó ser significativa ( $p = 0$ ) en todos los cruces de duración silábica en el corpus. Antes de entrar en los detalles sociolingüísticos del comportamiento de la duración, se describirán las características generales de esta en el corpus.

Tras aplicar la prueba HSD de Tukey (los resultados completos se pueden ver en el apéndice 3 ), resalta el hecho de que la mayoría de los contrastes de medias se dan con aquellas sílabas que son núcleo o linde de frase entonativa, tanto intermedia (ip) como mayor (IP). Esto se puede observar en el cuadro 3.1, en donde las medias de duración de todas las sílabas de la columna A (relativamente cercanas dentro de un rango de 10.5 ms) contrastan estadísticamente con aquellas en la columna B. Teniendo en cuenta que la diferencia de duración apenas perceptible de dos sílabas adyacentes es de 30 ms (Toledo, 1998 y 2000), se puede decir que la diferencia de duración entre las sílabas de ambas columnas sería perceptible en la gran mayoría de los casos, pues la duración más larga de la columna A (primera sílaba y primera tónica simultáneamente) difiere de la duración más corta de la columna B (linde de IP) por 29.04 ms. Esto deja un margen de diferencia aún mayor (por encima de los 30 ms) para el resto de los casos. Por lo tanto, los valores elevados de duración en las sílabas de la columna B (que involucran núcleos o linde de ip e IP) reflejan la tendencia general dentro del corpus a producir las secciones finales de frases entonativas con mayor duración perceptible que el resto del enunciado.



Cuadro 3.1. Sílabas con medias de duración (ms) estadísticamente diferentes en la totalidad del corpus

Columna A		Columna B	
Sílabas	Media	Sílabas	Media
Cuerpo	139.45	Núcleo de IP	178.08
Primera sílaba	134.68	Linde de IP	174.63
Primera sílaba y primera tónica	145.20	Linde de ip	241.04
Primera tónica	141.11	Primera tónica, núcleo y linde de ip	293.38
		Núcleo de ip	187.52
		Núcleo y linde de ip	271.11

Además, se registraron diferencias significativas entre las sílabas de la columna B. En este grupo, las medias del núcleo de ip, núcleo de IP y linde de IP contrastan con las de las otras tres sílabas restantes, cada una de las cuales se encuentra en el linde de ip. La diferencia de al menos 53.5 ms entre estos dos grupos de contraste es suficiente para establecer una distancia perceptivamente significativa entre las medias de las sílabas situadas en el linde de ip como aquellas con mayor duración. Adicionalmente, la media de la sílaba linde de ip es estadísticamente diferente de las sílabas que son núcleo y linde de ip, simultáneamente; lo que sitúa a estas últimas en el grupo de media más elevadas de duración (junto con aquellas que son primera sílaba tónica, núcleo y linde de ip a la vez). A los datos mencionados hasta este punto, se unen las medias de las sílabas que son primera sílaba, primera tónica y núcleo de ip (112.5 ms) y las que son primera tónica y núcleo de ip (183.6 ms). Si bien las primeras representan la media más baja de todas y las segundas tienen una media cercana a los 200 ms, ambas contrastan significativamente solo con las sílabas con mayor media, es decir, aquellas que son núcleo y linde de ip simultáneamente y primera tónica, núcleo y linde de ip a la vez.

Teniendo en cuenta los resultados del cuadro 3.1 y con base en los tipos de sílabas que han resultado estadísticamente significativas, se plantea el siguiente orden de duración (de menor a mayor) estadísticamente significativo: (cuerpo, primera sílaba, primera sílaba y primera tónica, primera tónica) < (núcleo de ip, núcleo de IP, linde de IP) < (linde de ip) < (núcleo y linde de ip; primera tónica, núcleo y linde de ip).

De acuerdo con el orden propuesto, la duración silábica aumenta en la sección final de las frases entonativas (sean ip o IP). En los casos de las frases entonativas intermedias

(ip), la duración aumenta significativamente del núcleo al linde. En aquellos casos en los que una sílaba linde de ip posee otro rasgo adicional (como ser también el núcleo o la primera sílaba tónica), su duración aumenta considerablemente. No ocurre lo mismo con la sílaba núcleo de ip, ya que su duración no aumenta considerablemente e, incluso, puede llegar a reducirse si se encuentra cerca del inicio del enunciado (como ocurre con los núcleos que también son primera sílaba). Por lo tanto, la mayor duración en el final de ip se puede atribuir al hecho de que la sílaba se encuentra al final de dicha frase entonativa, especialmente si esta termina con una sílaba tónica, en cuyo caso la duración aumenta aún más.

De acuerdo con los datos, la duración de las sílabas de los enunciados analizados se puede dividir en tres tipos: corta (134 ms - 157 ms), media (158 ms - 193 ms) y larga (194 ms - 294 ms).<sup>18</sup> Entre las sílabas más cortas, están las primeras sílabas (tónicas o no) y las del cuerpo del enunciado. En cuanto a la duración media, tanto en ambas redes como en ambos sexos, se encuentran las sílabas nucleares de frase entonativa intermedia (ip), los núcleos y los lindes de frases entonativas mayores (IP). En este grupo de sílabas con duración media, los hablantes de la red social laboral incluyen también aquellas sílabas que son, simultáneamente, la primera tónica del enunciado y núcleo de ip, lo que marca una distinción con respecto a los hablantes de la red social vecinal, quienes la producen con menor duración, lo que aporta el primer indicio de la relevancia del núcleo de ip en términos de duración. Las sílabas con duración larga incluyen, en líneas generales, las que constituyen linde de ip y las que son núcleo y linde de ip al mismo tiempo. No obstante, la red laboral solo incluye en esta categoría las sílabas lindes de ip. Por su parte, las mujeres, además de las sílabas mencionadas previamente, también producen en este grupo de duración aquellas sílabas que son, a la vez, la primera sílaba tónica, núcleo y linde de ip. Dicha tendencia por parte de las mujeres coincide, en cierta forma, con el alargamiento realizado por la red social laboral en las primeras sílabas tónicas en el núcleo de ip. Finalmente, la red social laboral se caracteriza por presentar un cuarto nivel de duración (con una duración mayor que las anteriores), en el que se incluyen las sílabas que son núcleos y lindes de ip a la vez. La mayor duración de

---

<sup>18</sup> Los valores para identificar las duraciones silábicas como cortas, medias o largas se determinaron tomando en cuenta tanto las medias de duración de cada red como las medias generales del corpus. En este proceso, se calculó la diferencia entre el valor de duración máximo y el mínimo, el cual se dividió entre tres para definir el rango de cada tipo de duración. Dicho valor se sumó a la duración más baja para determinar el rango de duración corto, luego se sumó al valor máximo del rango bajo para establecer el rango medio y, finalmente, se siguió el mismo proceso para el rango de duración larga.

dichas sílabas es única de esta red social, lo que la distingue de todos los demás grupos del corpus, por lo que (como se confirmará más adelante, en §3.2) representa un rasgo prosódico distintivo de los miembros de la red laboral.

### 3.1.2. Intensidad en el enunciado

La intensidad, aunque igualmente presentó una diferencia de medias estadísticamente significativa según la prueba ANOVA ( $p=0$ ), no refleja mayor variación ni contrastes en las sílabas analizadas. En el plano general, solo los valores medios de tres sílabas contrastan de forma relevante tras ejecutar la prueba HSD de Tukey (ver apéndice 4): i) primera sílaba y primera tónica a la vez (73.7 dB), ii) núcleo de IP (69.32 dB) y iii) el linde de IP (60.22 dB). La media del primer tipo de sílaba difiere estadísticamente de la media de las sílabas del cuerpo del enunciado (70.68 dB), del núcleo de IP, linde de IP y aquellos casos en los que la sílaba es núcleo y linde de ip simultáneamente (68.38 dB). Debido a que la diferencia de medias entre la primera sílaba y primera tónica y las sílabas del cuerpo del enunciado es inferior a 4 dB, umbral de percepción a partir del cual se identifican variaciones en la intensidad (Dorta *et al.*, 2017), no se puede tomar en cuenta como un contraste lingüísticamente significativo, aunque sí lo sea en el sentido estadístico.

No ocurre lo mismo con el resto de los casos, pues la diferencia de medias es superior a 4 dB. Esto incluye las sílabas núcleo de IP (que también contrasta estadísticamente de la primera sílaba tónica, que tiene una media de 73.09 dB) y del linde de IP (que difiere estadísticamente de las medias de todas las demás sílabas), las cuales se ubican entre los valores medios más bajos de todas las sílabas. La diferencia perceptible en los valores de intensidad implica un descenso de intensidad desde la primera sílaba tónica hasta el final del enunciado.<sup>19</sup> Aunque la media de intensidad de las sílabas del linde de ip suele ser inferior a la de las sílabas al inicio y en el cuerpo del enunciado, la diferencia de intensidad no solo carece de relevancia estadística, sino también perceptiva, pues no supera el umbral de percepción. Por lo tanto, no se puede afirmar que la disminución de la intensidad en las sílabas al final de las frases entonativas intermedias (ip) sea un factor relevante. Esta disminución puede deberse a que las ip no siempre están seguidas por una pausa o algún otro

---

<sup>19</sup> Aunque la diferencia de los valores medios de intensidad entre la primera sílaba tónica y el núcleo de frase entonativa mayor (IP) se aproxima al umbral de percepción (se encuentra 0.23 dB por debajo del umbral de percepción), la diferencia de intensidad entre la primera tónica y el linde de IP sí es perceptible.

elemento lingüístico que conlleve una reducción de intensidad. No obstante, la declinación general en la intensidad corresponde con los hallazgos de Jorge Trujillo *et al.* (2017) en el español de Canarias, con la de Martín Butragueño (2014, 2019) y con la noción general relacionada con el código biológico de producción (Gussenhoven, 2004), según la cual, la intensidad tiene su punto máximo al inicio del enunciado, desde donde presenta un descenso gradual hasta el final de este.

Todos los grupos analizados (tanto las dos redes sociales como los dos sexos) presentan patrones de intensidad bastante uniformes y en consonancia con los datos generales ya expuestos. Por tal motivo, no se considera que la intensidad represente un factor relevante en términos sociolingüísticos.

### 3.1.3. Frecuencia fundamental (F0) en el enunciado

En el análisis general, la prueba HSD de Tukey (ver apéndice 5) reveló que los contrastes significativos de medias de F0 ocurren con dos sílabas: núcleo de IP (80.19 st) y linde de IP (55.16 st). La media de F0 del núcleo de IP es significativamente diferente de aquellas de la primera sílaba del enunciado cuando es simultáneamente la primera tónica, la primera sílaba tónica y del linde de IP. Las dos primeras sílabas poseen medias que se encuentran entre los valores más altos (88.98 st y 87.96 st, respectivamente), únicamente superados por las sílabas lindes de ip que se encuentran cerca del inicio del enunciado (ya sea porque son al mismo tiempo la primera sílaba tónica o la primera sílaba en casos de ip monosilábicas). Además de que la diferencia de medias de F0 de las sílabas mencionadas con respecto a la del núcleo de IP es significativa en términos estadísticos, también lo es en términos perceptivos, pues supera el umbral de percepción tonal de 1.5 st (Pamies Bertrán *et al.*, 2002; Murrieta, 2016). La media de F0 en la sílaba nuclear (80.19 st) es perceptiblemente inferior que las del cuerpo del enunciado (83.97 st), lo que resalta el grado del descenso entonativo con respecto a la mayoría del enunciado. Por su parte, la media del linde de IP contrasta con todas las demás sílabas estudiadas, lo que la clasifica como la sílaba con la F0 significativamente más reducida. Cabe mencionar que lo expuesto en este párrafo se adecua a los postulados definidos por Navarro Tomás (1957) en su descripción de la entonación de los enunciados declarativos, ya que se observa una elevación tonal significativamente mayor en la primera

sílaba tónica y un descenso progresivo desde la última sílaba tónica hasta el final del enunciado (el descenso no necesariamente inicia en la última sílaba tónica).

En general, todos los datos de F0 apuntan hacia un descenso paulatino en la entonación de los enunciados desde la primera sílaba tónica, lo que coincide con lo expuesto por Martín Butragueño (2014, 2019) en su descripción de la prosodia de enunciados aseverativos en el español del centro de México. Esta tendencia es visible en los análisis llevados a cabo con las dos variables extralingüísticas, por lo que no se puede relacionar el patrón entonativo registrado con ninguna de ellas en concreto. No obstante, existen diferencias llamativas en lo relacionado con el campo tonal empleado por los hablantes. Los campos tonales de la red laboral y de las mujeres son considerablemente mayores que los de la red vecinal y los hombres.<sup>20</sup>

### 3.2. Primera sílaba tónica

El cuadro 3.2 permite observar que H\* es el acento tonal abstracto más frecuente y, por ende, el cognitivamente prototípico en esta posición de los enunciados aseverativos. La realización concreta central de este acento es un acento tonal alto (H\*), lo que deja al acento moderadamente alto (!H\*) en segundo lugar, en términos de frecuencia, no solo en relación con el acento fonológico H\*, sino con la totalidad de los acentos registrados del corpus. El segundo acento abstracto más frecuente y, por lo tanto, más cercano al prototipo es el acento ascendente L+H\*. Las representaciones concretas más comunes de este acento fonológico son las realizaciones con picos de F0 desplazados a la sílaba postónica (L+>H\* y L+>¡H\*, con excursiones tonales promedio de 1.8 st y 4.5 st, respectivamente), seguidas por aquellas con picos tempranos, es decir, alineados con el linde de la sílaba tónica (L+H\* y L+¡H\*, cuyas medias de ascenso son de 2.1 st y 4.8 st para cada caso) y en último lugar por los acentos con picos muy tempranos en los que el punto más alto de F0 se alcanza antes del final de la sílaba tónica (L+<H\* y L+<¡H\*, que presentan una elevación tonal promedio de 1.5 st y 4.2 st, respectivamente). Estos acentos con picos muy tempranos se encuentran entre los menos frecuentes, pues representan el 3.1 % de los acentos documentados, junto con

---

<sup>20</sup> Las diferencias en el campo tonal de ambas redes pueden explicarse debido a una mayor incidencia de fonación no modal (voz laringizada o susurrada), en las sílabas finales de la red vecinal y de las mujeres, lo cual afecta la altura de la F0 en esta parte del enunciado. Tal fenómeno puede constituir el tema de una investigación a futuro.

H+L\* (el menos frecuente de todos con un 2.2 % de frecuencia) y L\* (que representa el 3.9 % de los datos totales).

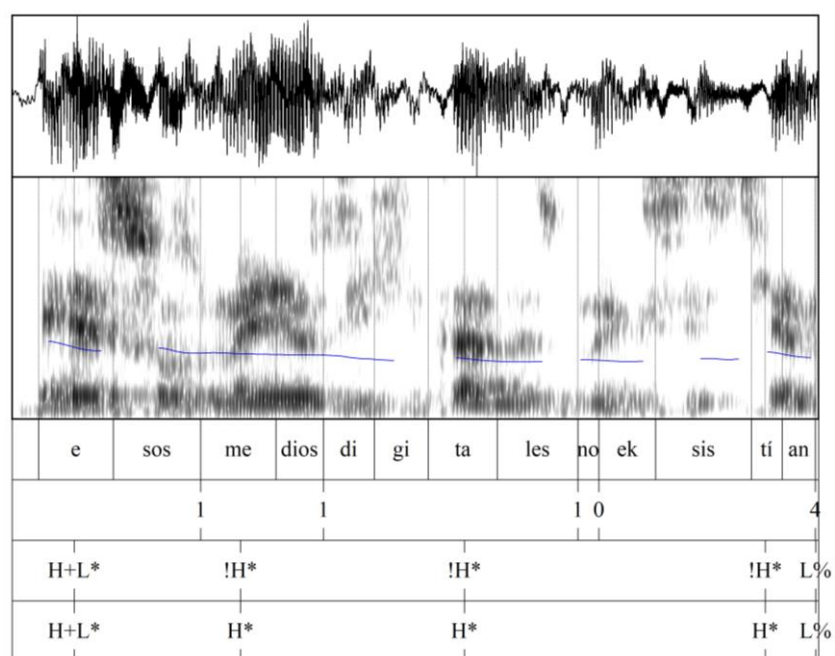
Cuadro 3.2. Recuento de acentos tonales de la primera sílaba tónica en los niveles concreto y abstracto. N=360

Acentos tonales concretos			Acentos tonales abstractos		
Acentos tonales	Casos	%	Acentos tonales	Casos	%
H*	125	34.7	H*	193	53.6
!H*	68	18.9			
L+H*	20	5.6	L+H*	145	40.3
L+ <sub>j</sub> H*	22	6.1			
L+<H*	5	1.4			
L+< <sub>i</sub> H*	6	1.7			
L+>H*	47	13.1			
L+> <sub>i</sub> H*	45	12.5			
L*	14	3.9	L*	14	3.9
H+L*	8	2.2	H+L*	8	2.2
Total	360	100	Total	360	100

Debido al bajo porcentaje de acentos tonales ascendentes con picos muy tempranos, descendentes y bajos, se decidió prestar especial atención a estos tres acentos. En relación con los acentos con picos muy tempranos, de los 11 casos registrados, 5 ocurren en casos en los que la palabra portadora de la primera sílaba tónica está focalizada (todas ellas producidas por hablantes de la red laboral) y las otras 5 en enunciados pronunciados de forma enfática, es decir, con un matiz expresivo (la mayoría de los cuales fueron generados por la red laboral). Por otra parte, 5 de las 8 realizaciones del acento descendente H+L\* se presentan en enunciados producidos con énfasis por parte de los hombres pertenecientes a la red social laboral (véase figura 3.3); mientras que 1 caso ocurre en una sílaba producida con voz laringizada en un enunciado emitido sin tipo alguno de énfasis. El hecho de que la mayoría de los casos de este acento tienen lugar en enunciados enfáticos permite inferir la existencia de una correlación entre ambos factores. Finalmente, el acento bajo (L\*) parece estar relacionado con la fonación o articulación de los hablantes, pues 8 de los 14 casos ocurren en contextos fónicos en los que se observa la presencia de voz laringizada (N=4) o debilitamiento vocálico (N=4). No obstante, tanto esta explicación como la relacionada con

los acentos descendentes y los ascendentes con picos muy tempranos no pueden ser tomadas como afirmaciones categóricas, pues los datos recolectados en relación con focos estrechos, énfasis, voz laringizada y debilitamientos vocálicos en la primera sílaba tónica del corpus son escasos y la mayoría de ellos son producidos por hombres y en hablantes de la red laboral. Por tal motivo, el estudio cualitativo de estos rasgos se abordará en el análisis sociolingüístico presentado en §4.2.

Figura 3.3. Oscilograma, espectrograma y F0 del enunciado Rs2H1-22: [(*Esos medios digitales no existían*)<sub>ip</sub>]<sub>IP</sub>



Si bien la red social fue la única variable extralingüística que reflejó una diferencia significativa en relación con los acentos tonales en la primera sílaba tónica, de acuerdo con el estadístico Exacto de Fisher ( $p = 0.007$ ), no se puede afirmar que esta diferencia tenga relación con la distribución de los acentos tonales producidos por ambas redes. Los únicos acentos que presentan una diferencia mayor a las 8 ocurrencias entre ambas redes son !H\* (con 12 casos de diferencia a favor de la red vecinal) y L+>H\* (con una diferencia de 11 casos a favor de la red laboral). Por otra parte, 8 de los 11 casos de picos muy tempranos son producidos por hablantes de la red social laboral, al igual que las realizaciones del acento descendente H+L\*. Por lo tanto, existe la posibilidad de que la diferencia estadísticamente

significativa esté más relacionada con la naturaleza pragmática o tipo de fonación en la enunciación. Como se mencionó en el párrafo anterior, las ocurrencias de estos acentos están relacionadas con la presencia de focos estrechos, énfasis, voz laringizada y debilitamientos, todos ellos dados principalmente en la red laboral, con muy poca producción de estos por parte de la red vecinal. Por ende, sería arriesgado establecer una correlación entre los fenómenos descritos y la red social laboral.

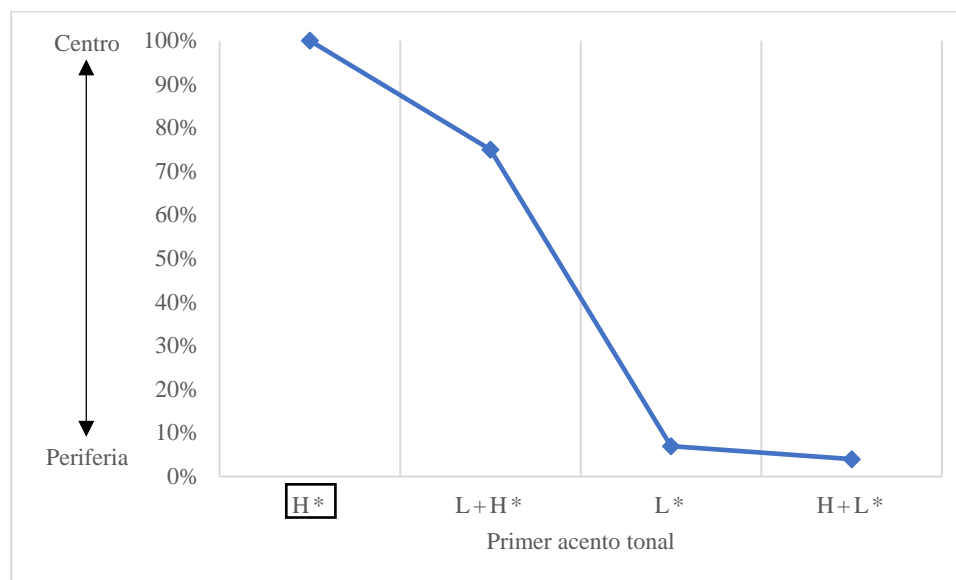
Teniendo en cuenta todo lo expuesto anteriormente, se puede proponer la organización cognitiva mostrada en la figura 3.4 para el primer acento tonal del enunciado.<sup>21</sup> Se maneja la noción de que el primer acento tonal es una categoría cognitiva debido a que tiene relevancia prosódica, pues, particularmente en la primera sílaba tónica, se han documentado fenómenos relacionados con valores pragmáticos del enunciado, como ocurre en los casos con focalización (De-la-Mota *et al.*, 2010). En la figura 3.4 se puede observar que el acento H\* es el prototipo de la categoría, seguido por L+H\* (el acento con mayor grado de centralidad después del prototipo), el cual está drásticamente separado de L\* y H+L\*, que están tan alejados del prototipo que se encuentran en la periferia de la categoría. Es importante tener en cuenta que la organización que se presenta en este punto debe ser asociada únicamente con enunciados aseverativos, concretamente, con aquellos carentes de énfasis al inicio y producidos con voz modal. Esto se debe a que los enunciados con características distintas pueden implicar un cambio importante en la frecuencia de uso de algunos de los acentos tonales presentados y, por ende, una diferencia en la centralidad (o prototipicidad) dentro de la categoría en cuestión.

---

<sup>21</sup> Para determinar el grado de centralidad o prototipicidad, se tomó como referencia el número de casos del esquema entonativo más frecuente (de acuerdo con la premisa de que el esquema más recurrente – el prototipo – se encuentra en el centro de la categoría). El valor de dicha frecuencia constituye el 100 % de centralidad, en relación con el cual se determina el grado de centralidad de los demás esquemas entonativos. Por ejemplo, si el esquema *x* cuenta con 200 casos de producción, el *y* con 100 y el *z* con 50, el grado de centralidad de *x* (el más frecuente) será de 100 %, el de *y* de 50 % y el de *z* de 25 %. de acuerdo con este método se generaron todas las figuras en las que se presenta el grado de centralidad en la categoría cognitiva estudiada en cada caso.



Figura 3.4. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente al primer acento tonal



En relación con la duración de la primera sílaba tónica, de acuerdo con la prueba ANOVA, la media de duración no resultó significativa entre los acentos tonales abstractos ( $p = 0.135$ ), pero sí lo fue en el caso de los acentos tonales en el nivel concreto. Como se puede ver en el cuadro 3.3, la media más alta es la del acento tonal  $L+<_iH^*$ , lo que tendría sentido al tener en cuenta el espacio temporal que puede ser necesario para que el tono ascienda un mínimo de 3st antes del final de la sílaba. Las otras duraciones medias más prolongadas, en orden descendente, pertenecen a los acentos  $L+_iH^*$ ,  $L+H^*$  y  $+L+<H^*$ , todos ellos con medias perceptiblemente inferiores a la de  $L+<_iH^*$  y superiores a las medias del resto de los acentos tonales, pues la diferencia en todos los casos supera los 30 ms, que se ha establecido como la diferencia de duración mínima perceptible (Toledo, 1998 y 2000). El resto de los acentos tonales tienen medias relativamente homogéneas, que van desde los 124.57 ms hasta los 142.59 ms, por lo que no superan la diferencia de duración media necesaria para distinguir su longitud en términos perceptivos. No obstante, este rango de duración se encontraría más cerca de lo que se consideraría la duración prototípica de la primera sílaba tónica, la cual se podría situar, teniendo en cuenta los datos, alrededor de los 140 ms.

Cuadro 3.3. Medias de duración (ms) de la primera sílaba tónica

Acentos tonales fonéticos	Media
H*	142.59
!H*	141.54
L+H*	181.35
L+ <sub>j</sub> H*	196.59
L+<H*	174.2
L+< <sub>j</sub> H*	288.67
L+>H*	131.3
L+> <sub>j</sub> H*	132.93
L*	124.57
H+L*	132
Total	147.1

p =0 (ANOVA)

La duración media que tiene la primera sílaba tónica en relación con los acentos tonales abstractos no resultó significativa en ninguna de las redes sociales, de acuerdo con la prueba ANOVA ( $p =0.502$ ), lo que es esperable, teniendo en cuenta que las medias de duración de los diferentes acentos tonales son relativamente uniformes. La mayoría de los casos tienen mayor duración en las emisiones de la red laboral. La diferencia más marcada está en la media de duración del acento tonal abstracto L\*, con un margen de 28.25 ms a favor de la red social vecinal, el cual no supera el umbral de 30 ms que permitiría percibir la diferencia de duración media. Por otra parte, la duración no resultó ser significativa en relación con el sexo de los informantes ( $p =0.710$ ). No obstante, llamó la atención que las medias de duración en los datos de los hombres de la red vecinal son similares a las de las mujeres de la red laboral, mientras que la de los hombres de la red laboral son similares a las de las mujeres de la red vecinal. Estas semejanzas entre unos y otros únicamente resultarían perceptivamente (y fonológicamente) significativa en el caso del acento tonal abstracto L\*, que presenta una diferencia de al menos 49 ms entre dichos grupos de hablantes (93 ms en los hombres de la red laboral y 102 ms en las mujeres de la vecinal), en contraste con la de las mujeres de la red laboral y los hombres de la vecinal (152 ms y 159 ms, respectivamente). ende acuerdo con estos resultados, se estima que existen rasgos entonativos y de duración suficientemente diferentes entre ambas redes como para considerarlos potenciales índices sociolingüísticos.

Por lo tanto, en §4.2 se presenta un análisis sociolingüístico de los datos recolectados en relación con la primera sílaba tónica.

### 3.3. Tonemas internos (intertonemas): Configuraciones nucleares de frase entonativa intermedia (ip)

En este apartado, se muestran los resultados del análisis de los intertonemas de las frases entonativas (ip), es decir, con las variaciones prosódicas que ocurren desde la última sílaba tónica hasta el final de la ip. Debido a que un conjunto de datos de las ip analizadas terminan con una sílaba tónica, el análisis se divide en dos partes: intertonemas no oxítonos (palabras esdrújulas y llanas) y oxítonos (palabras agudas). Esta división se lleva a cabo porque uno de los análisis está relacionado con la duración de cada tonema interno y, dado que los tonemas internos no oxítonos involucran dos sílabas, tanto los resultados estadísticos como la interpretación de los datos en general pueden verse afectados si se agrupan todos los intertonemas. Adicionalmente, esta decisión está respaldada por el hecho de que el resultado de la prueba ANOVA ( $p = 0$ ) realizada con las medias de ambos tipos de intertonemas demostró que existe una diferencia significativa entre sus respectivas duraciones totales.

#### 3.3.1 Tonemas internos (intertonemas) no oxítonos

El número total de intertonemas no oxítonos registrados en el nivel concreto es de 40, los cuales corresponden con 10 intertonemas abstractos, todos ellos visibles en el cuadro 3.4. Las configuraciones abstractas más registradas en el corpus son H\* H- (39.69 %) y L+H\* H- (23.71 %). Con respecto a la primera, que se puede establecer como el ejemplar central o prototípico de esta categoría dada su frecuencia, el 41.6 % de sus realizaciones concretas consisten en ascensos inferiores a 3 st (H-) y 39 % con ascensos de al menos 3 st (HH-) en la sílaba final del enunciado, esto con respecto al acento alto (H\*) o moderadamente alto (!H\*) producido en la sílaba nuclear de la ip. Ninguna de las otras realizaciones concretas del intertonema abstracto H\* H- supera los 8 casos, lo que refleja el predominio del empleo de ascensos marcados al final de las frases entonativas intermedias (ip) en general. Este hecho no resulta raro en las ip, pues concuerda con la elevación tonal al final de frase entonativa que caracteriza la continuidad o no conclusión del enunciado registrada en el español; (Navarro Tomás, 1957 y 1966; Quilis, 1981; Toledo, 2007) y es uno de los recursos

discursivos que ocurren a través del código de producción (Gussenhoven, 2004), para indicar que el enunciado continúa después de esta frase entonativa.

Cuadro 3.4. Recuento de tonemas internos no oxítonos en los niveles concreto y abstracto.

N=194

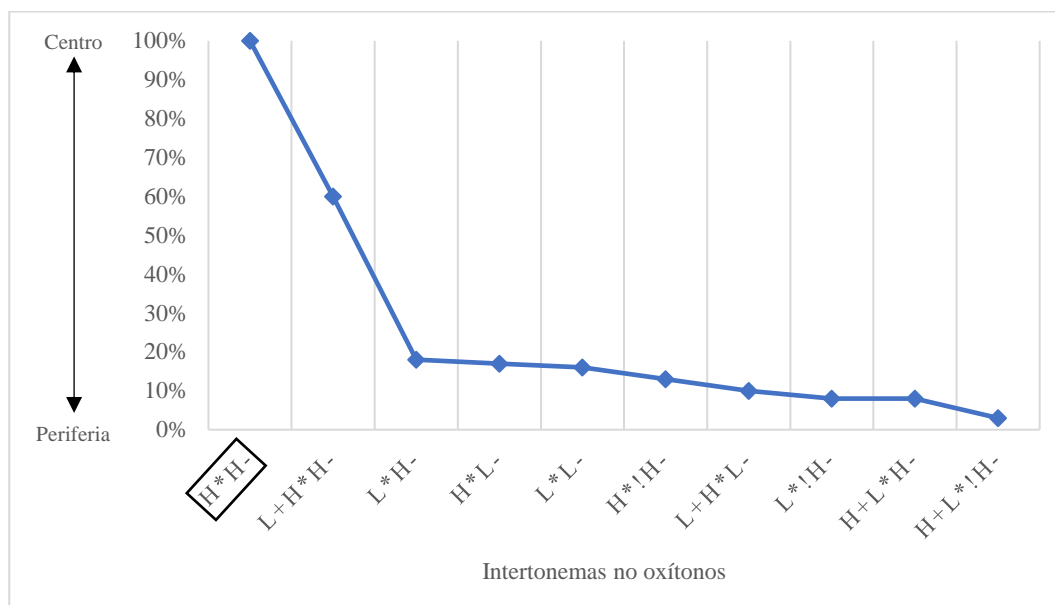
Intertonemas en el nivel concreto	Casos	%	Intertonemas en el nivel abstracto	Casos	%			
H* H-	11	5.67	H* H-	77	39.69			
H* =-	8	4.12						
H* HH-	17	8.76						
H* !HH-	2	1.03						
!H* H-	21	10.82						
!H* HH-	13	6.70						
!H* !HH-	4	2.06						
!H* LH-	1	0.52						
H* !H-	2	1.03				H* !H-	10	5.15
!H* =-	8	4.12	H* L-	13	6.70			
H* L-	4	2.06						
H* HL-	3	1.55						
!H* L-	3	1.55						
!H* HL-	3	1.55						
L+H* H-	4	2.06	L+H* H-	46	23.71			
L+H* =-	3	1.55						
L+H* HH-	16	8.25						
L+H* !HH-	1	0.52						
L+ <sub>j</sub> H* H-	3	1.55						
L+ <sub>j</sub> H* =-	1	0.52						
L+ <sub>j</sub> H* HH-	11	5.67						
L+ <sub>j</sub> H* !HH-	1	0.52						
L+<H* =-	1	0.52						
L+<H* HH-	2	1.03						
L+< <sub>j</sub> H* HH-	2	1.03						
L+< <sub>j</sub> H* !HH-	1	0.52						
L+H* L-	4	2.06				L+H* L-	8	4.12
L+H* HL-	1	0.52						
L+ <sub>j</sub> H* L-	2	1.03						
L+<H* L-	1	0.52						
L* H-	7	3.61	L* H-	14	7.22			
L* HH-	7	3.61						
L* !H-	5	2.58	L* !H-	6	3.09			

L* L!H-	1	0.52			
L* L-	10	5.15	L* L-	12	6.19
L* =-	2	1.03			
H+L* H-	2	1.03	H+L* H-	6	3.09
H+L* HH-	3	1.55			
H+L* !HH-	1	0.52			
H+L* !H-	2	1.03	H+L* !H-	2	1.03
Total				194	100

Ninguno de los otros tonemas internos en el nivel abstracto, como se muestra en el cuadro superior, supera el 7.5 % de los casos, lo que refleja su cualidad cognitiva periférica. Entre estos intertonemas, aquellos que poseen una juntura baja (L-) constituyen el 17.01 % del total de los casos, mientras que el conjunto de los que culminan con una juntura media (!H-) representan el 9.28 %. Al examinar en detalle las juntas L-, se pudo identificar que 18 de los 33 casos (54.6 %) son realizados con voz susurrada o laringizada. La posición de esta juntura, en el linde derecho de una ip en un enunciado producido sin interrupción en el discurso, implica que el 45.4 % de las juntas L- fueron realizadas con la intención de marcar el sentido de continuidad en la enunciación. A pesar de ello, es importante tener en cuenta que todas estas juntas L- (producidas con fonación modal o no) serían percibidas como tonos bajos en el linde de una ip, por lo que, probablemente, el oyente puede atribuir el sentido de continuidad de estas juntas bajas al percibir que el material lingüístico producido hasta ese punto no constituye una unidad de sentido completo, es decir, al inferir que falta información en el discurso emitido hasta dicho punto.

Con base en lo expuesto, en la figura 3.5 se plantea la organización de configuraciones entonativas de la categoría de tonemas internos no oxítonos, de acuerdo con los datos mostrados en el cuadro 3.4. En este gráfico, se puede ver que el esquema entonativo H\* H- tiene el mayor grado de centralidad, seguido por L+H\* H-; mientras que L\* H-, H\* L-, L\* L-, H\* !H-, L+H\* L-, L\* !H- y H+L\* H- se localizan en distintas posiciones cercanas a la periferia; y H+L\* !H- se ubica justo al margen periférico. Es preciso señalar que se considera que los tonemas internos constituyen una categoría cognitiva, debido a que tienen lugar en la sección final de la frase entonativa intermedia, donde ocurren fenómenos prosódicos relevantes, relacionados con la demarcación de la ip y con el sentido de no conclusión de la enunciación.

Figura 3.5. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los intertonemas no oxítonos



Respecto a la duración media de las sílabas que constituyen los tonemas internos (la última sílaba tónica de la ip junto con la sílaba final de ip), se observó que su duración media en los intertonemas concretos no es estadísticamente significativa, de acuerdo con la prueba ANOVA ( $p = 0.86$  para las medias de las sílabas nucleares de ip;  $p = 0.393$  en el caso de las sílabas finales de ip). Esto mismo ocurre en el caso de los intertonemas abstractos, pues las diferencias de medias en los valores de duración de la sílaba nuclear y final de estas configuraciones no son significativas en términos estadísticos. No obstante, el hecho de que la diferencia no sea significativa no quiere decir que no se puedan hacer observaciones generales con respecto a los datos obtenidos. Como se puede ver en el cuadro 3.5, las sílabas lindes de ip tienen una media de duración perceptiblemente mayor a las del núcleo (pues en casi todos los casos supera los 30 ms necesarios para identificar una diferencia de duración).

Cuadro 3.5. Medias de duración (ms) de las sílabas nucleares y lindes en los datos generales del corpus

Intertonemas abstractos	Duración del núcleo	Duración del linde
H* H-	189.01	242.17
H* !H-	180.9	265.8
H* L-	153.77	293.23
L+H* H-	192.67	230.57
L+H* L-	194.5	212.63
L* H-	188.57	250.14
L* !H-	192.83	191.67
L* L-	168.25	201.25
H+L* H-	210	272.5
H+L* !H-	135	323.5
Total	186.22	241.1

p =0.467 (ANOVA duración de las sílabas nucleares)

p =0.135 (ANOVA duración de las sílabas del linde)

Si bien las medias de las duraciones silábicas permiten tener una perspectiva bastante clara de la longitud de cada tipo de sílaba, estas impiden realizar observaciones con base en el uso, pues no dan a conocer con qué frecuencia la sílaba nuclear es más larga o más corta que la sílaba linde de la frase entonativa intermedia (ip). Por lo tanto, se consideró oportuno hacer un análisis que reflejara la tendencia de los hablantes a emitir una sílaba más larga que otra. Como se puede apreciar en el cuadro 3.6, los datos de frecuencia de uso indican que la mayoría de los intertonemas (67.5 %) son producidos con sílabas finales más largas que las nucleares. Esta tendencia es especialmente notoria en los dos intertonemas con acentos nucleares descendentes (H+L\*), los cuales no cuentan, en ninguna ocasión, con sílabas nucleares más largas que la final. Por ende, de acuerdo con los datos, el patrón de duración relativa prototípico de las configuraciones nucleares de las ip no oxítonas consiste en una mayor duración en la sílaba final en contraste con una menor en la nuclear.

Cuadro 3.6. Recuento de duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos. N=194

Intertonemas abstractos	Núcleo > Linde	Núcleo < Linde	Total
H* H-	26	51	77
H* !H-	4	6	10
H* L-	1	12	13
L+H* H-	16	30	46
L+H* L-	3	5	8
L* H-	4	10	14
L* !H-	2	4	6
L* L-	7	5	12
H+L* H-	0	6	6
H+L* !H-	0	2	2
Total	63	131	194

En líneas generales, la frecuencia de los patrones prosódicos identificados hasta este apartado coincide con los datos de ambas redes sociales. No obstante, se registraron diferencias significativas en la producción de algunos parámetros prosódicos por parte de las dos redes sociales que pueden considerarse índices sociolingüísticos. En otras palabras, entre dichas diferencias pueden existir patrones que sean potenciales portadores de significado social asociables a cada red. Por lo tanto, este aspecto se retomará en el análisis sociolingüístico (ver §4.3.1).

### 3.3.2. Tonemas internos (intertonemas) oxítonos

En los tonemas internos oxítonos se observa la misma tendencia que en los no oxítonos, en cuanto al mayor uso de las configuraciones abstractas H\* H- y L+H\* H-. En este caso, el intertonema prototípico es L+H\* H-, ya que su frecuencia de uso no solo es la más elevada, sino que supera por más del doble los casos de H\* H-, la segunda configuración más frecuente (ver cuadro 3.7). En el nivel concreto, la configuración más común es L+<sub>j</sub>H\* HH- (28 %), es decir, una elevación superior a 3 st, seguida por la configuración H\* H- (17.3 %). Debido a que el resto de los tonemas internos concretos no superan el 10 % de frecuencia, se puede establecer que las ip que culminan con palabras agudas se caracterizan (principalmente) por un pronunciado ascenso en la sílaba final o por un tono alto sostenido a lo largo de la sílaba final de la ip.



Cuadro 3.7. Recuento de tonemas internos oxítonos en los niveles concreto y abstracto.

N= 75

Intertonemas en el nivel concreto	Casos	%	Intertonemas en el nivel abstracto	Casos	%			
H* H-	13	17.33	H* H-	15	20.00			
!H* H-	2	2.67						
!H* !H-	7	9.33	H* !H-	7	9.33			
L+H* H-	4	5.33	L+H* H-	44	58.67			
L+H* HH-	5	6.67						
L+H* !HH-	1	1.33						
L+H* LH-	1	1.33						
L+ <sub>j</sub> H* H-	2	2.67						
L+ <sub>j</sub> H* HH-	21	28.00						
L+ <sub>j</sub> H* !HH-	3	4.00						
L+ <sub>j</sub> H* LH-	2	2.67						
L+<H* H-	1	1.33						
L+< <sub>j</sub> H* H-	2	2.67						
L+< <sub>j</sub> H* HH-	2	2.67						
L+H* !H-	1	1.33				L+H* !H-	3	4.00
L+H* L!H-	1	1.33						
L+ <sub>j</sub> H* !H-	1	1.33						
L* L-	3	4.00	L* L-	3	4.00			
H+L* HL-	3	4.00	H+L* L-	3	4.00			
Total				75	100			

Por otra parte, en estos tonemas internos se observan configuraciones abstractas que no se registraron en los no oxítonos: L+H\* !H- y H+L\* L-. Estos dos contornos, junto con L\* L-, se ubican en la sección de la periferia cognitiva más alejada del prototipo (todos ellos tienen un porcentaje de ocurrencia de 4 %.), aunque no muy lejos de H\* L-, que con un 9.33 % de frecuencia se encuentra ligeramente más cerca del prototipo. De hecho, los porcentajes de frecuencia mostrados en el cuadro 3.7 revelan la tendencia en cuanto a la frecuencia de uso de cada configuración, presentada en orden decreciente: L+H\* H- (58.67 %), H\* H- (20 %), H\* !H- (9.33 %), y L+H\* !H-, H+L\* L- y L\* L- (cada uno de ellos con un porcentaje de ocurrencias de 4 %). Esto permite observar que los tonemas internos abstractos oxítonos predominantes son aquellos que terminan en tonos altos (especialmente si ocurre un ascenso a lo largo de la sílaba final de la ip hasta la juntura H-), seguidos por los que poseen junturas

medias (!H-) y los que terminan con juntas bajas. La periferia llama la atención por la presencia de configuraciones que sostienen un tono medio a lo largo de la sílaba tónica final de ip (H\* !H-), como se muestra en la figura 3.6; y aquellas que presentan un ascenso que culmina en una junta media (L+H\* !H-), como el que se puede apreciar en la figura 3.7. Este hecho, junto con todos los demás datos mostrados previamente, apunta hacia una tendencia a favorecer los ascensos y los sostenimientos en el rango alto (principalmente) o en un rango medio, usando muy poco los ascensos hasta una junta media o los descensos al final de las frases entonativas intermedias que culminan con sílabas tónicas. Estos patrones no distan mucho de los registrados en los intertonemas no oxítonos, con la excepción de que, mientras el tonema interno prototípico en las ip oxítonas es L+H\* H-, la configuración más frecuente en las ip no oxítonas es H\* H-.

Figura 3.6. Oscilograma, espectrograma y F0 del enunciado Rs2M3-10: [(*Se me hace una cultura*)<sub>ip</sub> (*muy*)<sub>ip</sub> (*aislada*)<sub>ip</sub>]<sub>IP</sub>

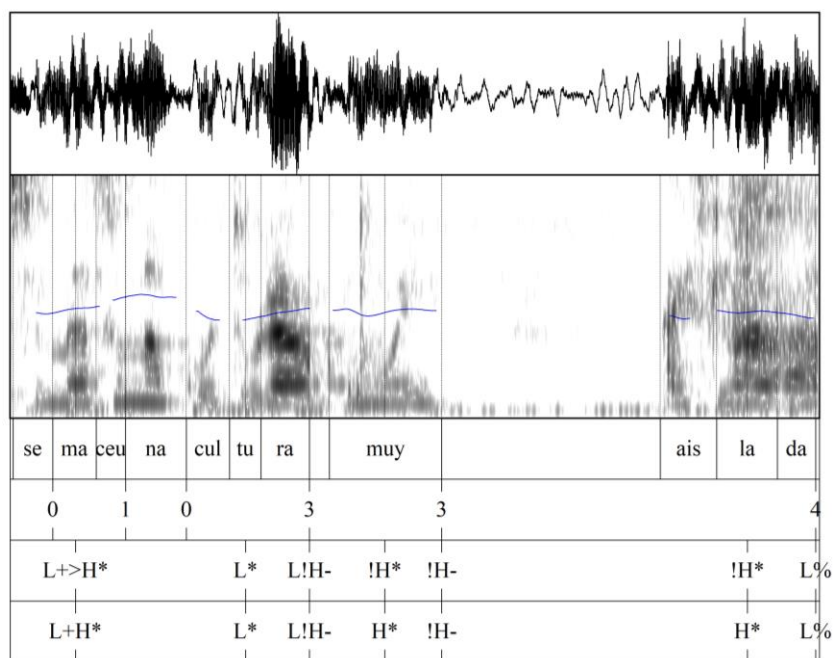
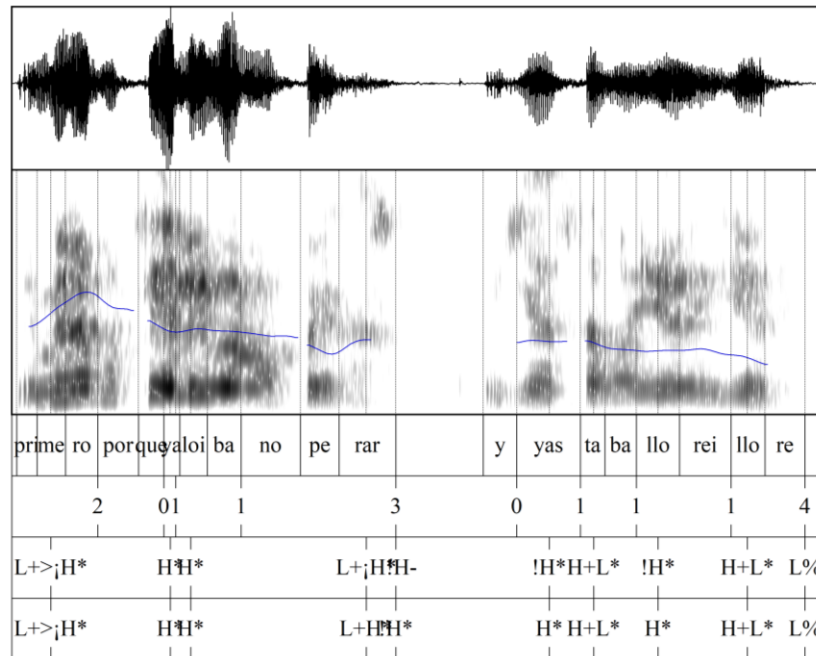
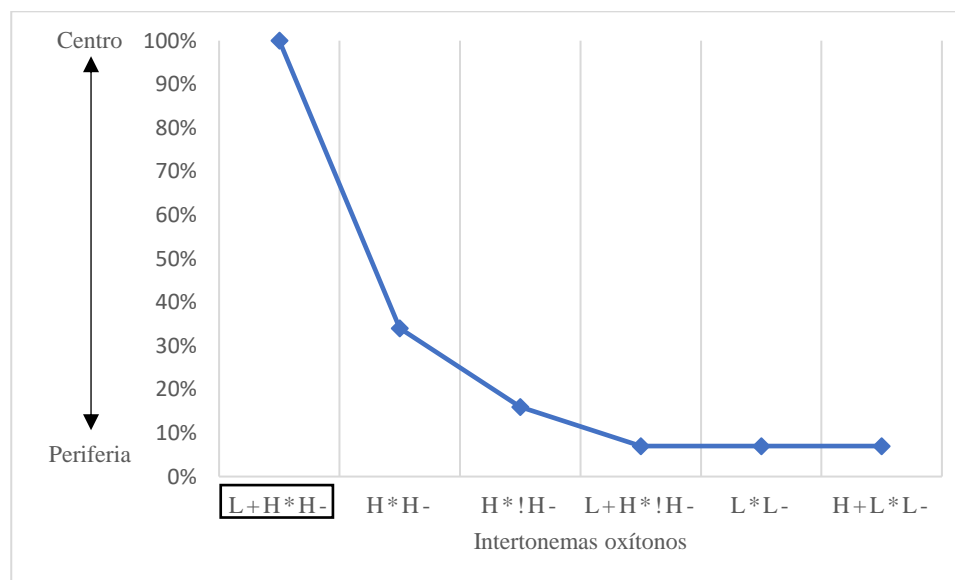


Figura 3.7. Oscilograma, espectrograma y F0 del enunciado Rs1M1-18: [(*Primero que ya lo iban a operar*)<sub>ip</sub> (y ya estaba llore y llore)<sub>ip</sub>]<sub>IP</sub>



Teniendo en cuenta todos los datos analizados, se considera que las configuraciones entonativas se distribuyen dentro de la categoría de tonemas internos oxítonos, de acuerdo con los grados de prototipicidad mostrados en la figura 3.8. En esta gráfica, se observa que hay una distancia considerable entre el esquema prototípico L+H\* H- y los demás ejemplares, de entre los cuales H\* H- es el más cercano al prototipo. De hecho, se puede ver cómo descende el grado de centralidad de las distintas configuraciones en la medida en que el tono de juntura es más bajo, lo que realza el papel de las juntas altas como portadoras del significado de continuidad.

Figura 3.8. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los intertonemas oxítonos



Ahora bien, el grado de centralidad de las configuraciones tonales al final de ip es solo una de las razones por las que se considera que ambos tipos de intertonemas pertenecen a categorías cognitivas diferentes, aunque no completamente independientes. Si bien ambos tipos de tonemas internos comparten la función de demarcadores de frases entonativas intermedias, tanto la constitución (en términos de las sílabas en las que se desarrollan) de las configuraciones tonales de los oxítonos y no oxítonos, como su frecuencia de uso son suficientemente distintas como para considerar que cada uno de ellos conforma una categoría por sí mismo. La razón por la que se piensa que esto es posible yace en el número de sílabas del que se dispone para almacenar la información prosódica cognitivamente, en cada caso: dos sílabas en los intertonemas no oxítonos y una sola en los oxítonos. En los no oxítonos, los elementos constituyentes del tonema interno se ubican en dos sílabas distintas, lo que implica que existe una separación de, al menos, un linde silábico entre ellos, por lo que el acento nuclear y el tono de juntura no son adyacentes. En cambio, en las instancias oxítonas no existe tal separación, lo que lleva a que el núcleo y la juntura sean adyacentes. Esta diferencia implica una restricción de las combinaciones de acentos tonales y juntas posibles en los intertonemas oxítonos, pues las juntas están condicionadas por el acento nuclear.

El espacio temporal que tienen los hablantes al momento de producir un tonema interno no oxítono es mayor (desde la última sílaba tónica de la ip hasta el linde derecho de la final), por lo que la complejidad de la configuración nuclear puede ser mayor y más variada. Por otra parte, el tiempo para ir del acento nuclear al tono de juntura es menor en los casos oxítonos (puesto que todo ocurre en una sola sílaba). Una prueba de esto es que el repertorio de intertonemas abstractos no oxítonos (cuadro 3.4) incluye configuraciones con acentos bitonales, seguidos por una juntura en el extremo opuesto del campo tonal (L+H\* L- y H+L\* H-); mientras que en los intertonemas oxítonos (cuadro 3.7) el tono de juntura no dista mucho del tono final del acento tonal. De hecho, en las configuraciones abstractas oxítonas, la juntura está siempre a la misma altura que el tono final del acento bitonal, lo que permite inferir que el acento tonal está cognitivamente enlazado con la juntura. Tal situación lleva a pensar que puede existir una restricción cognitiva que determina los movimientos tonales posibles en cada tipo de intertonema, con base en la distancia entre sus elementos constituyentes, y que crea la distinción entre las dos categorías de tonemas internos. Sin embargo, cabe señalar que no se descarta la posibilidad de registrar las configuraciones L+H\* L- y H+L\* H- en contextos oxítonos, pero resulta difícil concebir esta probabilidad sin que se produzca en una sílaba final de ip cuya duración prolongada (necesaria para llevar a cabo esta excursión tonal dentro de una sola sílaba) se perciba, al menos, como atípica.

En relación con la duración media de los intertonemas oxítonos (cuadro 3.8), la media total es perceptiblemente mayor que la de los intertonemas no oxítonos (241.1 ms), pues supera el umbral de diferencia apenas perceptible de 30 ms (Toledo 1998 y 2000). De hecho, en algunas configuraciones, la diferencia resulta bastante elevada, como es el caso de la media del tonema interno abstracto H+L\* L-, que supera los 300 ms. Esta situación también se registra en varias realizaciones concretas del intertonema abstracto L+H\* H-, especialmente en aquellas con un ascenso tonal mayor a los 3 st (se destaca la media de 448 ms de la configuración concreta L+<sub>1</sub>H\* H-). Si bien podría pensarse que la mayor duración está relacionada con los acentos nucleares bitonales, esta idea pierde peso al observar que el intertonema fonológico L+H\* !H- tiene la menor duración media de todas las configuraciones, pues ninguna de sus realizaciones fonéticas supera los 190 ms de duración. Por lo tanto, existe la posibilidad de que este alargamiento superior a los 300 ms esté relacionado con las junturas altas (H-) y bajas (L-) al final de una ip.

Cuadro 3.8. Medias de duración (ms) de los intertonemas oxítonos abstractos

Intertonema abstractos	Duración media
H* H-	283.47
H* !H-	290
L+H* H-	275.7
L+H* !H-	161
L* L-	240
H+L* L-	353
Total	275.67

En términos sociolingüísticos, los datos de los intertonemas oxítonos resultan particularmente llamativos. La diferencia observable en ambas redes en relación con la duración de las sílabas tónicas que culminan una ip se considera suficientemente sobresaliente como para ser considerada un potencial índice sociolingüístico para esta red. En §4.3.2 se muestra el análisis detallado de este parámetro, junto con las cualidades entonativas de ambas redes sociales y la forma en la que ambos elementos prosódicos se entrelazan.

### 3.3.3. Funciones sintáctico-discursivas en los tonemas internos (intertonemas)

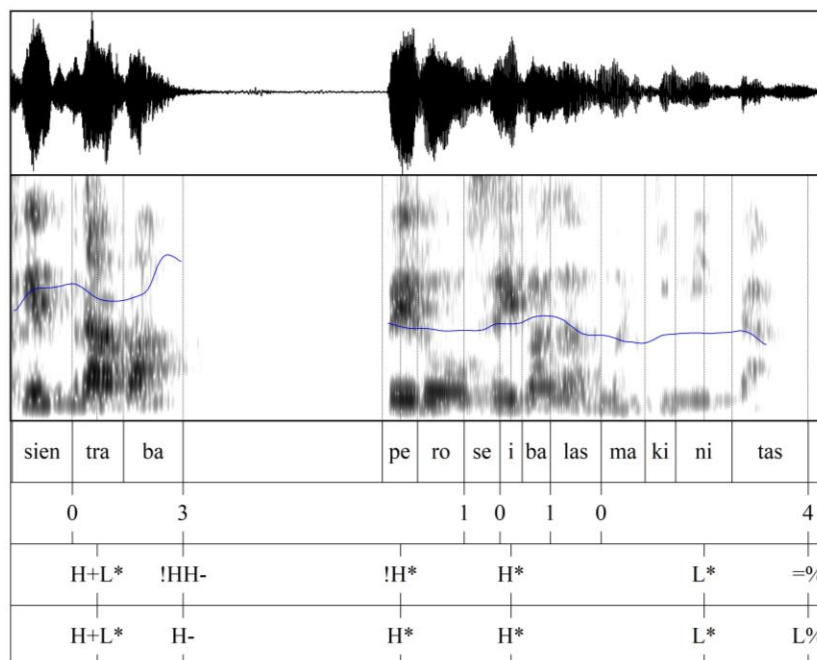
En los apartados anteriores, se presentó un análisis de la interacción entre los tonemas internos y la duración sin abordar otros elementos lingüísticos que pudieran influir en el uso de los recursos prosódicos descritos. Por lo tanto, con el fin de ahondar en la función o característica particular del empleo de las configuraciones mencionadas hasta este punto, se presenta, a continuación, un análisis conjunto de los intertonemas y de las variables: foco y expresividad, como elementos pragmáticos; condicionalidad y yuxtaposición, como rasgos sintácticos; y explicación (cuando el hablante aclara algo dicho o por decir en el enunciado) y extraoracionalidad (tema, iniciador, orientador de escenario y orientador de tiempo) como característica discursiva. Cabe señalar que tanto las variables sintácticas como las discursivas se consideran mutuamente excluyentes, por lo que una frase intermedia (ip) clasificada dentro de alguna de estas variables no puede pertenecer simultáneamente a la otra. Sin embargo, no ocurre lo mismo con los elementos pragmáticos (foco y expresividad), debido a que se consideran compatibles (ya que una ip producida con expresividad puede tener un foco estrecho, por ejemplo) y con otras variables.

En líneas generales, según se desprende del cuadro 3.9, la distribución de intertonemas abstractos relacionados con las variables descritas en el párrafo anterior no se aleja mucho de la tendencia observada hasta ahora: la mayoría de los casos consisten en las configuraciones H\* H- y L+H\* H- (cuyas realizaciones concretas más frecuentes son H\* H- y L+;H HH-, respectivamente). Aunque el volumen de casos es relativamente reducido en la mayoría de las variables, se pueden hacer observaciones en cuanto a los datos obtenidos. Los intertonemas con acentos nucleares descendentes (H+L\*) únicamente se registraron en foco estrecho y elementos extraoracionales (véase figura 3.9). Si bien es cierto que estas son las configuraciones nucleares internas con menos casos (junto con L+H\* !H-), llama la atención que los casos de H+L\* H- del corpus ocurren en ip con foco estrecho (principalmente) y en aquellas ip que conforman un elemento extraoracional. Esta misma tendencia se observa con el tonema interno abstracto L+H\* !H-, lo que sugiere que ambas configuraciones pueden estar conectadas con focos estrechos y elementos extraoracionales.

Cuadro 3.9. Recuento de tonemas internos abstractos de ip en cuanto a variables pragmáticas, sintácticas y discursivas. N= 210

Intertonemas abstractos	Foco estrecho	Expresividad	Prótasis	Preyuxtapuesta	Explicación	Extraoracional	Total
H* H-	47	12	5	4	5	16	84
H* !H-	9	2	1	0	0	0	12
H* L-	1	2	0	1	0	0	4
L+H* H-	35	7	9	4	5	11	66
L+H* !H-	2	0	0	0	0	1	3
L+H* L-	5	2	0	0	0	0	7
L* H-	6	1	1	0	3	1	9
L* !H-	4	2	0	0	0	0	6
L* L-	6	1	1	1	0	1	10
H+L* H-	4	0	0	0	0	2	6
H+L* !H-	1	0	0	0	0	0	1
H+L* L-	1	0	0	0	0	1	2
Total	121	29	17	10	13	33	210

Figura 3.9. Oscilograma, espectrograma y F0 del enunciado Rs1M2-23: [(*Sí entraba*)<sub>ip</sub> (*pero se iba a jugar a las maquinitas*)<sub>ip</sub>]<sub>IP</sub>



Llama la atención que existen variables que parecen operar con solo un tipo de intertonema abstracto. Tal es el caso de las dos variables sintácticas, es decir, de las ip que son prótasis de un enunciado condicional o cláusulas preyuxtapuestas, ya que sus casos se concentran en las dos configuraciones más frecuentes (L+H\* H- y con H\* H-), mientras que el resto de las instancias de cada variable ocurren con un tonema diferente en cada ocasión. Por otra parte, resalta el caso de las ip con función explicativa, pues ocurren únicamente con acentos nucleares monotonaes o ascendentes seguidos por una juntura alta (H\* H-, L+H\* H- y L\* H-), con una mayor tendencia a ser representados por acentos nucleares altos (H\*) o ascendentes (L+H\*) que por el bajo (L\*). Esto parece dejar claro el hecho de que las ip que aportan algún tipo de explicación acerca de otro elemento o aspecto del enunciado se destacan por emplear contornos melódicos únicamente ascendentes hasta el rango tonal superior del enunciado.

Al observar aquellos casos en los que un foco estrecho coincide con un matiz expresivo (véase cuadro 3.10), se pudo notar que los intertonemas con predominio de uso en cada caso son los mismos que se habían identificado como prototípicos en las ip, concretamente en aquellas que culminan con una palabra no oxítone (H\* H- y L+H\* H-).



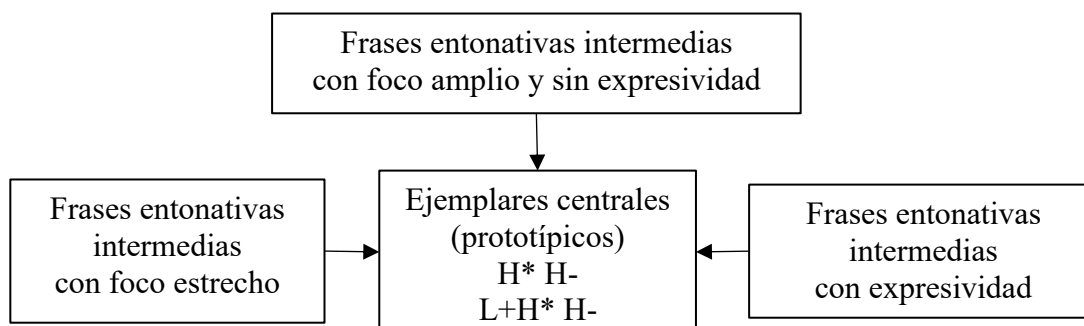
Ahora bien, el número de ip con foco estrecho, expresividad y ambas a la vez, delimitadas por estas dos configuraciones, es considerablemente elevado. Sin embargo, tras remover estos casos del total de ip del corpus (incluyendo terminaciones con palabras oxítonas y no oxítonas), las dos configuraciones mencionadas siguen manteniendo su predominio en términos de uso: H\* H- presenta 38 casos sin foco estrecho ni expresividad y las producciones de L+H\* H- suman 51 ocurrencias; mientras que la configuración tonal más cercana (H\* L- ) acumula solo 12 instancias.

Cuadro 3.10. Recuento de tonemas internos abstractos de ip en concordancia con la presencia de foco estrecho, expresividad y el conjunto de ambas. N= 136

Intertonemas abstractos	Foco estrecho	Expresividad	Foco estrecho con expresividad	Total
H* H-	42	7	5	54
H* !H-	7	0	2	9
H* L-	0	1	1	2
L+H* H-	32	4	3	39
L+H* !H-	2	0	0	2
L+H* L-	4	1	1	6
L* H-	6	1	0	7
L* !H-	2	0	2	4
L* L-	6	1	0	7
H+L* H-	4	0	0	4
H+L* !H-	1	0	0	1
H+L* L-	1	0	0	1
Total	107	15	14	136

El hecho de que estos intertonemas se mantengan como elementos centrales (aunque con frecuencia de uso variable, según la posición de la última sílaba tónica de la ip) en esta categoría sin importar la presencia de un foco estrecho o un matiz expresivo, permite cuestionar el hecho de que las variables pragmáticas desempeñen un papel determinante en la estructura de los esquemas cognitivos de la categoría de intertonemas de enunciados aseverativos. En este sentido, lo que puede ocurrir es que las frases entonativas intermedias (ip) producidas con foco amplio y sin expresividad comparten sus elementos centrales (es decir, las mismas configuraciones entonativas prototípicas) con las ip que poseen ambas cualidades pragmáticas. Esta idea se puede ver ejemplificada en la figura 3.10.

Figura 3.10. Diagrama representativo de la centralidad (prototipicidad) compartidos por tres categorías cognitivas.



A pesar de que esta noción de centralidad compartida se apoya en los datos, no se puede pasar por alto que las tendencias observadas en las funciones sintáctico-discursivas no se alejan de lo ya observado en los datos generales del corpus, en las secciones finales de ip. Por lo tanto, se considera que una opción posible es que, en estos casos, se empleen las configuraciones nucleares internas abstractas  $H^* H-$  y  $L+H^* H-$ , porque se localizan en la sección final de una frase entonativa intermedia (ip). En tal sentido, los hablantes realizarían estos ascensos en concordancia con el código biológico de producción (Gussenhoven, 2004), para transmitir el significado de no conclusión de la enunciación y no con el código biológico de esfuerzo (Gussenhoven, 2004), lo que implicaría un ascenso para focalizar prosódicamente un elemento en el discurso. En otras palabras, el nivel prosódico de la ip en la jerarquía prosódica parece imponerse sobre la expresión de la función informativa del foco.

En el análisis de los elementos extraoracionales (cuadro 3.11), se revela el predominio del uso de las configuraciones nucleares internas  $H^* H-$  y  $L+H^* H-$ . Tanto las tematizaciones como los orientadores temporales se inclinan principalmente por la primera de estas configuraciones, aunque un buen número (33.3 %) de estos orientadores también ocurren con la segunda configuración. Los elementos extraoracionales iniciadores se producen principalmente con el intertonema  $L+H^* H-$ . Por su parte, los orientadores locativos concentran dos tercios de sus casos en las dos configuraciones mencionadas. En cuanto al resto de los intertonemas abstractos, del total de 33 elementos extraoracionales estudiados, únicamente 6 de ellos son producidos con otras configuraciones.

Cuadro 3.11. Recuento de casos de los intertonemas abstractos en elementos extraoracionales. N= 33

Intertonemas abstractos	Tematización	Iniciador	Orientador Locativo	Orientador Temporal	Total
H* H-	4	1	1	10	16
L+H* H-	1	3	1	6	11
L+H* !H-	0	1	0	0	1
L* H-	0	0	0	1	1
L* L-	0	1	0	0	1
H+L* H-	1	0	0	1	2
H+L* L-	0	0	1	0	1
Total	6	6	3	18	33

El hecho de que el 81.8 % de las funciones sintáctico-discursivas hayan sido emitidas con las configuraciones H\* H- y L+H\* H- realza, una vez más, su carácter central. Sin embargo, el reducido número de casos no permite realizar un análisis estadístico que apoye una afirmación categórica. No se descarta la posibilidad de que un estudio más profundo de las cualidades entonativas de los elementos extraoracionales, que cuente con un mayor número de producciones, pueda aportar información que contradiga algunos de los resultados presentados en esta revisión de datos.

#### 3.4. Tonemas: configuraciones nucleares de frase entonativa mayor (IP)

En este apartado, se exponen los resultados del análisis de los tonemas de IP. En primera instancia, se muestra el recuento de todas las configuraciones nucleares registradas en el corpus, se presentan su distribución en los enunciados y su relación con los valores de duración silábica. Posteriormente, se analiza la relación entre los tonemas no descendentes, es decir, aquellos tonemas que no culminan con un tono de juntura bajo (L%), sino con altos (H%) o medios (!H%), con otros elementos pragmáticos, para establecer la función que pueden estar desarrollando en los enunciados aseverativos.

##### 3.4.1. Tonemas

La tendencia que se observa en los tonemas de las frases entonativas mayores (IP) consiste en un descenso tonal, como se puede ver en el cuadro 3.12, lo que resulta esperable de acuerdo con las descripciones del español general (Navarro Tomás, 1957 y 1966; Quilis,

1981) y del español de México (De-La-Mota *et al.*, 2010; Martín Butragueño 2019). Los tonemas abstractos descendentes (denominados así porque culminan con una juntura baja), organizados de más a menos frecuente, son los siguientes: L\* L% (33.33 %), H\* L% (27.5 %), L+H\* L% (18.89 %) y H+L\* L% (8.06 %). La frecuencia del resto de los tonemas abstractos, de forma individual, no supera el 5 %. No obstante, debido a que estos tonemas no descendentes (que no terminan en una juntura L%) no son típicos en esta posición, de acuerdo con las descripciones de la variedad del español de la Ciudad de México, estos serán analizados en §3.4.2.

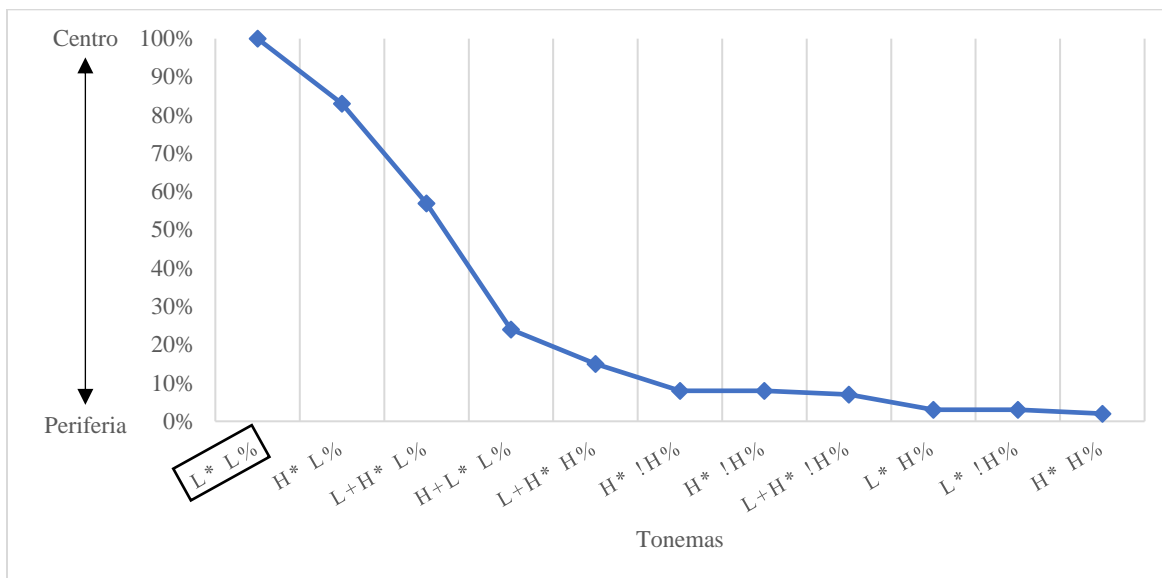
Cuadro 3.12. Recuento de tonemas en los niveles concreto y abstracto. N=360

Tonemas en el nivel concreto	Casos	%	Tonemas en el nivel abstracto	Casos	%
H* =%	2	0.56	H* H%	2	0.56
H* !H%	1	0.28	H* !H%	9	2.50
!H* =%	8	2.22			
H* L%	53	14.72	H* L%	99	27.50
H* HL%	5	1.39			
!H* L%	38	10.56			
!H* HL%	3	0.83			
L+H* =%	5	1.39	L+H* H%	18	5.00
L+ <sub>j</sub> H* =%	5	1.39			
L+<H* =%	2	0.56			
L+< <sub>j</sub> H* =%	6	1.67			
L+H* !H%	3	0.83	L+H* !H%	8	2.22
L+ <sub>j</sub> H* !H%	2	0.56			
L+< <sub>j</sub> H* !H%	3	0.83			
L+H* L%	23	6.39	L+H* L%	68	18.89
L+H* HL%	3	0.83			
L+ <sub>j</sub> H* L%	11	3.06			
L+ <sub>j</sub> H* HL%	6	1.67			
L+<H* L%	8	2.22			
L+< <sub>j</sub> H* L%	16	4.44			
L+< <sub>j</sub> H* HL%	1	0.28			
L* H%	4	1.11	L* H%	4	1.11
L* !H%	3	0.83	L* !H%	3	0.83
L* L%	64	17.78	L* L%	120	33.33
L* =%	56	15.56			
H+L* L%	21	5.83	H+L* L%	29	8.06
H+L* =%	8	2.22			
Total				360	100

En el nivel concreto, las realizaciones del tonema prototípico ( $L^* L\%$ ) son dos: una descendente desde el acento nuclear hasta la juntura ( $L^* L\%$ ) y otra que presenta un sostenimiento tonal desde el acento nuclear hasta la juntura ( $L^* =\%$ ). Si bien el tonema concreto más frecuente es el descendente  $L^* L\%$  (64/120 casos), el sostenimiento también cuenta con un número elevado de ocurrencias ( $N=56$ ). Considerando que ambos tienen frecuencias cercanas, no se puede asegurar que el tonema concreto descendente  $L^* L\%$  represente el prototipo, sino que ambas realizaciones ( $L^* =\%$  y  $L^* L\%$ ) pertenecientes al tonema abstracto  $L^* L\%$  poseen los mayores rasgos de representatividad o centralidad dentro del continuum de prototipicidad. En otras palabras, de acuerdo con los datos generales del corpus, el que la  $F_0$  descienda en la sílaba final de IP o se sostenga al mismo nivel desde la sílaba nuclear hasta el tono de juntura no parece ser un rasgo cognitivamente determinante, mientras que el hecho de que la juntura abstracta sea  $L\%$  (independientemente de su realización concreta) sí lo es.

Con base en los datos obtenidos, la figura 3.11 muestra la propuesta de los niveles de prototipicidad de cada configuración entonativa. Cabe destacar que este gráfico representa los grados de centralidad de los esquemas entonativos pertenecientes a la categoría de tonemas de enunciados aseverativos de foco amplio, ya que se debe tener en cuenta que tanto los esquemas como su grado de centralidad pueden variar en enunciados con focos estrechos en la sección nuclear. En la figura, se puede observar que los tonemas descendentes  $L^* L\%$ ,  $H^* L\%$  y  $L+H^* L\%$  son los que tienen mayor grado de centralidad, encabezados por el esquema prototípico  $L^* L\%$ . Desde la configuración circunfleja ( $L+H^* L\%$ ) se observa una distancia considerable hasta la configuración  $H+L^* L\%$  (la descendente más alejada del centro). De hecho, se puede decir que esta separación entre  $L+H^* L\%$  y  $H+L^* L\%$  marca el inicio de la periferia de la categoría cognitiva, donde se localizan todas las configuraciones no descendentes.

Figura 3.11. Grado de centralidad (prototipicidad) de los elementos entonativos de la categoría cognitiva correspondiente a los tonemas



Respecto a las medias de duración de las sílabas que integran los tonemas, el cuadro 3.13, refleja que, de manera general, ambas sílabas tienen una duración bastante cercana. Al prestar atención a las medias de cada tonema abstracto, se puede constatar que solo en la mitad de ellos la diferencia de medias entre ambas sílabas es mayor a 30 ms, es decir, supera la diferencia de duración apenas perceptible de acuerdo con Toledo (1998 y 2000). Estos casos consisten en aquellos tonemas con acentos nucleares monotonaes, seguidos por junturas medias (H\* !H% y L\* !H%), los que cuentan con ascensos en la sílaba nuclear seguidos por una juntura distinta a L% (L+H\* H% y L+H\* !H%) y la configuración nuclear con acento nuclear descendente (H+L\* L%).

Cuadro 3.13. Medias de duración (ms) de las sílabas nucleares y lindes en los datos generales del corpus

Tonemas abstractos	Duración del núcleo	Duración del linde
H* H%	212	201.5
H* !H%	201.67	157.67
H* L%	183.64	170.57
L+H* H%	157.89	190.72
L+H* !H%	181.5	222
L+H* L%	179.18	185.01
L* H%	144.25	161.25
L* !H%	184.67	153
L* L%	171.02	175.61
H+L* L%	191.72	144.59
Total	178.08	174.63

En términos de la frecuencia de duración relativa, el cuadro 3.14 revela que casi todos los tonemas abstractos son producidos principalmente con sílabas nucleares más largas que las finales de IP, lo cual implica que este comportamiento es el prototípico en el núcleo de las frases entonativas mayores (IP). Las excepciones a esta tendencia son las configuraciones abstractas L+H\* H% y L+H\* !H%, lo cual llama la atención debido a que se esperaría que la excursión tonal de al menos 1.5 st en la sílaba nuclear requiriera más tiempo que la ejecución del movimiento hasta el tono de juntura. En todo caso, debido a que estas dos configuraciones se encuentran entre las no descendentes, se les prestará especial atención en la sección siguiente (§3.4.2).

Cuadro 3.14. Recuento de duración relativa de la sílaba nuclear con respecto a la sílaba linde en la totalidad de los tonemas fonológicos del corpus. N=360

Tonemas abstractos	Núcleo > Linde	Núcleo < Linde	Total
H* H%	1	1	2
H* !H%	7	2	9
H* L%	59	40	99
L+H* H%	7	11	18
L+H* !H%	3	5	8
L+H* L%	39	29	68
L* H%	1	3	4
L* !H%	3	0	3
L* L%	61	59	120
H+L* L%	24	5	29
Total	205	155	360

En relación con las variables sociales, la red laboral emplea de forma casi exclusiva los tonemas abstractos descendentes, ya que solo 8 de los 180 tonemas siguen una orientación distinta. La red vecinal, aunque también se inclina principalmente por los tonemas descendentes, presenta una tendencia secundaria (ya que ocurre en el 20 % de los tonemas de esta red) por el uso de configuraciones abstractas no descendentes. Dichas configuraciones representan una quinta parte de los datos recolectados en esta red, que contrasta fuertemente con el uso casi nulo que hacen de ellos los hablantes de la red laboral. Por otra parte, la red vecinal tiende a producir tonemas más largos, con mayor elongación en la sílaba final, mientras que la red laboral presenta tonemas más cortos en los que la sílaba nuclear es la más prolongada. En cuanto a la variable sexo, se observó que las mujeres siguen las tendencias principales relacionadas con el uso de tonemas descendentes. Por su parte, los hombres hacen mayor empleo de tonemas no descendentes, lo cual es especialmente marcado en los hombres de la red laboral (7 de 8 casos). En general, las diferencias registradas en relación con las variables extralingüísticas, especialmente en la red social, permiten inferir que ciertas características de los tonemas pueden ser consideradas índices lingüísticos en potencia (véase el análisis sociolingüístico en §4.4)



### 3.4.2. Tonemas no descendentes

Como se mencionó en el inciso anterior, con el término “tonema no descendente” se hace referencia a las configuraciones nucleares de IP que no culminan con un descenso hasta un tono de juntura bajo (L%), sino que terminan con una juntura media (!H%) o alta (H%). Debido a que estas configuraciones representan casos poco comunes, especialmente, teniendo en cuenta que no son característicos de los enunciados aseverativos en la variedad del español de la Ciudad de México, se decidió prestarles especial atención en esta sección.

Antes de proseguir, vale la pena mencionar que todos los tonos de juntura en los tonemas H\* H% y L+H\* H% corresponden, en términos fonéticos, con un sostenimiento tonal (=%) desde el punto alto, en el que concluye la sílaba nuclear del enunciado (figura 3.12). Por tal motivo, conviene destacar que la juntura abstracta H%, en estos casos, no corresponde con un ascenso tonal posterior a la sílaba nuclear. Este hecho fonético es relevante debido a que el sostenimiento, en lugar de un ascenso, puede funcionar como un elemento prosódico diferenciador entre las categorías cognitivas prosódicas, correspondientes a los tonos de juntura de los enunciados aseverativos y los interrogativos. Por otra parte, la configuración nuclear L\* H% sí representa un ascenso tonal en la sílaba final de IP, en contraste con el tono bajo en la nuclear. No obstante, este ascenso no es continuo a lo largo de la sílaba final, sino que consiste en una elevación tonal que se mantiene en una meseta alta a lo largo de la sílaba final de IP (figura 3.13). Este caso contrasta con el de los enunciados interrogativos con configuraciones nucleares L\* HH% y L\* LH% descritos por De-la-Mota *et al.* (2010), ambos con una elevación en la última sílaba del enunciado (HH% y LH%), la cual es especialmente prominente en el caso de LH%.

Figura 3.12. Oscilograma, espectrograma y F0 del enunciado Rs1H1-18: [(*El francés ya no es la México Nuevo*)<sub>ip</sub>]<sub>IP</sub>

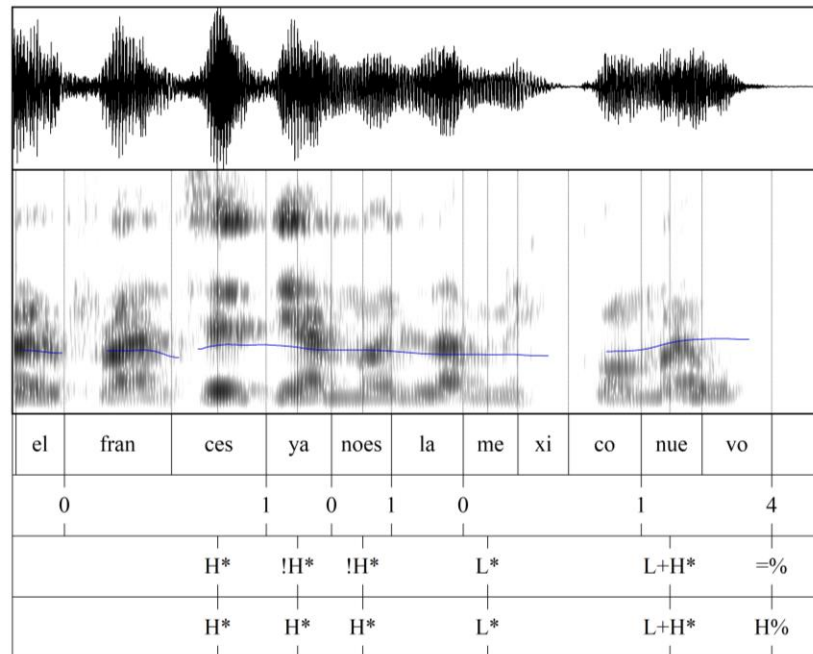
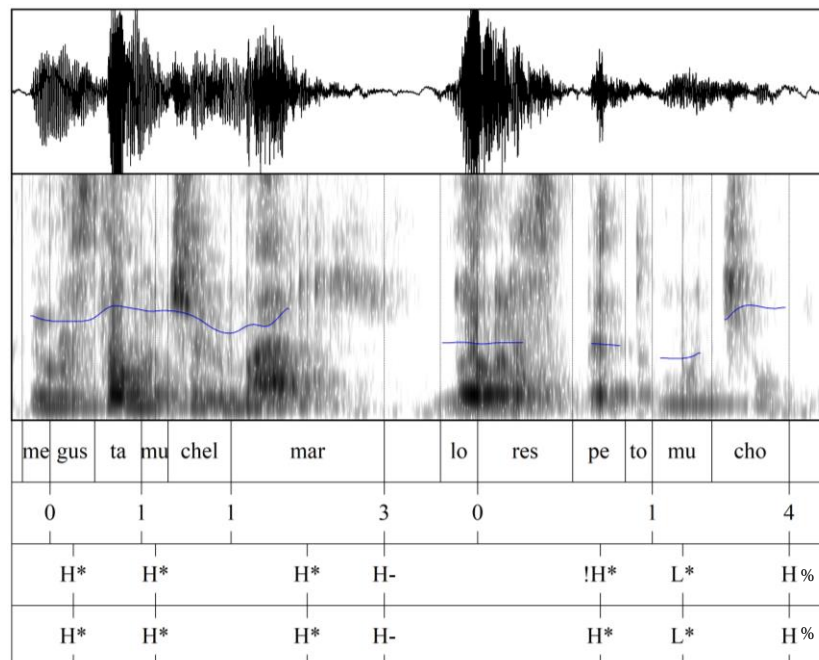


Figura 3.13. Oscilograma, espectrograma y F0 del enunciado Rs2H3-33: [(*Me gusta mucho el mar*)<sub>ip</sub> (lo respeto mucho)]<sub>ip</sub><sub>IP</sub>



El cuadro 3.15 muestra que, al contrastar los tonemas abstractos no descendentes con la expresividad al final del enunciado, se obtiene una diferencia estadísticamente significativa ( $p = 0.032$  de acuerdo con el estadístico Exacto de Fisher). La mayoría de los casos de estas configuraciones nucleares se producen en enunciados neutros. En enunciados con carga expresiva, las configuraciones nucleares más empleadas son las que presentan un acento tonal ascendente ( $L+H^*$ ), con independencia del tono de juntura que les siga, y acentos nucleares bajos ( $L^*$ ) y junturas altas ( $H\%$ ).

Cuadro 3.15. Recuento de tonemas no descendentes de acuerdo con la expresividad en la enunciación.  $N=44$

Tonemas abstractos	Expresividad				Total
	Neutro		Expresivo		
	Casos	%	Casos	%	
$H^* H\%$	1	3.23	1	7.69	2
$H^* !H\%$	9	29.03	0	0	9
$L+H^* H\%$	14	45.16	4	30.77	18
$L+H^* !H\%$	4	12.90	4	30.77	8
$L^* H\%$	1	3.23	3	23.08	4
$L^* !H\%$	2	6.45	1	7.69	3
Total	31	100	13	100	44

Por otra parte, el recuento de casos de cada tonema con respecto al tipo de foco (ver cuadro 3.16), cuya distribución de casos es significativa de acuerdo con el estadístico Exacto de Fisher ( $p = 0.032$ ), permite observar el mayor uso de los mismos tonemas abstractos que predominan en los casos con expresividad. Sin embargo, el porcentaje de frecuencia en este caso permite ver una tendencia mucho más orientada al uso de los tonemas con núcleos ascendentes, especialmente la configuración que culmina con un tono de juntura alto ( $L+H^* H\%$ ). Debido a que la mayoría de los tonemas  $L+H^* H\%$ ,  $L+H^* !H\%$  y la totalidad de  $L^* H\%$  se producen en contextos con foco estrecho, es posible inferir que existe una correlación entre todos ellos y dicho tipo de focalización.

Cuadro 3.16. Tonemas no descendentes y tipo de foco. N=44

Tonemas abstractos	Foco				Total
	Amplio		Estrecho		
	Casos	%	Casos	%	
H* H%	2	11.76	0	0	2
H* !H%	6	35.29	3	11.11	9
L+H* H%	5	29.41	13	48.15	18
L+H* !H%	2	11.76	6	22.22	8
L* H%	0	0	4	14.81	4
L* !H%	2	11.76	1	3.70	3
Total	17	100	27	100	44

Tanto la expresividad del enunciado como la presencia de foco estrecho conllevan el uso mayoritario de los mismos tonemas. Ahora bien, al observar los casos en los que se da un foco estrecho expresivo, se puede notar que 2 de los 18 casos de L+H\* H% ocurren con esta condición, al igual que 3 de las 8 realizaciones de L+H\* !H% y 3 de las 4 ocurrencias de la configuración L\* H%. Este último caso resulta llamativo, pues 75 % de sus producciones se dan con una convergencia de foco estrecho y expresividad, lo que presenta la posibilidad de que este tonema tenga una correlación estrecha con ambas funciones. Esto tendría sentido en cuanto a que, en el nivel concreto, el tonema L\* H% es el único que presenta un ascenso real desde un tono bajo en el linde derecho de la sílaba nuclear hasta una juntura alta, puesto que, como ya se había mencionado anteriormente, las juntas de los tonemas con acentos nucleares L+H\* y H\* no presentan un ascenso posterior en la última sílaba del enunciado, sino que el tono se sostiene en el rango tonal alto desde el pico tonal, o meseta tonal, dado en la sílaba nuclear. No obstante, el reducido número de casos documentados no permite presentar este hecho como categórico.

### 3.4.3. Picos muy tempranos en frases entonativas mayores (IP)

Durante el estudio de los tonemas de frase entonativa mayor (IP) se encontraron casos de acentos nucleares en los que se produce un ascenso de al menos 1.5 st, que fueron agrupados bajo la única forma del tonema abstracto L+H\* (ver cuadro 3.17) y que, de acuerdo con Martín Butragueño (2019), suelen estar asociados, aunque de forma no categórica, con focos estrechos y expresividad en la enunciación cuando son seguidos por un tono de juntura bajo

(L%). En algunos de los acentos incluidos en la categorización abstracta L+H\*, la frecuencia fundamental (F0) alcanza el pico tonal antes del final de la sílaba, fenómeno que se denomina “pico muy temprano” y se transcribe en este trabajo como L+<H\*. En estos casos, de acuerdo con las descripciones que se han realizado de dicho acento tonal en el español del centro de México, la F0 alcanza su punto máximo en algún lugar cercano al punto medio de la sílaba (Martín Butragueño y Mendoza, 2018), aunque también se ha reportado que estos picos se suelen dar una vez transcurrido aproximadamente el 80 % de la sílaba (Martín Butragueño, 2015a).

Dado que existen 36 casos de picos muy tempranos en los enunciados estudiados, se consideró hacer un estudio de sus características concretas, no solo para aportar una descripción de la evolución de F0 a lo largo de la sílaba, sino para identificar posibles índices sociolingüísticos que puedan estar presentes en este tipo de acentos nucleares. Para ello, se tomaron los parámetros: duración de la sílaba nuclear (expresada en milisegundos), punto de la sílaba en el que culmina el pico tonal (medida en milisegundos y en el porcentaje de la sílaba restante antes del final de la sílaba),<sup>22</sup> movimiento de F0 (referente tanto a un sostenimiento tonal como a un descenso desde el pico tonal alcanzado en la sílaba nuclear) y tono de juntura del enunciado (que refleja el movimiento de F0 en la sílaba final del enunciado). Cabe señalar que el análisis se dividió en casos en los que la sílaba nuclear tiene mayor duración que la final (cuadro 3.17) y en instancias en las que la duración de la sílaba linde de IP es menor que la de la nuclear (cuadro 3.18), debido a que se consideró oportuno estudiar si la duración relativa de estas sílabas puede estar relacionada con las características fonéticas de los acentos nucleares.

---

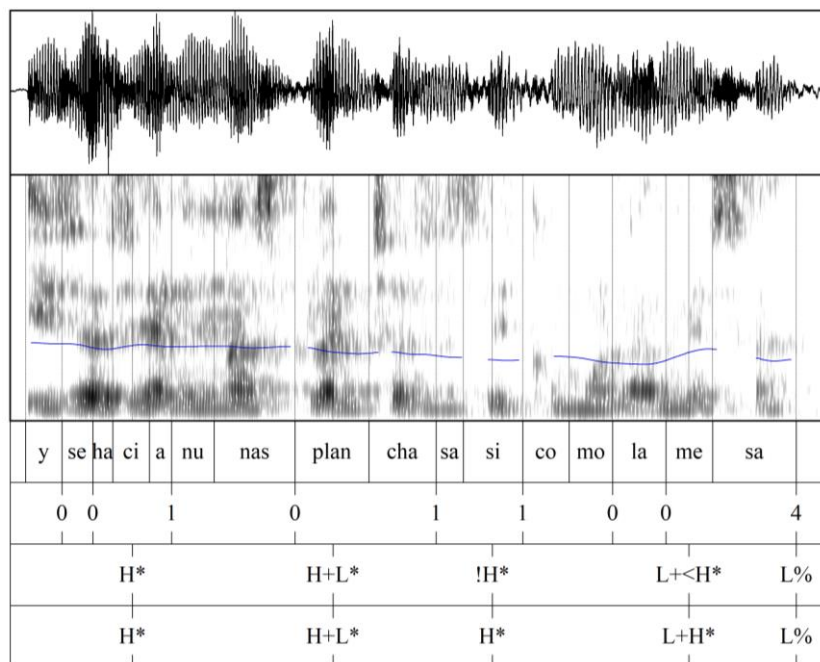
<sup>22</sup> Para identificar la posición del pico muy temprano en la sílaba, se emplea como punto de referencia el extremo final de la sílaba, es decir, se contabiliza el porcentaje que le resta a la sílaba desde el punto en el que se encuentra el pico de F0 hasta el punto en el que la sílaba concluye. Por ejemplo, si una sílaba dura 200 ms y el pico se encuentra en el milisegundo 150, se dirá que este ocurre 25% antes del linde silábico.

Cuadro 3.17. Tonema de IP con acento nuclear L+<H\* en los que la duración de la sílaba nuclear es menor que la del linde (Núcleo < Linde) y parámetros prosódicos

Nº de caso	Duración total (ms)	Pico de F0 (ms)	Pico de F0 (%)	Movimiento de F0	Juntura
1	205	102	50.24	Sostenido	!H%
2	199	114	42.71	Sostenido	=%
3	111	50	54.95	Sostenido	L%
4	139	84	39.57	Descendente	HL%
5	212	141	33.49	Sostenido	L%
6	107	73	31.78	Sostenido	L%
7	160	104	35.00	Sostenido	H%
8	106	73	31.13	Sostenido	L%
9	147	102	30.61	Sostenido	!H%
10	184	81	55.98	Sostenido	L%
11	111	30	72.97	Sostenido	H%
12	141	70	50.35	Descendente	L%

Teniendo en cuenta que las sílabas finales tienen mayor duración que las nucleares y que casi todos los tonos de juntura son bajos (L%), se puede inferir que existe una tendencia a ubicar el pico tonal, generalmente, alrededor de un punto ligeramente posterior al medio de la sílaba nuclear (44.07 % antes del linde silábico) y sostener el tono hasta el final de dicha sílaba. El empleo de este sostenimiento, en lugar de un descenso dentro de los límites de la misma sílaba nuclear, puede deberse a una necesidad de iniciar el descenso una vez culminada la sílaba nuclear, de forma tal que esta disminución melódica hasta la base del rango tonal del enunciado pueda ser prolongada a lo largo de la sílaba final (ver ejemplo en la figura 3.14, donde se registró con ascenso de 3.2 st en la sílaba nuclear y un posterior descenso de 3.4 st en sílaba final). Cabe señalar que la tendencia descrita es casi exclusiva de la red vecinal, pues el 91.67 % (casos del 1 al 11) de las ocurrencias de picos muy tempranos en sílabas nucleares más cortas que las finales tiene lugar en cinco hablantes de esta red. Esto ocurre en tres hombres y dos mujeres, lo que da pie a pensar que la variable sexo no es significativa en este caso.

Figura 3.14. Oscilograma, espectrograma y F0 del enunciado Rs1H3-21: [(*Cuando ya estaba*)<sub>ip</sub> (*agarrabas un bote lo llenabas*)<sub>ip</sub> (*y se hacían unas planchas así como la mesa*)<sub>ip</sub>]<sub>IP</sub>



Por otra parte, se registraron 24 casos en los que los picos muy tempranos ocurren en sílabas nucleares más largas que las del linde de las frases entonativas mayores (IP), todos ellos desglosados en el cuadro 3.18. La ubicación promedio del pico tonal en este grupo tiene lugar 43.31 % antes del linde silábico, aunque en la mitad de los casos se sitúa entre la porción de la sílaba ubicada entre el 40 % y el 60 % antes del final de la sílaba. Tanto la ubicación promedio del pico tonal en estos casos como los que se presentaron en el cuadro 3.17 confirman lo señalado por Martín Butragueño y Mendoza (2018) en la descripción que hacen de este acento tonal pues, en ambos grupos de datos, la cumbre de F0 se localiza poco después del punto medio de la sílaba sin que la longitud relativa de las sílabas del tonema desempeñe un papel determinante.

Cuadro 3.18. Tonema de IP con acento nuclear L+<H\* en los que la duración de la sílaba nuclear es mayor que la del linde (Núcleo > Linde) y parámetros prosódicos

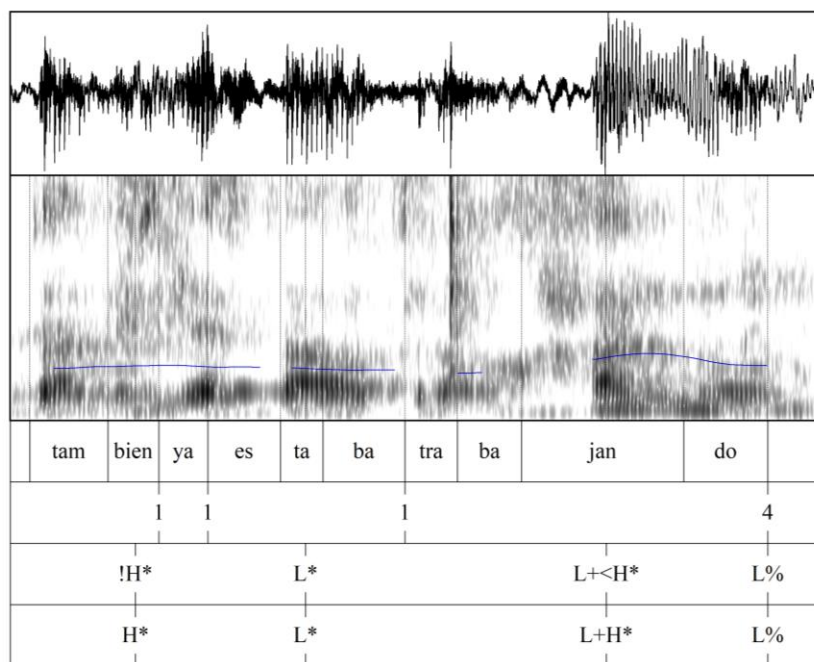
Nº de caso	Duración (ms)	Pico de F0 (ms)	Pico de F0 (%)	Movimiento de F0	Juntura
13	240	103	57.08	Descendente	L%
14	209	133	36.36	Sostenido/Descenso	!H%
15	156	93	40.38	Sostenido	H%
16	162	80	50.62	Sostenido	H%
17	138	96	30.43	Descendente	H%
18	144	81	43.75	Sostenido	L%
19	307	242	21.17	Sostenido/Descenso	L%
20	277	224	19.13	Sostenido	L%
21	206	133	35.44	Descendente	!H%
22	181	120	33.70	Sostenido	L%
23	206	133	35.44	Descendente	L%
24	236	72	69.49	Sostenido/Descenso	L%
25	114	52	54.39	Sostenido	L%
26	124	61	50.81	Sostenido/Descenso	L%
27	142	71	50.00	Sostenido	=%
28	279	134	51.97	Descendente	L%
29	244	92	62.30	Sostenido	L%
30	248	94	62.10	Sostenido/Descenso	L%
31	209	124	40.67	Sostenido/Descenso	L%
32	171	77	54.97	Descendente	L%
33	139	104	25.18	Sostenido	L%
34	176	93	47.16	Sostenido	L%
35	157	93	40.76	Sostenido	L%
36	126	93	26.19	Sostenido	L%

Al igual que en el primer conjunto de sílabas analizado en esta sección (cuadro 3.17), el sostenimiento del tono desde el pico de F0 resulta ser el movimiento tonal predominante (48 % de las emisiones). Sin embargo, en el cuadro 3.18, se observa un mayor número de ocurrencias en los datos de los hablantes de la red social laboral; quienes, en 13 oportunidades (de los 14 casos totales), emplean tonos de frontera bajos (L%). Los hablantes de la red vecinal presentan una tendencia menos uniforme pues, entre los 10 casos correspondientes a esta red, se registraron tres tonos de juntura altos (H%) y dos medios (!H%). Adicionalmente, se registran sostenimientos tonales hasta el final de la sílaba y movimientos descendentes luego del pico de F0 que, al igual que en los primeros casos analizados, se producen en sílabas



que inician con consonantes sordas. No obstante, también se registran casos en los que el tono se sostiene desde el pico de F0 y luego desciende antes de que culmine la sílaba nuclear, como se muestra en la figura 3.15.

Figura 3.15. Ejemplo de acento nuclear L+<H\* con descenso tonal en la misma sílaba en la sección final del enunciado Rs1H3-13:[(Cuando yo estaba en la secundaria)<sub>ip</sub> (también ya estaba trabajando)<sub>ip</sub>]<sub>IP</sub>



En aquellas instancias en las que los descensos tonales inician antes de que culmine la sílaba nuclear (cuadro 3.19) se observa que el pico promedio de F0 suele ocurrir 46.77 % antes del linde silábico, lo que hace persistente la idea de que el punto tonal más elevado se encuentra cercano al centro de la sílaba. Luego del pico de F0, el tono suele sostenerse por un promedio del 12.49 % de la longitud total de la sílaba, aunque la mayoría de las duraciones del sostenimiento tonal se encuentran por encima del 14 % y solo dos casos no superan el 10 % de la longitud de la sílaba. Posterior a este sostenimiento, el tono desciende alrededor del 34.28 % (en promedio) antes del final de la sílaba, espacio que queda disponible para que se lleve a cabo el descenso promedio de 1.45 st hasta el final de la sílaba. Ahora bien, este descenso puede empezar mucho después del valor medio presentado como referencia, como se refleja en los casos 14 y 19, o mucho antes, como en el caso 24, sin que esto repercuta

necesariamente en el grado del descenso. En el caso 31, el hablante presenta un descenso de 1.8 st en una porción silábica restante de 26.32 %; mientras que el caso 24 evidencia un descenso tonal de 1.1 st en un margen restante del 52.12 % de la duración silábica. Estos ejemplos y lo observable en el resto de los casos apuntan hacia una posible ausencia de correlación entre el punto en el que inicia el descenso y la cantidad de semitonos que lo conforman.

Cuadro 3.19. Tonemas de IP con acento nuclear L+<H\* con sostenimiento y posterior descenso de F0

Nº de caso	Pico de F0 (%)	Sostenimiento (%)	Descenso (%)	Descenso (st)	Juntura
14	36.36	14,83	21.53	1.011	!H%
19	21.17	4,56	16.61	0.998	L%
24	69.49	17,37	52.12	1.133	L%
26	50.81	8,06	42.74	2.679	L%
30	62.10	15,73	46.37	1.137	L%
31	40.67	14,35	26.32	1.851	L%

En lo referente al descenso de F0 en la sílaba nuclear y su continuidad en la sílaba final, se observa que, en cinco de los siete casos registrados, la F0 prolonga su descenso hasta culminar con una juntura baja (L%) al final del enunciado. En el caso restante, la juntura final es media (!H%), lo que indica que el tono no desciende mucho más en la sílaba final, en contraste con lo que ocurre con las junturas L%. En términos perceptivos, se encuentran solo dos casos (26 y 31) que superan el umbral de percepción mencionado por Pamies Bertrán *et al.* (2002). Los demás descensos, que van desde 0.9 st hasta 1.1 st, no serían percibidos dentro de los límites de esta sílaba, pero sí lo serían una vez que se prolongue la disminución tonal en la sílaba siguiente hasta alcanzar la juntura L%. Por lo tanto, los semitonos descendidos en la sílaba nuclear se sumarían a los de la sílaba final, aumentando así la excursión tonal en la declinación tonal. Esto quiere decir que el descenso final del enunciado inicia en la sílaba nuclear, lo que implica un adelantamiento de la juntura final a la sílaba nuclear.

Como último punto del análisis de los acentos nucleares prealineados, se presentan los resultados de los siete casos de descenso en sílabas que inician con consonante sorda prevocálica. Estos casos son llamativos debido a que 6 de las 8 realizaciones presentan

descensos tonales desde el inicio de la vocal, como refleja el ejemplo en la figura 3.16 (descenso de 3.12 st desde el inicio del diptongo [we] hasta el final de la sílaba). Sin embargo, no todas las sílabas nucleares con una consonante sorda prevocálica presentan un descenso progresivo desde el comienzo de la sonoridad, pues existen cuatro casos (2, 25, 34 y 35) en los que el tono se sostiene e incluso presenta una circunflexión a partir del punto en el que empieza la sonoridad (ver ejemplo en figura 3.17).

Figura 3.16. Ejemplo de sílaba nuclear con consonante sorda prevocálica en el enunciado Rs1H1-15: [(*Había un padre*)<sub>ip</sub> (*que se casó con una Miss Venezuela*)<sub>ip</sub>]<sub>IP</sub>

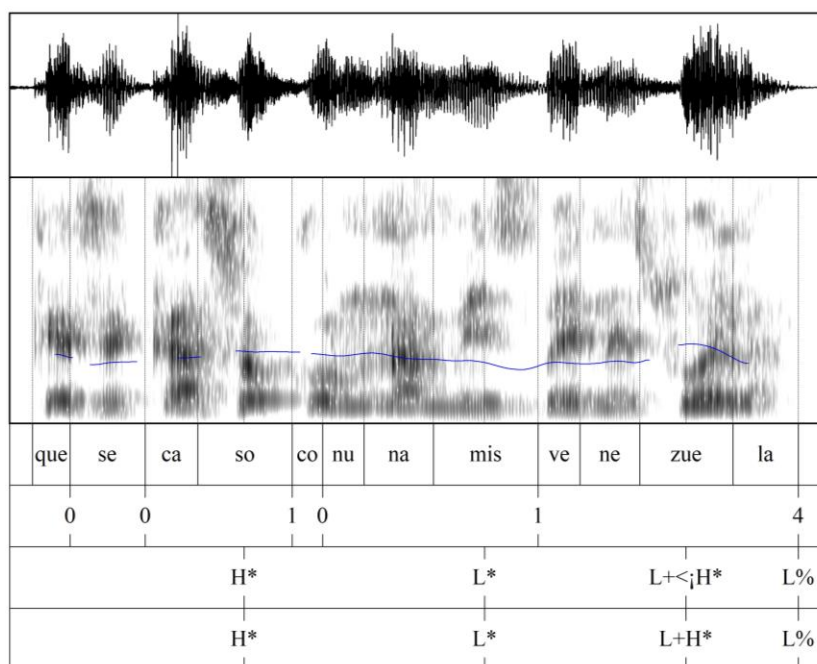
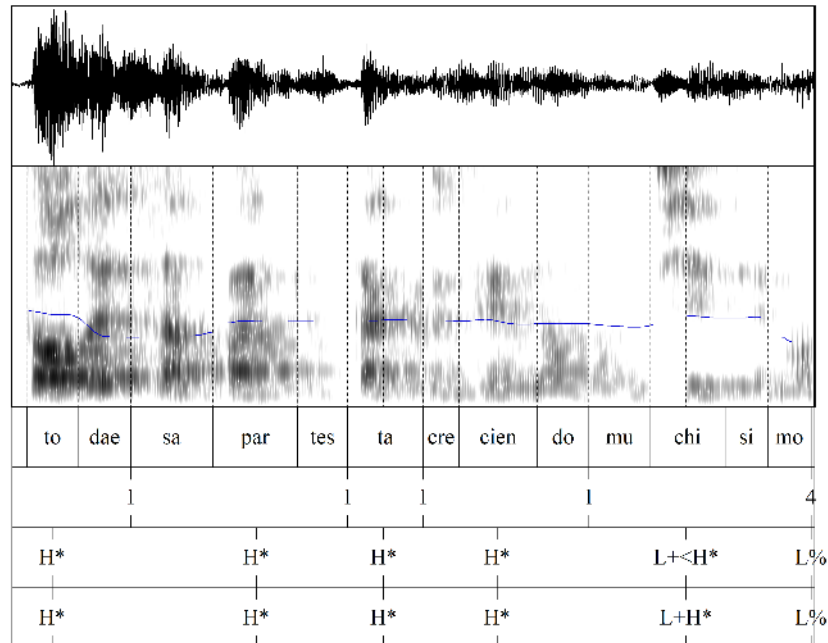


Figura 3.17. Ejemplo de sílaba nuclear con consonante sorda prevocálica en el enunciado Rs2M1-06: [(*Toda esa parte está creciendo muchísimo*)<sub>ip</sub>]<sub>L</sub>IP



Cuadro 3.20. Tonemas de IP con acento nuclear L+<H\* en los que la sílaba nuclear inicia con consonante fricativa u oclusiva sorda

Núcleo < Linde				
Nº de caso	Pico de F0 (%)	Descenso (st)	Contexto previo	Juntura
4	39.57	0.681	/t/	HL%
12	50.35	3.974	/t/	L%
Núcleo > Linde				
Nº de caso	Pico de F0 (%)	Descenso (st)	Contexto previo	Juntura
13	57.08	3.12	/s/	L%
17	30.43	1.014	/t/	H%
21	35.44	3.237	/kr/	!H%
23	35.44	1.537	/s/	L%
28	51.97	3.072	/s/	L%
32	54.97	2.595	/k/	L%

De acuerdo con los datos mostrados en el cuadro 3.20, el pico tonal se ubica, en promedio, a 44.41 % antes del linde silábico, lo que implica que la sonoridad de la vocal suele iniciar alrededor del medio de la sílaba. La media del descenso que ocurre a partir de este punto es de 2.4 st (como el que se muestra en la figura 3.18), por lo que el descenso sería perceptible

dentro de los límites de la sílaba nuclear. No obstante, se observan dos casos (4 y 17) en los que el descenso tonal sería imperceptible, pues es inferior a 1.5 st. En cuanto a la configuración general del tonema, se observa que este existe solo en el caso 17, que no presenta una circunflexión (como se puede ver en la figura 3.19, donde el descenso desde la sílaba nuclear hasta la juntura es de 1.3 st) ni siquiera parcial, como ocurriría con una juntura media (!H\*). Esto es contrario a lo que ocurre en el caso 4 (visible en la figura 3.20), en el que el descenso (imperceptible dentro de la misma sílaba) sí se concreta en la sílaba final, con una disminución total de 2.4 st.

Figura 3.18. Oscilograma, espectrograma y F0 del enunciado Rs2H3-35:

[(*Queríamos hacer un viaje a Cuba*)<sub>ip</sub>]<sub>IP</sub>

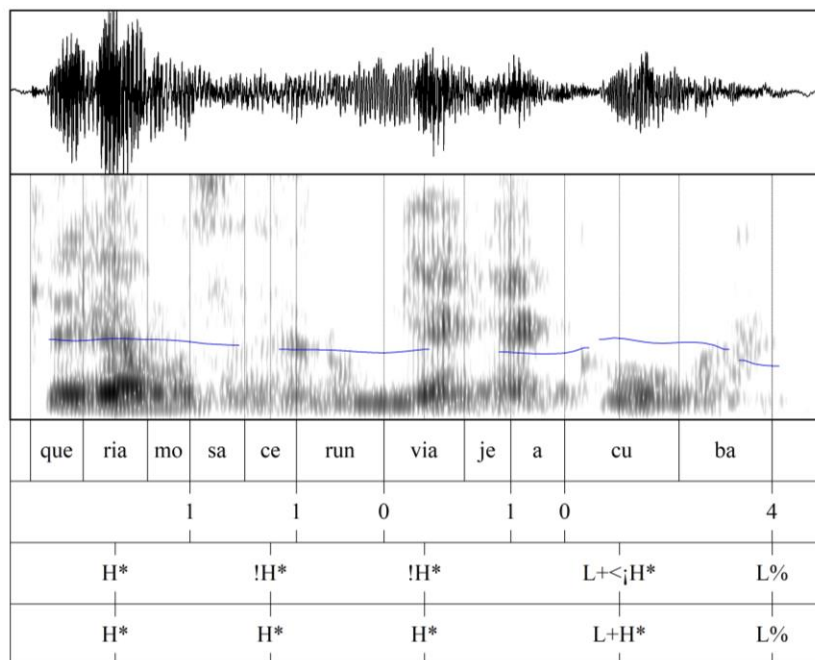


Figura 3.19. Oscilograma, espectrograma y F0 del enunciado Rs1H2-29: [(*Dicen que de tal palo tal astilla*)<sub>ip</sub>]<sub>IP</sub>

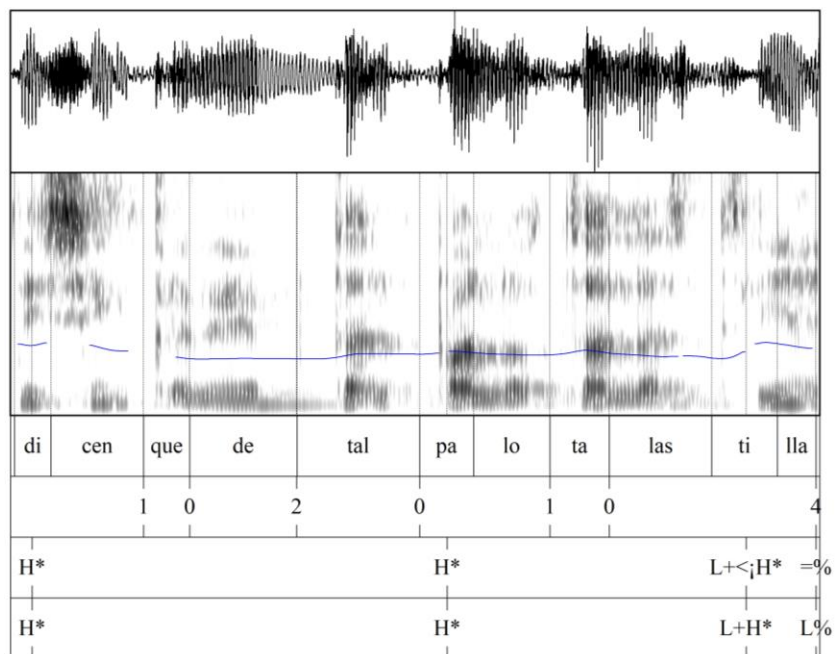
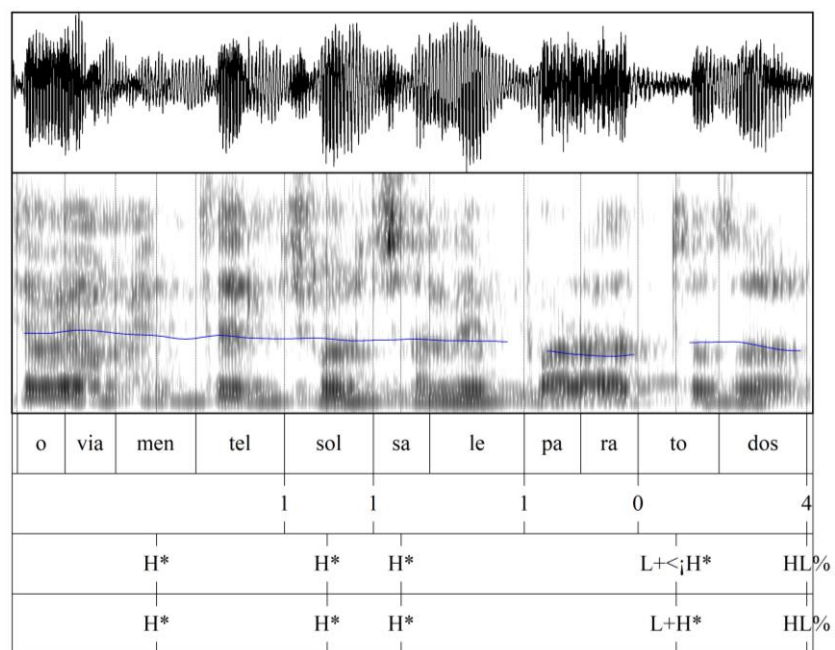


Figura 3.20. Oscilograma, espectrograma y F0 del enunciado Rs1H2-05: [(*Obviamente el sol sale para todos*)<sub>ip</sub>]<sub>IP</sub>



### 3.5. Recapitulación del análisis fonético-fonológico

La relación de los rasgos prosódicos de duración, intensidad y frecuencia fundamental (F0) presenta patrones bastante consistentes en el corpus. En cuanto al patrón de duración, con base en los resultados estadísticos, se pudo realizar una agrupación que refleja el aumento progresivo de la duración en las sílabas finales de las frases entonativas (sin distinción entre intermedias o mayores): (cuerpo primera sílaba, primera sílaba y primera tónica, primera tónica) < (núcleo de ip, núcleo de IP, linde de IP) < (linde de ip) < (núcleo y linde de ip; primera tónica, núcleo y linde de ip). Tanto la intensidad como la frecuencia fundamental (F0) presentan una declinación a lo largo del enunciado, es decir, una disminución progresiva desde el pico, localizado en la primera sílaba tónica, hasta el final del enunciado. Esta declinación es esperable, pues coincide con las características postuladas para el español desde los estudios clásicos de Navarro Tomás (1957 y 1966) hasta los trabajos más recientes del español del centro de México (Martín Butragueño, 2019).

El acento tonal abstracto más frecuente (el prototípico, de acuerdo con el enfoque basado en el uso), es H\*, seguido por L+H\*. Entre las realizaciones concretas de este último se encuentran aquellas con picos muy tempranos (L+<H\*), de las que el 45.5 % corresponde con la presencia de un foco estrecho y otro 45.5 % ocurre en enunciados con cualidad expresiva. Por otra parte, se encuentran los acentos abstractos L\* y H+L\*, que se consideran periféricos debido a que son los menos frecuentes. Entre estos, la mayoría de H+L\* se registran en enunciados enfáticos (producidos por hombres de la red social laboral), lo que sugiere una correlación entre el descenso melódico y el énfasis en los hablantes de esta red. En cuanto al acento tonal abstracto L\*, la mayoría de los casos se deben a que la primera sílaba tónica fue emitida con voz laringizada o con debilitamiento vocálico. En lo relativo a la duración, el acento tonal concreto L+<H\* posee la mayor media de duración, lo cual es explicable basándose en el mayor margen de tiempo necesario para realizar el ascenso superior a 3 st. En un segundo lugar, se encuentran los acentos concretos L+H\*, L+H\* y L+<H\*, que poseen una duración media perceptiblemente inferior a la de L+<H\* y mayor a la del resto de los acentos concretos.

Los intertonemas no oxítonos y oxítonos en las frases entonativas intermedias (ip) presentan cualidades similares, aunque difieren en algunas características. En ambos casos, las configuraciones entonativas centrales son H\* H- y L+H\* H-; entre las cuales, la primera

es la prototípica de las ip no oxítonas y la segunda para las oxítonas. En cuanto a los tonos de juntura, la juntura concreta más usual en la configuración H\* H- es H-, mientras que la de L+H\* H- es HH-. Por otra parte, no se puede afirmar que las junturas L- registradas en estos intertonemas carezcan de un sentido de continuidad en la enunciación, pero se puede considerar que la no conclusión puede ser inferida discursivamente, pues la información en la ip no constituye una unidad de sentido completo y a través de la mayor duración de la sílaba final de ip. En el caso de los intertonemas oxítonos, la configuración entonativa abstracta más recurrente es L+H\* H-, cuya realización concreta predominante consiste en un ascenso mayor a 3 st en la última sílaba tónica de la ip (L+;H\* HH-); mientras que la configuración H\* H- queda en una segunda posición de acuerdo con su frecuencia. Por otra parte, las configuraciones periféricas consisten en L\* !H-, H+L\* H- y H+L\* !H- en los intertonemas no oxítonos y en L+H\* !H-, H+L\* L- y L\* L- en los tonemas oxítonos. En relación con la duración, los intertonemas no oxítonos reflejan una mayor media de duración en la sílaba final de ip, lo cual no solo es relevante en términos absolutos, sino también relativos, ya que la sílaba final es alargada en detrimento de la nuclear en la gran mayoría de los casos. Por su parte, la duración media de los intertonemas oxítonos (275.67 ms) es, a grandes rasgos, perceptiblemente mayor que la de los no oxítonos (241.1 ms).

Al observar las diferencias entre los resultados obtenidos en las configuraciones entonativas de los dos tipos de intertonemas, es posible proponer que cada uno representa una categoría cognitiva relativamente independiente de la otra. No se puede decir que sean totalmente independientes, porque ambos tonemas internos comparten la misma función en el discurso: delimitar un nivel dentro de la jerarquía prosódica, lo que da pie a que existan ciertas similitudes entre ambas, como el hecho de que las dos configuraciones centrales sean H\* H- y L+H\* H-. Sin embargo, el hecho de que los intertonemas no oxítonos se desarrollen en dos sílabas diferentes (la nuclear y la final de la ip) y los oxítonos en una sola (nuclear y final de ip simultáneamente) conlleva implicaciones fonéticas que pueden repercutir en la estructura fonológica y cognitiva. Un ejemplo de esto es que no se registraron casos de L+H\* L- ni H+L\* H- (con excursiones tonales considerables entre el tono final del acento bitonal y la juntura) en intertonemas oxítonos. Por lo tanto, se considera que el número de sílabas que conforman cada tonema interno (específicamente el linde silábico entre los constituyentes de los intertonemas no oxítonos) y las implicaciones que esto tiene en la



organización cognitiva de los rasgos prosódicos puede generar una división de los intertonemas en dos categorías cognitivas separadas (al menos parcialmente).

Las funciones sintáctico-discursivas analizadas mantienen un patrón de frecuencia similar al descrito en las ip: los intertonemas predominantes son H\* H- y L+H\* H- (cuyas realizaciones concretas más frecuentes son H\* H- y L+;H HH-, respectivamente). Los intertonemas con acentos nucleares descendentes (H+L\*) y la configuración L+H\* !H- únicamente ocurren en construcciones con foco estrecho y en elementos extraoracionales, lo que pudiera implicar una correlación entre estas funciones informativas, de no ser porque el reducido número de casos no permite postular conclusiones categóricas. Los focos estrechos y el matiz expresivo se expresan, principalmente, a través de las configuraciones H\* H- y L+H\* H-, lo cual coincide con lo observado en las tendencias principales del corpus. Las ip que son prótasis de un enunciado condicional o cláusulas preyuxtapuestas concentran la mayoría de sus producciones con los intertonemas fonológicos L+H\* H- y con H\* H-, mientras que las ip con función explicativa ocurren únicamente con configuraciones abstractas con acentos nucleares monotonaes o ascendentes, seguidos por una juntura en la parte alta del rango tonal (H\* H-, L+H\* H- y L\* H-); juntura que parece ser una característica de estas construcciones. En cuanto a los elementos extraoracionales, los casos registrados se concentran en el empleo de los dos intertonemas abstractos prototípicos (H\* H- y L+H\* H-) y los pocos casos documentados en los otros tonemas dificultan la observación de alguna otra tendencia por parte de los hablantes.

La tendencia observada en los tonemas concuerda con lo que se había establecido previamente para el español en general (Navarro Tomás, 1957 y 1966; Quilis, 1981) y para el del centro de México (De-La-Mota *et al.*, 2010) específicamente. Los tonemas abstractos predominantes, organizados desde el más prototípico hacia el menos central, son los siguientes: L\* L% (33.33 %), H\* L% (27.5 %), L+H\* L% (18.89 %) y H+L\* L% (8.06 %). Aunque en menor proporción, se registraron casos de tonemas no descendentes, entre los que destaca la configuración L+H\* H% como la más usual. Es importante señalar que, en este caso, al igual que en el tonema H\* H%, todas las junturas concretas corresponden con un sostenimiento (=%) desde el tono final de la sílaba nuclear; y no con un ascenso que supera el tono de dicha sílaba. Este hecho realza la importancia de una transcripción prosódica en dos niveles (concreto y abstracto), basada en la transcripción de cualidades fonéticas; y una

posterior agrupación de acuerdo con los rasgos centrales invariables, ya que gracias a esta precisión fonética se establece el carácter aseverativo de la juntura abstracta H%. De lo contrario, la interpretación de dicha juntura pudiese ser asociada con una elevación tonal en la sílaba final de IP, similar a la que ocurre en algunos enunciados interrogativos (De-la-Mota *et al.*, 2010).

Los tonemas no descendentes, especialmente H\* !H%, L+H\* H% y L+H\* !H%, se registraron en focos estrechos y de expresividad en la enunciación. Adicional a esto, se identificaron 36 casos de acentos nucleares con picos muy tempranos (L+<H\*), que presentan las siguientes características: una vez alcanzado el pico tonal, generalmente alrededor del punto medio de la sílaba nuclear (casi siempre más larga que la final de IP), en la mayoría de los casos, la F0 se sostiene hasta el final de dicha sílaba, desde donde desciende hasta una juntura baja (L%). Por otra parte, en la mayoría de las ocasiones las medias de duración de las sílabas del tonema suelen ser similares, aunque la sílaba nuclear es más larga que la final. Tal frecuencia refleja la centralidad (prototipicidad) de una sílaba nuclear más prolongada que la final, en la estructura de la categoría cognitiva de enunciados aseverativos.

En términos sociolingüísticos, el análisis fonético-fonológico sugiere la relación de los patrones entonativos y de duración de la primera sílaba tónica, los intertonemas y los tonemas, que apuntan hacia la existencia de índices sociolingüísticos. El análisis sociolingüístico de los parámetros prosódicos y entonativos identificados en esta sección se desarrollará en el capítulo siguiente.

## Capítulo IV

### Aspectos sociolingüísticos en la fase de producción

En este capítulo, se muestra el análisis sociolingüístico de los tres parámetros que en el análisis fonético-fonológico presentaron diferencias prosódicas y estadísticas que sugieren la idea de índices sociolingüísticos: la primera sílaba tónica de los enunciados, los tonemas internos (intertonemas) de las frases entonativas intermedias (ip) y los tonemas de las frases entonativas mayores IP. Cabe mencionar que, debido a que ambas redes sociales presentan un comportamiento bastante uniforme en términos de la duración, intensidad y frecuencia fundamental (F0), en el análisis de las sílabas ubicadas en distintos puntos de los enunciados del corpus, presentado en §3.1 del capítulo anterior, estos aspectos no se someterán a un estudio sociolingüístico más minucioso en este capítulo.

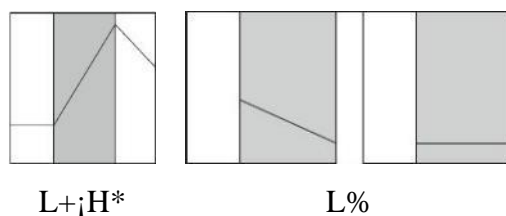
En relación con los índices sociolingüísticos en la red vecinal, se considera que estos pueden subyacer en las siguientes características: i) ascensos superiores a 3 st en la primera sílaba tónica en acentos bitonales (L+<sub>j</sub>H y L+><sub>j</sub>H\*); ii) intertonemas abstractos no oxítonos L+H\* H- (con juntura concreta HH-); iii) tonemas internos abstractos H\* H- como marcador de focos estrechos; iv) tonemas de IP en los que la combinación de los valores medios de las sílabas que los constituyen se encuentren alrededor de los 376 ms; y v) tonemas abstractos no descendentes (especialmente L+H\* H%, junto con sus respectivas realizaciones concretas). Por otra parte, los índices sociolingüísticos que se plantean para la red laboral constan de: i) acentos tonales H+L\* en la primera sílaba tónica con matiz enfático; ii) acentos tonales L+H\*, L+<sub>j</sub>H\* y L+<<sub>j</sub>H\* en la primera sílaba acentuada con una duración de 263 ms; iii) tonema interno abstracto no oxítono H\* H- con juntura concreta HH-; iv) tonemas internos oxítonos por encima de los 300 ms; v) tonemas internos abstractos oxítonos L+H\* H- (con tonos de juntura concretos HH-); y vi) tonemas con acento nuclear bitonal descendente H+L\* L% en los que la sílaba nuclear es más larga que la final.

#### 4.1. Antecedentes de trabajos prosódicos y sociolingüísticos

Actualmente, se cuenta con cuantiosa información, relacionada con el español de México, que va desde sus características generales, hasta aspectos más puntuales de su caracterización. Con respecto a las características generales del español mexicano, Sosa

(1999) y Martín Butragueño (2006a) corroboran lo ya mencionado por Matluck (1952) y Kvavik (1979), al señalar que la entonación circunfleja en el tonema es el rasgo más característico y distintivo del habla del centro del país, caracterizada por Sosa (1999) como  $L+H^* L\%$ . Martín Butragueño (2004, 2006b, 2011), explica que lo que provoca la percepción de la circunflexión parece ser el acento nuclear  $L+;H^*$  seguido por algún descenso, sea  $L\%$  o  $L-$ , lo que genera las realizaciones  $L+;H^* L\%$  (como las de la figura 1.1) y  $L+;H^* L- H\%$ . Asimismo, comenta que estas configuraciones son favorecidas por los hombres con un bajo grado de instrucción y que, además, dicha circunflexión parece estar relacionada con la estructura informativa de los enunciados.

Figura 4.1. Representación gráfica del tonema circunflejo  $L+;H^* L\%$



Si bien la influencia de elementos pragmáticos en la prosodia de los enunciados aseverativos no es el punto focal de este trabajo, es pertinente tocar este tema debido a que es abordado en el análisis. De acuerdo con Martín Butragueño (2021), existen valores pragmáticos que influyen en la entonación circunfleja ( $L+H^* L\%$ ): la prominencia informativa, la asertividad o contundencia de la información y la emotividad de la expresión. El uso de acentos nucleares ascendentes (como el del tonema circunflejo) está más relacionados con la expresividad que con la focalización; y un mayor grado en el ascenso acompañado por una juntura compleja indica vernacularidad (asociada con el habla urbana) y está relacionado con hombres, personas mayores y aquellos con bajo nivel de instrucción. Tal relación entre circunflexión y expresividad también fue confirmada desde el estudio prosódico con base en el uso, llevado a cabo por Martín Butragueño (2015a), quien afirma que los enunciados expresivos poseen elementos prototípicos de acuerdo con el subtipo expresivo al que pertenecen (ánimo o desánimo), que tienen características prosódicas específicas: los enunciados que expresan ánimo manifiestan los acentos nucleares  $L+;H^*$  y  $L+H^*$ , que infringen la restricción NO RISE (planteada desde la teoría de la optimidad) para enunciados no interrogativos en

español, mientras que los enunciados con desánimo presentan los acentos nucleares prototípicos (!)H\* y (H+!)L\*, que respetan la ya mencionada restricción.

En el ámbito de la indexación, uno de los primeros trabajos en los que se analizó el significado social indexado en variables lingüísticas es la investigación de Eckert (1988). En este trabajo, la autora estudió el cambio vocálico en un grupo de 52 estudiantes de una escuela preparatoria pública de los suburbios de Detroit, con base en la premisa de que la adquisición de las variables fonológicas locales está íntimamente involucrada en la creación de la identidad social, construida mediante el desarrollo en una estructura social dentro de un grupo etario. Como resultado de este estudio, Eckert (1998) señala que el grado de afiliación de los individuos a la red tiene incidencia en el proceso de cambio vocálico estudiado, pero ni la clase socioeconómica de los padres ni el sexo de los hablantes influyen en dicho proceso.

Si bien en la presente investigación no se estudia un fenómeno segmental, se comparte la premisa de que la variación de un elemento lingüístico está relacionada con la creación y la manifestación de la identidad social. De igual modo, se considera que la noción de mercado lingüístico puede ser un factor determinante en la producción de distintos rasgos prosódicos que pueden llevar carga social. Zhang (2005), quien analizó el uso de cuatro variables fonológicas en dos grupos profesionales de Beijing, pudo constatar que los trabajadores de empresas extranjeras, a diferencia de los que trabajan en empresas locales, estarían usando una nueva forma de mandarín, con características del mandarín de Beijing y del mandarín insular de Hong Kong y Taiwán, para construir un estilo cosmopolita. Esto hace sospechar, desde el planteamiento inicial del análisis, que existen variaciones prosódicas, basándose en la participación de los miembros de ambas redes de este trabajo en el mercado lingüístico.

Finalmente, se considera necesario mencionar el trabajo de Campbell-Kibler (2007), quien explora la relación existente entre tres variantes de (ING) en inglés y dos acentos distintos (el sureño y el *gay*) en el razonamiento sociolingüístico de estudiantes universitarios estadounidenses. En esta investigación, la autora realizó un análisis sociolingüístico y una prueba de percepción que buscó conocer el significado social identificado en cada una de las variantes que estudió, lo que le permitió concluir que existe una densa conexión entre las diferentes formas de asociación de una ejecución lingüística y las reacciones sociales que esta puede generar. Ahora bien, a diferencia de Campbell-Kibler (2007), en este trabajo no se plantean de antemano las variantes que se consideran portadoras de significado social, ya

que dichas variantes se tomaron de los resultados obtenidos a través del análisis sociolingüístico. Sin embargo, trabajos como el de esta autora ponen en relieve la pertinencia de la ejecución de pruebas de percepción para confirmar las hipótesis generadas en torno a las variantes lingüísticas y su significado social potencial.

#### 4.2. Primera sílaba tónica

Según la distribución de acentos tonales anclados en las primeras sílabas tónicas producidas por ambas redes (cuadro 4.1), ambos grupos se orientan principalmente hacia el empleo del acento tonal H\* (especialmente por la realización concreta H\*), que es el ejemplar central (prototípico) de esta categoría, como ya se había definido en el capítulo III. En segundo lugar de frecuencia, y como el ejemplar más cercano al central, se encuentra L+H\*, entre cuyas realizaciones concretas destacan el uso de la configuración L+>H\* en la red social vecinal y L+>H\* en la laboral (como se puede ver en las figuras 4.2 y 4.3 , respectivamente), lo cual implica que la primera red prefiere ascensos melódicos de al menos 3 st hasta la primera sílaba postónica del enunciado, mientras que la red laboral no privilegia estos ascensos. De hecho, esta tendencia a producir ascensos superiores a 3 st por parte de la red vecinal se acentúa con el hecho de que doblan en número a las producciones de dichos ascensos de la red social laboral.

Cuadro 4.1. Acentos tonales en la primera sílaba tónica en cada red social. N=360

Acentos tonales abstractos	Red vecinal	Red laboral	Total
H*	101	92	193
L+H*	75	70	145
L*	4	10	14
H+L*	0	8	8
Total	180	180	360

p = 0.007 (Estadístico Exacto de Fisher)

Figura 4.2. Oscilograma, espectrograma y F0 del enunciado Rs1M1-15: [(*Nosotros no los necesitamos*)<sub>ip</sub>]IP

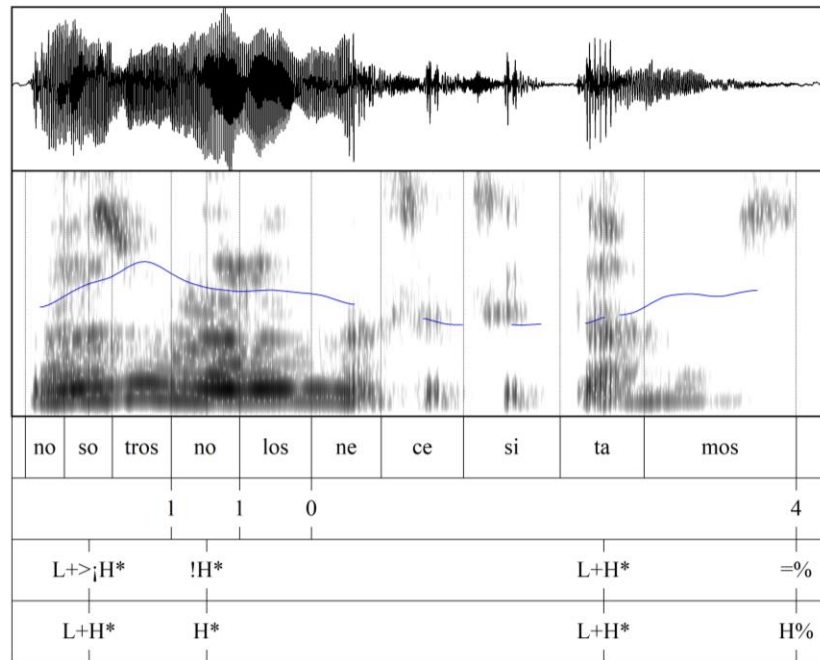
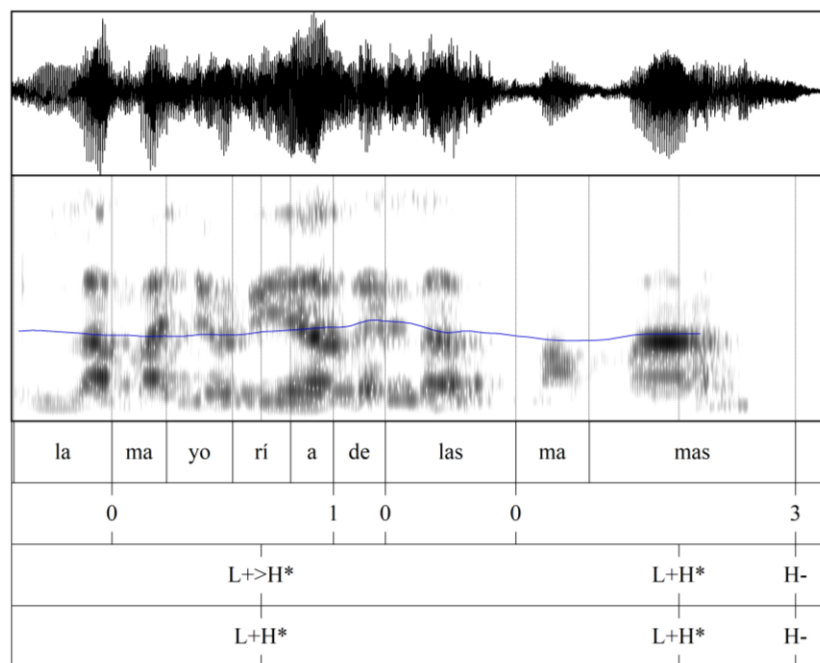


Figura 4.3. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M1-24: [(*La mayoría de las mamás*)<sub>ip</sub> (*de la escuela de mi hija*)<sub>ip</sub> (*no trabajan*)<sub>ip</sub>]IP



Entre los acentos menos comunes en la periferia de la categoría se encuentran L\* y H+L\*. La distribución de estos acentos revela que el primero es más empleado por la red laboral, mientras que el segundo es empleado exclusivamente por dicha red. Ahora bien, la aparición poco esperada de los acentos tonales L\* y H+L\* en la primera sílaba tónica, no solo contradice la declinación desde tonos altos al inicio del enunciado (Navarro Tomás 1957 y 1966; De-La-Mota *et al.*, 2010), sino también las premisas del código biológico de producción (Gussenhoven, 2004). Por tal motivo, se presta atención a estos casos. De las 4 producciones documentadas de L\* en la red social vecinal, una corresponde con un debilitamiento en la producción de la vocal, otra con un efecto de laringización y las otras dos, ambas producidos por el hablante Rs1H3, sí corresponden con realizaciones de fonación plena, (figura 4.4). En cuanto a la red laboral, 4 de las realizaciones del acento tonal L\* ocurren en contextos de debilitamiento vocálico, 2 en fonación laringizada; mientras que los 4 casos restantes suceden en vocales plenas y con fonación modal. Esto no solo implicaría un mayor empleo de fonación no modal por parte de los integrantes de la red laboral, sino también un uso ligeramente mayor del acento L\* pleno (es decir, sin debilitamiento ni laringización) por parte de la red laboral. En cualquier caso, se considera que la frecuencia de los acentos L\*, sea debido a fenómenos de fonación no modal o no, no representa una diferencia substancial en el habla y, por ende, en la disposición cognitiva de elementos prosódicos de ambas redes.

Por otra parte, con respecto a las producciones del acento H+L\* en la red laboral, es necesario precisar que todas ellas ocurren en contextos en los que los hablantes iniciaban el enunciado de forma enfática (como es el caso en la figura 4.5, donde el descenso tonal en la parte estable de la vocal es de 2.7 st). A diferencia de lo observado con el acento L\*, las producciones de H+L\* sí podrían implicar una diferencia relevante en términos sociolingüísticos, pues es un rasgo considerablemente extendido en la red laboral, empleado por 5 de sus 6 miembros. Sin embargo, esta idea debe ser considerada como una primera visión de este acento en enunciados que inician de manera enfática, pues el número de casos es muy reducido como para hacer cualquier afirmación categórica al respecto.



Figura 4.4. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H3-7: [(*Pues mientras no caigan las cabezas*)<sub>ip</sub> (*los peones siguen trabajando*)<sub>ip</sub>]<sub>IP</sub>

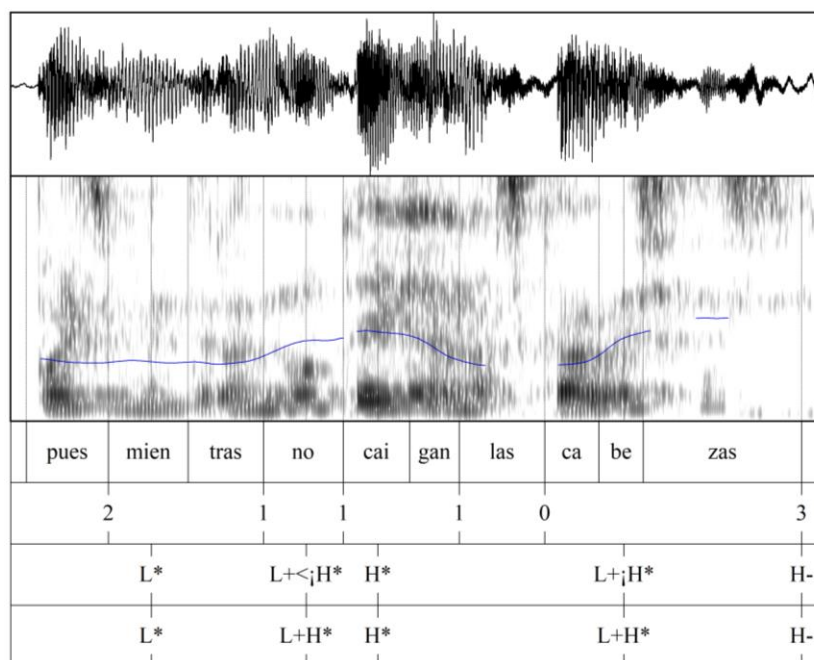
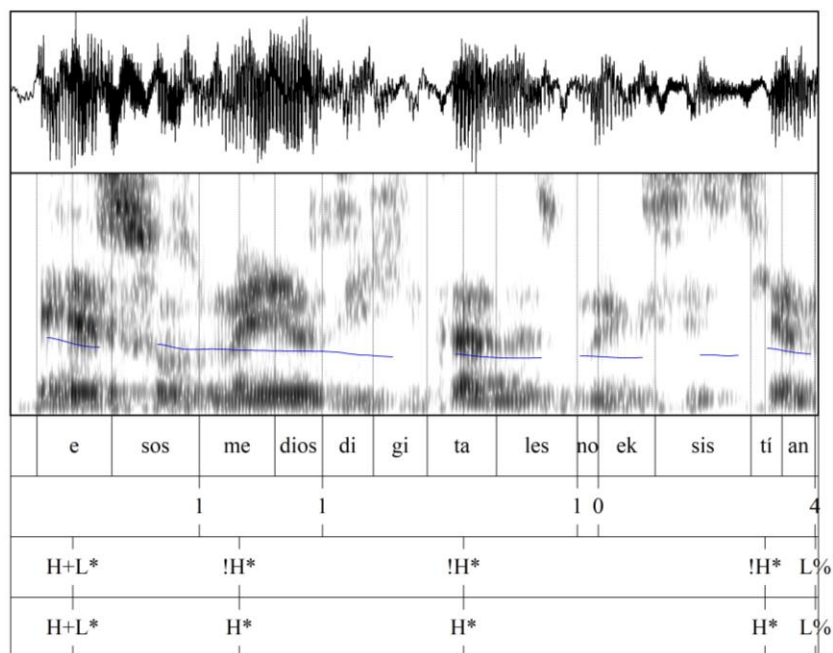


Figura 4.5. Oscilograma, espectrograma y F0 del enunciado Rs2H1-22: [(*Esos medios digitales no existían*)<sub>ip</sub>]<sub>IP</sub>



Los que revelan los datos mostrados en el cuadro 4.1 es, en términos cognitivos, que la categoría cognitiva correspondiente con el primer acento tonal es bastante similar en ambas redes y que corresponde con lo mostrado en la figura 3.4 del capítulo III. El hecho de que no se observe una diferencia considerable en el grado de centralidad de los esquemas fonológicos, correspondientes con los distintos acentos tonales en las dos redes, permite inferir que la organización de los esquemas entonativos dentro de esta categoría cognitiva puede concordar con patrones cognitivos de la comunidad de habla de la Ciudad de México. Cabe señalar que esta afirmación es solo una hipótesis que surge sobre la base de la similitud de los datos de frecuencia de producción de ambas redes, la cual queda pendiente por ser probada. Por otra parte, la ausencia de casos del acento H+L\* en la red vecinal no necesariamente implica que este esquema tenga una posición periférica dentro de esta red, sino que el esquema, incluso, podría no existir en su estructura cognitiva, lo cual podría ser un indicador de una diferencia sociocognitiva. Sin embargo, no se puede afirmar la ausencia de dicho acento en las estructuras cognitivas de esta red, hasta llevar a cabo un estudio con un número mucho mayor de enunciados con énfasis en el inicio, especialmente teniendo en cuenta que el acento H+L\* tampoco cuenta con muchos casos en la red laboral.

En cuanto al parámetro de duración, la duración media que tiene la primera sílaba tónica en relación con los acentos tonales abstractos no resultó estadísticamente significativa en ninguna de las redes sociales, lo que es esperable teniendo en cuenta que las medias de duración de los diferentes acentos tonales son relativamente uniformes, como se puede ver en los datos mostrados en el cuadro 4.2. La mayoría de los casos comparables (en los que se registran los mismos acentos en ambas redes) tienen mayor duración en las emisiones de la red laboral (excluyendo el acento tonal H+L\*, que no cuenta con casos por parte de la red vecinal). Sin embargo, las diferencias de medias entre ambas redes son bastante reducidas en la mayoría de los casos. La única diferencia considerable es la media de duración del acento tonal abstracto L\*, que, con un margen de 28.25 ms a favor de la red social vecinal, no supera la diferencia apenas perceptible de 30 ms (Toledo 1998 y 2000).

Cuadro 4.2. Medias de duración (ms) de la primera sílaba tónica de acuerdo con los acentos tonales producidos por cada red social

Acentos tonales abstractos	Red vecinal	Red laboral
H*	139.72	144.97
L+H*	151.28	162.31
L*	144.75	116.5
H+L*	-	132
Total	144.65	149.56

$p = 0.474$  (ANOVA red vecinal);  $p = 0.201$  (ANOVA red laboral);

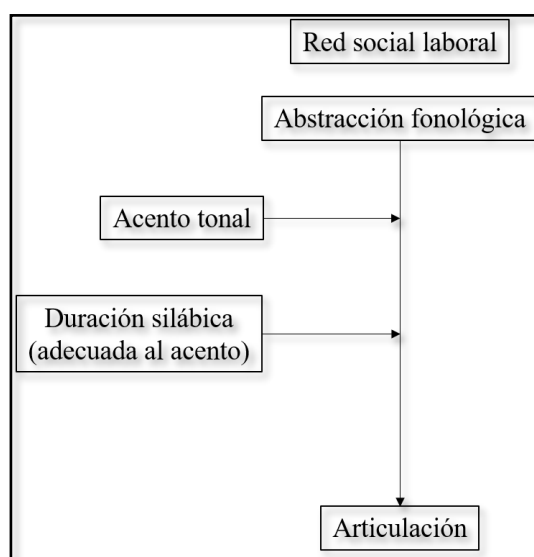
$p = 502$  (ANOVA al nivel intergrupar)

En el nivel concreto, la frecuencia de los acentos tonales L+H\*, L+<sub>i</sub>H\* y L+<<sub>i</sub>H\* presentan diferencias notorias entre los hablantes de las dos redes. Las duraciones medias de estos tres acentos en la red vecinal van de 166 ms a 182 ms (media de 175.3 ms);, mientras que en la red laboral van de 218 ms a 344 ms (media de 263 ms). Debido a que estos patrones son consistentes en ambas redes, pues los valores medios de cada hablante de ambas redes no distan mucho de los expuestos previamente, se puede decir que los datos mostrados en el cuadro 4.2 son representativos de cada grupo.

La mayor duración observada en los miembros de la red laboral con respecto a los acentos concretos L+H\*, L+<sub>i</sub>H\* y L+<<sub>i</sub>H\* puede estar relacionada con que los hablantes de esta red consideren necesario utilizar una mayor cantidad de tiempo para producir la elevación de F0 hasta su punto máximo dentro de los límites de la primera sílaba tónica. Expuesto en términos cognitivos, en la producción lingüística de los hablantes de la red laboral, la asignación del acento tonal L+ (<)(<sub>i</sub>)H\* tiene lugar en una etapa más temprana que la asignación de la duración silábica en el enunciado (como se ejemplifica en la figura 4.6). La duración, designada posteriormente, se determina con base en el acento tonal, con la finalidad de otorgar el tiempo necesario para la ejecución del movimiento tonal predefinido en la fase prearticulatoria (previa a la oralidad). Dicha relación de causalidad no parece estar presente en la estructura cognitiva de los miembros de la red vecinal, lo que da pie a pensar que los rasgos sociales propios de las redes pueden implicar que sus miembros estén expuestos a distintos fenómenos en sus experiencias lingüísticas, que configuren sus estructuras cognitivas no solo en relación con el uso del lenguaje (en términos de frecuencia, al menos), sino también con el orden de los factores involucrados en la construcción

lingüística. Ahora bien, es posible que el tipo y el volumen de casos analizados no sean los indicados para realizar esta afirmación de forma categórica; sin embargo, los datos observados son suficientes como para crear una duda acerca de la posibilidad de que los individuos de una misma comunidad de habla presenten diferencias cognitivas, relacionadas con el orden de integración de rasgos prosódicos en un enunciado de acuerdo con las características de las redes sociales a las que pertenecen.

Figura 4.6. Orden de asignación de rasgos prosódicos (acento tonal y duración) en el sistema cognitivo de la red laboral



Al analizar la frecuencia de los acentos tonales en relación con la variable sexo, estadísticamente significativa (cuadro 4.3), se observa que el acento tonal abstracto es prototípico de los hombres es H\*, pues cuenta con una diferencia de casos considerable con respecto al acento tonal abstracto más cercano en frecuencia (L+H\*). En el plano concreto, el ejemplar más frecuente corresponde con el acento H\*, ya que 73 de los 104 casos (70.2 %) del acento abstracto H\* son producidos a través de esta realización. En las mujeres, en cambio, no se observa un acento tonal que pueda ser catalogado como prototípico. Aunque la frecuencia total del acento L+H\* difiere por 23 casos entre mujeres y hombres, las ocurrencias de L+H\* y de H\* no difieren mucho entre los grupos femeninos. Ahora bien, al comparar las frecuencias de los datos de las mujeres de acuerdo con la red social a la que

pertenecen, se observa que las integrantes de la red vecinal favorecen claramente el uso del acento tonal H\* (figura 4.7), mientras que las de la red laboral tienden al uso de L+H\* (figura 4.8).

Cuadro 4.3. Recuento de acentos tonales en la primera sílaba tónica en relación con la variable sexo de acuerdo con la red social. N=180

Acentos tonales fonológicos	Hombres			Mujeres		
	Red vecinal	Red laboral	Total	Red vecinal	Red laboral	Total
H*	49	55	104	52	37	89
L+H*	38	23	61	37	47	84
L*	3	6	9	1	4	5
H+L*	0	6	6	0	2	2
Total	90	90	180	90	90	180

p = 0.050 (Estadístico Exacto de Fisher únicamente con la variable sexo)

p = 0.046 (Estadístico Exacto de Fisher combinando variables sexo y red social)

Figura 4.7. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1M3-33: [(Conocí a un chavo con el que empecé a salir)<sub>ip</sub> (pero me causó muchos problemas porque era muy celoso)<sub>ip</sub>]<sub>IP</sub>

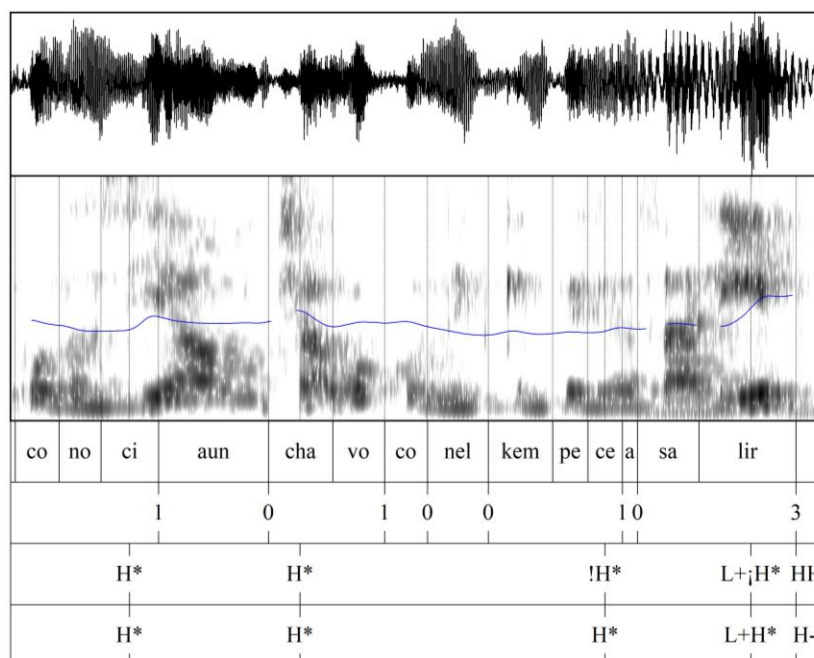
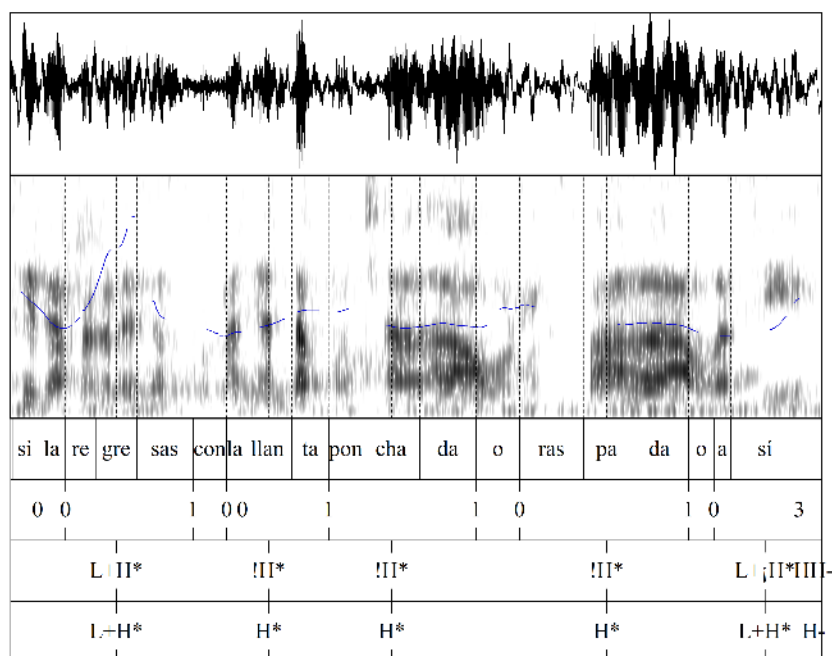


Figura 4.8. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M3-19: [(Si la regresas con la llanta ponchada o raspada o así)<sub>ip</sub> (pues te cobro)<sub>ip</sub>]<sub>IP</sub>



En cuanto a las realizaciones concretas de los acentos tonales por parte de las mujeres, se observa que la representación concreta del acento H\* es H\* (al igual que en los hombres, aunque con 21 casos menos que ellos), mientras que las realizaciones más frecuentes del acento L+H\* son las que cuentan con el pico tonal desplazado a la sílaba postónica (L+><sub>i</sub>H\* y L+>H\*, con 30 y 27 emisiones, respectivamente), las cuales han sido documentadas como un acento característico de los enunciados aseverativos de foco amplio en la Ciudad de México (De-la-Mota *et al.*, 2010). Si bien las realizaciones L+><sub>i</sub>H\* y L+>H\* también son las más favorecidas por los hablantes masculinos dentro del esquema entonativo L+H\*, las frecuencias registradas en los datos de las mujeres los superan en ambos casos, especialmente en L+><sub>i</sub>H\* (duplican el número de emisiones masculinas). Esto reafirma la evidencia a favor de la tendencia secundaria<sup>23</sup> femenina, más marcada que la masculina, al uso de acentos tonales ascendentes con pico tonal desplazado a la sílaba postónica, principalmente con un

<sup>23</sup> Con el término “tendencia secundaria” se hace referencia a un patrón recurrente que, aunque menos frecuente que el patrón predominante o primario, es suficientemente productivo como para ser reconocible entre la totalidad de los datos.

incremento de 3 o más st. En concreto, se observa una diferencia bastante marcada entre el uso de acentos tonales ascendentes de las mujeres de la red laboral y el del resto de los grupos (todos ellos emplean principalmente el acento tonal abstracto H\*), especialmente en contraste con las producciones de los hombres de su misma red.

Los valores de la duración media de las primeras sílabas tónicas (cuadro 4.4), presentan variación con respecto al sexo de los informantes. Aunque los valores medios totales en los datos de hombres y mujeres son bastante cercanos, se registran diferencias en función del acento tonal. En todos los acentos tonales, con excepción del acento tonal H+L\* (que solo se produce en la red laboral), se observa que las medias de duración de los hombres de la red vecinal son similares a las de las mujeres de la red laboral; mientras que la de los hombres de la red laboral son similares a las de las mujeres de la red vecinal. Sin embargo, las medias de todos los grupos en lo referente a los acentos H\* y L+H\* se pueden considerar uniformes, debido a que no superan el umbral para alcanzar una diferencia perceptible. No ocurre lo mismo en el caso del acento L\*, donde la diferencia de duración entre los hombres de la red vecinal y las mujeres de la red laboral y los demás grupos supera el margen de 30 ms, valor mínimo para que dicha diferencia sea percibida. Esta similitud de medias superiores registrada en las mujeres de la red laboral y los hombres de la vecinal podría ser analizada desde la noción del prestigio encubierto (§1.3.4). Es posible que este grupo de mujeres produzca la primera sílaba tónica con una mayor duración cuando en ella se ancla el acento L\*, debido a que asocian este rasgo con cualidades masculinas de la clase trabajadora que pueden ayudar a reforzar su posición de autoridad dentro de su entorno laboral (fuerza, carácter, dominio, poder, etc.). Esta idea se apoya en que tanto los hombres de la red laboral como las mujeres de la vecinal cuentan con medias de duración al menos 50 ms más cortas. No obstante, la propuesta de prestigio encubierto se plantea como hipótesis debido a que el número de ocurrencias del acento en cuestión (L\*) es bastante reducido y, principalmente, porque en este trabajo no se realizaron los procedimientos metodológicos para confirmar dicha hipótesis. Así, se considera contemplar este aspecto como un eje de estudio en un trabajo posterior.

Cuadro 4.4. Medias de duración (ms) de la primera sílaba tónica de acuerdo con los acentos tonales en relación con la variable sexo

Acentos tonales abstractos	Hombres			Mujeres		
	Red vecinal	Red laboral	General	Red vecinal	Red laboral	General
H*	139	152.95	146.38	140.4	133.11	137.37
L+H*	161.5	152.52	158.11	140.78	167.11	155.51
L*	159	93	115	102	151.75	141.8
H+L*	-	136.67	136.67	-	118	118
Total	149.17	147.76	148.46	140.13	151.36	145.74

p = 0.710 (ANOVA al nivel intergrupar)

En el nivel concreto, los hombres mantienen su tendencia a generar las primeras sílabas tónicas más prolongadas que las mujeres, aunque en las mujeres se observa un comportamiento llamativo en relación con las representaciones concretas del acento tonal L+H\*. En las realizaciones de este acento con picos muy tempranos L+<H\* y L+<H\*, las medias de los datos de las mujeres son perceptiblemente mayores que las masculinas, con 361.5 ms y 217, respectivamente, en el caso de las mujeres y 143 ms y 163.5 ms en el caso de los hombres. Teniendo en cuenta que 10 de los 11 casos de picos muy tempranos en la primera sílaba tónica del corpus están relacionados con énfasis prosódico (no necesariamente asociados con una estructura focalizadora sintáctica, como la mostrada en la figura 4.9, sino con un tipo de énfasis expresivo, como el que ocurre en la figura 4.10), se puede inferir que las mujeres refuerzan este énfasis prosódico al añadir una mayor duración a la primera sílaba acentuada, refuerzo que no está presente en las producciones masculinas. En las demás realizaciones concretas de este acento tonal, excepto L+>H\* (aunque por una diferencia muy reducida de apenas 4.44 ms), los hombres presentan mayor duración media.



Figura 4.9. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2H2-32: [(En ese instante)<sub>ip</sub> (para mí desapareció la ventana)<sub>ip</sub>]<sub>IP</sub>

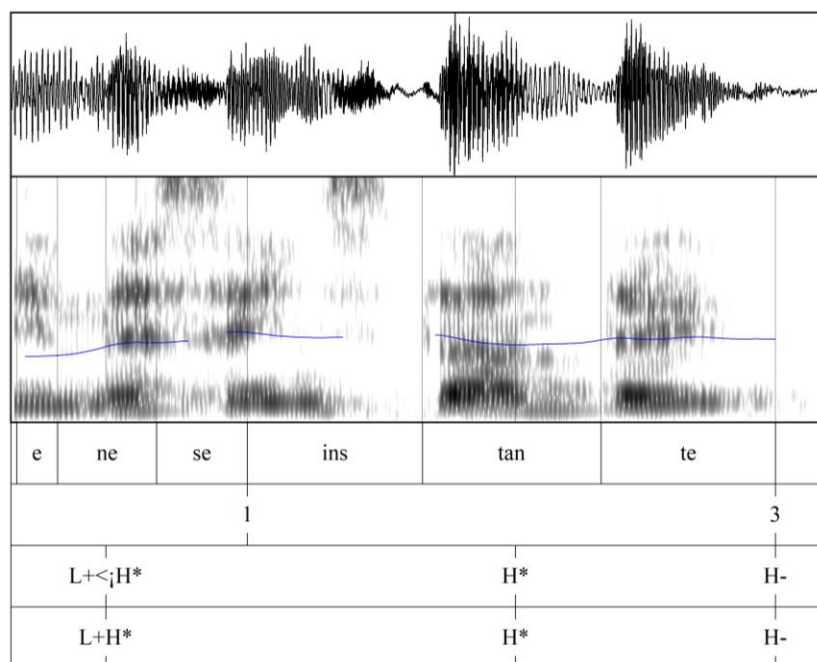
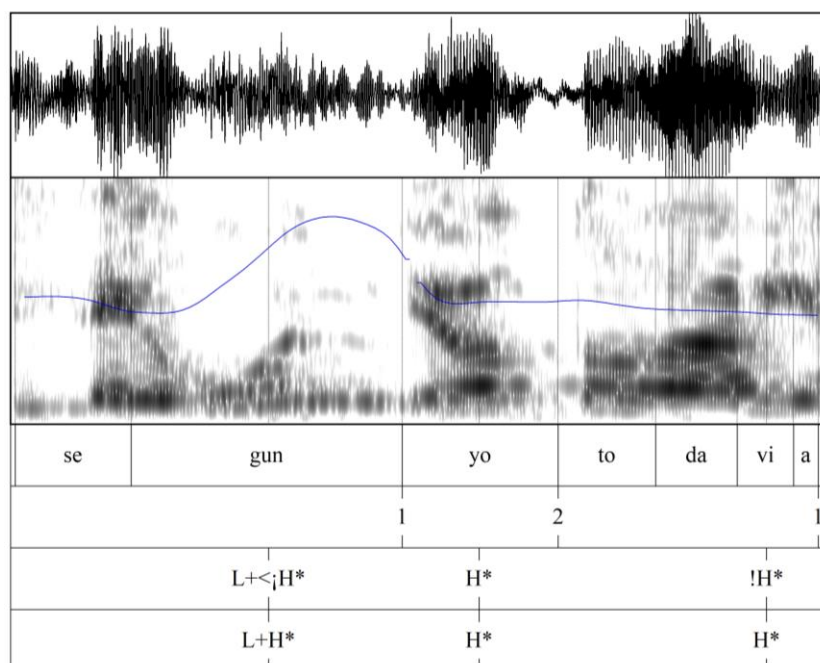


Figura 4.10. Oscilograma, espectrograma y F0 de las tres primeras palabras fonológicas del enunciado Rs2M2-28: [(Según yo todavía en aquel entonces no estaban como ahorita los niños)<sub>ip</sub>]<sub>IP</sub>



En conclusión, con base en la frecuencia de uso, los esquemas entonativos de la red social vecinal y la laboral presentan un nivel de centralidad similar. En ambos casos, el grado de prototipicidad correspondería, en orden descendente, con los acentos tonales H\*, L+H\*, L\* y H+L\*. En este punto, es conveniente discutir el carácter periférico del acento tonal H+L\* en la red social vecinal, ya que no se registra ninguna producción de este en dicha red. Esta ausencia de casos en los hablantes de la red vecinal no necesariamente indicaría que este esquema es un elemento periférico en las primeras sílabas tónicas de un enunciado, pues el carácter periférico, de acuerdo con los postulados de la teoría basada en el uso (Bybee, 2001), dependería de que la frecuencia de producción se encuentre en el rango inferior de ocurrencias en relación con la de los demás elementos. Ahora bien, dado que este esquema no se registra dentro del repertorio entonativo de dicha red, con base únicamente en los datos recolectados para este análisis, no podría considerarse como un elemento periférico de la categoría de movimientos melódicos en la primera sílaba tónica de un enunciado. Por lo tanto, se puede plantear la hipótesis de que el acento tonal H+L\* no esté presente dentro de la categoría cognitiva correspondiente a la primera sílaba tónica, razón por la cual no tendría ningún estatus de centralidad en ella. Dicha hipótesis, sin embargo, no puede ser confirmada o descartada hasta que se realice un estudio con un mayor número de casos de producción de enunciados con énfasis en el inicio del enunciado.

Por otra parte, lo observado en el nivel concreto revela cierto nivel de variación que puede implicar indexación sociolingüística. La consideración de que un rasgo prosódico es portador de significado social relacionado con uno u otro grupo no necesariamente viene dada porque el elemento que presenta la variación sea el prototípico en una u otra red. El hecho de que la diferencia de productividad en una red con respecto a la otra es suficiente como para crear una brecha en términos de frecuencia entre ambas redes. Este es el caso, por ejemplo, de los acentos concretos L+<sub>i</sub>H\* y L+><sub>i</sub>H\*, los cuales, a pesar de no ser los más frecuentes en la red social vecinal, poseen una cantidad de ocurrencias considerablemente elevada por parte de los hablantes de esta red que supera a aquella de los hablantes de la red laboral.

#### 4.3. Tonemas internos (intertonemas): Configuraciones nucleares de frase entonativa intermedia (ip)

En los siguientes incisos se muestran los análisis sociolingüísticos relacionados con los tonemas internos de las frases entonativas (ip). Al igual que en el capítulo III, este apartado se divide en el análisis por separado de los intertonemas no oxítonos y los oxítonos debido a dos motivos: en primer lugar, el primero de estos intertonemas abarca dos sílabas y el segundo abarca solo una, por lo que un análisis unificado resulta inconveniente al momento de estudiar la duración de los intertonemas; en segundo lugar, en la duración de los intertonemas oxítonos se identificó una diferencia significativa entre ambas redes sociales.

##### 4.3.1. Tonemas internos (intertonemas) no oxítonos

En los datos mostrados en el cuadro 4.5, se observa que en cada red predominan los intertonemas abstractos  $H^* H-$  y  $L+H^* H-$  (figuras 4.11 y 4.12, respectivamente). Como se puede ver en la figura 4.13,<sup>24</sup> el primero de ellos es el esquema prototípico en la categoría cognitiva de intertonemas no oxítonos de la red laboral, por su mayor frecuencia de uso. La red vecinal, en cambio, no parece tener un ejemplar central igual de definido que la red laboral, pues en la red vecinal se observa que ambos esquemas entonativos tienen frecuencias de uso bastante cercanas, aunque resalta el hecho de que el porcentaje de ocurrencias de la configuración  $L+H^* H-$  en esta red es mayor que en la red laboral. Por lo tanto, se puede inferir que, dada la diferencia en porcentajes, los intertonemas no oxítonos  $H^* H-$  pueden ser más representativos (y por ende índices sociolingüísticos) de la red social laboral. Por otra parte, la configuración  $L+H^* H-$ , mucho más frecuente en la red vecinal, podría transmitir un significado sociolingüístico relacionado con los rasgos sociales de esta red, lo cual explicaría por qué su uso es tan reducido entre los trabajadores ejecutivos.

---

<sup>24</sup> Todas las figuras en las que se representa el grado de centralidad o prototipicidad de un esquema cognitivo presentadas en este capítulo se crearon siguiendo el mismo procedimiento que en el capítulo III. Se tomó como referencia la frecuencia absoluta del esquema entonativo más recurrente, el cual constituye el 100% de centralidad, partiendo de la premisa de que el valor más elevado de frecuencia corresponde con el mayor grado de centralidad (prototipicidad).

Cuadro 4.5. Tonemas internos no oxítonos en cada red social. N=194

Tonemas internos abstractos	Red vecinal		Red laboral		Total	
	Casos	%	Casos	%	Casos	%
H* H-	29	34.12	48	44.04	77	39.69
H* !H-	6	7.06	4	3.67	10	5.15
H* L-	4	4.71	9	8.26	13	6.70
L+H* H-	26	30.59	20	18.35	46	23.71
L+H* L-	5	5.88	3	2.75	8	4.12
L* H-	6	7.06	8	7.34	14	7.22
L* !H-	1	1.18	5	4.59	6	3.09
L* L-	2	2.35	10	9.17	12	6.19
H+L* H-	5	5.88	1	0.92	6	3.09
H+L* !H-	1	1.18	1	0.92	2	1.03
Total	85	100	109	100	194	100

p = 0.044 (Estadístico Exacto de Fisher)

Figura 4.11. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M2-1: [(*Finalmente estudiamos todo*)<sub>ip</sub> (y decidimos que)<sub>ip</sub> (*comunicación era lo mío*)<sub>ip</sub>]<sub>IP</sub>

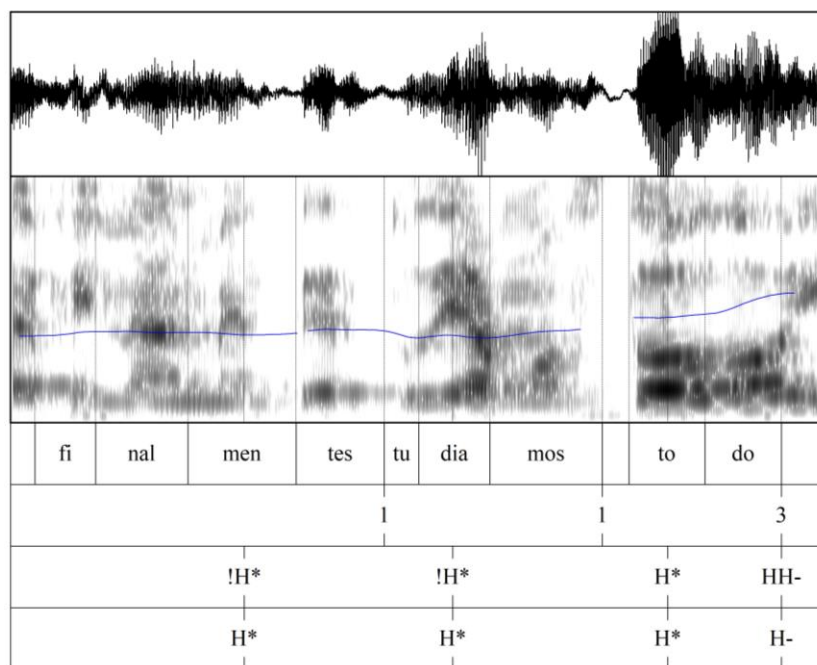


Figura 4.12. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1M2-24: [(*Mejor me dieron informes*)<sub>ip</sub> (*de mi sobrino*)<sub>ip</sub>]<sub>IP</sub>

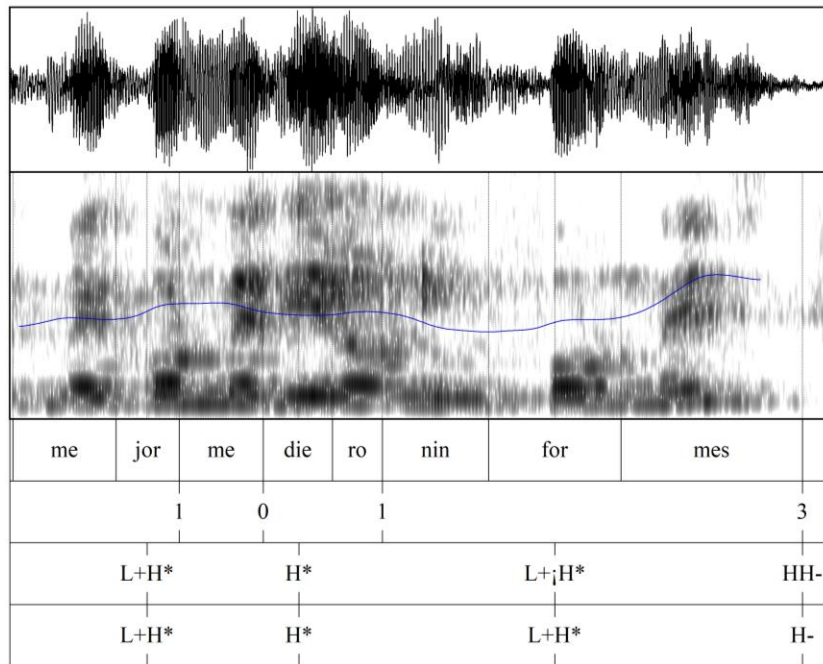
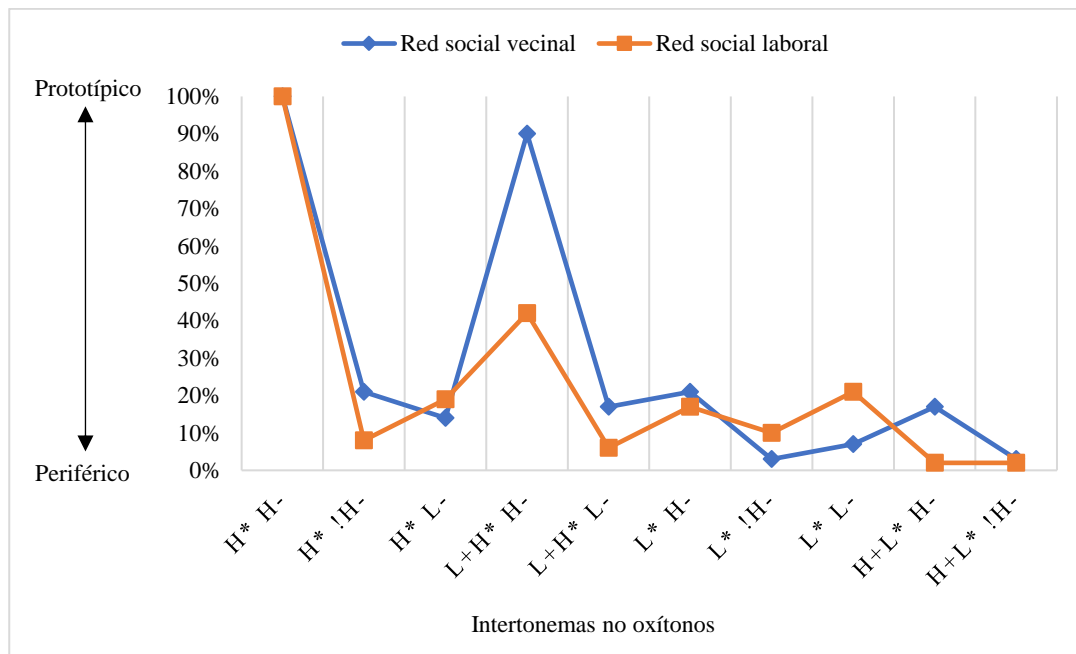
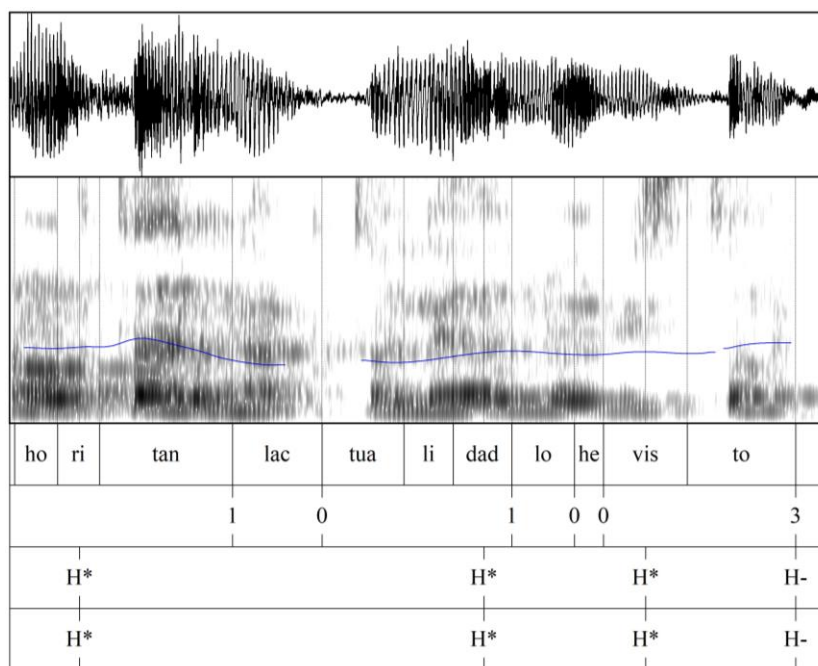


Figura 4.13. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente a los intertonemas no oxítonos en ambas redes sociales



El tonema interno abstracto H\* H- se realiza como cuatro configuraciones principales: !H\* H\*, H\* H-, !H\* HH- y H\* HH-. Entre estos intertonemas concretos, los que tienen juntura H-, que representa un ascenso entre los 1.5 st y 3 st (figura 4.14), son empleados por ambas redes en un número similar de ocasiones (14 en la red vecinal y 18 en la laboral), al tiempo que los que presentan juntura HH- (mayor a 3 st, como los que se observan en las figuras 4.11 y 4.12) son mucho más empleados por la red social laboral que por la vecinal. Con relación al tonema interno abstracto L+H\* H-, las realizaciones concretas L+¡H\* HH- y L+H\* HH- destacan sobre las demás por su elevado número de ocurrencias (27 en total). Ambas configuraciones son producidas con mayor frecuencia por parte de la red vecinal, pero la diferencia no es mayor a tres casos, por lo que no se puede considerar significativa.

Figura 4.14. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H3-33: [(Ahorita en la actualidad lo he visto)<sub>ip</sub> (ya ves a chavos drogándose)<sub>ip</sub>]<sub>IP</sub>



El resto de tonemas internos abstractos se podrían considerar esquemas periféricos de la categoría de intertonemas no oxítonos, pues ninguno de ellos supera el 9.5 % de ocurrencias. En el caso de la red vecinal, los esquemas entonativos más alejados del prototipo son aquellos que implican un ascenso melódico, desde un acento nuclear bajo hasta una juntura media (L\*

!H- y H+L\* !H-), ya que, para ambas instancias, se registra únicamente un caso. Llama la atención el hecho de que el intertonema abstracto H+L\* H- cuenta con 5 casos en la red vecinal y solo uno en la laboral, lo que podría resaltarlo como un posible índice sociolingüístico entre los intertonemas periféricos, de no ser porque la gran mayoría de sus realizaciones provienen de un solo hablante, lo que implica que esta configuración no necesariamente está asociada con el comportamiento prosódico de la red. En el caso de la red laboral, los tonemas internos periféricos son aquellos que poseen un acento tonal H+L\* seguido de un ascenso hasta una juntura alta (H-) o media (!H-). Esta red presenta un número de producciones del intertonema abstracto L\* L- que contrasta con las pocas producciones de este en la red vecinal. Sin embargo, no puede tomarse como un índice sociolingüístico debido a que 7 de las 10 producciones de esta configuración son emitidas únicamente por dos hablantes que las realizan con voz laringizada (en 5 de los 7 casos).

En concordancia con la tendencia general y con la de cada red social, los intertonemas más frecuentes en informantes masculinos y femeninos son igualmente H\* H- y L+H\* H- (cuadro 4.6), cuya distribución de casos según el sexo de los informantes no es estadísticamente significativa. Al igual que en lo observado en el análisis de las redes sociales, el tonema interno abstracto predominante en hombres y mujeres es H\* H-, aunque las mujeres de la red vecinal representan la excepción a la tendencia mayoritaria, por el uso de la configuración L+H\* H-. En el nivel más concreto, el ejemplar más frecuente del esquema abstracto H\* H- es H\* H-, seguido por H\* =- en ambos sexos. En el caso de L+H\* H-, los hombres muestran preferencia por la emisión de L+<sub>i</sub>H\* HH-; al tiempo que las mujeres tienen un número de casos similar entre dicha configuración y L+H HH-, que cuenta con menor elevación en el acento nuclear.

Cuadro 4.6. Tonemas internos abstractos no oxítonos y la variable sexo. N= 194

Tonemas internos abstractos	Hombres			Mujeres		
	Red vecinal	Red laboral	Total	Red vecinal	Red laboral	Total
H* H-	20	21	41	9	27	36
H* !H-	4	3	7	2	1	3
H* L-	4	5	9	0	4	4
L+H* H-	11	13	24	15	7	22
L+H* L-	3	0	3	2	3	5
L* H-	1	4	5	5	4	9
L* !H-	0	1	1	1	4	5
L* L-	2	6	8	0	4	4
H+L* H-	0	1	1	5	0	5
H+L* !H-	0	0	0	1	1	2
Total	45	54	99	40	55	95

$p = 0.134$  (Estadístico Exacto de Fisher únicamente con la variable sexo)

$p = 0.003$  (Estadístico Exacto de Fisher al combinar variables sexo y red social)

Las demás configuraciones nucleares internas son menos frecuentes que las mencionadas anteriormente y presentan mayor variación entre ambos sexos. Los hombres, por una parte, favorecen levemente al empleo de intertonemas abstractos no oxítonos que implican un descenso melódico desde un acento nuclear H\* hasta el tono de juntura (H\* !H- con 7.07 % de frecuencia y H\* L- con 9.09 %), y a la sucesión de tonos bajos (L\* L- con una frecuencia de 8.08 %). Las mujeres, en cambio, continúan demostrando una tendencia al uso de incrementos tonales hasta una juntura alta (L\* H- con una frecuencia de uso del 9.47 %), mientras que el resto de las configuraciones nucleares internas supera el 5.23 %. En referencia a los intertonemas abstractos con acentos nucleares descendentes, es preciso señalar que no se toman en cuenta como una diferencia sustancial entre ambas redes, debido a que 3 de las ocurrencias de la configuración H+L\* H- y la única de H+L\* !H- documentadas en esta red pertenecen a una sola hablante (Rs1M3). Aunque esta hablante es la que comparte más vínculos con los integrantes de la red social, no se observa que este rasgo sea compartido por la mayoría de los miembros de la red (únicamente es utilizado por las otras dos informantes femeninas en una ocasión), por lo que se considera que el empleo de estas configuraciones no necesariamente representan un rasgo característico de las mujeres de la red vecinal.



En cuanto a la relación entre los tonemas internos y la duración en cada red social, se puede afirmar que no existe relación significativa entre las medias de la duración de cada sílaba con respecto a la red social y a los tonemas internos abstractos registrados (ANOVA = 0.650 para núcleos y ANOVA = 0.643 para lindes en la red vecinal; ANOVA = 0.483 para núcleos y ANOVA = 0.230 para lindes en la red laboral). No obstante, estos datos permiten presentar algunas anotaciones relacionadas con lo observado en cada red social. De acuerdo con los datos que se presentan en el cuadro 4.7, las medias de duración de la red laboral son mayores a las de la red vecinal, tanto en las sílabas nucleares como en las del linde de las frases entonativas intermedias (ip). Sin embargo, al observar las medias de duración de estas sílabas en cada tonema interno, se percibe el siguiente patrón: la red laboral produce sílabas nucleares más largas, excepto cuando el acento nuclear es descendente (H+L\*) y sílabas finales más prologadas, con excepción de aquellos casos en los que los acentos nucleares son bitonales, independientemente del movimiento tonal (L+H\* o H+L\*). Ahora bien, es preciso aclarar que las variaciones de medias entre ambas redes superan la diferencia perceptible de 30 ms (Toledo 1998, 2000) en muy pocos casos. Por otra parte, en ambas redes ocurre una sola emisión de la configuración nuclear interna H+L\* !H-, por lo que los valores registrados en las sílabas que conforman esta configuración podrían variar si se contara con más casos de este tonema interno.

Cuadro 4.7. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos en cada red social

Tonemas internos abstractos	Duración del núcleo		Duración del linde	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	171	199.9	236.52	245.58
H* !H-	167.5	201	265.83	265.75
H* L-	154.5	153.44	195	336.89
L+H* H-	184.65	203.1	242.31	215.3
L+H* L-	179.6	219.33	232	180.33
L* H-	187.83	189.13	243.83	254.88
L* !H-	186	194.2	190	192
L* L-	161	169.7	149.5	211.6
H+L* H-	225	135	296.8	151
H+L* !H-	144	126	338	309
Total	178.65	192.13	240.8	241.33

p = 0.111 (ANOVA del núcleo al nivel intergrupar);

p = 0.967 (ANOVA del linde al nivel intergrupar)

Al analizar estadísticamente los datos de duración relativa en los tonemas internos de ambas redes sociales, se comprobó que el recuento de casos de mostrado en el cuadro 4.8 es significativo. Lo primero que salta a la vista es que los tonemas internos con acentos bitonales descendentes (H+L\* H- y H+L\* !H-) son siempre producidos con mayor duración en la sílaba linde de ip por ambas redes. También se observa que, en ambas redes, las sílabas finales de ip son más frecuentemente producidas con una duración mayor que la sílaba nuclear. Esto se relaciona con lo expuesto anteriormente, en cuanto a que las sílabas del linde de ip tienen una media de duración mayor que las del núcleo, lo cual va de la mano con la mayor ocurrencia de este patrón y lo posiciona como el prototípico en las configuraciones nucleares de las frases intermedias (ip). Por lo tanto, debido a la uniformidad en los datos de ambas redes, no se puede considerar que algunos patrones de duración en los intertonemas no oxítonos representen un índice sociolingüístico.

Cuadro 4.8. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos en cada red social. N=194

Tonemas internos abstractos	Núcleo > Linde				Núcleo < Linde			
	Red vecinal		Red laboral		Red vecinal		Red laboral	
	Casos	%	Casos	%	Casos	%	Casos	%
H* H-	8	32.00	18	47.37	21	35.00	30	42.25
H* !H-	3	12.00	1	2.63	3	5.00	3	4.23
H* L-	1	4.00	-	-	3	5.00	9	12.68
L+H* H-	7	28.00	9	23.68	19	31.67	11	15.49
L+H* L-	1	4.00	2	5.26	4	6.67	1	1.41
L* H-	3	12.00	1	2.63	3	5	7	9.86
L* !H-	-	-	2	5.26	1	1.67	3	4.23
L* L-	2	8.00	5	13.16	-	-	5	7.04
H+L* H-	-	-	-	-	5	8.33	1	1.41
H+L* !H-	-	-	-	-	1	1.67	1	1.41
Total	25	100	38	100	60	100	71	100

p = 0.043 (Estadístico Exacto de Fisher)

En relación con los esquemas abstractos en particular, en todos los casos se observa que siguen las mismas tendencias, lo que dificulta identificar un patrón de comportamiento que pueda representar una clara diferencia. Adicionalmente, entre aquellos elementos en los que existe una diferencia superior a dos casos, no se identifica un patrón diferenciado para cada red que permita establecer alguna relación entre las configuraciones nucleares internas y la duración relativa de las sílabas que las constituyen.

En cuanto a la variable sexo, los datos mostrados en los cuadros 4.9 y 4.10 (aunque no son estadísticamente significativos a nivel intergrupar, ANOVA del núcleo con  $p = 0.225$  y del linde con  $p = 0.780$ ), reflejan que los hablantes de ambos sexos producen sílabas lindes de ip más largas que las nucleares, lo que concuerda con todo lo descrito hasta este punto. No obstante, las mujeres tienden a emitir ambas sílabas del tonema interno con una media de duración general mayor que la de los hombres, pero esta diferencia es inferior a los 30 ms, lo que la hace imperceptible. Ahora bien, el hecho de que los hombres no presenten ningún caso de la configuración nuclear interna H+L\* !H- puede dar la impresión de que los valores generales de duración podrían cambiar si la situación fuese diferente; sin embargo, tras determinar la media general de las mujeres en ambas sílabas, excluyendo el tonema interno H+L\* !H-, se encontró que los valores medios totales de las mujeres no varían más de 1.8 ms

en relación con las medias totales previas (las medias totales sin el intertonema H+L\* !H- fueron de 195.65 ms en la sílaba nuclear y 241.19 ms en la del linde). Por lo tanto, la interpretación de los datos se puede mantener sin modificaciones. Por otra parte, si bien son las mujeres las que presentan medias de duración general más largas, no ocurre lo mismo al observar los datos en relación con variables específicas. Los hombres presentan mayor media de duración de la sílaba nuclear en los tonemas internos H\* !H- y L+H\* H-; aunque la diferencia en el segundo con respecto a la media de las mujeres es bastante reducida (no supera los 2 ms). En cuanto a la duración de la sílaba en el linde de ip, los hombres poseen mayor media en los casos en los que el acento nuclear es ascendente L+H\* y con la configuración descendente L\* L-.

Cuadro 4.9. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de los hombres de ambas redes sociales

Tonemas internos abstractos	Duración del núcleo		Duración del linde	
	Hombres		Hombres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	171.15	188.05	244.7	237.95
H* !H-	183.75	215.67	230.5	270.33
H* L-	154.5	148.6	195	345.2
L+H* H-	184.18	202.31	251.45	222.85
L+H* L-	185.33	196.5	239.67	235.5
L* H-	159	188	158	134
L* !H-	-	-	-	-
L* L-	161	170.67	149.5	249.83
H+L* H-	-	135	-	151
H+L* !H-	-	-	-	-
Total	174.2	187.07	234.18	243.65

p = 0.609 (ANOVA de duración del núcleo).

p = 0.788 (ANOVA de duración del linde).

Cuadro 4.10. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de las mujeres de ambas redes sociales

Tonemas internos abstractos	Duración del núcleo		Duración del linde	
	Mujeres		Mujeres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	170.67	209.11	218.33	251.52
H* !H-	135	157	336.5	252
H* L-	-	159.5	-	326.5
L+H* H-	185	204.57	235.6	201.29
L+H* L-	171	219.33	220.5	180.33
L* H-	193.6	181.75	261	274.25
L* !H-	186	195.75	190	206.5
L* L-	-	168.25	-	154.25
H+L* H-	225	-	296.8	-
H+L* !H-	144	126	338	309
Total	183.65	197.09	248.25	239.05

p = 0.663 (ANOVA de duración del núcleo).

p = 0.06 (ANOVA de duración del linde).

El recuento de casos de la duración relativa de tonemas internos de acuerdo con el sexo de los informantes (mostrado en los cuadros 4.11 y 4.12) presenta una distribución no significativa cuando se toma en cuenta únicamente el sexo de los informantes ( $p = 0.247$  según el estadístico Exacto de Fisher), pero dicha distribución sí resulta significativa cuando se conjugan las variables del sexo y la red social ( $p = 0.017$  según estadístico Exacto de Fisher). De acuerdo con los datos, tanto hombres como mujeres tienden a producir la sílaba linde de ip con mayor longitud que la nuclear. Otro hecho común para ambos sexos es que los tonemas internos más frecuentes (H\* H- y L+H\* H-) presentan la mayoría de sus realizaciones con una duración más prologada en la sílaba final de ip (como se puede apreciar en las figuras 4.15 y 4.16).

Cuadro 4.11. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos de los hombres de ambas redes sociales. N=99

Tonemas internos abstractos	Núcleo > Linde		Núcleo < Linde	
	Hombres		Hombres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	5	8	15	13
H* !H-	3	1	1	2
H* L-	1	0	3	5
L+H* H-	3	5	8	8
L+H* L-	1	0	2	0
L* H-	1	1	0	3
L* !H-	2	1	0	0
L* L-	0	2	0	4
H+L* H-	0	0	0	1
H+L* !H-	0	0	0	0
Total	16	18	29	36

p = 0.533 (Estadístico Exacto de Fisher)

Cuadro 4.12. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas internos no oxítonos de las mujeres de ambas redes sociales. N=95

Tonemas internos abstractos	Núcleo > Linde		Núcleo < Linde	
	Mujeres		Mujeres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	3	10	6	17
H* !H-	0	0	2	1
H* L-	4	0	11	4
L+H* H-	0	4	0	3
L+H* L-	0	2	2	1
L* H-	2	0	3	4
L* !H-	0	1	1	3
L* L-	0	3	0	1
H+L* H-	0	0	5	0
H+L* !H-	0	0	1	1
Total	9	20	31	35

p = 0.002 (Estadístico Exacto de Fisher)

Figura 4.15. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs2M2-8: [(*Ahorita tienes programas bien divertidos*)<sub>ip</sub> (y la verdad es que sí me hubiera visto ahí perfecto)<sub>ip</sub>]<sub>IP</sub>

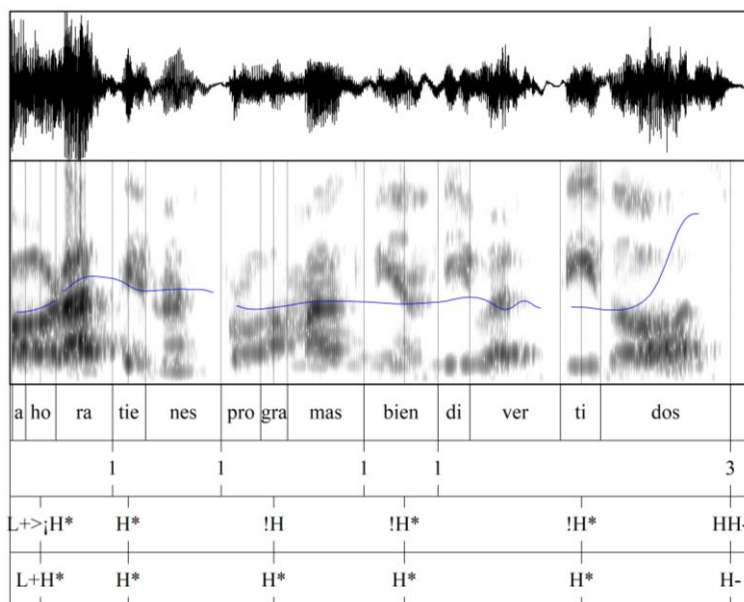
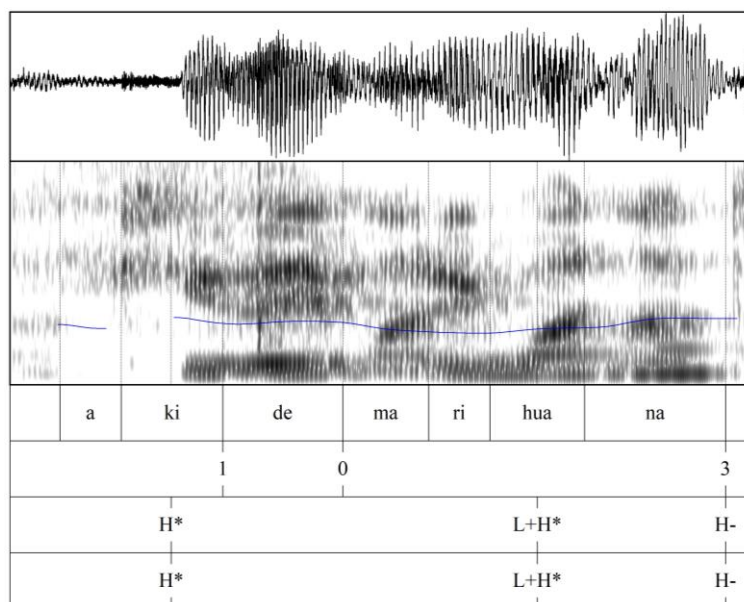


Figura 4.16. Oscilograma, espectrograma y F0 de la primera frase entonativa intermedia del enunciado Rs1H2-26: [(*Aquí de marihuana*)<sub>ip</sub> (*saldría esto*)<sub>ip</sub> (*quince pesos*)<sub>ip</sub>]<sub>IP</sub>



Entre los datos de los demás tonemas internos, la distribución no es significativa en términos estadísticos. En el caso de las mujeres, las configuraciones con acento nuclear L\* seguidas por junturas medias (!H-) o altas (H-) tienden a ser emitidas con sílabas finales más largas que las nucleares. Adicionalmente, las configuraciones que implican un descenso desde un acento nuclear alto hasta una juntura media o baja (H\* !H- y H\* L-) son siempre producidas con sílabas finales más largas por parte de las mujeres; mientras que los hombres presentan una tendencia parecida únicamente en el segundo de estos tonemas internos (H\* L-). Los demás intertonemas presentan diferencias de frecuencias muy reducidas, de tal manera que no se identifica algún patrón entonativo distintivo tanto de hombres como de mujeres.

#### 4.3.2. Tonemas internos (intertonemas) oxítonos

Al detallar las ocurrencias de cada tonema interno oxítono en ambas redes sociales (cuadro 4.13), se percibe, a primera vista, que la red laboral emplea principalmente las dos configuraciones más frecuentes del corpus (H\* H- y L+H\* H-), especialmente en sus realizaciones concretas H\* H-, L+H\* !HH- y L+H\* HH-, las dos primeras con la misma frecuencia de uso. La red vecinal sigue la misma tendencia (con una producción predominante del tonema interno concreto L+;H\* HH-), pero destaca principalmente por ser la única en la que se registran casos de la configuración L+H\* !H-. Por otro lado, las diferencias entre las frecuencias de las configuraciones registradas en cada red son bastante reducidas en algunas instancias, lo que explica el hecho de que el recuento de casos presentado en el cuadro no sea estadísticamente significativo.

Cuadro 4.13. Tonemas internos oxítonos en cada red social. N=75

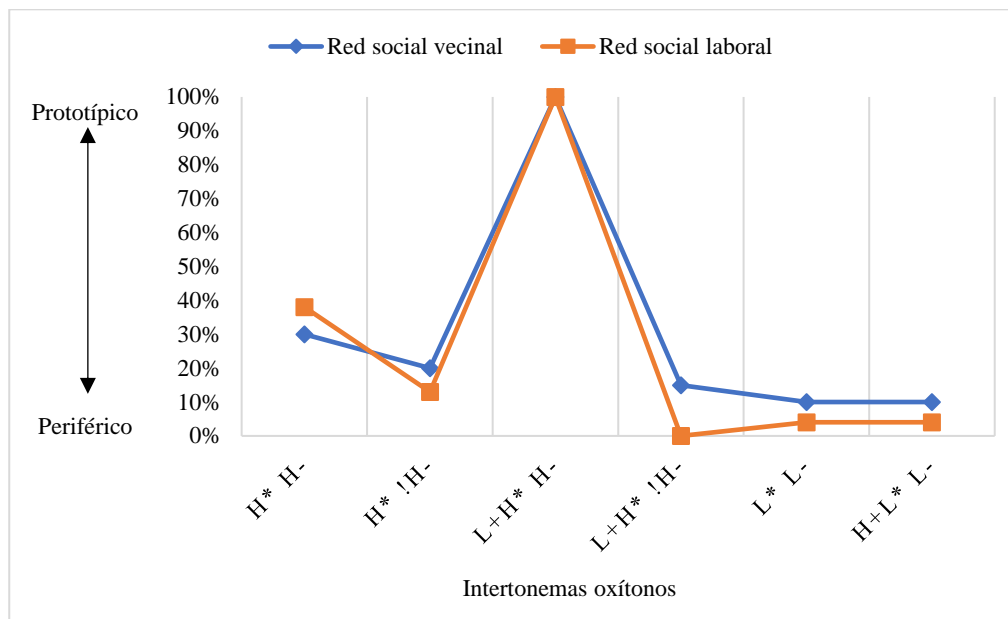


Tonemas internos abstractos	Red vecinal		Red laboral		Total	
	Casos	%	Casos	%	Casos	%
H* H-	6	16.22	9	23.68	15	20.00
H* !H-	4	10.81	3	7.89	7	9.33
L+H* H-	20	54.05	24	63.16	44	58.67
L+H* !H-	3	8.11	0	0	3	4.00
L* L-	2	5.41	1	2.63	3	4.00
H+L* L-	2	5.41	1	2.63	3	4.00
Total	37	100	38	100	75	100

p = 0.501 (Estadístico Exacto de Fisher)

Al analizar el grado de centralidad de las configuraciones en la categoría cognitiva de intertonemas oxítonos, se observa que no hay mucha variación entre ellas, como se puede ver en la figura 4.17. En este gráfico se perciben solo dos diferencias considerablemente llamativas. La primera es la distancia que existe entre los esquemas abstractos H\* H- y H\* !H- en las estructuras cognitivas de las dos redes. En este sentido, el grado de centralidad (o prototipicidad) del primero con respecto al segundo es más marcado en la red laboral, pues el H\* H- está considerablemente más cerca del prototipo y H\* !H- más cerca de la periferia. Esto permite ver la importancia que tienen las junturas H- en la construcción categorial de esta red, pues los esquemas que comparten este tono de juntura poseen un mayor grado de prototipicidad, mientras que los demás son ubicados en la periferia. En la red vecinal no se observa tanta distancia entre los intertonemas periféricos, pues se encuentran en un rango que va del 30 % al 10 % de las emisiones. Sin embargo, se puede ver que la centralidad guarda cierta relación con el tono de juntura, pues los esquemas entonativos se alejan de la periferia en la medida en que tienen una juntura más alta. Por otra parte, la segunda diferencia es que el esquema L+H\* !H- pudiese encontrarse fuera de la categoría de intertonemas oxítonos de la red laboral. El hecho de que no se hayan documentado ocurrencias de esta configuración tonal interna hace pensar que puede existir la posibilidad de que (con base en la muestra de datos) no exista dentro de los límites de esta categoría. Sin embargo, este planteamiento no debe ser considerado como algo más que una hipótesis en este primer acercamiento desde la perspectiva sociocognitiva.

Figura 4.17. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente los intertonemas oxítonos en ambas redes sociales



El análisis de los tonemas internos oxítonos con respecto a la variable sexo (cuya distribución de casos no es significativa), mostró que los hombres tienden a favorecer el uso casi exclusivo de los tonemas internos oxítonos abstractos prototípicos de las ip (H\* H- y L+H\* H-), especialmente la configuración L+H\* H- (cuadro 4.14). Las realizaciones concretas correspondientes a estas dos contornos son H\* H-, como segunda opción más frecuente, y L+<sub>i</sub>H\* HH-, como el elemento más central. Esto deja constancia, una vez más, de la relevancia de los ascensos prominentes en los tonemas internos de las ip. Las mujeres, en cambio, incluyen en su repertorio todas las configuraciones nucleares internas registradas, aunque se concentran más en el empleo del tonema interno abstracto L+H\* H-. En relación con las configuraciones L+H\* !H- y L\* L- en las mujeres, cabe señalar que estas son producidas únicamente por las dos hablantes de la red vecinal de mayor edad (50 y 57 años) que también son las únicas cuyas actividades laborales están directamente relacionadas con el mercado de la comunidad (una de ellas vende productos fuera del mercado los fines de semana y la otra acude al mercado diariamente a comprar los insumos necesarios para su puesto de comida). La otra mujer de la red vecinal, cuya edad es de 33 años y acude muy poco a dicho mercado, no genera ningún caso de L+H\* !H- ni de L\* L-. Por tal motivo, la producción de estas configuraciones se puede vincular con dos motivos posibles: al hecho de

que pertenecen a un grupo etario femenino de esta comunidad, mayor a los 50 años o a que tengan relación frecuente y directa con el mercado de la comunidad. Por otra parte, se observa que la tendencia hacia el mayor o menor uso de los intertonemas documentados está más relacionada con la variable sexo que con la de redes sociales, pues la tendencia femenina a presentar mayor cantidad de patrones periféricos contrasta con la reducida preferencia masculina hacia estos contornos entonativos en los lindes tónicos de ip. Por lo tanto, en términos de frecuencia de uso, se puede plantear que las mujeres favorecen el empleo de una mayor variedad de configuraciones melódicas al final de las ip que culminan con sílabas tónicas.

Cuadro 4.14. Tonemas internos oxítonos en variables sexo y red social. N=75

Tonemas internos abstractos	Hombres			Mujeres		
	Red vecinal	Red laboral	Total	Red vecinal	Red laboral	Total
H* H-	3	6	9	3	3	6
H* !H-	1	1	2	3	2	5
L+H* H-	11	14	25	9	10	19
L+H* !H-	0	0	0	3	0	3
L* L-	0	1	1	2	0	2
H+L* L-	0	0	0	2	1	3
Total	15	22	37	22	16	38

p = 0.106 (Estadístico Exacto de Fisher únicamente con la variable sexo)

p = 0.483 (Estadístico Exacto de Fisher al combinar variables sexo y red social)

El análisis de las medias de duración de cada red social en los tonemas internos oxítonos permite ver una diferencia estadísticamente significativa. Como reflejan los datos del cuadro 4.15, las medias de duración de las sílabas nucleares de ip producidas por la red laboral son, en términos perceptibles, mucho más largas que las de la red vecinal, pues todos los tonemas internos oxítonos producidos por esta red superan los 300 ms (exceptuando L+H\* !H-, que no cuenta con casos en la red laboral). Se puede considerar que este rasgo en la duración de la sílaba tónica final de ip (como se muestra en la primera ip de la figura 4.18) es característico de esta red, debido a que está ampliamente extendido entre todos sus miembros, pues la gran mayoría de las producciones de esta sílaba superan los 300 ms (excepto en los hablantes Rs2H1 y Rs2M2). Debido a la elevada frecuencia con la que se registra este rasgo

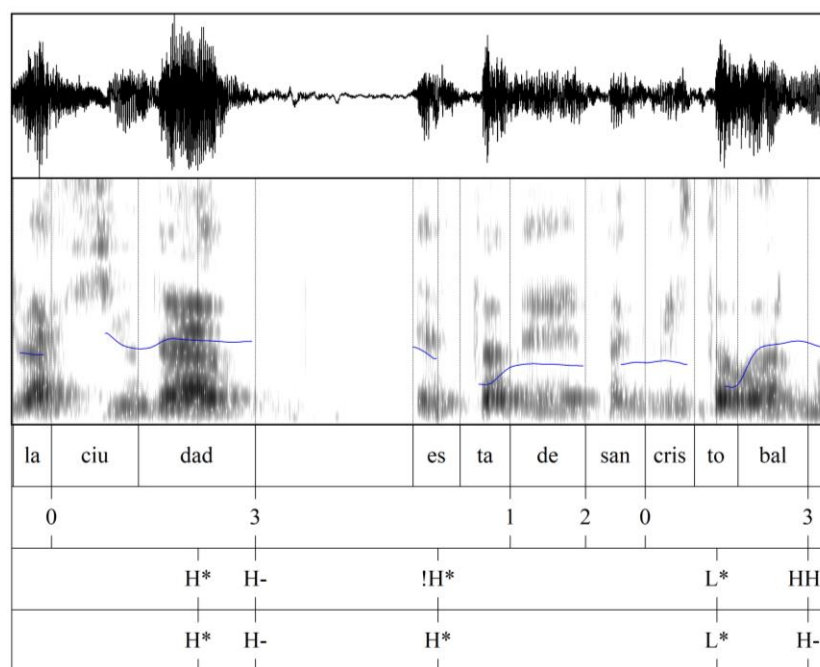
prosódico, se considera que es prototípico en la estructura cognitiva de los miembros de la red laboral (al menos en la de la mayoría).

Cuadro 4.15. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos oxítonos en cada red social

Tonemas internos abstractos	Red vecinal	Red laboral
H* H-	253.33	303.56
H* !H-	277.5	306.67
L+H* H-	243.65	302.42
L+H* !H-	161	-
L* L-	164	392
H+L* L-	361.5	336
Total	244.24	306.26

p = 0.007 (ANOVA)

Figura 4.18. Oscilograma, espectrograma y F0 de las dos primeras frases entonativas intermedias del enunciado Rs2H3-32: [(*La ciudad*)<sub>ip</sub> (*esta de San Cristóbal*)<sub>ip</sub> (*me gustó mucho*)<sub>ip</sub>]<sub>IP</sub><sup>25</sup>



<sup>25</sup> Se transcribió el acento tonal en *Cristóbal* como bajo (L\*) debido al efecto de la laringización en esta sílaba.

Por otra parte, solo el intertonema abstracto oxítono H+L\* L- producido por la red vecinal supera los 300 ms. Cabe señalar que en la red vecinal se registraron medias superiores a los 300 ms, en algunas configuraciones concretas, pero esto no es un rasgo frecuente en los hablantes de dicha red, pues se observó solamente en la informante Rs1M2 y en una de las configuraciones producidas por el hablante Rs1H2. Excluyendo estos casos específicos, las configuraciones internas oxítonas de ip emitidas por la red vecinal presentan medias bastante irregulares que van desde los 151 ms hasta los 254 ms. Por tal motivo, la prolongación de la sílaba tónica final de ip es una de las características distintivas más notorias en la producción de enunciados, pues la diferencia en los datos de ambas redes es marcada por la superioridad de las medias de la red social laboral, las cuales son relativamente uniformes en cuanto a la base mínima de duración de 300 ms.

Al sustituir la variable de la red social e incluir la del sexo en el análisis estadístico, en cambio, no se obtuvo un resultado significativo (significación de ANOVA con  $p = 0.142$  en el caso de los hombres y con  $p = 0.064$  en el de las mujeres). Sin embargo, se puede reconocer un patrón en aquellos tonemas internos oxítonos comparables, es decir, los que cuentan con ocurrencias por parte de los hablantes de ambos sexos (lo que descarta las configuraciones L+H\* !H- y H+L\* L-, únicamente producidas por mujeres). Como se puede apreciar en el cuadro 4.16, los hombres y mujeres de la red laboral presentan una mayor tendencia a producir sílabas tónicas finales de ip con medias de duración superiores a los 300 ms. Esta tendencia únicamente es observable en la red social vecinal en la media de la configuración H\* !H- producida por hombres y en la de H+L\* L- emitida por mujeres.

Cuadro 4.16. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas internos no oxítonos de acuerdo con el sexo de los informantes y con la red social

Tonemas internos abstractos	Hombres		Mujeres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H-	237	332.17	269.67	246.33
H* !H-	345	290	255	315
L+H* H-	223	276.14	268.89	339.2
L+H* !H-	0	0	161	0
L* L-	0	392	164	0
H+L* L-	0	0	361.5	336
Total	233.93	297.32	251.27	318.56

p = 0.699 (ANOVA al nivel intergrupar)

En cuanto a los miembros de la red laboral, los hombres presentan medias superiores a los 300 ms en los tonemas internos oxítonos con núcleo monotonal (aunque es preciso señalar que la media registrada para la configuración L+H- H- es considerablemente mayor en los hombres de la red laboral que en los de la vecinal, aunque no supere los 300 ms), mientras que las mujeres superan este valor de duración en todos los casos, excepto H\* H- y H+L\* L-. Por lo tanto, a pesar de que la duración promedio de los hombres y mujeres de la red laboral no siempre supera los 300 ms, se puede afirmar que esta tendencia tiene mayor definición en los hablantes de esta red.

#### 4.3.3. Funciones sintáctico-discursivas

Las afirmaciones realizadas en esta sección, relacionadas con las diferencias identificadas en la prosodia de las funciones sintáctico-discursivas de ambas redes sociales, no se pueden considerar categóricas debido a que la producción de elementos que desempeñan estas funciones es reducida (su escasez se debe a que, al tomar los datos de habla a través de entrevistas sociolingüísticas, no se controló la producción de enunciados con estas características, sino que se extrajeron los datos directamente de la entrevista semidirigida); sin embargo, no existe un impedimento para que se consideren como variables en una primera aproximación al análisis de los rasgos prosódicos de estos elementos en ambas redes sociales.

El foco estrecho se realiza principalmente con los intertonemas abstractos H\* H- y L+H\* H-, específicamente, con las realizaciones concretas !H H- y H\* H- para el primero de

ellos y L+<sub>i</sub>H\* HH- para el segundo, todos con prácticamente la misma frecuencia en ambas redes. La configuración H\* H- es más empleada por la red social vecinal para transmitir este significado pragmático, mientras que la red laboral utiliza ambas configuraciones con la misma frecuencia (20 casos). Debido a que la mayoría de los informantes de la red laboral usan ambas configuraciones con frecuencias muy cercanas, no se puede ver que alguna de ellas tenga una posición central en la categoría de focalización estrecha al final de una ip. De igual modo, debido a que ambas redes emplean la configuración H\* H- con casi la misma frecuencia, no se puede decir que el uso de este tonema interno para marcar focos estrechos sea un índice sociolingüístico propio de la red vecinal. En relación con la expresividad local, específicamente en el final de las ip, aunque los casos no son abundantes en cada red, se observa que 6 de las 9 emisiones expresivas de la red laboral ocurren con H\* H-, mientras que la red vecinal concentra 4 de sus 5 producciones en el intertonema L+H\* H-. Por otra parte, la red vecinal emplea mayoritariamente esta misma configuración en la sección nuclear de las cláusulas prótasis, mientras que la red laboral utiliza el tonema interno L+H\* H-.

El resto de las funciones sintáctico-discursivas (focos estrechos con expresividad, cláusulas preyuxtapuestas, cláusulas con sentido explicativo y los distintos elementos extraoracionales) cuentan con muy pocos casos por parte de cada red como para poder realizar un análisis sociolingüístico en términos cuantitativos. Sin embargo, los datos obtenidos en ambas redes (cuadros 4.17 y 4.18) permiten llevar a cabo una aproximación cualitativa de la distribución del uso de los distintos intertonemas abstractos en relación con las funciones sintáctico-discursivas.

Cuadro 4.17. Intertonemas abstractos en las distintas funciones sintáctico-discursivas en la red vecinal

Intertonemas abstractos	Res social 1						
	Prótasis	Preyuxtapuesta	Explicación	Tematización	Iniciador	Orientador locativo	Orientador temporal
H* H-	2	2	3	4	0	0	6
H* !H-	1	0	0	0	0	0	0
H* L-	0	1	0	0	0	0	0
L+H* H-	5	1	2	1	1	1	6
L+H* !H-	0	0	0	0	1	0	0
L+H* L-	0	0	0	0	0	0	0
L* H-	1	0	0	0	0	0	1
L* !H-	0	0	0	0	0	0	0
L* L-	0	1	0	0	1	0	0
H+L* H-	0	0	0	1	0	0	1
H+L* !H-	0	0	0	0	0	0	0
H+L* L-	0	0	0	0	0	1	0
Total	9	5	5	6	3	2	14

Cuadro 4.18. Intertonemas abstractos en las distintas funciones sintáctico-discursivas en la red laboral

Intertonema abstractos	Red laboral						
	Prótasis	Preyuxtapuesta	Explicación	Tematización	Iniciador	Orientador locativo	Orientador temporal
H* H-	3	2	2	0	1	1	4
H* !H-	0	0	0	0	0	0	0
H* L-	0	0	0	0	0	0	0
L+H* H-	4	3	3	0	2	0	0
L+H* !H-	0	0	0	0	0	0	0
L+H* L-	0	0	0	0	0	0	0
L* H-	0	0	3	0	0	0	0
L* !H-	0	0	0	0	0	0	0
L* L-	1	0	0	0	0	0	0
H+L* H-	0	0	0	0	0	0	0
H+L* !H-	0	0	0	0	0	0	0
H+L* L-	0	0	0	0	0	0	0
Total	8	5	8	0	3	1	4

Como se puede observar en los cuadros 4.17 y 4.18, ambas redes sociales prefieren el uso del intertonema abstracto L+H\* H- al final de cláusulas que desempeñan la función de



prótesis; entre cuyas realizaciones concretas, la red vecinal se decanta por la configuración L+H\* HH- y la red laboral emplea tanto dicha configuración como su variante L+;H\* HH-, con un ascenso más pronunciado en el acento nuclear. Tanto en las ip que contienen cláusulas preyuxtapuestas como en las explicativas, la red vecinal tiende, principalmente, al empleo del intertonema concreto H H- y la laboral por el L+H\* H-, aunque esta última red también emplea las configuraciones H\* H- y L\* H- en las ip explicativas, lo que implica la ausencia de una configuración prototípica en frases entonativas con esta función.

En cuanto a la función tematizadora, no se cuenta con casos en la red laboral, por lo que no se puede establecer un contraste con las producciones del intertonema H\* H- predominante en la red vecinal en este caso. Una situación similar tiene lugar en las ip con función de orientación locativa, ya que se cuenta con solo un caso en la red laboral (con el intertonema H\* H-) y 2 en la vecinal (que ocurren con las configuraciones L+H\* H- y H+L\* L-). En relación con los elementos extraoracionales iniciadores, mientras la red laboral presenta una sutil mayoría de producciones con la configuración abstracta L+H\* H-, la red vecinal no tiene un elemento central claro, ya que emplea ascensos (L+H\* H-), circunflexiones no descendentes (L+H\* !H-) y sucesión de tonos bajos (L\* L-). Finalmente, algo parecido ocurre en las ip con función de orientación temporal, pues la red vecinal favorece, con igualdad de casos, el uso de dos configuraciones ascendentes (H\* H- y L+H\* H-) y la red laboral solo hace uso de H\* H-.

Si bien los datos en relación con algunas funciones sintáctico-discursivas pueden ser escasos, el hecho de que la gran mayoría de ellas ocurran principalmente con los intertonemas abstractos H\* H- y L+H\* H- en ambas redes hace pensar que la función que prima al momento de producir un tonema interno es la de marcar el sentido de continuidad de la enunciación. Expuesto en términos de los códigos biológicos de Gussenhoven (2004), el código de producción (relacionado con la elevación melódica que expresa la no conclusión del enunciado) es priorizado ante el código de esfuerzo (que tiene que ver con movimientos entonativos asociados con focos estrechos o expresividad). Por lo tanto, teniendo en cuenta los pocos casos recolectados en relación con estas funciones sintáctico-discursivas, resulta difícil identificar un patrón de producción claro en alguna red social, que permita establecer una tendencia suficientemente definida como para reconocer índices sociolingüísticos, más allá de los ya señalados en el estudio de los casos generales.

#### 4.3.4. Resumen de resultados de los tonemas internos de ip

De acuerdo con los datos de producción de enunciados, se sugiere que las configuraciones tonales de las frases entonativas intermedias pueden aportar un valor de indexación sociolingüística (Eckert, 2018). Lo que promueve esta idea es la diferencia entre la frecuencia de uso de determinados rasgos prosódicos de una y otra red, es decir, el grado de prototipicidad o centralidad cognitiva. Tómese como ejemplo el intertonema no oxítono abstracto L+H\* H- que, a pesar de ser el segundo más frecuente en las producciones de ambas redes, es mucho más utilizado por la red vecinal (al punto de estar prácticamente en el centro de la categoría cognitiva, junto con H\* H-) que por la laboral, en la que se observa mucho más orientado hacia la periferia cognitiva debido a la gran distancia que lo separa de la configuración más productiva (H\* H-).

Por otra parte, al momento de constituir un índice sociolingüístico, los distintos elementos prosódicos no necesariamente deben coincidir. En los intertonemas no oxítonos no se percibe que la duración sea un factor determinante en la estructura del significado social. Sin embargo, lo contrario ocurre en el caso de los tonemas internos oxítonos, en los que el parámetro de duración marca una diferencia clara entre ambas redes, independientemente de la configuración entonativa en el intertonema. En este caso, la tonicidad de la última sílaba de la frase entonativa intermedia (ip) es la que da pie a que se produzca con una mayor duración por parte de la red laboral. Dicha influencia es mucho más notoria en las producciones de la red laboral que en las de la vecinal, en la que no se registró un alargamiento tan prominente. Esto lleva a considerar la idea de que la duración y la entonación pueden operar de forma independiente en la generación de índices sociolingüísticos.

#### 4.4. Tonemas: configuraciones nucleares de frase entonativa mayor (IP)

El análisis del núcleo de las frases entonativas mayores (IP) revela la existencia de otros elementos prosódicos (configuraciones nucleares no descendentes y duración del tonema, por ejemplo) con un probable significado social. A continuación, se expone el análisis sociolingüístico de los datos relacionados con los tonemas en la totalidad del corpus, de los

tonemas no descendentes y, finalmente, de aquellos que poseen un acento nuclear con el pico tonal prealineado.

#### 4.4.1. Tonemas de IP en conjunto con las variables duración, red social y sexo

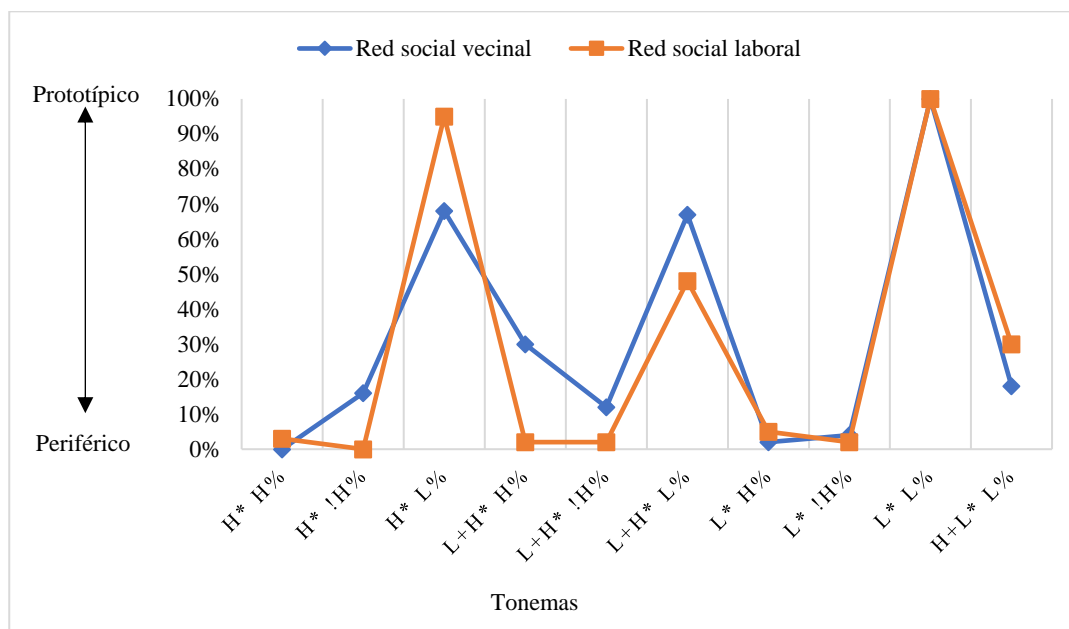
En términos de la centralidad de los esquemas entonativos que integran la categoría de tonemas de IP, en los datos del cuadro 4.19 se observa que la configuración prototípica (más frecuente) de la red vecinal es L\* L%; mientras que la red laboral no parece tener un elemento prototípico claro, ya que las configuraciones descendentes L\* L% y H\* L% se producen con una frecuencia bastante cercana, aunque, en cualquier caso, superior a la de los demás tonemas, como refleja el gráfico en la figura 4.19. Por otra parte, el segundo lugar de centralidad en esta categoría es ocupado por el tonema circunflejo L+H\* L% en las emisiones de la red social laboral. Por su parte, la red vecinal, en esta oportunidad, presenta una frecuencia de casos casi iguales en la configuración circunfleja y en H\* L%.

Cuadro 4.19. Tonemas en cada red social. N=360

Tonemas abstractos	Red vecinal	Red laboral	Total
H* H%	0	2	2
H* !H%	9	0	9
H* L%	39	60	99
L+H* H%	17	1	18
L+H* !H%	7	1	8
L+H* L%	38	30	68
L* H%	1	3	4
L* !H%	2	1	3
L* L%	57	63	120
H+L* L%	10	19	29
Total	180	180	360

p = 0 (Estadístico Exacto de Fisher)

Figura 4.19. Grado de centralidad (prototipicidad) de los esquemas entonativos de la categoría cognitiva correspondiente los tonemas en ambas redes sociales



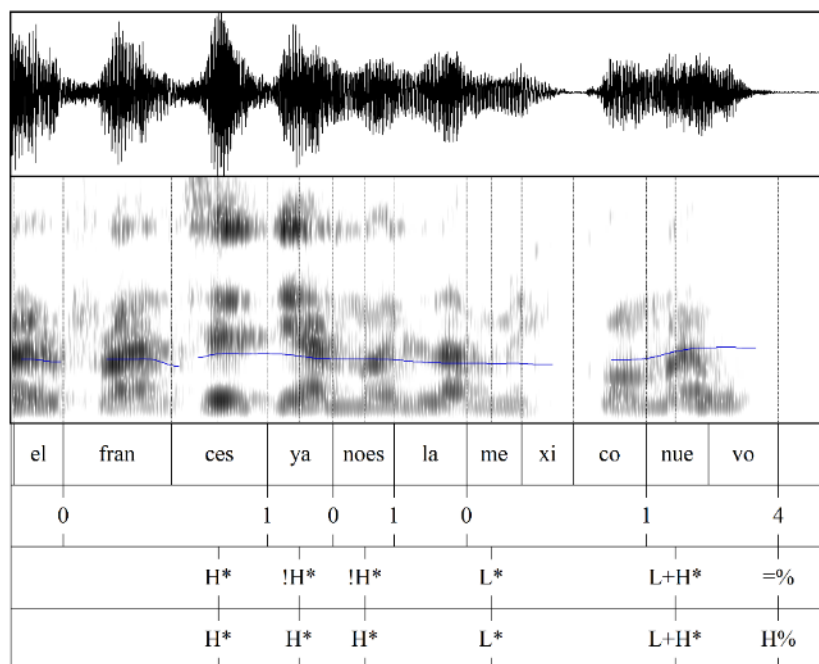
De acuerdo con la metodología basada en el uso, la distribución mostrada en la figura 4.19 implicaría que existe una diferencia en la estructura categorial de los tonemas de enunciados aseverativos de cada red social, en relación con los esquemas entonativos prototípicos. Si bien el esquema prototípico de ambas redes es una configuración descendente (L\* L%), en la red laboral el esquema H\* L% tiene casi el mismo grado de centralidad que L\* L%. Esto abre la posibilidad de que, con base en la frecuencia de uso, dos esquemas puedan compartir la posición central de una categoría (la del tonema, en este caso). En cualquier instancia, el rasgo constante en los dos esquemas y el que parece definir las cualidades entonativas del prototipo categorial es un descenso hasta una juntura en la base tonal del enunciado (L%), desde un tono bajo sostenido en la sílaba nuclear. Por otra parte, la red vecinal presenta un caso similar, aunque no con el esquema prototípico, sino con los dos más cercanos al prototipo. En esta red, dichos esquemas están a una distancia bastante similar del centro de la categoría, por lo que se puede decir que ambos ocupan el segundo lugar en términos de prototipicidad.

Los resultados obtenidos en este estudio coinciden parcialmente con los mostrados por Martín Butragueño (2004, 2011 y 2021), en los que el autor señala que el tonema circunflejo (L+H\* L%) es más común en el habla de hombres con bajo nivel de instrucción

y que el acento nuclear descendente ( $H+L^*$ ) ocurre más en personas con grado de instrucción alto (Martín Butragueño, 2011). Sin embargo, los datos mostrados en el cuadro 4.19 se orientan hacia una tendencia mayoritaria en ambas redes a producir los tonemas  $L^* L\%$ , principalmente, y  $H^* L\%$ , en segunda instancia. No obstante, no hay que perder de vista que los hablantes de la red vecinal (con nivel educativo no mayor a la preparatoria culminada) favorecen el uso del tonema circunflejo (especialmente cuando el ascenso nuclear es mayor a 3 st, como en los casos de las configuraciones concretas con acentos nucleares  $L+\underset{j}{H^*}$  y  $L+\underset{i}{H^*}$ ), y los de la red laboral (quienes culminaron estudios universitarios) emplean la configuración nuclear abstracta  $H+L^* L\%$ . Esto implica que, si bien ambas redes comparten tendencias generales, la frecuencia de uso de ciertas configuraciones entonativas nucleares en una u otra red puede guardar relación con las características sociales propias de cada una. Es decir, la pertenencia a una red (o grupo) social puede condicionar, en cierta medida, una mayor frecuencia en el uso de ciertos contornos entonativos ( $L+\underset{j}{H^*} L\%$  o  $H+L^* L\%$ , por ejemplo).

Si bien en los datos generales se observa un predominio de los tonemas descendentes, los hablantes de la red social laboral emplean casi exclusivamente estas configuraciones, ya que solo 8 de las 180 producciones de esta red no culminan con tonos de juntura  $L\%$ . En la red social vecinal, en cambio, aunque resulta notoria la tendencia principal por el uso de los tonemas descendentes, también se observa una tendencia secundaria (en el 20 % de los tonemas) por el uso de tonemas abstractos no descendentes, especialmente  $L+H^* H\%$  (figura 4.20). Cabe señalar que, en el nivel concreto, todos los casos de este tonema no descendente de la red vecinal (al igual que el único registrado en la red laboral) ocurren con un sostenimiento desde el tono más alto alcanzado en el ascenso en la sílaba nuclear. Dicho de otra forma, una vez alcanzado el pico tonal correspondiente con el acento  $L+H^*$ , la curva melódica no se sigue elevando en la sílaba final de IP, sino que se sostiene en la misma altura del pico tonal hasta concluir con un tono de juntura igualmente alto ( $=\%$ ).

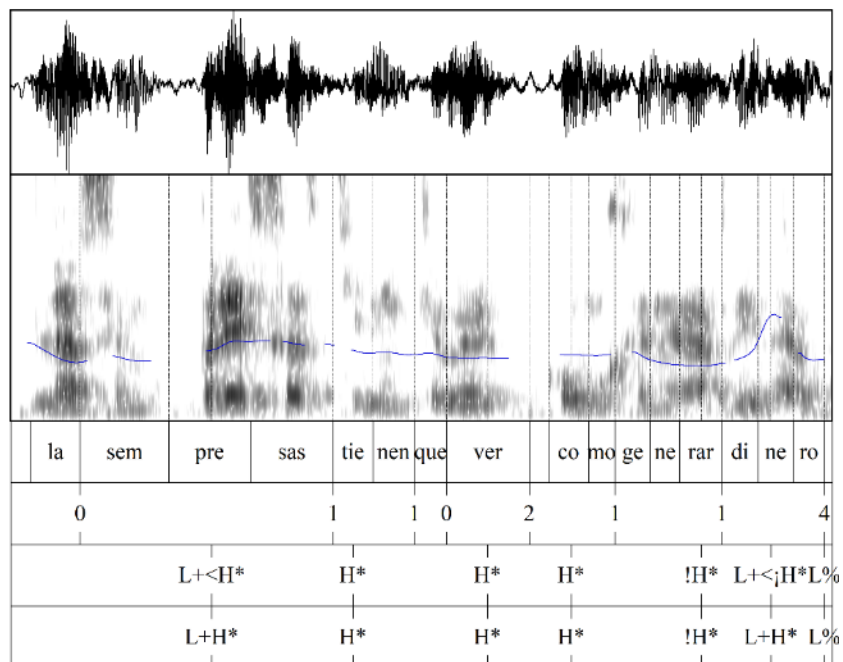
Figura 4.20. Oscilograma, espectrograma y F0 del enunciado Rs1H1-18: [(El francés ya no es la México Nuevo)<sub>ip</sub>]IP



Las realizaciones concretas presentan variación en los parámetros fonéticos de los tonemas descendentes. En las tres configuraciones más frecuentes de cada red, las variaciones tienen lugar principalmente en la excursión tonal, es decir, el grado del ascenso o descenso de la F0. Mientras que la red vecinal favorece el uso de los tonemas concretos (mencionados en orden de frecuencia descendente) L\* L%, !H\* L% y L+H\* L%, la red laboral se inclina más por las configuraciones L\* =%, H\* L%, L+H\* L% y L+<¡H\* L%,<sup>26</sup> entre las que tanto el último como el penúltimo tonema (visible en la figura 4.21) cuentan con frecuencias bastante cercanas.

<sup>26</sup> El análisis sociolingüístico relacionado con los picos muy tempranos, o prealineados, se presenta en §4.4.3.

Figura 4.21. Oscilograma, espectrograma y F0 del enunciado Rs2H1-18: [(Las empresas tienen que ver cómo generar dinero)<sub>ip</sub>]<sub>IP</sub>



Uno de los aspectos que llama la atención es que el patrón más frecuente de los miembros de la red vecinal es el descenso progresivo del tono hasta el final del enunciado (L\* L%), mientras que el de la red laboral es el sostenimiento del tono bajo de la sílaba nuclear. Si bien un análisis cognitivo está más relacionado con los aspectos abstractos que con los concretos de esta diferencia, lo observado da pie a sospechar que puede existir una restricción cognitiva en la red vecinal, relacionada con la sílaba donde se alcanza el piso tonal en el enunciado. En la red vecinal parece prevalecer una restricción cognitiva que previene que el piso tonal se alcance antes de la sílaba final, por lo que la producción primaria del tonema abstracto L\* L% consiste, acústicamente, en un descenso tonal progresivo hasta llegar al punto más bajo, al final del enunciado. Los hablantes de la red laboral, en cambio, no parecen tener esta restricción, pues su tendencia principal es la de alcanzar el punto más bajo del campo tonal en la sílaba nuclear. Sin embargo, este planteamiento acerca de dicha restricción no quiere decir que exista necesariamente una variación sociocognitiva entre ambas redes, pues las dos realizaciones concretas (L\* L% y L\* =%) están relacionadas con un mismo elemento abstracto subyacente (L\* L%), el cual es invariablemente prototípico en ambas redes.

Otro punto de contraste que llama la atención es el tonema abstracto descendente H+L\* L%, puesto que su frecuencia de uso por parte de ambas redes lo postula como un posible índice sociolingüístico. En el caso de la red vecinal, esta configuración presenta una relación de producción relativamente equitativa de las dos realizaciones concretas registradas (4 consisten en H+L\* L% y 6 de H+L\* =%), por lo que no se puede establecer una tendencia principal en esta red. Sin embargo, la configuración concreta H+L\* L% es claramente predominante en la red laboral, pues sus 17 casos superan por mucho los 2 del tonema H+L\* =% en esta misma red. Por ende, debido al predominio de la configuración concreta H+L\* L% en la red laboral y a la considerable diferencia de casos que posee con las producciones de la red vecinal, se considera que este tonema concreto puede ser un índice sociolingüístico.

Al observar la distribución de los casos con respecto a la variable sexo (cuadro 4.20), al igual que en las redes sociales, es notorio que los tonemas abstractos predominantes siguen siendo los descendentes. De hecho, la frecuencia de uso sigue el mismo patrón documentado en los datos generales del corpus: mayor presencia del descenso progresivo representado por L\* L%, pasando por H\* L% y L+H\* L%, hasta llegar a H+L\* L% (el menos frecuente de todos los tonemas descendentes). Ahora bien, este orden decreciente se apega a los datos totales de ambos sexos y a las tendencias de las mujeres de ambas redes sociales, mas no describe en detalle lo que ocurre con los hombres.

Por una parte, los hombres de la red laboral presentan la misma cantidad de casos en los tonemas H\* H% y L+H\* H%, por lo que ambas configuraciones son igualmente predominantes en los hablantes de este grupo. Por la otra, los hombres de la red social vecinal tampoco siguen el patrón general, pues producen la misma cantidad de tonemas L\* L% que de L+H\* L%, lo que refleja que, en este grupo, tanto el tonema descendente como el circunflejo tienen el mismo grado de representatividad prototípica. Sin embargo, aunque esta configuración es de las más frecuentes en los hombres de la red vecinal, comparte la primera posición en frecuencia con el tonema L\* L%, lo que implica que este grupo masculino no necesariamente se aparta de lo que se puede considerar la norma común, sino que, tanto el tonema prototípico de los hablantes en general (L\* L%) como la configuración circunfleja, coexisten como rasgos entonativos prototípicos en los hombres de la red vecinal.



Cuadro 4.20. Tonemas de acuerdo con la variable sexo y red social. N=360

Tonemas abstractos	Hombres			Mujeres		
	Red vecinal	Red laboral	Total	Red vecinal	Red laboral	Total
H* H%	0	1	1	0	1	1
H* !H%	6	0	6	3	0	3
H* L%	19	28	47	20	32	52
L+H* H%	10	1	11	7	0	7
L+H* !H%	6	1	7	1	0	1
L+H* L%	22	18	40	16	12	28
L* H%	1	3	4	0	0	0
L* !H%	0	1	1	2	0	2
L* L%	22	28	50	35	35	70
H+L* L%	4	9	13	6	10	16
Total	90	90	180	90	90	180

$p = 0.036$  (Estadístico Exacto de Fisher únicamente con la variable sexo)

$p = 0$  (Estadístico Exacto de Fisher al combinar variables sexo y red social)

Los tonemas no descendentes ocurren en ambos sexos, aunque la variable que parece favorecer la presencia o ausencia de dichas configuraciones es la misma red. En el caso de las mujeres, aquellas pertenecientes a la red vecinal son las que más producen este tipo de configuraciones, principalmente L+H\* H% (figura 4.22), pues las mujeres de la red social laboral presentan un solo caso de tonemas H\* H% (figura 4.23). Los hombres, en contraste, registran más casos de dichos tonemas abstractos que las mujeres, la mayoría se concentran en la red vecinal. Este grupo de hombres se inclina principalmente por el uso del tonema abstracto L+H\* H% y, en segundo lugar, por la emisión de L+H\* !H% y H\* !H% (figuras 4.24 y 4.25, respectivamente). Los hombres de la red laboral, por su parte, presentan menos casos de configuraciones no descendentes que los de la red vecinal y dichos casos se concentran en el tonema abstracto L\* H%.

Figura 4.22. Oscilograma, espectrograma y F0 del enunciado Rs1M1-6: [(*Está tomando el café apenas*)<sub>ip</sub>]<sub>IP</sub>

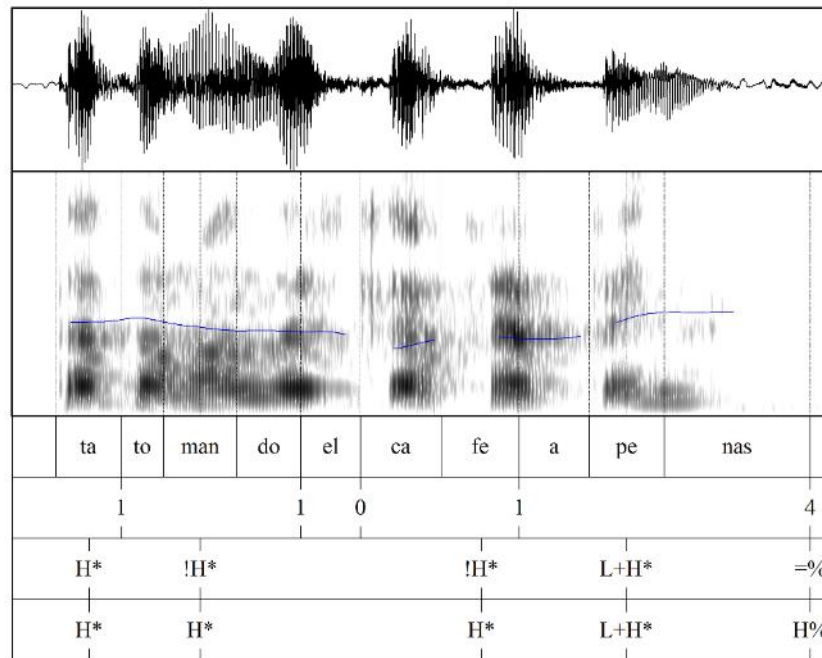


Figura 4.23. Oscilograma, espectrograma y F0 del enunciado Rs2M3-20: [(*Con un portuñol espectacular nos comunicábamos perfecto*)<sub>ip</sub>]<sub>IP</sub>

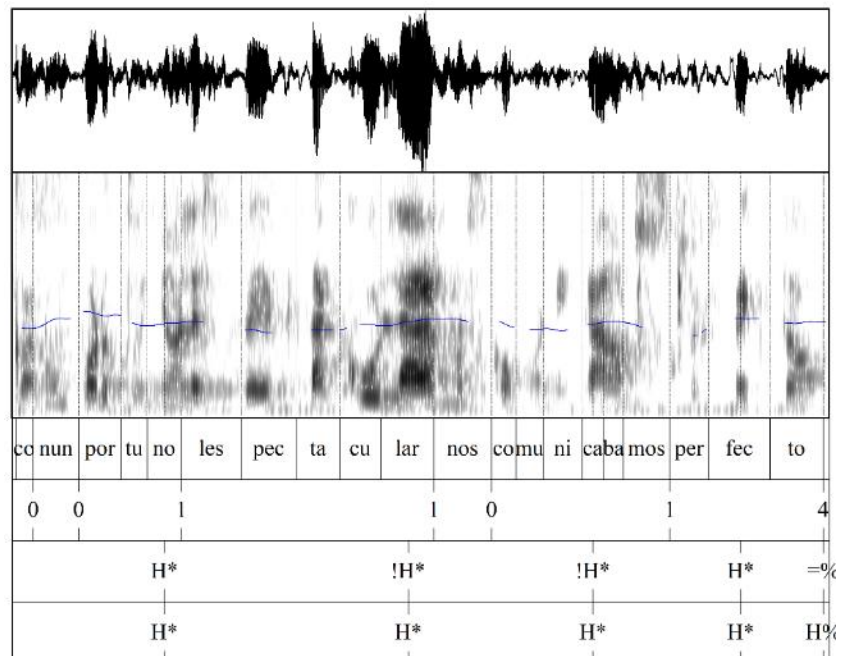


Figura 4.24. Oscilograma, espectrograma y F0 del enunciado Rs1H1-6: [(*Tendría diez once años*)<sub>ip</sub>]<sub>IP</sub>

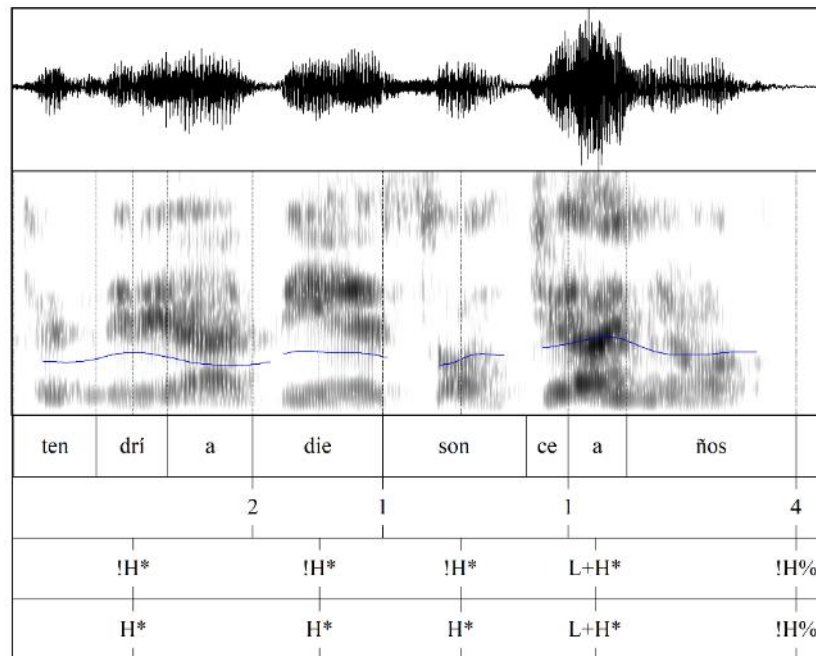
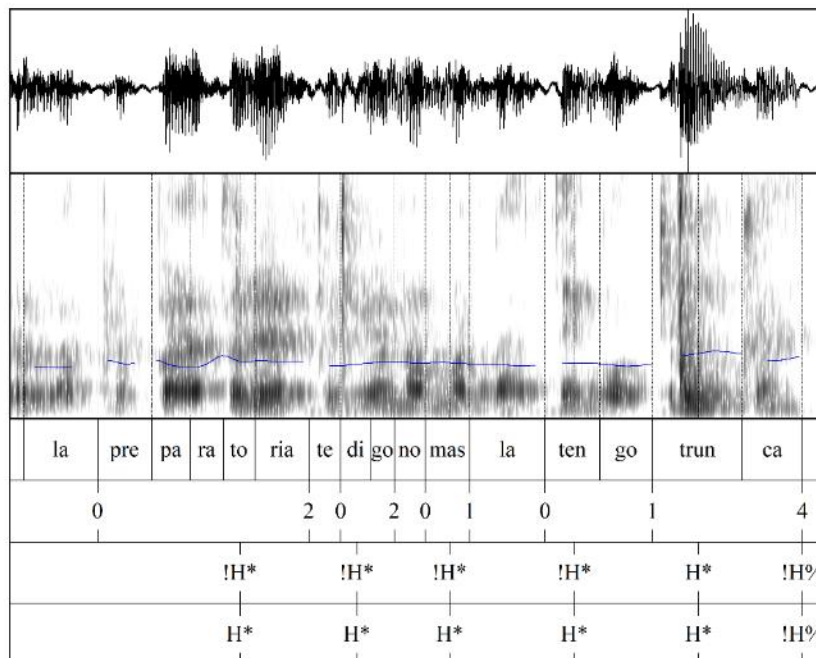


Figura 4.25. Oscilograma, espectrograma y F0 del enunciado Rs1H3-10: [(*La preparatoria te digo nada más la tengo trunca*)<sub>ip</sub>]<sub>IP</sub>



Respecto al parámetro de duración, las medias mostradas en el cuadro 4.21 revelan una tendencia por parte de los integrantes de la red vecinal a producir tonemas más largos, ya que las medias de duración totales en las dos sílabas que los conforman son mayores que las de la red laboral. La mayor duración por parte de la red vecinal (de 376 ms, en promedio) se considera un potencial índice sociolingüístico de esta red. Esto se debe a que dicho promedio difiere estadística (ANOVA con  $p = 0$ ) y perceptivamente de la media del tonema de la red vecinal (330 ms), pues la variación supera el umbral de 30 ms necesario para percibir la diferencia de duración (Toledo 1998 y 2000). Otro contraste entre ambas redes se establece en la sílaba del tonema que cada red produce con mayor duración. La red vecinal refleja una mayor media total de duración en la sílaba linde con respecto a la nuclear, lo cual es especialmente notorio en los tonemas con ascensos tonales en la sílaba nuclear (L+H\*), ya que todos se producen con mayor duración en dicha sílaba final. Por otra parte, las medias totales de la red laboral evidencian que esta presenta un comportamiento opuesto, pues las sílabas nucleares de todos los tonemas son producidas con una mayor duración promedio (excepto la del tonema L\* H%, cuya media de duración es más larga en la sílaba linde de IP).

Cuadro 4.21. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas en cada red social

Tonemas abstractos	Duración del núcleo		Duración del linde	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H%	0	212	0	201.5
H* !H%	201.67	0	157.67	0
H* L%	198.28	174.12	181.95	163.17
L+H* H%	158.82	142	195.76	105
L+H* !H%	181.71	180	230.86	160
L+H* L%	183.24	174.03	206.63	157.63
L* H%	110	155.67	128	172.33
L* !H%	198.5	157	153	153
L* L%	178.37	164.37	196.46	156.75
H+L* L%	195.1	189.95	162	135.42
Total	183.93	172.23	192.02	157.25

$p = 0.036$  (ANOVA del núcleo al nivel intergrupar);

$p = 0$  (ANOVA del linde al nivel intergrupar)

En términos de duración relativa, como se puede ver en el cuadro 4.22, lo primero que resulta llamativo es que la red vecinal tiene la misma cantidad de ocurrencias en ambos tipos de duración relativa, tanto en los casos en los que la sílaba nuclear es más larga que la final como en las instancias donde ocurre lo contrario. La primera idea que se sugiere es que no existe un patrón de duración prototípico en las sílabas nucleares de IP. Sin embargo, esta idea se cuestiona al observar una relación entre la duración relativa y los movimientos melódicos, ya que las sílabas finales de IP son producidas con una mayor duración que las nucleares cuando los tonemas tienen acento nuclear ascendente (L+H\*), con independencia del tono de juntura, y cuando se trata del tonema L\* L%. Por lo tanto, la estructura cognitiva de esta red, de acuerdo con los datos de frecuencia de uso, parece establecer una asociación entre la configuración tonal y la duración en el tonema, ya que esta tendencia no solo se refleja en los datos generales, sino también en las emisiones individuales de los 6 integrantes de esta red (los únicos que se desvían ligeramente de este patrón son los hablantes masculinos más jóvenes).

Cuadro 4.22. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por cada red social. N=360

Tonemas abstractos	Núcleo > Linde				Núcleo < Linde			
	Red vecinal		Red laboral		Red vecinal		Red laboral	
	Casos	%	Casos	%	Casos	%	Casos	%
H* H%	0	0	1	0.87	0	0	1	1.54
H* !H%	7	7.78	0	0	2	2.22	0	0
H* L%	26	28.89	33	28.70	13	14.44	27	41.54
L+H* H%	6	6.67	1	0.87	11	12.22	0	0
L+H* !H%	2	2.22	1	0.87	5	5.56	0	0
L+H* L%	18	20.00	21	18.26	20	22.22	9	13.85
L* H%	0	0	1	0.87	1	1.11	2	3.08
L* !H%	2	2.22	1	0.87	0	0	0	0
L* L%	22	24.44	39	33.91	35	38.89	24	36.92
H+L* L%	7	7.78	17	14.78	3	3.33	2	3.08
Total	90	100	115	100	90	100	65	100

p = 0.018 (Estadístico Exacto de Fisher)

La red laboral presenta una inclinación diferente a la de la red vecinal pues, en este caso, se ve una clara tendencia a concluir los enunciados con sílabas nucleares más largas que las

finales, lo cual ocurre en prácticamente la totalidad de los tonemas. Debido a que el tonema H+L\* L% es un posible índice de esta red, llama la atención que la mayor duración de la sílaba nuclear no solo es notoria en los valores medios (54.53 ms más larga que la final), sino también en términos de frecuencia, por lo que se considera que la mayor duración de la sílaba nuclear en este tonema es un posible índice sociolingüístico. Por otra parte, dado que la configuración nuclear de la IP no parece ser decisiva al momento de definir el patrón de duración del tonema, la mayor duración de la sílaba nuclear (en detrimento de la final) parece ser el rasgo de duración prototípico de esta red, en los tonemas de IP aseverativas. Dicho de otra forma, lo que parece ser más relevante en el constructo cognitivo de los hablantes de esta red es la prolongación de la sílaba nuclear en detrimento de la final sin que la configuración del tonema influya en la prolongación de una u otra sílaba. Por lo tanto, el elemento prosódico que se relaciona con la mayor duración es la tonicidad de la sílaba, y no la configuración tonal. No está de más señalar que se registraron casos individuales en los que la tendencia principal es la de prolongar más la sílaba final, pero este patrón parece ser aleatorio, ya que se presentó en distintas configuraciones en algunos hablantes que, de igual manera, siguen la tendencia principal en la gran mayoría de configuraciones tonales nucleares.

Las medias de duración de las sílabas del tonema, de acuerdo con el sexo de los informantes (cuadros 4.23 y 4.24), no resultaron ser significativas a nivel intergrupar (ANOVA del núcleo con  $p = 0.159$  y del linde con  $p = 0.892$ ). Los datos reflejan un patrón que permite asociar la duración con las redes sociales a las que pertenecen los hablantes, pues los miembros de la red vecinal producen ambas sílabas con medias generalmente mayores que las de los miembros de la red laboral. Al comparar las medias de duración de hombres y mujeres de la misma red social, se percibe que los hombres de la red vecinal tienden a producir sílabas nucleares y lindes de IP más largas que las mujeres de la misma red. En la red laboral se puede ver que la situación es inversa, ya que las mujeres tienden a presentar medias de duración mayores que los hombres en ambas sílabas.

Cuadro 4.23. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas de los hombres de ambas redes sociales

Tonemas abstractos	Duración del núcleo		Duración del linde	
	Hombres		Hombres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H%	-	149	-	161
H* !H%	226.33	-	169.67	-
H* L%	204.16	148.21	182.05	147.25
L+H* H%	166.2	142	185.5	105
L+H* !H%	173.67	180	250.33	160
L+H* L%	183.41	178.5	211.95	168.44
L* H%	110	155.67	128	172.33
L* !H%	-	157	-	153
L* L%	163.77	163.71	195.41	159.46
H+L* L%	208.5	197.89	177	131.22
Total	183.59	164.7	195.91	154.41

p = 0.024 (ANOVA de duración del núcleo).

p = 0 (ANOVA de duración del linde).

Cuadro 4.24. Medias de duración (ms) de las sílabas nucleares y lindes de los tonemas de las mujeres de ambas redes sociales

Tonemas abstractos	Duración del núcleo		Duración del linde	
	Mujeres		Mujeres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H%	-	275	-	242
H* !H%	152.33	-	133.67	-
H* L%	192.7	196.78	181.85	177.09
L+H* H%	148.29	-	210.43	-
L+H* !H%	230	-	114	-
L+H* L%	183	167.33	199.31	141.42
L* H%	-	-	-	-
L* !H%	198.5	-	153	-
L* L%	187.54	164.89	197.11	154.57
H+L* L%	186.17	182.8	152	139.2
Total	184.28	179.77	188.12	160.09

p = 0.549 (ANOVA de duración del núcleo).

p = 0.009 (ANOVA de duración del linde).

Los datos de la duración relativa en los hablantes de ambos sexos (cuadros 4.25 y 4.26) presentan una distribución significativa de acuerdo con el estadístico Exacto de Fisher, tanto al tomar en cuenta únicamente la variable sexo ( $p = 0.005$ ) como al incluir la variable de la red social ( $p = 0$ ). Los hombres de la red social vecinal suelen presentar más casos de alargamientos de la sílaba nuclear (en detrimento de la final) que las mujeres de esta misma red, quienes hacen lo mismo únicamente con los tonemas abstractos  $H^* L\%$  y  $L^* L\%$ . En el caso de los hablantes de la red laboral, se registra que los hombres producen la sílaba linde del tonema  $L+H^* L\%$  con mayor duración que las mujeres de su red. En el resto de los casos, las producciones de hombres y mujeres son bastante similares en relación con la red social de los hablantes.

Cuadro 4.25. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por hombres de ambas redes sociales.  $N=180$

Tonemas abstractos	Núcleo > Linde		Núcleo < Linde	
	Hombres		Hombres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
$H^* H\%$	0	0	0	1
$H^* !H\%$	5	0	1	0
$H^* L\%$	12	16	7	12
$L+H^* H\%$	5	1	5	0
$L+H^* !H\%$	1	1	5	0
$L+H^* L\%$	11	12	11	6
$L^* H\%$	0	1	1	2
$L^* !H\%$	0	1	0	0
$L^* L\%$	4	16	18	12
$H+L^* L\%$	3	8	1	1
Total	41	56	49	34

$p = 0$  (Estadístico Exacto de Fisher únicamente con hablantes masculinos)



Cuadro 4.26. Duración relativa de la sílaba nuclear con respecto a la sílaba linde en los tonemas producidos por mujeres de ambas redes sociales. N=180

Tonemas abstractos	Núcleo > Linde		Núcleo < Linde	
	Mujeres		Mujeres	
	Red vecinal	Red laboral	Red vecinal	Red laboral
H* H%	0	1	0	0
H* !H%	2	0	1	0
H* L%	14	17	6	15
L+H* H%	1	0	6	0
L+H* !H%	1	0	0	0
L+H* L%	7	9	9	3
L* H%	0	0	0	0
L* !H%	2	0	0	0
L* L%	18	23	17	12
H+L* L%	4	9	2	1
Total	49	59	41	31

P = 0.054 (Estadístico Exacto de Fisher únicamente con hablantes femeninos)

En general, los patrones tonales y de duración parecen estar más relacionados con la red social a la que pertenecen los hablantes que con la variable sexo. Esto se observa tanto en la distribución de configuraciones nucleares como en el recuento de casos, de acuerdo con la duración relativa de las sílabas que integran el tonema. En estos casos, los hablantes de ambos sexos se apegan mucho más a los comportamientos prototípicos de las redes sociales a las que pertenecen que a algún otro factor. Esto da pie a pensar que la variable sexo no desempeña un papel particularmente relevante en la percepción de las experiencias lingüísticas y sociales que conllevan la creación de esquemas cognitivos relacionados con la prosodia del tonema.

#### 4.4.2. Tonemas no descendentes

Al extraer los tonemas abstractos no descendentes y analizarlos por separado, se confirma que su uso es una característica más sobresaliente de la red vecinal que de la laboral. Como reflejan los datos del cuadro 4.27, estadísticamente significativos, el número de tonemas abstractos no descendentes en la red vecinal supera el de la laboral. El tonema no descendente más utilizado por la red vecinal es el L+H\* H%, realizado, principalmente, en igualdad de

ocasiones, a través de las configuraciones concretas  $L+H^* =\%$ ,  $L+;H^* =\%$  y  $L+<;H^* =\%$  (figuras 4.20, 4.26 y 4.27, respectivamente), lo que deja la configuración  $L+<H^* =\%$  en la posición menos frecuente de este grupo. A este tonema abstracto, en orden descendente, le siguen las configuraciones  $H^* !H\%$  (principalmente bajo la forma fonética  $!H^* =\%$ ) y  $L+H^* !H\%$  (realizado mayoritariamente como  $L+H^* !H\%$  y  $L+<;H^* !H\%$ , en igualdad de casos). Esto deja en evidencia la preferencia por parte de los hablantes de la red vecinal a emplear sostenimientos tonales desde la sílaba nuclear ( $=\%$ ), como realización concreta del tono de juntura abstracto alto ( $H\%$ ). Por otra parte, en relación con los casos registrados en la red laboral, es necesario precisar que 5 de los tonemas no descendentes (todas las producciones de  $L^* H\%$  y  $L^* !H\%$ , junto con una de  $H^* H\%$ ) son producidas por el hablante Rs2H3, mientras que un caso de  $H^* H\%$  es producido por la hablante Rs2M3 y las ocurrencias de  $L+H^* H\%$  y  $L+H^* !H\%$  por el hablante Rs2H2. Por lo tanto, no se puede considerar que los tonemas no descendentes en esta red sean representativos de ella, sino que se deben principalmente a un rasgo individual.

Cuadro 4.27. Tonemas no descendentes y red social. N=44

Tonemas abstractos	Red vecinal	Red laboral	Total
$H^* H\%$	0	2	2
$H^* !H\%$	9	0	9
$L+H^* H\%$	17	1	18
$L+H^* !H\%$	7	1	8
$L^* H\%$	1	3	4
$L^* !H\%$	2	1	3
Total	36	8	44

p = 0.003 (Estadístico Exacto de Fisher)

Figura 4.26. Oscilograma, espectrograma y F0 del enunciado Rs1M1-1: [(A Óscar le tocó negra)<sub>ip</sub>]<sub>IP</sub>

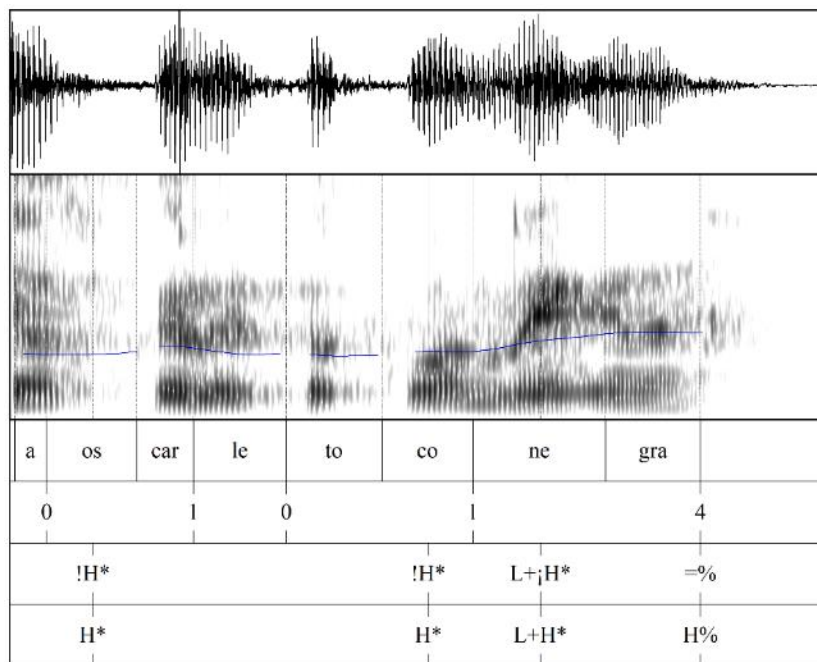
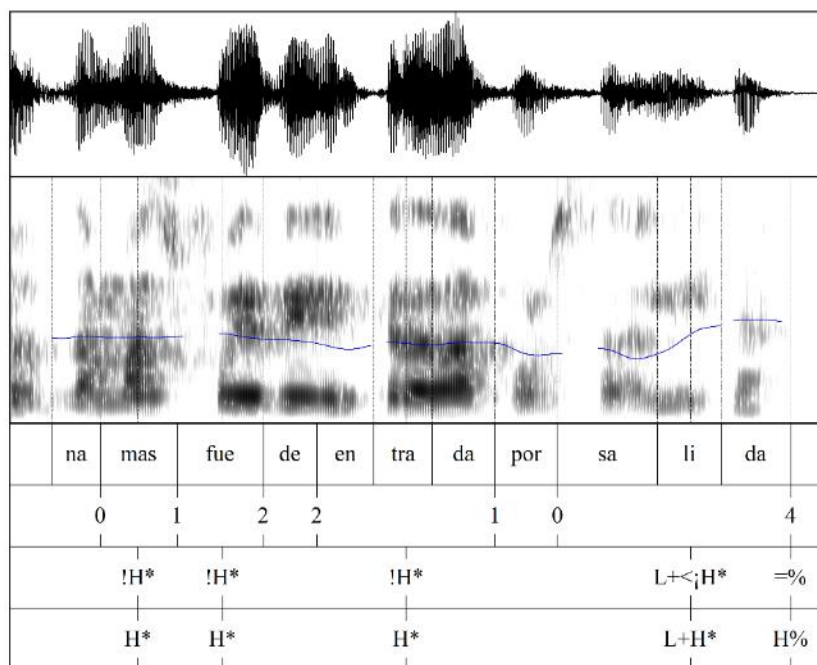


Figura 4.27. Oscilograma, espectrograma y F0 de la última frase entonativa intermedia del enunciado Rs1M1-14: [(No le salió tampoco tan caro al Carlos)<sub>ip</sub> (nada más fue de entrada por salida)<sub>ip</sub>]<sub>IP</sub>



La distribución de casos relacionados con la variable sexo (cuadro 4.28) no resultó ser significativa. Sin embargo, se puede decir que los hombres de la red social vecinal presentan mayor tendencia al empleo de los tonemas no descendentes. Aunque las mujeres también emplean de forma relativamente elevada estos tonemas, lo hacen con una frecuencia que se reduce a poco más de la mitad que la de los hombres. Esto podría ser un indicador de que las configuraciones no descendentes están relacionadas con rasgos de masculinidad en el habla, pues se observa en hombres pertenecientes a las dos redes sociales. Sin embargo, esta hipótesis no podría confirmarse hasta contar con un mayor volumen de datos.

Cuadro 4.28. Tonemas no descendentes y variable sexo. N=44

Tonemas abstractos	Hombre	Mujer	Total
H* H%	1	1	2
H* !H%	6	3	9
L+H* H%	11	7	18
L+H* !H%	7	1	8
L* H%	4	0	4
L* !H%	1	2	3
Total	30	14	44

p = 0.849 (Estadístico Exacto de Fisher)

Al llevar a cabo el recuento de casos de tonemas no descendentes en contraste con las variables de foco y expresividad, se obtuvieron los cuadros 4.29 y 4.30, que muestran los datos en relación con las redes sociales y el sexo de los hablantes, respectivamente. La mayoría de los tonemas abstractos no descendentes de la red vecinal ocurre mayoritariamente en coincidencia con focos estrechos sin expresividad, donde resalta el número de ocurrencias de la configuración L+H\* H%, principalmente con las realizaciones concretas L+<sub>j</sub>H\* =%, !H\* =% (figura 4.28), L+<<sub>j</sub>H\* =% y L+<H\* =% (figura 4.29), la primera de las cuales cuenta solo con un caso más que las demás. En enunciados no expresivos y con foco amplio, producidos por la red vecinal, los cuales representan el 22.2 % de las configuraciones no descendentes, resaltan como los más frecuentes !H\* =% y L+<<sub>j</sub>H\* =%, con tres casos cada uno, solo uno más que en los casos con foco estrecho. La red laboral, en cambio, presenta muy pocos casos de estas configuraciones, ninguno de ellos con más de una ocurrencia. Es importante tener en cuenta que 5 de los 8 casos registrados en la red laboral son producidos

por un único hablante masculino (Rs2H3), mientras que 2 son producidos por otro hombre y 1 por una mujer, por lo que no se puede establecer ninguna conclusión basándose en los datos de la red laboral, más allá del hecho de que estas configuraciones no representan un rasgo representativo de esta red.

Cuadro 4.29. Tonemas no descendentes de acuerdo con la expresividad y tipo de foco en cada red social. N=44

Tonemas abstractos	Red vecinal				Red laboral			
	Neutro		Expresivo		Neutro		Expresivo	
	Amplio	Estrecho	Amplio	Estrecho	Amplio	Estrecho	Amplio	Estrecho
H* H%	0	0	0	0	0	1	1	0
H* !H%	3	3	2	1	0	0	0	0
L+H* H%	5	8	2	2	0	0	1	0
L+H* !H%	0	3	2	2	0	0	1	0
L* H%	0	1	0	0	1	1	1	0
L* !H%	0	0	2	0	0	1	0	0
Total	8	15	8	5	1	3	4	0

Cuadro 4.30. Tonemas no descendentes de acuerdo con la expresividad y tipo de foco en relación con el sexo de los informantes. N=44

Tonemas abstractos	Hombre				Mujer			
	Neutro		Expresivo		Neutro		Expresivo	
	Amplio	Estrecho	Amplio	Estrecho	Amplio	Estrecho	Amplio	Estrecho
H* H%	0	1	0	0	0	0	1	0
H* !H%	2	2	1	1	1	1	1	0
L+H* H%	3	5	2	1	2	3	1	1
L+H* !H%	0	3	3	1	0	0	0	1
L* H%	1	2	1	0	0	0	0	0
L* !H%	0	1	0	0	0	0	2	0
Total	6	14	7	3	3	4	5	2

Figura 4.28. Oscilograma, espectrograma y F0 de la última ip del enunciado Rs1M3-19: [(*Me acosté*)<sub>ip</sub> (*prendí puse una película*)<sub>ip</sub> (*y*)<sub>ip</sub> (*y me quedé dormida*)<sub>ip</sub>]<sub>IP</sub>

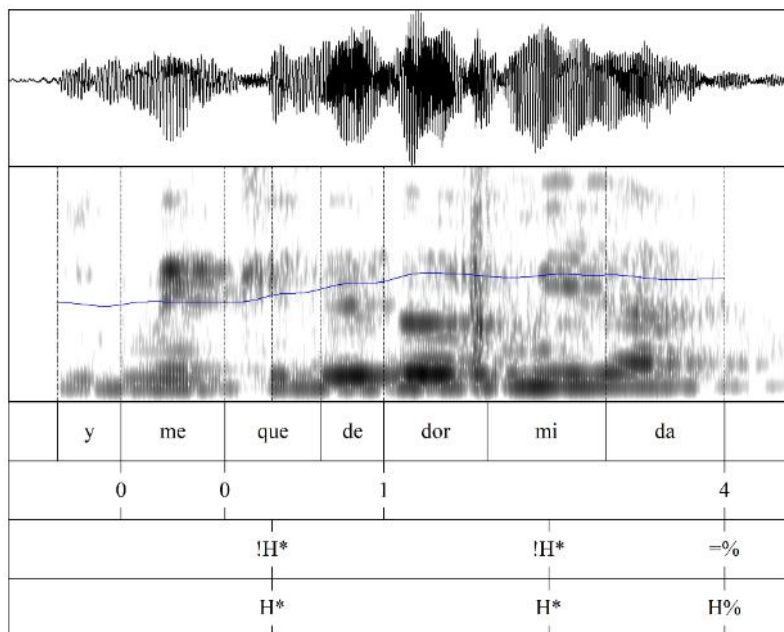
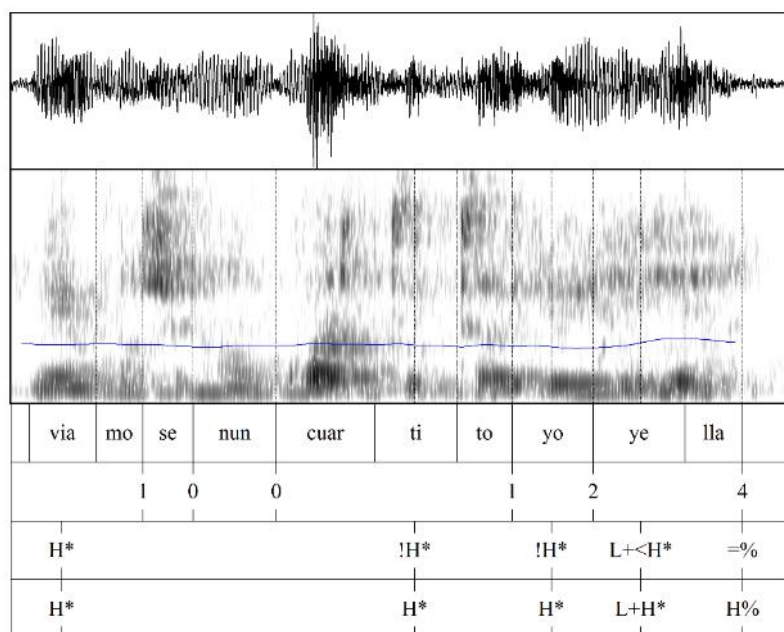


Figura 4.29. Oscilograma, espectrograma y F0 del enunciado Rs1H2-6: [(*Vivíamos en un cuartito yo y ella*)<sub>ip</sub>]<sub>IP</sub><sup>27</sup>



<sup>27</sup> La primera sílaba de *vivíamos* es elidida por el hablante; es por ello que la transcripción del enunciado comienza con la sílaba tónica.

Los datos de estas configuraciones en función de la variable sexo permiten ver que tienen lugar principalmente en focos estrechos no expresivos emitidos por hombres, mayoritariamente con el tonema abstracto L+H\* H%. Los casos documentados en las mujeres, en cambio, se concentran en enunciados de foco amplio y con matiz expresivo, producidos con la configuración nuclear abstracta L\* !H%; aunque el tonema que concentra la mayoría de casos es L+H\* H% en los enunciados con foco estrecho sin expresividad. Teniendo en cuenta que la gran mayoría de los casos ocurren en la red vecinal, se puede decir que estas tendencias representan particularmente a los hombres y mujeres de esta red. Dichas tendencias son observables en todos los miembros de la red vecinal, excepto por un hombre (Rs1H3) y una mujer (Rs1M2), quienes no producen tonemas no descendentes.

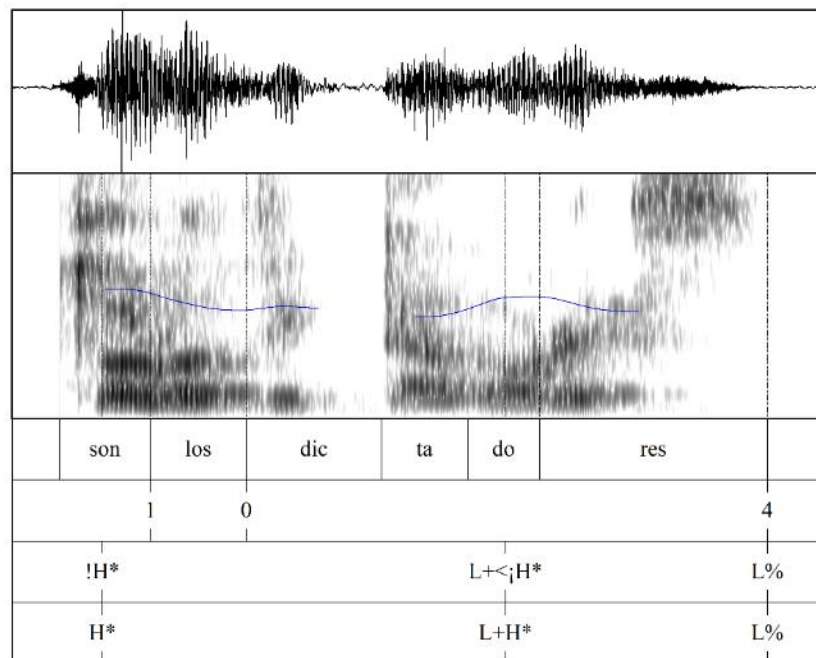
Es importante mencionar que estos dos hablantes son los únicos que establecen contacto laboral con personas de distintos grupos socioeconómicos que habitan fuera del área geográfica en la que viven todos los miembros de la red vecinal. Debido a que sus actividades consisten en la prestación de servicios, es posible que su uso de configuraciones tonales no descendentes haya sido evitado durante la entrevista sociolingüística por ser considerado (por ellos) como un rasgo no apropiado en ciertos contextos conversacionales con individuos no originarios de la zona en la que habitan o que tienen características sociales distintas a las de la red (como es el caso del entrevistador), cuyo uso puede afectar su captación o mantenimiento de clientes y, por ende, sus ingresos económicos. De hecho, ambos informantes reportaron en el cuestionario sociolingüístico que le dan bastante importancia al uso correcto del lenguaje y que hablan de forma diferente a lo usual ante “todo aquel que amerite respeto” (Rs1H3) y ante “clientes y personas importantes” (Rs1M2). Por lo tanto, la ausencia de este tipo de tonemas se puede explicar a través de la teoría del mercado lingüístico, la cual explica la existencia de ciertas conductas lingüísticas que dependen de las actividades socioeconómicas de los individuos, que puede ocasionar que algunos de ellos se apeguen a la norma lingüística y otros no, aunque todos compartan unos mismos rasgos socioeconómicos (Sankoff y Laberge, 1978). Teniendo en cuenta que el uso de configuraciones no descendentes está suficientemente extendido en la red vecinal y que existe la posibilidad de que su uso puede ser evadido por (presuntamente) no ser bien

aceptado por miembros de otros grupos sociales, se considera que los tonemas no descendentes representan un índice sociolingüístico de la red social vecinal.

#### 4.4.3. Picos muy tempranos en frases entonativas mayores (IP)

Al igual que los tonemas no descendentes, los picos muy tempranos (figura 4.30) se pueden asociar más con la red vecinal que con la laboral, pues la distribución de casos entre ambas redes resultó ser significativa (estadístico Exacto de Fisher = 0.045). Las realizaciones de prealineamientos tonales en la red vecinal son bastante uniformes en todos los miembros, excepto en uno. En cambio, la mayoría de los casos en la red laboral (N=9) son realizados por un solo hablante masculino, mientras que 3 casos son producidos por otro hombre; y los 3 restantes son emitidos por las tres mujeres de esta red (cada una de ellas produce un tonema con prealineamiento tonal). Esto implica que los picos muy tempranos registrados en la red laboral, a diferencia de lo que ocurre en la vecinal, no son representativos de todos sus hablantes.

Figura 4.30. Oscilograma, espectrograma y F0 del enunciado Rs1H1-29: [(La cámara de diputados)<sub>ip</sub>(son los dictadores)<sub>ip</sub>]<sub>IP</sub>





El 71 % de los prealineamientos tonales en la red vecinal ocurren de forma equitativa en los tonemas  $L+\langle_j H^* =\%$ ,  $L+\langle H^* L\%$  y  $L+\langle_j H^* L\%$ , lo que apunta hacia un uso mayoritario de circunflexión con ascenso variable en el acento nuclear, que puede ser entre 1.5 y 3 st ( $L+\langle H^*$ ) y superior a 3 st ( $L+\langle_j H^*$ ). La red laboral, en cambio, produce el 73 % de los picos muy tempranos con la realización fonética  $L+\langle_j H^* L\%$ . Adicionalmente, como se puede ver en los cuadros 4.31 y 4.32, la mayoría de los prealineamientos tonales suceden principalmente en enunciados con foco amplio, carentes de expresividad, en ambas redes. Aunque los casos son escasos, en algunas instancias de focalización y expresividad se puede observar que la red vecinal hace uso de tonos de juntura sostenidos ( $=\%$ ) para expresar la focalización; y de una circunflexión con un ascenso prominente ( $L+\langle_j H^*$ ) para la expresividad. Sin embargo, no hay que perder de vista que se registraron más casos de  $L+\langle_j H^* =\%$  y  $L+\langle_j H^* L\%$  en contextos de foco amplio y sin expresividad, respectivamente, lo que pone en duda cualquier relación entre estas configuraciones y dichas funciones. Por otra parte, aunque el tonema  $L+\langle_j H^* L\%$  también es el más empleado por la red laboral para transmitir expresividad, se registró principalmente en enunciados carentes de expresividad y con foco amplio, lo que puede implicar que la red laboral no relaciona esta configuración con expresividad.

Cuadro 4.31. Tonemas con picos muy tempranos en la red vecinal de acuerdo con el tipo de foco y expresividad. N=21

Tonemas concretos	Red vecinal				Total
	Foco estrecho	Foco estrecho con expresividad	Expresivo	Foco amplio sin expresividad	
$L+\langle H^* =\%$	2	0	0	0	2
$L+\langle H^* L\%$	1	0	0	4	5
$L+\langle_j H^* =\%$	2	0	0	3	5
$L+\langle_j H^* !H\%$	1	1	1	0	3
$L+\langle_j H^* L\%$	0	0	2	3	5
$L+\langle_j H^* HL\%$	0	0	0	1	1
Total	6	1	3	11	21

Cuadro 4.32. Tonemas con picos muy tempranos en la red laboral de acuerdo con el tipo de foco y expresividad. N=15

Tonemas concretos	Red laboral				Total
	Foco estrecho	Foco estrecho con expresividad	Expresivo	Foco amplio sin expresividad	
$L+<H^* =\%$	0	0	0	0	0
$L+<H^* L\%$	0	0	1	2	3
$L+<_iH^* =\%$	0	0	1	0	1
$L+<_iH^* !H\%$	0	0	0	0	0
$L+<_iH^* L\%$	1	1	3	6	11
$L+<_iH^* HL\%$	0	0	0	0	0
Total	1	1	5	8	15

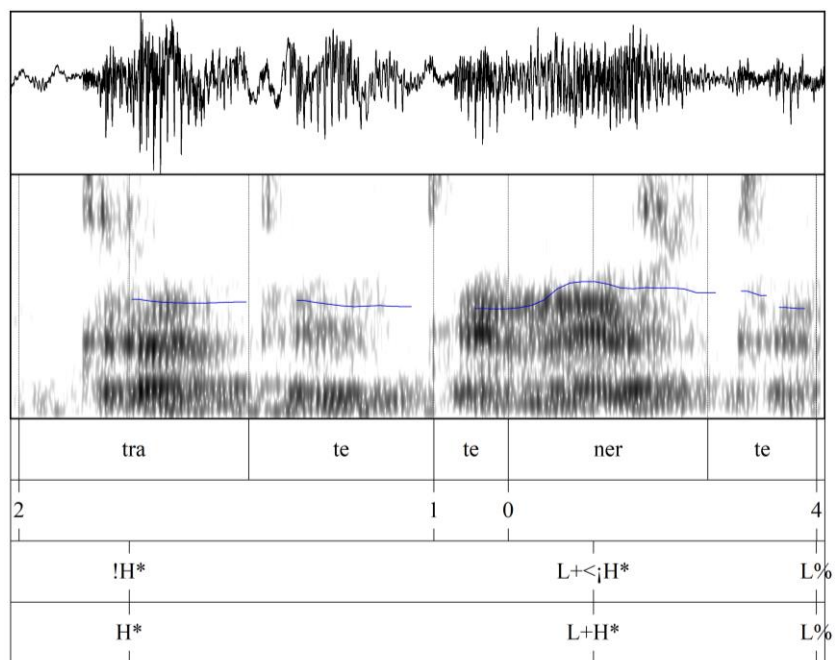
En relación con la duración, a diferencia de la red social laboral que produce los picos muy tempranos casi siempre en sílabas nucleares más largas que la final de la IP, la red vecinal presenta casi la misma cantidad de picos prealineados en sílabas nucleares más cortas y más largas que la final (la diferencia es de una ocurrencia). Esto indica que la duración de la sílaba nuclear, en relación con la de la final, no es un factor determinante en la producción de prealineamientos tonales en esta red. Sin embargo, la duración de la sílaba nuclear sí parece estar relacionada con la duración de las sílabas finales en la red laboral.

Con base en los datos de frecuencia de uso, se puede sugerir que la estructura cognitiva de los hablantes de la red social laboral está configurada de forma tal que los prealineamientos tonales están asociados con mayor duración, lo cual se puede deber a que los picos muy tempranos presentan ascensos más prominentes en el acento nuclear ( $L+<_iH^*$ ). En contraste, la estructura cognitiva de los informantes de la red vecinal parece tener una configuración distinta, ya que los prealineamientos tonales no están relacionados con algún patrón de duración relativa de las sílabas del tonema. Esto se puede explicar partiendo de la premisa de que la información almacenada en el sistema cognitivo viene de la realidad que experimentan los sujetos, lo que implica, en este contexto, que los patrones de producción de los hablantes de ambas redes son un reflejo de las configuraciones de duración a las que han estado expuestos; y que han almacenado en su sistema durante las interacciones sociales con otros miembros de las redes a las que pertenecen y con personas de su entorno cercano.

En relación con el movimiento tonal después de la culminación del pico de F0, el sostenimiento del tono alto hasta el final de la sílaba nuclear es la situación más recurrente

en ambas redes. Otra realización es con un descenso en la sílaba final hasta alcanzar un tono de juntura bajo (L%). En el caso de la red social vecinal, los sostenimientos tonales en la sílaba nuclear son especialmente predominantes cuando dicha sílaba es más corta que la final (solo hay un caso en el que ocurre lo contrario). Estos sostenimientos suelen continuar con un posterior descenso en la sílaba posnuclear hasta una juntura L%; aunque se registraron 4 instancias (19 %) en las que el tono desciende levemente hasta una juntura media (!H%), un caso en el que el tono se mantiene sostenido hasta el final del enunciado (=%) y 4 casos (19 %) en los que el tono asciende hasta una juntura H%. Esto va de la mano con la tendencia de la red vecinal, en relación con el uso de tonemas no descendentes. En contraste, casi todos los picos muy tempranos producidos por la red social laboral (con excepción de un caso de sostenimiento tonal hasta la juntura) presentan un posterior descenso hasta una juntura baja (L%). En aquellos casos en los que no se observa un sostenimiento de F0 en la sílaba nuclear una vez alcanzado el pico tonal, la red vecinal suele iniciar el descenso melódico inmediatamente después de alcanzar la altura tonal máxima en la sílaba nuclear.

Figura 4.31. Oscilograma, espectrograma y F0 de las últimas sílabas del enunciado  
Rs2H1-6: [(*Siempre va a haber algo que trate de detenerte*)<sub>ip</sub>]<sub>IP</sub>



Por otra parte, la red laboral evidencia una mayor tendencia a sostener el tono a la misma altura del pico tonal (en promedio, 14 % del tiempo total de duración de la sílaba nuclear), antes de iniciar el descenso dentro de los lindes de la misma sílaba. Como ejemplo, se puede ver la figura 4.31, donde se registró un breve sostenimiento de F0 durante el 15 % de la duración total de la sílaba. Una vez alcanzado el pico tonal (86.23 st), la F0 se mantiene a la misma altura con variaciones menores a los 0.1 st, hasta que inicia su descenso dentro de la misma sílaba nuclear. Al final de dicha sílaba, la F0 ha descendido hasta los 85.26 st, lo que implica un descenso de 0.97 st posterior al sostenimiento. Este tipo de descensos, posteriores a un sostenimiento de F0 tras un pico muy temprano, podría representar otra diferencia en la estructura cognitiva de ambas redes sociales, la cual quedará pendiente de ser confirmada en una investigación posterior.

#### 4.4.4. Resumen de resultados de los tonemas de IP

Las junturas de IP más recurrentes se producen con el tono de juntura bajo (L%), con una recurrencia de la configuración L\* L%. El tonema circunflejo es el tercero más frecuente y, aunque es producido principalmente por la red social vecinal, presenta un número considerable de casos en la red social laboral. La frecuencia de casos de circunflexión puede asociarse con su caracterización como un rasgo dialectal del español del centro de México (Martín Butragueño, 2017). Sin embargo, en la red vecinal se registran más casos de ascensos mayores a 3 st (L+;H\*) que en la red laboral, lo cual marca la diferencia entre ambas redes en el grado de ascenso previo al descenso de la F0. Adicionalmente, estos datos concuerdan con el grado de instrucción bajo, asociado con dicho ascenso prominente, señalado por Martín Butragueño (2021). La mayoría de las circunflexiones registradas en la red laboral son producidas por un solo hablante, cuyo acento tonal ascendente prototípico es L+<;H\*, el cual presenta el ascenso superior a 3 st mencionado previamente. Esto lleva a pensar que la interacción que este hablante ha sostenido a lo largo de su vida con otros grupos sociales difiere de la de los demás integrantes de la red laboral, pues su esquema cognitivo de acentos nucleares ascendentes presenta este elemento como prototípico, mientras que no sucede lo mismo con los otros miembros.

Por otra parte, al igual que en los datos de Martín Butragueño (2021), los tonemas circunflejos registrados en la red vecinal son más frecuentes en hombres que en mujeres y,

adicionalmente, son producidos con una media de duración (390 ms), no solo mayor a la media general de esta red (375.95 ms), sino también mucho mayor a la media de la red laboral en esta misma configuración (332 ms). Cabe señalar que los tonemas circunflejos ocurren en mayor medida en dos hablantes jóvenes (el hombre tiene 32 años y la mujer 33), quienes son los únicos de esta red que se encuentran entre los 25 y los 40 años de edad y que, adicionalmente, han sido pareja por algunos años. Esto genera la duda acerca de si ambos comparten la frecuencia de esta configuración por ser contemporáneos o si se debe al tiempo de convivencia entre ellos, durante el cual alguno pudo haber aumentado el uso de la circunflexión por influencia del habla de su pareja.

Otro rasgo que distingue a la red social vecinal es su mayor empleo de tonemas no descendentes, en comparación con los escasos casos documentados en la red laboral. Se puede inferir que dichos tonemas están asociados con los rasgos sociales de la red vecinal, debido a que no tienen mucha presencia en la red laboral y a que es evitado por los dos únicos hablantes de esta red (Rs1H3 y Rs1M2) que prestan servicios a personas de distintos niveles socioeconómicos. Ambos hablantes reportaron que hablan de forma diferente a la habitual ante clientes y personas a las que desean mostrar respeto, por lo que podrían apegarse más a la norma y evitar el uso de rasgos prosódicos asociables con el grupo social al que pertenecen. Esta posibilidad se ve aumentada al tener en cuenta que la informante Rs1M3, quien comparte vínculos familiares con Rs1M2 y Rs1H3<sup>28</sup> y trabaja dentro de la zona habitada por los miembros de la red, sí hace uso de las configuraciones nucleares no descendentes. Lo expuesto previamente implica que los tonemas no descendentes constituyen un rasgo instaurado en la estructura lingüística cognitiva de los hablantes de la red vecinal (pues la exposición a estas configuraciones fue lo suficientemente constante como para fijarla en las estructuras cognitivas de estos hablantes) y que puede ser inhibida por la influencia del mercado lingüístico.

#### 4.5. Recapitulación del análisis sociolingüístico

En términos de la primera sílaba tónica, ambas redes muestran un mayor uso del acento tonal H\*, por lo que se puede considerar el acento tonal prototípico de esta posición en los enunciados aseverativos. El segundo acento más frecuente en ambas redes es L+H\*, cuya

---

<sup>28</sup> La informante Rs1M3 es hija de Rs1M2 y esposa de Rs1H3.

distribución de casos refleja una diferencia en la configuración cognitiva de ambas redes, ya que es principalmente producido con un ascenso superior a 3 st por la red social vecinal y con un ascenso entre los 1.5 st y los 3 st en la red laboral. Adicional a esto, la red laboral hace uso del acento tonal H+L\* en la primera sílaba tónica en contextos enfáticos, situación que no ocurre en la red vecinal, en la que el énfasis se expresa mediante el acento abstracto L+H\*. No obstante, debido a que no se registraron casos en la red vecinal, no se puede decir que este acento sea un elemento periférico en la categoría de acentos tonales de primeras sílabas tónicas en enunciados aseverativos. En este caso, se plantea la hipótesis de que un esquema entonativo, como el acento H+L\*, pueda no formar parte de una categoría cognitiva en un grupo social, debido a su ausencia en el uso. Dicha hipótesis, sin embargo, queda pendiente de confirmación mediante el análisis de un mayor número de casos de énfasis al inicio de los enunciados distribuidos equitativamente entre ambas redes.

El análisis de la primera sílaba tónica en relación con la variable sexo reveló que la media de duración de dicha sílaba con un acento tonal L\* es similar entre los hombres de la red vecinal y las mujeres de la laboral (en el sentido de que es considerablemente mayor que la de las mujeres y hombres de sus respectivas redes). Dicha similitud hace sospechar que la mayor duración en presencia de un acento L\* puede ser una cualidad propia del habla masculina de la red vecinal, que posee prestigio encubierto entre las mujeres de la red laboral. Sin embargo, es necesario investigar acerca de las actitudes lingüísticas en la red laboral antes de poder afirmar esta idea de manera categórica. De confirmarse la hipótesis del prestigio encubierto en este caso, ellas podrían estar haciendo uso de este recurso para transmitir cualidades relacionadas con la masculinidad de la clase trabajadora (fuerza, dominio, poder, etc.).

En la totalidad de la muestra de datos analizada se observa un predominio de los intertonemas abstractos no oxítonos H\* H- y L+H\* H-. La red vecinal no parece tener un ejemplar central claro, ya que ambas configuraciones tienen frecuencias de uso bastante cercanas en esta red. Sin embargo, en la red laboral se nota una clara centralidad de H\* H- en esta categoría cognitiva. Entre las realizaciones concretas de esta configuración, aquellas con juntura HH- (mayor a 3 st) tienen una frecuencia mayor en la red laboral (en contraste con las producciones de la vecinal), que permiten considerar que estas juntas precedidas de un acento abstracto H\* pueden ser un índice sociolingüístico. Por otra parte, los tonemas

internos abstractos más frecuentes de los hablantes de ambos sexos son igualmente H\* H- y L+H\* H-. El primero es el más común en todos los hablantes, excepto por las mujeres de la red social vecinal, quienes tienden más al empleo de la configuración L+H\* H-. De igual modo, llama la atención que las mujeres hacen uso de un repertorio entonativo mayor que el de los hombres, pues ellas emplean una mayor variedad de configuraciones nucleares internas. En lo relacionado con la duración de los intertonemas, no se presentan variaciones estadísticamente significativas.

En lo relacionado con los intertonemas oxítonos, la red laboral emplea principalmente las configuraciones abstractas H\* H- y L+H\* H-, cuyas realizaciones concretas más frecuentes, en orden descendente, son H\* H-, L+H\* !HH- (con igualdad de casos) y L+H\* HH-. La red social vecinal sigue la misma tendencia en el nivel abstracto, aunque en términos de realización concreta emplea predominantemente la configuración L+¡H\* HH-. La duración media de las sílabas tónicas finales de ip es uno de los indicadores más marcados de significado social, pues casi todos los casos documentados en la red laboral superan los 300 ms. En relación con la variable sexo, se observa que los hombres muestran una mayor preferencia por el uso de las configuraciones H\* H- y L+H\* H-, mientras que las mujeres manejan un repertorio más amplio de configuraciones nucleares oxítonas de ip. Esto es especialmente notorio en las mujeres de la red social vecinal, quienes son las únicas en incluir en sus repertorios las configuraciones L+H\* !H- y L\* L-, aunque probablemente pueda atribuirse a un rasgo femenino generacional, ya que las dos informantes que producen estos intertonemas abstractos superan los 50 años de edad.

Las funciones sintáctico-discursivas de foco estrecho y expresividad en las secciones nucleares de las frases entonativas intermedias (ip) se manifiestan principalmente con las configuraciones abstractas H\* H- y L+H\* H-. La primera de ellas es la más empleada por la red social vecinal para marcar focos estrechos y expresividad; mientras que la red social laboral emplea ambas configuraciones con la misma frecuencia para marcar focos estrechos (por lo que no se puede decir que alguna de ellas sea prototípica) y favorece el uso de L+H\* H- para agregar un matiz expresivo en sus emisiones. Aunque se considera que los datos en relación con algunas funciones pueden ser escasos, el hecho de que los intertonemas abstractos H\* H- y L+H\* H- sean los más empleados en todas las instancias (independientemente de la función sintáctico-discursiva que esté en uso) lleva a concluir que

la característica que predomina en la sección del intertonema de ip es la marcación del sentido de continuidad (o no conclusión) del enunciado. Expuesto en términos de los códigos biológicos (Gussenhoven, 2004), la elevación melódica con sentido de no conclusión (código de producción) se superpone a cualquier otra configuración tonal asociada con focos estrechos o expresividad (código de esfuerzo). Sin embargo, esta afirmación no es categórica, debido a los pocos casos registrados en relación con las funciones sintáctico-discursivas analizadas.

Los tonemas descendentes ( $L^* L\%$ ,  $H^* L\%$ ,  $L+H^* L\%$  y  $H+L^* L\%$ ) son los más frecuentes en ambas redes sociales, entre los que destaca el tonema  $L^* L\%$  como claro elemento central en la categoría de tonemas de enunciados aseverativos en ambas redes sociales. De acuerdo con las investigaciones de la prosodia en español y con los códigos biológicos (Gussenhoven, 2004), esto es esperable debido al tipo de enunciado que se está analizando. Se identificó una brecha en el uso del tonema  $H+L^* L\%$ , que indica que esta configuración puede ser un elemento indexador de la red social laboral. El índice sociolingüístico de la red vecinal, por su parte, serían los tonemas no descendentes, no solo porque esta red es la que los utiliza con una frecuencia superior a la de la red laboral, sino porque casi todos los casos documentados en esta segunda red son producidos por un solo hablante. Cabe señalar que los tonemas abstractos no descendentes más frecuentes en la red vecinal ( $L+H^* H\%$  y  $H^* !H\%$ ) son realizados, en prácticamente todos los casos, como un sostenimiento tonal desde de la sílaba nuclear ( $=\%$ ). Lo anterior pone en relieve la importancia de esta juntura concreta, pues es utilizada para marcar el significado aseverativo (mediante el uso de un tono alto no ascendente en la sílaba final) y diferenciarlo de enunciados con junturas altas ascendentes ( $H\%$ ) que aportan otro tipo de significado.

Solo hay dos integrantes de la red social vecinal que no producen ningún caso de tonemas no descendentes, son los únicos que desarrollan actividades laborales fuera del ámbito de la red social y de la zona geográfica en donde esta se encuentra. Debido a que ambos son los únicos que prestan servicios profesionales a personas de distintos niveles socioeconómicos, es posible que estos informantes hayan reducido su producción de configuraciones nucleares no descendentes por considerarlas como no pertenecientes a la norma o no gozar de un prestigio abierto dentro de ciertos sectores del mercado lingüístico. Por lo tanto, y teniendo en cuenta la presencia de estos tonemas en el habla de los demás



hablantes de la red (que no desempeñan actividades laborales fuera de la zona en la que residen), se considera que estas configuraciones son un índice sociolingüístico que identifica a esta red.

En relación con la duración en el tonema de las frases entonativas mayores (IP), las medias absolutas de las sílabas que integran el tonema reflejan que la red social vecinal produce sílabas nucleares, en general, con mayor duración que las del linde, mientras que los datos de la red social laboral reflejan una tendencia opuesta. Por otra parte, el análisis de la duración relativa en ambas redes refleja la existencia de una diferencia relacionada con una aparente correlación entre distintos elementos prosódicos en las dos redes: en la red vecinal se ve una relación entre una mayor duración de la sílaba final de IP (en comparación con una disminución en la longitud de la sílaba nuclear) y el tonema  $L^* L-$ , al igual que con los tonemas que cuentan con acentos nucleares ascendentes ( $L+H^*$ ); mientras que en la red laboral la relación ocurre entre una mayor duración en la sílaba nuclear (en detrimento de la duración de la sílaba final) y la tonicidad de la sílaba, sin que el acento nuclear influya en dicha duración.

Respecto a la duración relativa en los tonemas, se observa que los patrones de producción están más relacionados con la variable de las redes sociales que con la del sexo. Tal vínculo se debe a que tanto hombres como mujeres de una red producen patrones bastante homogéneos que contrastan con los de la otra red. Ello difiere de lo registrado en el análisis de los tonemas internos, donde las mujeres utilizan un repertorio entonativo más amplio que el de los hombres. Esto puede ser señal de una mayor influencia de la red social en la construcción cognitiva del núcleo prosódico del enunciado, lo cual puede deberse a que el enunciado constituye la frase entonativa mayor (IP), el nivel más alto de la jerarquía prosódica (denominado “enunciado fonológico” por Nespor y Vogel, 2007). En un nivel inferior, como la frase entonativa intermedia (ip), existe mayor variabilidad entre los integrantes de las redes sociales.

Con base en todo lo expuesto en este capítulo y en los postulados de la teoría basada en el uso (Bybee, 2001), se puede decir que los elementos prosódicos que resultaron ser centrales o prototípicos (acento  $H^*$  en la primera sílaba tónica, el intertonema no oxítono  $H^* H-$  con mayor duración en la sílaba linde de ip, el intertonema oxítono  $L+H^* H-$  y el tonema  $L^* L\%$ , que no cuenta con un patrón de duración relativa central) indexan las características

sociales comunes a ambas redes. Esto se debe a que, en líneas generales, estos elementos también son los prototípicos en la estructura cognitiva de las dos redes sociales. Lo más parecido a una discrepancia en cuanto al estatus cognitivo de los elementos más frecuentes es que dos de ellos tengan un número de ocurrencias lo suficientemente cercanos como para no poder identificar un elemento central, por lo que ambos comparten la cualidad de prototipo, como es el caso de los intertonemas abstractos no oxítonos H\* H- y L+H\* H- en la red vecinal. En cualquier caso, estos dos tonemas internos son los más frecuentes en los datos generales del corpus (donde el más empleado es H\* H-, seguido por L+H\* H-), por lo que no se puede decir que los miembros de la red vecinal empleen patrones prosódicos periféricos dentro de la comunidad de habla, sino que ambos pueden ser prototípicos en sus estructuras cognitivas. Esto se debe a que ambas configuraciones son las más frecuentes en esta red, con un margen de diferencia bastante reducido, lo que disuade la idea de que alguno de ellos pueda ser un elemento secundario. Ahora bien, no se descarta la posibilidad de que el uso de uno u otro responda a la presencia de variables no contempladas en este estudio (como la presencia de un elemento focalizado poco antes del núcleo de la frase entonativa intermedia), por lo que sería necesario confirmar esta hipótesis en investigaciones posteriores.

Sin embargo, la centralidad en el empleo de ambas configuraciones nucleares de ip en la red vecinal puede ser un elemento portador de significado sociolingüístico relacionado con las características sociales de esta red. La brecha entre la mayor frecuencia con la que la red vecinal produce el tonema interno L+H\* H- y la menor frecuencia con la que lo emplea la red laboral, aunque en ambas es uno de los elementos más frecuentes, es suficientemente amplia como para asociar esta configuración con un rasgo prosódico de la red vecinal. Sin embargo, este tipo de brechas también ocurre con esquemas entonativos que, en relación con los datos generales del corpus, se pueden considerar periféricos en la estructura cognitiva debido a su baja productividad. Tal es el caso de los tonemas no descendentes, producidos de forma casi exclusiva por la red social vecinal (82 % de las producciones de estos tonemas pertenecen a esta red). Con vista en los datos generales del corpus, únicamente el 12.2 % de las configuraciones totales son no descendentes, lo que resalta su cualidad periférica. No obstante, justamente esto puede ser lo que haga resaltar su papel como posible índice sociolingüístico, ya que, a diferencia del intertonema abstracto L+H\* H-, este rasgo es tan

poco común en el corpus que posiblemente capture inmediatamente la atención de un miembro de la comunidad de habla; quien podría establecer la conexión entre el movimiento melódico y el significado social sin mucha dificultad. De hecho, es posible que los miembros de la red vecinal que brindan servicios a grupos sociales con modos de vida que implican mayor poder adquisitivo y grado de instrucción eviten producir estas configuraciones (o al menos reduzcan sustancialmente su uso) precisamente para distanciarse del significado social con el que pueden ser fácilmente asociadas.

Ahora bien, algunos factores que pueden asociarse a simple vista con elementos periféricos grupales pueden deberse a factores individuales, como es el caso de los tonemas no descendentes en el hablante Rs2H3, quien produce 5 de los 8 casos de las configuraciones con esta cualidad en la red laboral. La frecuencia de junturas medias (!H%) y altas (H%) no implica que este hablante posea características sociales que se puedan relacionar con la red social vecinal (de hecho, los datos socioeconómicos, obtenidos a través de la encuesta sociolingüística, confirman que no posee ninguna relación aparente con rasgos sociales de la red vecinal), sino que adquirió dicho rasgo prosódico por motivos que van más allá del alcance del instrumento de recolección de datos (lo que imposibilita una hipótesis acerca de su producción de tonemas no descendentes) y que estos tonemas se posicionaron en la periferia de su estructura cognitiva. En contraste, los elementos centrales en el uso de este hablante coinciden en gran medida con los del resto de los miembros de su red, por lo que no se puede percibir un distanciamiento entre las cualidades prosódicas del hablante Rs2H3 y el resto de los miembros de su red, más allá de elementos periféricos individuales. Por lo tanto, la única conclusión a la que se puede llegar, en cuanto a la presencia de configuraciones nucleares de IP no descendentes en el habla de Rs2H3, es que estas pudiesen representar un índice sociolingüístico relacionado con otra red social a la que pertenece este individuo, la cual comparte un grado de similitud considerable con la red social laboral.

Teniendo en cuenta lo observado en el análisis sociolingüístico, se puede concluir que los elementos prosódicos prototípicos de ambas redes sociales coinciden en la gran mayoría de los casos. En aquellos casos en los que una de las redes parece tener dos elementos prototípicos en alguna de las categorías estudiadas (como ocurre en la red vecinal con los intertonemas abstractos no oxítonos H\* H- y L+H\* H-), uno de ellos coincide con el elemento central de la otra red (que en el caso mencionado corresponde con H\* H-). Esto lleva a

concluir, al menos dentro del marco de esta investigación, que los prototipos prosódicos corresponden con características relacionadas con la comunidad de habla y no con los subgrupos que la constituyen. En otras palabras, los elementos prototípicos (los de uso más frecuente) consisten en rasgos prosódicos afines a la comunidad de habla y no a características específicas de los distintos grupos sociales. Los elementos prosódicos que se podrían considerar distintivos de cada grupo social son aquellos más frecuentes en un grupo que en otro, aunque no sean necesariamente exclusivos del primero. En otras palabras, los rasgos prosódicos propios de un grupo social dentro de una comunidad de habla son aquellos que consisten principalmente en elementos prosódicos periféricos, es decir, patrones con menor frecuencia de uso que los elementos centrales. Por lo tanto, lo que se puede ver en la muestra de datos es una serie de patrones prosódicos prototípicos, comunes a la comunidad de habla de la Ciudad de México, y una serie de rasgos periféricos que ocurren con una frecuencia considerablemente mayor en una red en particular, que son los que se pueden considerar como portadores de significado sociolingüístico.

Los datos y análisis presentados hasta este punto (tanto en el capítulo III como en este) representan los resultados y avances de la fase de producción. En el capítulo posterior (capítulo V) se abordará la metodología seguida durante el diseño de la prueba de percepción, junto con el análisis de los datos obtenidos mediante la aplicación de dicha prueba.

## Capítulo V

### Análisis de los resultados de la prueba de percepción

En este capítulo, se presentan todos los aspectos relacionados con la fase de percepción de este estudio. En primer lugar, se muestran algunos antecedentes de trabajos en los que se llevaron a cabo pruebas de percepción. Posteriormente, se describen los aspectos metodológicos relacionados con la fase de percepción. En tercer lugar, se exponen, analizan y discuten los resultados relacionados con los datos obtenidos en ambas partes de la prueba de percepción. Finalmente, se desarrolla una serie de consideraciones relacionadas con los aspectos más relevantes de los resultados de la prueba de percepción, sus repercusiones cognitivas y anotaciones metodológicas. Cabe señalar que los aspectos cognitivos del análisis se plantean en la sección final de este capítulo, y no en el análisis. Esto se debe a que, como se verá a lo largo de las secciones siguientes, las respuestas de los jueces, en la mayoría de los casos, no son muy diferentes en relación con las tendencias principales, lo que lleva a una posible explicación desde una perspectiva cognitiva que aplicaría para todos ellos simultáneamente y para la cual hay que tener en cuenta el panorama general.

#### 5.1. Antecedentes de trabajos de percepción

La percepción de los hablantes es un tema que ha cobrado auge en el campo de la lingüística en las últimas dos décadas, principalmente por la facilidad de acceso a los recursos tecnológicos que posibilitan este tipo de trabajos. Algunos de los estudios que se pudieron revisar emplearon estímulos consistentes en señales de audio con poca o ninguna manipulación; mientras que otros ejercen un mayor grado de control sobre los estímulos. En ambos casos, el procedimiento va acorde con las necesidades específicas de la investigación que los autores deseaban realizar.

En el grupo de trabajos que utilizaron señales de audio con poco o ningún control sobre los estímulos, se pueden encontrar las contribuciones de Abdurrahman (2014), Yi (2015), Pinillos (2016) y Martín Butragueño (2017). De entre estos trabajos, Martín Butragueño (2017) llevó a cabo su prueba de percepción con muestras de habla tomadas del Corpus sociolingüístico de la Ciudad de México (Martín Butragueño y Lastra, 2011), para hacer un primer acercamiento a las repercusiones del contacto dialectal en el nivel entonativo en el marco de la fonología variable; lo que lo ubica como el estudio que emplea estímulos

menos controlados y el que más privilegia la espontaneidad en el habla. No muy alejados de esta metodología, se encuentran procedimientos como el seguido por Abdurrahman (2014), quien utilizó fragmentos de un corpus leído para evaluar las actitudes lingüísticas hacia el inglés hablado por afroamericanos en Minnesota. Si bien este método busca restringir un poco la presencia de determinadas variables que se pueden dar en el habla semiespontánea, tampoco se puede considerar que ejerce un grado de control sustancial en la producción de los enunciados empleados como estímulos. Tal es la razón por la que este tipo de estímulos no se consideraron apropiados para cumplir los objetivos de la presente investigación pues, en este caso, era necesario incluir características prosódicas específicas (previamente identificadas durante la fase de análisis sociolingüístico) en los estímulos.

Existen otros diseños de pruebas de percepción en los que los investigadores dan instrucciones específicas a sus voluntarios al momento de grabar las señales de audio. Como ejemplo de esto, se pueden ver los trabajos de Pinillos (2016), quien estudió la duración de los sistemas vocálicos del español limeño y madrileño; Yi (2015), autora que empleó la técnica *matched guise* en su análisis de los factores lingüísticos y extralingüísticos que influyen en la elevación de la vocal /o/ junto con su significado social, y Lee (2016), cuyo trabajo consiste en el estudio de la prominencia de la voz laringizada y los índices sociolingüísticos evocados en su uso. En todos estos casos, los autores orientaron a sus voluntarios para que produjeran determinados rasgos lingüísticos, lo que puede implicar cierta falta de precisión en caso de que se desee estudiar un aspecto bastante concreto. Sin embargo, aunque los procedimientos seguidos por estos autores no se consideran los más efectivos para los planteamientos de este trabajo (al igual que los casos mencionados en el párrafo anterior), sí fueron eficientes para el cumplimiento de los objetivos particulares de cada una de las investigaciones.

Por otra parte, se pueden encontrar estudios en los que los autores ejercieron un mayor grado de control sobre los estímulos empleados en la prueba de percepción. Entre estos autores se pueden encontrar Face (2006 y 2007), Díaz-Campos y Ronquest (2007), Murrieta (2016), Ortega (2019) y Olivar Espinosa (en prensa). El autor que incluyó el menor número de alteraciones en sus estímulos fue Ortega (2019), pues en su caracterización de la entonación de distintos tipos enunciativos en Riohacha (Colombia) manipuló tres enunciados tomados de un corpus, para incluir solo cinco modificaciones consistentes en las variables

que deseaba analizar. Por su parte, Face (2006), al examinar cuatro factores cognitivos que pueden influenciar la percepción del acento en español, alteró digitalmente distintas palabras creadas *ad hoc*, para que todas las sílabas tuviesen la misma duración y frecuencia fundamental (F0). Un año después, este mismo autor (Face, 2007) creó estímulos consistentes en enunciados parciales (reproducidos mediante un sistema de esclusas) y totales, en los que el patrón de F0 fue modificado para contener todas las combinaciones posibles de puntos de inflexión de acuerdo con el tipo de enunciado. Díaz-Campos y Ronquest (2007), en cambio, con el propósito de determinar si los acentos tonales L+H\* y L\*+H tienen un carácter fonético o fonológico en español, manipularon la frecuencia fundamental (F0) en los acentos tonales, de tal forma que coincidieran con L+H\* en los casos prenucleares y con L\*+H en los nucleares, al tiempo que el resto de la F0 se alteró para que fuese uniforme (es decir, sin ascensos ni descensos). Murrieta (2016), quien se propuso analizar el comportamiento del umbral de percepción de las variaciones en la F0, incluyó en su prueba estímulos en diferentes contextos lingüísticos (fonemático, silábico y frástico), con variaciones tonales de una sílaba a otra. Finalmente, Olivar Espinosa (en prensa) modificó la F0 de distintos enunciados, para que la F0 fuese lineal y, sobre esta base, ajustó los valores de frecuencia de los acentos tonales H\*, L+>H\* y L+H\* en las sílabas prominentes de acuerdo con las medias obtenidas en una etapa previa, lo que resultó en la generación de ocho patrones melódicos distintos.

Este último grupo de trabajos guarda una relación más estrecha con la presente investigación, pues en todos ellos se consideró necesario tener un mayor grado de control experimental sobre las características prosódicas que debían estar presentes en los estímulos. Como se pudo observar, el número y grado de las modificaciones llevadas a cabo en cada caso fueron distintos. Esto no responde a un criterio concreto o ideal que deba estar presente en todas las pruebas de percepción, sino a las necesidades particulares que tiene cada autor para poder cumplir con los objetivos de sus respectivos trabajos. En cualquier caso, todos ellos consideraron indispensable sacrificar un poco de la espontaneidad en el habla en pro del estudio de la percepción de eventos prosódicos bastante específicos que, posiblemente, fuesen difíciles de encontrar en habla espontánea o de replicar mediante la ayuda de un voluntario.

Antes de pasar a la sección siguiente, es importante tener en cuenta que, tal como ocurre con muchos otros aspectos metodológicos, las pruebas de percepción no siguen un marco rígido en su diseño. Como se puede concluir tras observar cada uno de los trabajos mencionados en este apartado, el procedimiento seguido en la creación de una prueba de percepción debe estar orientado por los objetivos de cada investigación. Por lo tanto, el empleo de un mayor o menor grado de control experimental depende de las características de cada estudio. Por ejemplo, mientras que Martín Butragueño (2017) trabajó con un corpus de hablantes nativos de sus regiones de origen para su estudio geolectal, Face (2006) utilizó palabras creadas *ad hoc* en su estudio del español general.

## 5.2. Metodología de la fase de percepción

La percepción, junto con la frecuencia de las producciones fónicas, desempeña un papel fundamental en el establecimiento de categorías y prototipos, pues de acuerdo con ello se generan las posibilidades de variación (Moreno Fernández, 2012). Por tal motivo, se consideró necesario llevar a cabo un análisis perceptivo. El objetivo principal de este estudio de percepción es determinar si las características prosódicas, identificadas como posibles índices prosódicos durante la fase de producción, son percibidos como estos por parte de los jueces oyentes. De esta forma, se busca contrastar los datos cuantitativos del análisis sociolingüístico con la percepción de los miembros de la comunidad de habla. Con base en los datos obtenidos, se pretende confirmar la hipótesis de que los índices prosódicos no están constituidos únicamente por un único elemento, sino por el efecto conjunto de, al menos, dos parámetros (concretamente, la duración y la entonación).

La hipótesis fue puesta a prueba a través de un experimento con estímulos que se diseñaron a través de programa Praat 6.1.06 (Boersma y Weenink, 2019). Estos estímulos consisten en modificaciones de duración y F0 en la sílaba que contiene el intertonema oxítono en frases entonativas intermedias (ip) y en las sílabas que conforman el tonema de las frases entonativas mayores (IP) (ver §5.2.2).<sup>29</sup> Debido a que los estímulos consistieron en

---

<sup>29</sup> No se tomaron en cuenta para la prueba los intertonemas no oxítonos debido a que no se encontraron indicios de posibles índices sociolingüísticos en los patrones de duración entre las dos sílabas que los contienen. De igual modo, no se consideraron para estas pruebas las primeras sílabas tónicas ya que, a pesar de que en ellas sí se identificaron posibles índices sociolingüísticos relacionados con la entonación y la duración, resulta difícil poder evaluar la percepción de estos índices en términos prácticos.



enunciados aislados (es decir, carentes de un contexto lingüístico) y a que la percepción entonativa se basa en contrastes tonales con respecto al entorno prosódico (pues no se puede saber que un tono es alto hasta que se cuente con otro tono con el cual se pueda establecer una referencia), no se podría contar con suficientes puntos de referencia tonales, previos a la primera sílaba tónica, que les permitieran a los jueces tener una idea del campo tonal del enunciado hasta ese punto (lo que no es un problema en una situación comunicativa real o en un contraste de pares). Por tal motivo, se considera posible que los jueces no hubiesen sido capaces de identificar exitosamente los índices ubicados en esta posición de los estímulos. Adicionalmente, alterar las cualidades del material previo a la primera sílaba tónica no es algo imposible en términos técnicos, pero el resultado de dicha manipulación sería un enunciado con un inicio que resultaría, muy posiblemente, altamente artificial para los jueces y que no representaría la realidad lingüística del español del centro de México. Por lo tanto, un estímulo que no sea posible relacionar con el habla de esta ciudad no aportaría ningún beneficio ni dato relevante para la presente investigación.

Con los datos obtenidos en esta fase del estudio, se pretende dar respuesta a la siguiente interrogante: ¿las respuestas de los jueces confirman las cualidades indexicales los rasgos identificados como potenciales índices sociolingüísticos durante el análisis de los datos de habla? Con ello, se complementa la información obtenida en la fase de producción, se puede profundizar en la identificación del estatus cognitivo (prototípicos o periféricos de cada red social) de los rasgos prosódicos con significado social y se asienta una base metodológica para el estudio de índices prosódicos con métodos estadísticos y perceptivos.

#### 5.2.1. Variables

La variable dependiente es netamente lingüística, pues está constituida por los rasgos prosódicos existentes en los estímulos. Dichos rasgos consisten en una serie de combinaciones de duración y de contornos melódicos nucleares de frases entonativas (explicados detalladamente *infra* en §5.2.2.1 y §5.2.2.2), obtenidos a través del análisis sociolingüístico realizado durante la fase de producción. Por otra parte, la variable independiente consta de las características sociales con las que los jueces asocian cada uno de los estímulos presentados. Dichas características se representarán a través de una serie de

indicadores directamente relacionados con los distintos modos de vida tomados en cuenta en esta investigación (§2.1.2).

### 5.2.2. Estímulos

En la prueba de percepción, se pusieron a prueba distintos tipos de patrones prosódicos, tanto en frases entonativas intermedias (ip) como en mayores (IP). En las intermedias, se estudiaron los intertonemas oxítonos, es decir, las configuraciones nucleares de ip que culminan con sílaba tónica; mientras que en las mayores se analizaron los tonemas (configuraciones nucleares de IP). Para estudiar los dos tipos de frases entonativas se emplearon los enunciados mostrados en (1) y (2), extraídos del corpus recolectado (en la fase de producción). En cada uno, se buscó la identificación de los índices sociolingüísticos explicados en los apartados §5.2.2.1 y §5.2.2.2. Cabe señalar que la selección de los enunciados no fue un proceso aleatorio, pues estos debían ser neutros en términos de género, es decir, que no ofrecieran ninguna pista acerca del sexo de la persona que los produjo. Adicionalmente, no debían contener el nombre del informante, de algún conocido de ellos o de la empresa para la que trabajaban. De igual modo, se descartaron aquellos en los que se diera algún tipo de información relacionada con el trabajo, nivel de estudios o alguna otra característica de los hablantes que pudiese asociarse con sus rasgos sociales, como es el caso de los enunciados Rs1M3-15 (*[(Ya no continué con la prepa)<sub>ip</sub>]*IP), Rs1H2-31 (*[(Lo que sí me pegó mucho fue)<sub>ip</sub> (la marihuana)<sub>ip</sub>]*IP), Rs2H1-1 (*[(Es una moneda)<sub>ip</sub> (que te garantiza la desintermediación)<sub>ip</sub> (de un tercero)<sub>ip</sub>]*IP) o Rs2H3-19 (*[(Cuando entramos al banco)<sub>ip</sub> (hubo un apoyo)<sub>ip</sub>]*IP). Todas estas medidas se tomaron en cuenta, debido a que se consideró que factores como los mencionados podían condicionar las respuestas de los jueces durante su ejecución de la prueba de percepción.

Tras evaluar cada una de las opciones posibles de acuerdo con los criterios ya mencionados, los enunciados seleccionados fueron los siguientes:

- 1) *[(La mezcla cultural de ese lugar)<sub>ip</sub> (lo hacía muy divertido)<sub>ip</sub>]*IP
- 2) *[(Defendía mucho a mi hermano)<sub>ip</sub>]*IP

El enunciado (1) es producto de una modificación realizada al enunciado Rs2M3-5 del corpus: *[(Esa mezcla como de cultural [sic] incluso en el mismo lugar)<sub>ip</sub> (lo hacía muy divertido)<sub>ip</sub>]*IP. La decisión de no utilizar el enunciado tal como fue producido por la

informante se debe a que la señal original contiene un poco de ruido del ambiente en el que fue grabada la entrevista y la informante produce la sección nuclear de la IP con voz laringizada. Esta laringización afecta directamente el resultado de la modificación posterior de frecuencia fundamental (F0) y duración en Praat, en cuanto a que el resultado suena sumamente artificial. Adicional a esto, la primera ip del enunciado original tiene un error sintáctico y tiene una extensión de 19 sílabas, lo que la excluye de la longitud promedio de la unidad entonativa del español, que se ubica entre cinco y diez sílabas (Navarro Tomás, 1966; Gili Gaya, 1975; Sosa, 1999; y Martín Butragueño, 2006). Con las modificaciones implementadas en la primera ip, se elimina el error sintáctico y se ajusta la extensión a 10 sílabas. Por otra parte, la selección de este enunciado del corpus responde a que la cesura entre el sujeto sintáctico de la oración (generalmente, como en este caso, en forma de sintagma nominal) es altamente recurrente en el corpus.

En cuanto al enunciado (2), este es el resultado de una pequeña modificación en el enunciado Rs1M1-18: [(*Yo defendía mucho a mis hermanos*)<sub>ip</sub>]IP. La modificación del enunciado original en este caso se debe a los siguientes motivos: se observó debilitamiento vocálico en la sílaba final del enunciado (lo que afecta el resultado después de la manipulación de la señal de audio), la presencia del pronombre personal de la primera persona del singular (*yo*) al inicio de la cláusula puede presentar un matiz focalizador en términos sintácticos; y se consideró pertinente preservar la sonoridad al final de la IP al eliminar los morfemas de pluralidad (fonéticamente representados como [s]) en *mis hermanos*.

Ambos enunciados mencionados fueron grabados con una voz femenina. Previo a la grabación, la voluntaria tuvo oportunidad de leer los enunciados tantas veces como consideró necesarias para garantizar una mayor fluidez y naturalidad al momento de la grabación. Posteriormente, se le indicó que leyera los dos enunciados de la forma más natural posible, prestando atención a realizar la cesura indicada luego de la primera frase entonativa intermedia (ip) en el primero de ellos (cesura que fue marcada con una coma en el texto que se le entregó). De igual manera, se le pidió que prestase especial atención a la pronunciación plena de las sílabas finales de las frases entonativas, con lo cual se pretendió evitar la pronunciación del enunciado con debilitamientos vocálicos, voz laringizada o susurrada. Adicionalmente, debido a que uno de los índices de la sílaba final del enunciado que se

analiza en esta prueba consiste en una duración prolongada de esta (como se explica en el §5.2.2.2), se le pidió a la voluntaria que produjera cinco repeticiones del enunciado [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub> con una mayor duración en la sílaba final, con lo cual se buscó evitar que dicha sílaba sonara altamente artificial tras realizar las modificaciones que se explican en §5.2.2.2. Una vez que se hizo la grabación de los enunciados, se revisaron en Praat (Boersma y Weenink, 2019) para confirmar la producción plena de todas las sílabas del enunciado y, debido a la presencia de voz laringizada y debilitamientos vocálicos en algunas sílabas (especialmente en las finales de los enunciados), se procedió a grabar los enunciados una segunda vez. Tras revisarlos en Praat, se pudo confirmar que no existía fonación no modal ni debilitamientos, por lo que se dio por finalizada la etapa de grabación. Finalmente, todos los enunciados fueron manipulados a través del programa Praat (Boersma y Weenink, 2019). para obtener los estímulos que se emplearon en cada una de las pruebas.

Además de lo ya mencionado, en cada estímulo se modificaron los parámetros globales de frecuencia, para que cada uno pareciera producido por dos personas diferentes: un hombre y una mujer. De tal forma, un mismo estímulo sin ningún tipo de modificación adicional en cuanto a las características específicas que se describirán en §5.2.2.1 y §5.2.2.2 fue reproducido con dos timbres de voz diferentes: el original y uno masculino. Con esto, al igual que en el ttabajo de Díaz-Campos y Killam (2012), aunque con un método diferente, se evitó la familiarización de los jueces con la voz de un solo hablante. Dicha alteración del timbre de voz original de los estímulos no representaría un problema, teniendo en cuenta los resultados del estudio de Clopper *et al.* (2005), quienes determinaron que los patrones de reconocimiento no pierden su integridad con independencia del sexo del hablante.

Antes de continuar con la descripción del proceso para la creación de los estímulos, es necesario hacer una precisión metodológica. Si bien se puede considerar que el empleo de enunciados producidos naturalmente durante las entrevistas sociolingüísticas es más oportuno para esta tarea, pues habrían sido producidos con tanta espontaneidad y autenticidad como la situación lo hubiese permitido, el objetivo de este trabajo llama a tomar consideraciones especiales en este sentido. Los datos obtenidos en la búsqueda de índices sociolingüísticos a través del enfoque basado en el uso (que se detallan en §4.2, §4.3 y §4.4) son bastante concisos en términos de los patrones prosódicos con potencial significado social. Esto implica que los elementos sometidos a la prueba de percepción deben ser igualmente

detallados, por lo que se consideró necesario controlar tanto el material léxico como el prosódico de los enunciados, para favorecer el alcance del objetivo sociolingüístico y evitar que variables no consideradas dentro de los márgenes de este estudio pudiesen interferir en los resultados finales. De esta forma, tras tomar las medidas que se juzgaron oportunas para eliminar otras variables que también pudiesen contener o ayudar a generar significado social, el foco de la prueba se acota únicamente a las variables prosódicas que se consideran, hasta este punto, potencialmente relevantes en términos sociolingüísticos. Sin embargo, el procedimiento mismo de la selección y modificación del material lingüístico para la prueba de percepción fue sometido a análisis a la luz de los resultados y las consideraciones al respecto se presentan en las conclusiones de esta investigación.

#### 5.2.2.1. Intertonemas oxítonos en ip

Los estímulos para evaluar la percepción de los intertonemas oxítonos consistieron en una serie de modificaciones del enunciado [(*La mezcla cultural de ese lugar*)<sub>ip</sub> (*lo hacía muy divertido*)<sub>ip</sub>]<sub>IP</sub>. Se crearon dos baterías de estímulos sobre la base de este enunciado. La primera cuenta con un reducido número de cambios con respecto a la grabación original; mientras que la segunda posee un número más elevado de alteraciones que buscan un mayor grado de control experimental en las características prosódicas del enunciado. En esta etapa, se crearon ambas baterías de estímulos con la intención de someterlas a una prueba de validación (§5.2.5), cuyos datos permitieron determinar cuál de las dos podía resultar más eficiente para alcanzar el objetivo planteado en esta fase de la investigación.

En el proceso de creación de la primera batería de estímulos, las modificaciones prosódicas que se llevaron a cabo involucran únicamente los rasgos suprasegmentales considerados como índices sociolingüísticos. Esto quiere decir que únicamente se modificó la F0 y la duración de la sílaba en la que se encuentra el intertonema oxítono de acuerdo con los parámetros descritos *infra*. Sin embargo, debido a que se consideró necesario aportar un punto de referencia claro para la percepción del contraste entonativo en dicha sílaba, también se alineó la F0 de la sílaba previa en el centro del rango tonal medio del enunciado. De esta manera, los ascensos de F0 presentes en la última sílaba de la primera ip, todos en el rango alto del campo tonal, tendrían una transición más fluida. El resultado de estas modificaciones se ejemplifica en las figuras 5.1 y 5.2, en las que se pueden ver las cualidades de la señal de

audio original, en contraste con una señal de audio base sobre la que se hicieron modificaciones posteriores, de acuerdo con los criterios mencionados más adelante en este mismo inciso.

Figura 5.1. Oscilograma, espectrograma y F0 del enunciado original [(*La mezcla cultural de ese lugar*)<sub>ip</sub> (*lo hacía muy divertido*)<sub>ip</sub>]<sub>IP</sub>

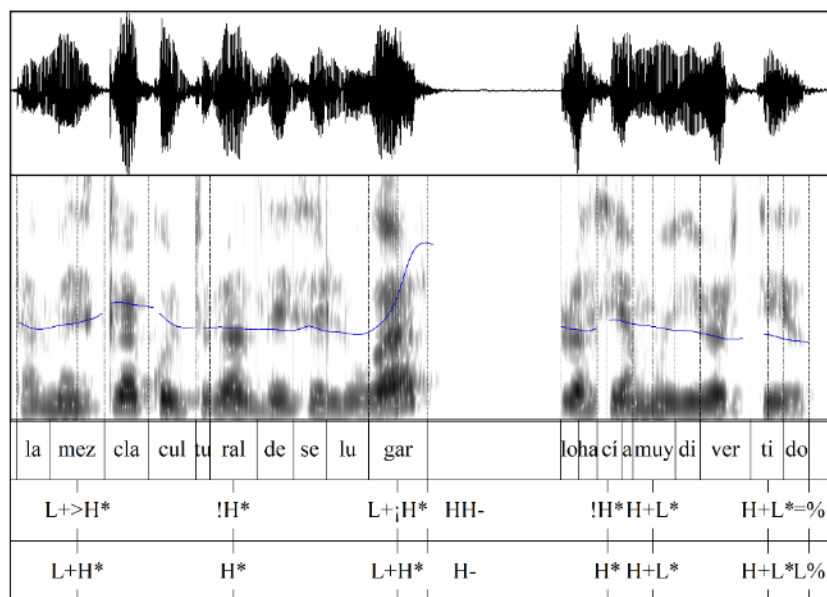
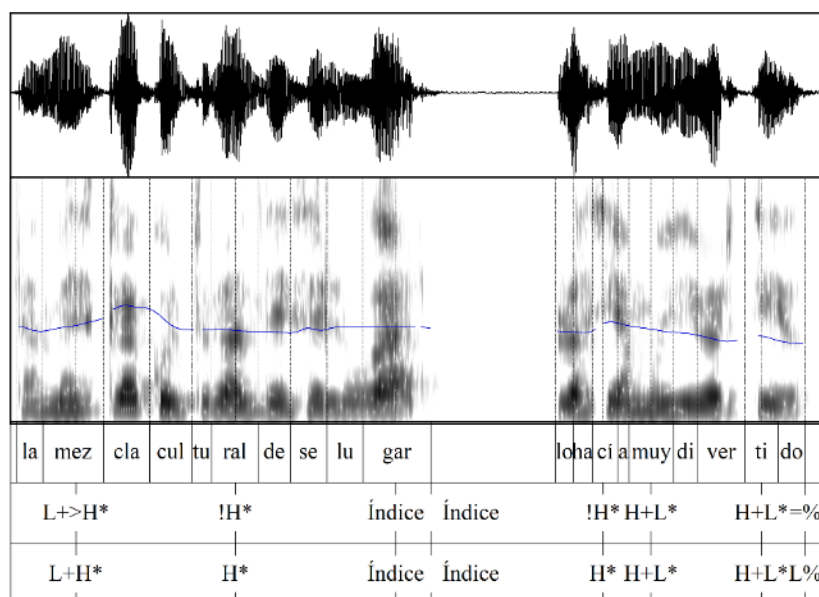


Figura 5.2. Oscilograma, espectrograma y F0 del enunciado base [(*La mezcla cultural de ese lugar*)<sub>ip</sub> (*lo hacía muy divertido*)<sub>ip</sub>]<sub>IP</sub> según criterios de la primera batería de estímulos

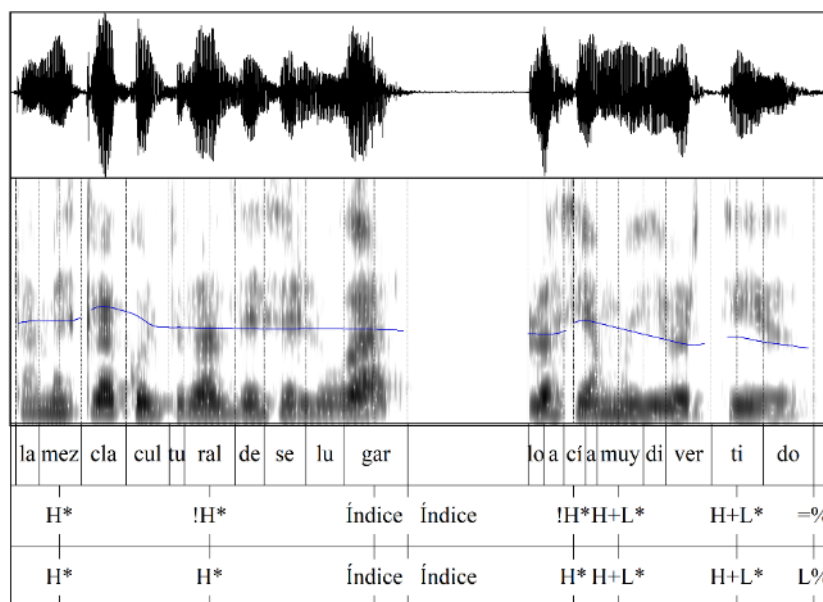


El proceso de creación de la segunda batería de estímulos puede ser dividido en dos etapas. En la primera, se dividió el campo tonal del enunciado en tres rangos: bajo, medio y alto. Para ello, se calculó la diferencia entre el techo y el piso tonal y esta se dividió en tres, lo que permitió obtener el valor de rango tonal (*vrt*). De tal manera, al sumar el *vrt* al valor del suelo tonal se identificó el rango bajo, al repetir la operación con el *vrt* y el límite superior del rango bajo se determinó el rango medio, y al sumar el *vrt* al tono más alto del rango medio se calculó el rango alto. Una vez que se contó con estas secciones del campo tonal, se construyó la señal de audio base, es decir, una señal de audio que solo recibiría modificaciones adicionales en la sección que corresponde con los índices sociolingüísticos que se estudiaron (que en este caso es la sílaba donde se localiza el intertonema).

Para construir la base, primero se modificó la F0 de la primera sílaba tónica y de todo el material fonético previo al punto medio del rango tonal alto, de manera que la primera sílaba tónica fuese portadora del acento tonal H\*. Además, se manipuló la duración de dicha sílaba para que fuese de 147 ms (con esto, los rasgos de la sílaba se ajustan al acento prototípico y a la duración promedio en el corpus, como se menciona en §4.3). Las dos sílabas siguientes (*cla* y *cul*) preservaron sus tonos originales; pero en todas las sílabas posteriores,

hasta el final de la penúltima de la primera ip, se ubicó la F0 en el punto medio del rango tonal medio. Adicionalmente, se modificó la duración de la sílaba previa al intertonema a 141 ms, valor que corresponde con la media de duración de esta sílaba en el corpus. En la segunda ip, se preservó la F0 del enunciado original, pues no se presentó ningún fenómeno distinto a los movimientos tonales más frecuentes del corpus, ya que cuenta con una declinación tonal a lo largo de la ip que concluye con el tonema L\* L%. Sin embargo, las duraciones de las sílabas nuclear y final del enunciado se ajustaron a 178 ms y 175 ms, respectivamente, para que coincidieran con las duraciones medias registradas en el corpus para cada sílaba. En la figura 5.3 se presenta la señal de audio resultante de las modificaciones descritas en este párrafo (comparar con la señal original en la figura 5.2).

Figura 5.3. Oscilograma, espectrograma y F0 del enunciado base [(La mezcla cultural de ese lugar)<sub>ip</sub> (lo hacía muy divertido)<sub>ip</sub>]<sub>IP</sub> según criterios de la segunda batería de estímulos



En la segunda etapa, se modificaron los rasgos prosódicos de duración y entonación en las sílabas que constituyen los intertonemas oxítonos, que son el punto focal de esta prueba. La primera alteración que se realizó en esta sílaba consistió en el ajuste de la F0 para generar las configuraciones concretas: H\* H-, H\* !H-, L+H\* H- y L+<sub>i</sub>H\* HH-. Aunque la única configuración nuclear que se identificó como un posible índice sociolingüístico



(correspondiente a la red social laboral) es  $L+\uparrow H^* HH-$ , las demás configuraciones fueron tomadas en cuenta debido a que representan otras tendencias principales en el corpus (tal es el caso de  $H^* H-$ ) y otras que se encuentran en la periferia ( $cH^* !H-$ ) que cuentan con más de 3 casos de ocurrencia en cada red y que, por lo tanto, se considera que pueden cobrar relevancia al conjugarse con los patrones de duración asociados con cada red social. Por lo cual, la presencia de estas otras configuraciones responde a dos propósitos: el primero es el de cumplir con el papel de distractores y el segundo es el de ayudar a confirmar (o no) la hipótesis operativa de las pruebas, según la cual los índices prosódicos están conformados por una combinación de rasgos entonativos y de duración.

La construcción de las configuraciones en términos fonéticos consistió en una manipulación de la  $F_0$  en la sílaba tónica al final de la ip (que engloba el núcleo y el linde de la frase entonativa intermedia). En el caso de la configuración  $H^* H-$ , la altura de  $F_0$  se ubicó 2 st por encima del valor de la sílaba prenuclear. Por otra parte, la manipulación acústica de  $L+H^* H-$  consistió en un ascenso de 2 st desde el final de la sílaba prenuclear hasta el final del enunciado. Por último, se realizó una elevación tonal de 4.3 st desde el final de la sílaba prenuclear hasta el final de la frase entonativa intermedia (ip) para presentar la configuración  $L+\uparrow H^* HH-$ , pues ese es el valor promedio de la elevación tonal registrada en el corpus, en este intertonema oxítono.

Debido a que los rasgos melódicos estudiados en los intertonemas oxítonos ocurren en la última sílaba de la ip, las modificaciones en los valores de duración se realizaron desde la sílaba prenuclear (átona). Esto se debe a que se considera necesario controlar la duración de la sílaba prenuclear, para poder brindar un punto de referencia prosódico con el cual se pueda contrastar la longitud de la sílaba nuclear. Los patrones de duración se describen a continuación, en los que  $D$  representa un valor de 141 ms, que es la media de duración de las sílabas del cuerpo del enunciado:<sup>30</sup>

- $D \wedge D$  (ambas sílabas tienen la misma duración)
- $D \wedge D + 103$  (de esta forma se alcanza la media de duración de la sílaba nuclear de los hablantes de la red vecinal, que es de 244 ms)

---

<sup>30</sup> Los patrones están escritos de acuerdo con las convenciones matemáticas, según las cuales  $\wedge$  se lee como la conjunción y, a la vez que la suma se expresa a través de una cruz griega (+). En relación con las sílabas de las configuraciones tonales, aquello previo al símbolo  $\wedge$  se atribuye a la sílaba previa a la nuclear y lo posterior a la sílaba asociada con el intertonema.

- $D \wedge D + 165$  (lo que quiere decir que la sílaba nuclear durará 306 ms, la media de duración de esta sílaba en las producciones de la red laboral).

El total de combinaciones resultantes de configuraciones melódicas y patrones de duración, cada una de las cuales constituye un estímulo, fue de 12 en cada batería de estímulos. Cada estímulo fue repetido 2 veces, una de las repeticiones con el timbre de voz original de la voluntaria y la otra con el rango tonal modificado para que pareciera producida por un hablante masculino. En total, se contó con una batería de 24 estímulos.

#### 5.2.2.2. Tonemas de IP

Antes de iniciar con la descripción del procedimiento de manipulación del enunciado, es necesario acotar que, como se mencionó previamente, se grabaron dos versiones del enunciado [*Defendía mucho a mi hermano*]<sub>IP</sub>: una de ellas con una duración normal en la sílaba final (es decir, sin que la voluntaria recibiera instrucciones específicas en este sentido) y otra con un alargamiento intencional en la sílaba final. Esto se debe a que uno de los índices corresponde con una duración prolongada en dicha sílaba. Esta duración prolongada resultó ser difícil de alcanzar partiendo de la modificación artificial de una sílaba con duración relativamente normal sin que sonara altamente artificial. Por ello, se optó por solicitarle a la voluntaria que produjera este enunciado con alargamiento, con lo que se obtuvo una calidad considerablemente más natural, tras reducir la duración producida por ella, para que se ajustara a los parámetros que se explican un poco más adelante en este mismo inciso.

De la misma manera que con los intertonemas, se realizaron dos baterías de estímulos. La primera de ellas consistió en modificaciones de F0 y de duración únicamente en las sílabas en las que se encuentran los índices sociolingüísticos, es decir, la nuclear y la final del enunciado, mientras que el resto del enunciado preservó los valores prosódicos de la grabación original. Los enunciados originales y las bases generadas con este método se presentan en las figuras 5.4, 5.5, 5.6 y 5.7. La segunda batería, en cambio, se construyó en dos etapas. En la primera, se dividió el campo tonal de las dos señales de audio en tres rangos, de acuerdo con el procedimiento descrito en §2.3.2.1. Luego, se ubicó la F0 de la primera sílaba tónica y del material previo al punto medio del rango tonal alto y se ajustó la duración de la sílaba tónica a 147 ms (el valor medio de la duración de esta sílaba en el corpus). El resto del enunciado, con excepción de las dos sílabas en la que se ubica el tonema, se mantuvo

con los mismos valores de F0 y duración del enunciado original. Los oscilogramas, espectrogramas y F0 de las dos bases creadas se pueden ver en las figuras 5.8 y 5.9.

Figura 5.4. Oscilograma, espectrograma y F0 del enunciado original [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub> con duración normal en la sílaba final

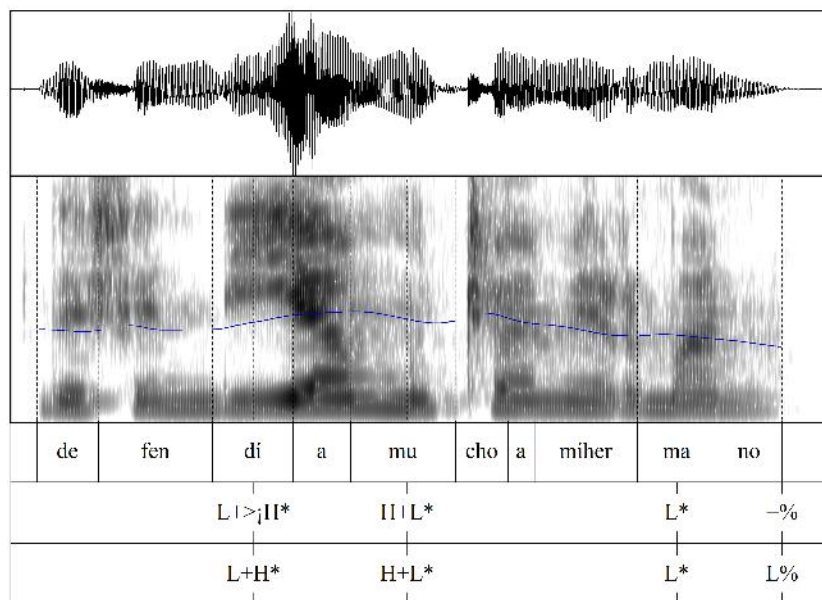


Figura 5.5. Oscilograma, espectrograma y F0 del enunciado base [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub> según criterios de la primera batería de estímulos con duración normal en la sílaba final

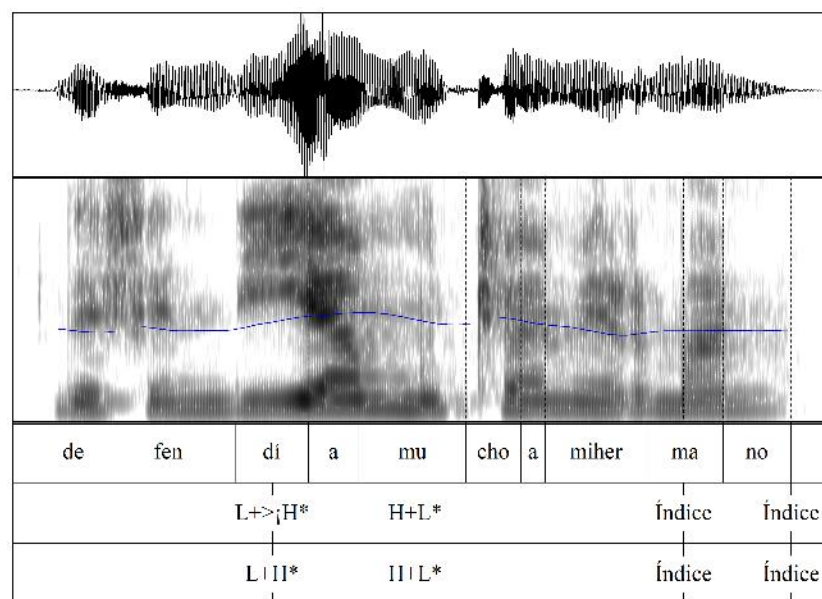


Figura 5.6. Oscilograma, espectrograma y F0 del enunciado original [(Defendía mucho a mi hermano)<sub>ip</sub>]IP con duración modificada en la sílaba final

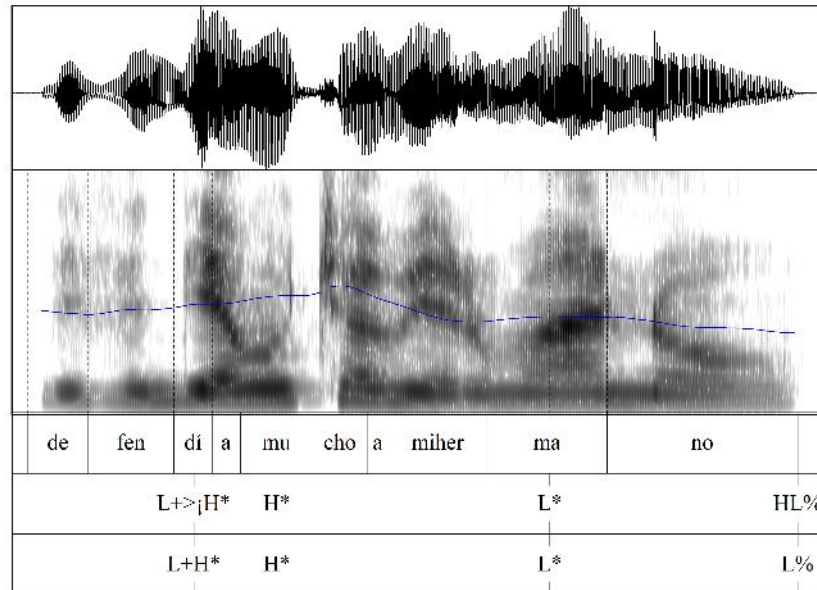


Figura 5.7. Oscilograma, espectrograma y F0 del enunciado base [(Defendía mucho a mi hermano)<sub>ip</sub>]IP según criterios de la primera batería de estímulos con duración modificada en la sílaba final

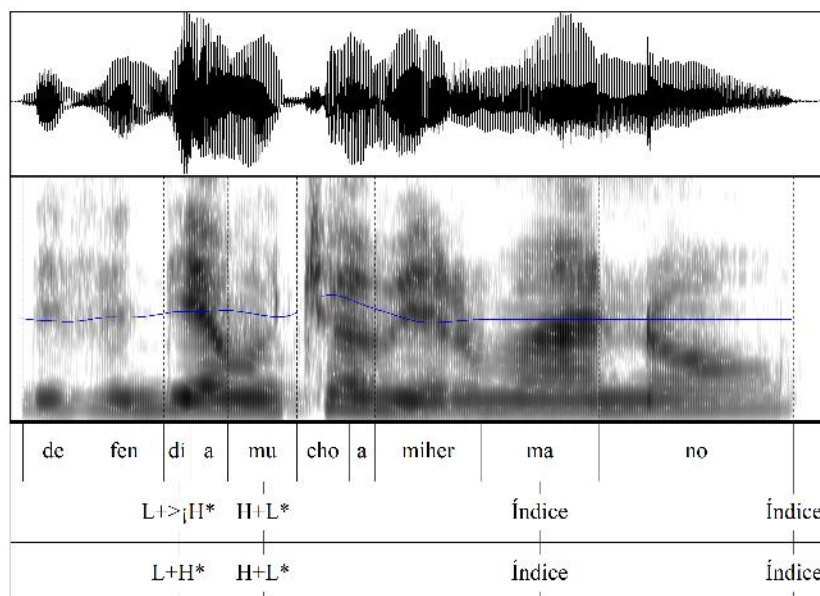


Figura 5.8. Oscilograma, espectrograma y F0 del enunciado base [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub> según criterios de la segunda batería de estímulos con duración normal en la sílaba final

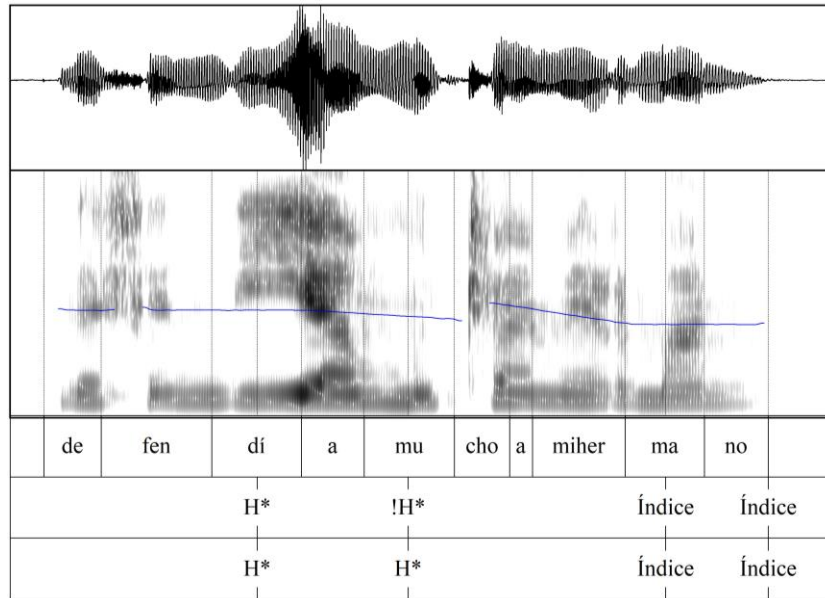
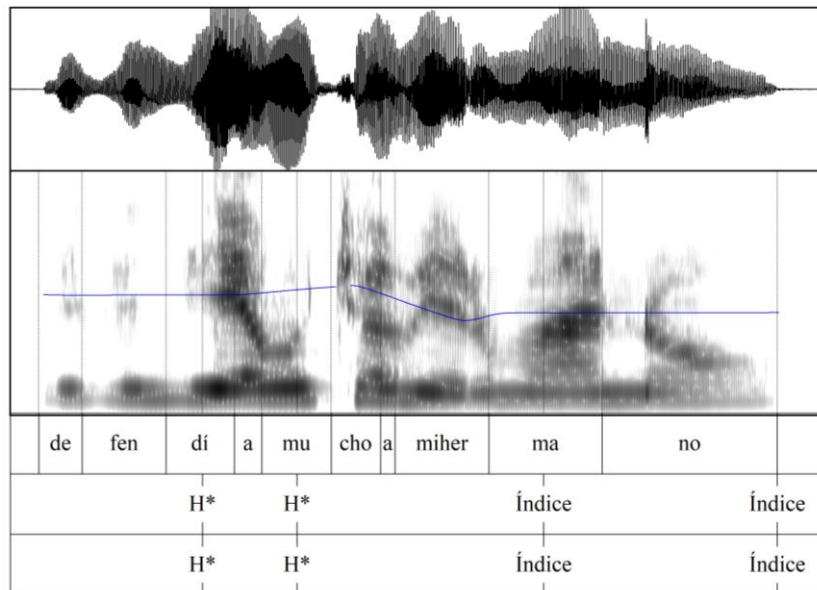


Figura 5.9. Oscilograma, espectrograma y F0 del enunciado base [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub> según criterios de la segunda batería de estímulos con duración modificada en la sílaba final



Las dos últimas sílabas del enunciado grabado para esta prueba, que corresponden con el tonema, se modificaron para corresponder con las siguientes configuraciones melódicas concretas:

- a. Configuraciones no descendentes: !H\* =%, L+H\* =%.
- b. Configuraciones descendentes: L+H\* L%, H\* L%, L\* L% y H+L\* L%.

Las configuraciones no descendentes (que no concluyen con una juntura L%) se han considerado, hasta este punto, como posibles índices prosódicos de la red social vecinal, por lo que se seleccionaron aquellas realizaciones concretas más frecuentes; mientras que la configuración H+L\* L% se presume como un índice de la red laboral. Los otros 3 tonemas (L+H\* L%, H\* L% y L\* L%) conforman las tres configuraciones más frecuentes en ambas redes sociales, por lo que se emplearon como distractores que ayudaron, a su vez, a dilucidar si estos rasgos generales pudiesen llegar a ser identificados como índices sociolingüísticos de alguna red junto con algún patrón de duración.

Las modificaciones tonales que se llevaron a cabo en el tonema, para que la F0 coincidiera con los tonemas concretos descritos previamente, son los siguientes: i) el acento nuclear L\* tuvo una altura 2 st inferior en relación con el tono de la sílaba anterior; ii) el acento tonal ascendente L+H\* y el descendente H+L\* tuvieron un movimiento melódico de 2 st dentro de la sílaba nuclear de acuerdo con su respectiva orientación; iii) la juntura L% se ubicó 2 st por debajo del acento nuclear L\* y 4 st de H\* y L+H\* (así se evitó que la juntura baja quedase a la misma altura que el nivel medio del enunciado) ; iv) se mantuvo sostenido el tono con el que culmina la sílaba nuclear hasta el final del enunciado en aquellos casos en los que la juntura era =%.

Por otra parte, en esta prueba se manejaron 3 patrones de duración. El primero de ellos presenta un rasgo prototípico de la red social laboral (de acuerdo con los datos del análisis sociolingüístico), el cual consiste en una mayor duración de la sílaba nuclear con respecto a la final, por lo que se ajustaron los valores de estas sílabas a 212 ms y 157 ms, respectivamente. Esto se debe a que la media de la duración de la última sílaba del enunciado en los datos de la red laboral corresponde con 157 ms, en tanto que la sílaba nuclear (en el promedio de los casos en los que la duración de esta es mayor) es 55 ms más larga que la de la final. Las otras dos configuraciones, que requieren medidas menos específicas y en las

que  $D$  es igual a 178 ms (el valor medio de la duración de la sílaba nuclear en el corpus), son las siguientes:<sup>31</sup>

- $D \wedge D$  (la sílaba nuclear y la final tienen la misma duración)
- $D \wedge (D \cdot 2) + 20$  (la sílaba final tiene un valor igual al doble del de la nuclear más 20 ms, lo que es igual a 376 ms, que es la media de la duración de las sílabas finales producidas por la red vecinal).

Con estos tres patrones de duración, se incluyeron en la prueba dos elementos que se consideran índices sociolingüísticos de las dos redes sociales (el primero para la red laboral y el último para la vecinal) y uno adicional. Este último desempeñó el papel de distractor y puso a prueba la hipótesis operativa mencionada previamente, según la cual los patrones de duración únicamente representan un índice sociolingüístico cuando se conjugan con las configuraciones melódicas nucleares.

El total de configuraciones posibles tras la combinación de los tonemas y los patrones de duración fue de 18 elementos para cada batería de estímulos. Al igual que con los intertonemas oxítonos, cada uno de los estímulos resultantes fue reproducido 2 veces, una vez con la voz original de la voluntaria y otra con una modificación para que simule una voz masculina, lo que suma un total de 36 estímulos.

### 5.2.3. Jueces

Debido a que con estas pruebas se buscó describir un fenómeno dentro de una población, el número de voluntarios requeridos para la muestra no debía ser necesariamente significativo en términos estadísticos (Strauss y Corbin, 2002). De igual forma, si se deseara tomar una muestra representativa del universo poblacional de la comunidad de habla que se está estudiando, su dimensión resultaría imposible de manejar en términos prácticos, lo mismo que podría ocurrir con el volumen de los datos recolectados en ese caso. Por lo tanto, para determinar el valor de la muestra se apeló al concepto de “saturación teórica”, de acuerdo con el cual se alcanza la saturación cuando se considera que la teoría ya está construida y que

---

<sup>31</sup> Al igual que en los patrones de duración descritos para la P1, los símbolos empleados en estos patrones corresponden con las convenciones matemáticas:  $\wedge$  representa la conjunción y, la cruz griega (+) implica una suma y el punto (.) una multiplicación. En la representación de las sílabas que constituyen el tonema, la información previa al símbolo  $\wedge$  es atribuible a la sílaba nuclear y lo posterior a la sílaba final de IP.

los datos adicionales no aportan información que contribuya a su desarrollo (Glaser y Strauss, 1967; Corbin y Strauss, 2002).

Para la ejecución de la prueba, se contó con la participación de 12 jueces. Teniendo en cuenta la cantidad de estímulos base en cada prueba (y que cada uno de ellos se presenta dos veces), cada patrón prosódico obtuvo un total de 24 datos de percepción, lo que generó un total de 720 respuestas (288 de las cuales contuvieron información relacionada con los intertonemas oxítonos y 432 de los tonemas). Tal cantidad de respuestas se consideró suficiente para observar una tendencia clara en los datos recolectados.

En relación con las características sociales de los jueces que constituyeron la muestra poblacional, todos ellos debían ser miembros de la comunidad de habla de la Ciudad de México (nacidos en ella o haber llegado a esta antes de los 5 años de edad) y tener una edad comprendida entre 22 y 57 años de edad (el mismo rango etario que el de los informantes del corpus oral). Debido al hecho de que se maneja este enfoque de comunidad de habla desde el planteamiento inicial del trabajo, para esta prueba no se tomó en cuenta únicamente a miembros de un grupo de dicha comunidad, sino que cualquier miembro de ella fue un juez potencial. Sin embargo, se manejó un sistema de cuota, según el cual se representan equitativamente los tres modos de vida, al igual que los dos sexos, planteados en este trabajo (§2.1.2). Por lo tanto, la muestra estuvo constituida por dos hombres y dos mujeres con características sociales compatibles con el modo de vida 1, al igual que con los modos de vida 2 y 3, como se muestra en el cuadro 5.1.

Cuadro 5.1. Distribución de los jueces de acuerdo con sus rasgos sociales

Sexo	Modos de vida			Total
	1	2	3	
Hombres	2	2	2	6
Mujeres	2	2	2	6
Total	4	4	4	12

Cada participante fue contactado directamente por el investigador, ya sea porque era una persona conocida o porque se compartía un contacto en común (alguien conocido por el investigador que conoce al potencial participante). Esto garantizó que los jueces de la prueba



contaran con los requisitos estipulados previamente, ya que se pudo obtener la información necesaria de primera mano, antes de solicitar explícitamente su participación en el experimento.

#### 5.2.4. Instrumento

La prueba de percepción se diseñó a través del programa OpenSesame (Mathôt *et al.*, 2012) en su versión 3.3. Este programa permite diseñar experimentos psicológicos a través de una interfaz visual y subirlos a internet a través de su extensión OSWeb. Esta extensión consiste en una biblioteca de JavaScript que interpreta y ejecuta los experimentos creados en OpenSesame, aunque no es compatible con todas las posibilidades que ofrece el *software* (al menos hasta el momento en el que se realizó esta investigación). Debido a que el programa emplea principalmente codificación en el lenguaje Python, la mayoría de las funciones con las que cuenta el programa solo pueden ser ejecutadas en una computadora en la que esté instalado el *software*. Para realizar experimentos en línea, como se plantea en esta investigación, las funciones están limitadas a aquellas que el programa puede “traducir” (por decirlo de alguna manera) al lenguaje Java, que es el empleado por los servidores JATOS (Lange *et al.*, 2015) en los que se carga el experimento para que pueda ser realizado por los participantes.

Desde el servidor JATOS se creó un enlace único para cada participante, es decir, un vínculo a la prueba que puede ser utilizado una sola vez. Los servidores JATOS ofrecen esta posibilidad como una herramienta que otorga mayor control a los investigadores, en relación con los sujetos que participarán en los experimentos que se cargan a la web. De esta manera, a través del control de la persona a la que se le envía cada enlace, se pudo saber exactamente a quién corresponden cada una de las respuestas recolectadas por el servidor. Cabe señalar que a cada participante se le informó que el enlace puede ser utilizado una sola vez y se le pidió que completara la prueba en un solo intento. Adicionalmente, se le solicitó a cada uno que informara al investigador en caso de que hubiese habido un problema durante su participación en el experimento (por ejemplo, que se cierre el navegador accidentalmente) para que se le pudiera enviar un nuevo enlace.

La prueba tuvo tres secciones: introducción, desarrollo y agradecimiento. En la introducción, primero que nada, se dio un agradecimiento inicial por aceptar participar en la

prueba y se mostró un breve resumen del proyecto de tesis, en el cual no se dio información específica acerca del objetivo del proyecto, para evitar condicionar a los jueces (la información contenida en esta parte se muestra en el apéndice 6). Luego, se dieron las instrucciones generales (véase apéndice 7), la dinámica de la prueba (ver apéndice 8) y se llevó a cabo una prueba de sonido que contó con dos reproducciones de una señal de audio distinta a cualquiera de las empleadas en la sección principal del experimento (véase apéndice 9). A esto le siguió, como parte final de la sección introductoria, una sección de práctica (las instrucciones se pueden ver en el apéndice 10), en la que cada participante debió seguir el mismo procedimiento que realizaría posteriormente en la parte principal del experimento, con la finalidad de que se familiarizara con la dinámica de la prueba y el sistema de respuestas. Cabe destacar que las preguntas que se realizaron en esta sección de práctica no fueron las mismas que se emplearon en el ejercicio principal, aunque sí tuvieron el mismo sistema de respuesta con teclado (véase apéndice 11).

El desarrollo de la prueba consistió en dos actividades. La primera tuvo como objetivo principal identificar las asociaciones que los oyentes establecían entre cada estímulo, especialmente aquellos que incluían índices sociolingüísticos, y los distintos modos de vida considerados en esta investigación. Al dar inicio a este primer ejercicio, se mostró un mensaje que indicaba la culminación de la práctica y el comienzo de la primera sección de la prueba principal (ver apéndice 12). En esta parte, cada juez debió apretar cualquier botón de su teclado para escuchar una señal de audio (seleccionada al azar por el *software*) una única vez. Posterior a esto, los participantes respondieron ante los siguientes planteamientos (A-C) a través de una escala de Likert; y contestaron la pregunta D mediante la selección de una de las posibles respuestas ofrecidas (los elementos entre paréntesis corresponden con las teclas asignadas para recolectar las respuestas):

A. *¿Qué tan de acuerdo estás con la siguiente afirmación?*

*Esta persona es de la Ciudad de México.*

Escala de Likert: En desacuerdo (1) - (2) - (3) - (4) - (5) - (6) - (7) De acuerdo.

B. *¿Qué tan de acuerdo estás con la siguiente afirmación?*

*Esta persona habla igual o de forma similar a tus familiares y amigos más cercanos.*

Escala de Likert: En desacuerdo (1) - (2) - (3) - (4) - (5) - (6) - (7) De acuerdo.

C. *¿Qué tan de acuerdo estás con la siguiente afirmación?*

*Esta persona habla igual o de forma similar que tú.*

Escala de Likert: En desacuerdo (1) - (2) - (3) - (4) - (5) - (6) - (7) De acuerdo.

D. *¿A cuál de las siguientes actividades crees que podría dedicarse esta persona?*

Posibles respuestas: (Q) *Vende en la calle* - (J) *Enseña en una secundaria* - (P) *Diseña sistemas informáticos en un banco* - (T) *Limpia coches* - (B) *Estudia último año en la universidad* - (D) *Gerencia una empresa internacional*

Antes de proceder a justificar la selección de estas preguntas, es pertinente hacer algunas aclaraciones relacionadas con los aspectos técnicos de esta parte. Debido a que la versión para web de OpenSesame (OSWeb) no permite poner los formularios de texto en orden aleatorio, las preguntas se presentaron en el mismo orden mostrado previamente. El orden de las preguntas se decidió en función de su grado de especificidad para evitar, en la medida de lo posible, que la respuesta dada en una de ellas condicione la respuesta de la siguiente. De tal manera, se inicia con una pregunta que hace referencia a la comunidad de habla (A), para continuar con una relacionada con el habla de personas del entorno cercano del informante (B), seguida por otra relacionada directamente con su estilo de vida (C) y culminar con una que solicita un mayor grado de especificidad en relación con una cualidad social del hablante que escucharon (D). Adicionalmente, se decidió emplear una escala de Likert con siete puntos, porque se considera que esto les permitirá a los jueces poder seleccionar una opción que represente su opinión con un mayor grado de detalle. Finalmente, se seleccionaron los botones de respuesta de la última pregunta, porque están suficientemente separadas en el teclado como para evitar que los jueces puedan pulsar una opción no deseada erróneamente y para aumentar el estado de alerta o consciencia durante el proceso de respuesta.

En relación con los motivos de selección de las interrogantes presentadas, se diseñó la primera afirmación (A) para poder corroborar si las características prosódicas de la señal de audio son relacionadas con la prosodia de la Ciudad de México. Este elemento tiene una función de confirmación debido a que también se realizó en la prueba de validación (explicada *infra* en §5.3.5). Sin embargo, se consideró oportuno incluirlo en la prueba final debido a que cumple la función de presentar un contexto situacional macro (comunidad de habla), antes de ahondar en las preguntas relacionadas con los modos de vida, que tienen un

carácter más específico y, en el caso de las dos preguntas siguientes, más personal para el participante.

Las afirmaciones B y C se formularon con la intención de saber si los jueces podían relacionar los rasgos prosódicos escuchados con cualidades sociales similares a las suyas, pues se relacionan tanto con su círculo social más cercano (B) como con su modo de vida (C). Aunque se podría pensar que existe cierto nivel de solapamiento entre estos dos elementos, cada uno de ellos busca conocer elementos distintos en cuanto a la percepción del oyente. Con la afirmación B, se buscó plantear una asociación de la prosodia escuchada con la de personas cercanas al oyente (su círculo de relaciones más cercano), es decir, se trata de establecer una conexión indirecta entre los patrones prosódicos escuchados y sus propios rasgos sociales, a través de la asociación del audio con un vínculo existente con sus personas más allegadas (quienes muy posiblemente compartan su mismo modo de vida). La afirmación C, en cambio, plantea una asociación directa entre lo escuchado y características sociales del propio oyente, ya que se trata de observar si el participante considera que la prosodia que escuchó puede corresponder con alguien con su mismo modo de vida. Con estas dos respuestas se busca observar si existe ausencia de reconocimiento, por parte del oyente, de los índices sociolingüísticos identificados –en la fase de producción– para su modo de vida cuando se trata de sí mismo o su propia habla, pero no cuando se trata de alguien más o del habla de alguien de su entorno. Por ejemplo, si un participante con modo de vida 1 responde con un 7 en la afirmación B y un 1 en la C, tras haber escuchado una señal de audio con índices sociolingüísticos estadísticamente relacionados con su modo de vida, es posible señalar que este reconozca dichos índices en el habla de sus allegados, pero no en su propia habla (lo que no resultaría totalmente extraño). En un caso como este, la última pregunta servirá para disipar alguna posible duda en cuanto a la asociación de los índices sociolingüísticos con respecto a los modos de vida.

Finalmente, la pregunta D pretende conocer la asociación entre los patrones prosódicos escuchados en las señales de audio con determinadas labores relacionadas con los distintos modos de vida. Esto se hace a través de un planteamiento que no implica al participante en el contexto de la pregunta, lo que genera una distancia entre sí mismo y los índices sociolingüísticos, lo suficientemente segura como para responder con honestidad sin sentirse involucrado. En este caso, se cuenta con tres distractores, en cierta forma, entre las

opciones que se ofrecen como respuestas. La razón por la que no se consideran distractores totales es que cada modo de vida está asociado a dos respuestas: (Q) *Vende en la calle* y (T) *Limpia coches* están relacionados con el modo de vida 1, (J) *Enseña en una secundaria* y (B) *Estudia último año en la universidad*, con el modo de vida 2 y (P) *Diseña sistemas informáticos en un banco* y (D) *Gerencia una empresa internacional*, con el modo de vida 3.

Al finalizar la primera parte, se mostró un mensaje en el que se indicó el fin del primer ejercicio y se dieron las instrucciones para el desarrollo de la segunda parte (véase apéndice 13). La segunda actividad de la prueba consistió en una prueba de contraste, cuyo objetivo es determinar cuáles son las pistas prosódicas más asociadas a la comunidad de habla de la Ciudad de México. En esta etapa, se presentó una señal de audio que contenía dos enunciados idénticos, cuya única diferencia radicaba en la configuración del intertonema o del tonema, según fuese el caso. Cabe destacar que los enunciados empleados en esta parte fueron los mismos utilizados para la actividad anterior y que los audios (cada uno consistente en una pareja de enunciados) fueron reproducidos en un orden aleatorio definido por el *software*. Como se puede ver en el cuadro 5.2, se tomaron en cuenta, para esta segunda fase, todos los patrones melódicos de intertonemas oxítonos empleados en el ejercicio previo ( $H^* !H^-$ ,  $H^* H^-$ ,  $L+H^* H^-$  y  $L+H^* HH^-$ ) y solo 4 de los tonemas ( $L+H^* =\%$ ,  $L+H^* L\%$ ,  $H+L^* L\%$  y  $L^* L\%$ ). La selección de estos tonemas se debe a que corresponden con dos posibles índices sociolingüísticos ( $L+H^* =\%$  y  $H+L^* L\%$ ), la configuración prototípica de todo el corpus, es decir, la más frecuente en ambas redes ( $L^* L\%$ ) y el tonema circunflejo ( $L+H^* L\%$ ), identificado como uno de los rasgos más característicos del habla del centro de México (Kvavik, 1979; Martín Butragueño, 2004, 2011, 2021; Matluck, 1952; Sosa, 1999).

Cuadro 5.2. Configuraciones entonativas incluidas en el ejercicio de contraste

Intertonemas oxítonos	
Pares de configuraciones	Tipo de voz
L <sub>+</sub> iH* HH- / L+H* H-	Femenina
H* !H- / L <sub>+</sub> iH* HH-	Masculina
H* H- / L <sub>+</sub> iH* HH-	Femenina
L+H* H- / H* !H-	Masculina
H* H- / L+H* H-	Femenina
H* !H- / H* H-	Masculina
Tonemas	
Pares de configuraciones	Tipo de voz
H+L* L% / L+H* =%	Femenina
L+H* L% / H+L* L%	Masculina
L* L% / H+L* L%	Femenina
L+H* =% / L+H* L%	Masculina
L* L% / L+H* =%	Femenina
L+H* L% / L* L%	Masculina

El patrón de duración seleccionado para todos los enunciados de la prueba fue  $D\Delta D$ . Tal selección se debe a que los otros patrones de duración se consideran índices sociolingüísticos, hasta este punto; mientras que  $D\Delta D$  se estima neutral, dado que no es asociable (estadísticamente) con ninguna de las dos redes sociales en concreto. Por otra parte, para evitar un efecto rutinario y monótono en los estímulos de esta actividad, la mitad de ellos fueron reproducidos con la voz original de la voluntaria, mientras que la otra mitad consistían en señales manipuladas para que sonaran como voces masculinas.

Luego de haber escuchado cada par de enunciados, se les preguntó a los jueces cuál audio consideraban que se parecía más al habla de la Ciudad de México. Para responder esta pregunta, los jueces pudieron escoger entre alguna de las siguientes opciones (las letras entre paréntesis corresponden con las teclas que debían pulsar para seleccionar su respuesta): *Primero (Q) / Segundo (L)*. Al igual que en la última pregunta del ejercicio anterior, se asignaron las teclas de respuestas en el teclado, considerando su ubicación en los extremos, con la finalidad de evitar que los jueces seleccionaran una opción no deseada de forma accidental.

Tras haber concluido con el desarrollo de la prueba, se presentó un mensaje en el que se agradeció la participación del oyente (visible en el apéndice 14). En general, el tiempo de

ejecución de toda la prueba, desde que se activó la primera pantalla hasta el mensaje de agradecimiento, varió entre 23 y 27 minutos.

#### 5.2.5. Validación y prueba piloto

Antes de completar la preparación de los estímulos e iniciar la construcción de la prueba en el *software*, se llevó a cabo una prueba de validación para evaluar las señales de audio manipuladas con Praat, en términos de asociación con el habla de la Ciudad de México, artificialidad y espontaneidad. Posteriormente, se realizó una prueba piloto con la finalidad de comprobar que la dinámica era apropiada para el cumplimiento de los objetivos que involucraron su diseño.

Para comprobar la calidad de los estímulos, se tomaron 6 señales de audio de cada batería de preguntas, 3 de ellos con la voz original de la voluntaria y 3 manipulados para que sonaran con voz masculina. De igual forma, 3 señales fueron variaciones del enunciado [(*La mezcla cultural de ese lugar*)<sub>ip</sub> (*lo hacía muy divertido*)<sub>ip</sub>]<sub>IP</sub> y las otras 3 fueron modificaciones de la frase entonativa mayor [(*Defendía mucho a mi hermano*)<sub>ip</sub>]<sub>IP</sub>. Que se emplearan 3 enunciados de cada tipo por cada batería responde al hecho de que se buscó determinar si cada una de las modificaciones de duración en las sílabas mencionadas en § 5.2.2 (3 patrones de duración en ip y 3 en IP) afectaba la naturalidad (en términos de artificialidad y espontaneidad) de los audios implementados en la prueba. En relación con los patrones de entonación, se seleccionaron 3 al azar, tanto para las secciones nucleares de ip como para las de IP. Adicional a esto, se incluyeron 6 grabaciones de enunciados adicionales que funcionaron como distractores. De estas señales adicionales, 3 consistieron en los mismos enunciados utilizados en los estímulos reales, pero con modificaciones acústicas en la duración y F0 para que sonaran claramente artificiales. Las otras 3 señales de audio estuvieron constituidas por los mismos enunciados previos, pero con algunas modificaciones para que se asemejaran a una variedad de español mexicano distinta de la del centro de México. De esta forma, la prueba estuvo compuesta por 18 estímulos distribuidos según se muestra en el cuadro 5.3.

Cuadro 5.3. Distribución de los estímulos en la prueba de validación de estímulos

Tipo de voz	Batería de estímulos 1		Batería de estímulos 2		Distractores		Total
	ip	IP	ip	IP	Claramente artificial	Ajeno a la Ciudad de México	
Voz masculina	2	1	2	1	2	1	9
Voz femenina	1	2	1	2	1	2	9
Total	3	3	3	3	3	3	18

Las modificaciones acústicas de todos los distractores se llevaron a cabo con el *software* Praat. Las alteraciones realizadas en el enunciado [(*La mezcla cultural de ese lugar*)<sub>ip</sub>] (*lo hacía muy divertido*)<sub>ip</sub>]IP para que sonara artificial constaron de un descenso de 2 st, desde el inicio del enunciado hasta el punto medio de la segunda sílaba tónica, desde donde se generó un ascenso de 2 st hasta la frontera de la primera ip. La segunda ip inicia con un tono a la misma altura que la del inicio del enunciado, punto desde el cual se formó un descenso de 6 st hasta el final de la segunda tónica de esta ip, que continúa con un ascenso de 4 st hasta el final del enunciado. Adicional a esto, se duplicó la duración de las dos primeras sílabas tónicas del enunciado, de la final de la primera ip y de todas las tónicas de la segunda frase entonativa intermedia, excepto la de la nuclear. El resultado de estas modificaciones se puede ver en la figura 5.10. Por otra parte, para lograr que el enunciado [(*Defendía mucho a mi hermano*)<sub>ip</sub>]IP sonara artificial, se provocó un descenso de 9 st desde el inicio del enunciado hasta el final de la primera sílaba tónica, desde donde se sostuvo el mismo tono hasta el final de la sílaba nuclear, para luego elevarlo 9 st desde dicho punto hasta el final del enunciado. Adicionalmente, se duplicó la duración de la primera sílaba, de las dos tónicas prenucleares, de la sílaba nuclear y de la final. En la figura 5.11 se puede ver el producto de todas las alteraciones realizadas en este enunciado.



Figura 5.10. Oscilograma, espectrograma y F0 del enunciado distractor notoriamente artificial [(La mezcla cultural de ese lugar)<sub>ip</sub> (lo hacía muy divertido)<sub>ip</sub>]<sub>IP</sub>

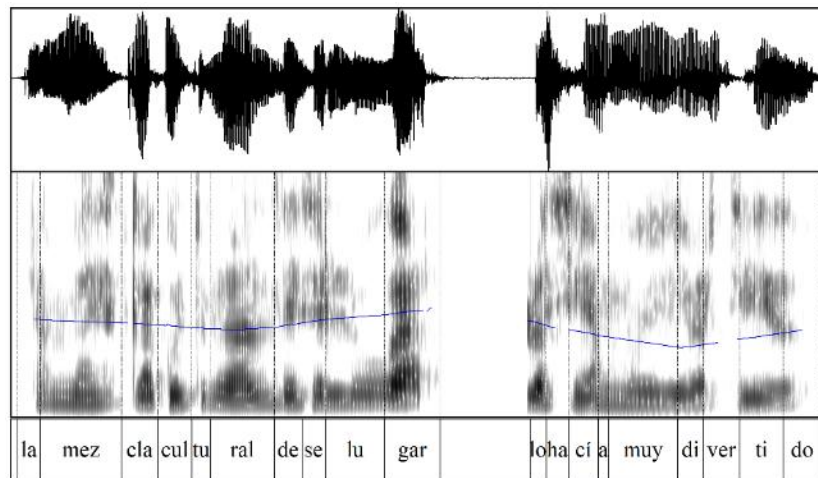
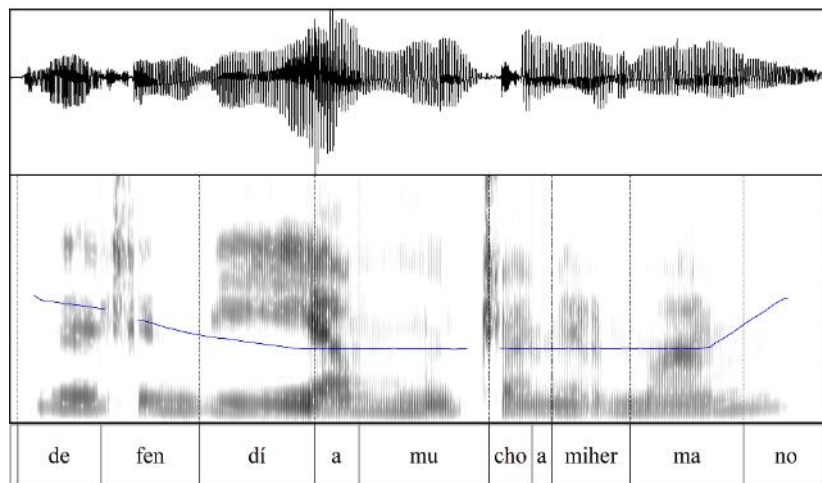


Figura 5.11. Oscilograma, espectrograma y F0 del enunciado distractor notoriamente artificial [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub>



Los distractores manipulados para sonar como si hubiesen sido producidos por un hablante no perteneciente a la Ciudad de México se construyeron sobre las bases de duración diseñadas para ambos enunciados descritas en §5.2.2.1 y §5.2.2.2. Estas bases no incluyen las modificaciones específicas de las secciones nucleares (tanto de ip como de IP) que representan el punto focal de la prueba de percepción. Las alteraciones de la F0 se realizaron con base en la curva tonal registrada en una muestra de habla del español venezolano (en concreto de Caracas). Para ello, se le pidió a un voluntario venezolano que pronunciara los



Figura 5.13. Oscilograma, espectrograma y F0 del enunciado distractor no comunitario [(Defendía mucho a mi hermano)<sub>ip</sub>]<sub>IP</sub>

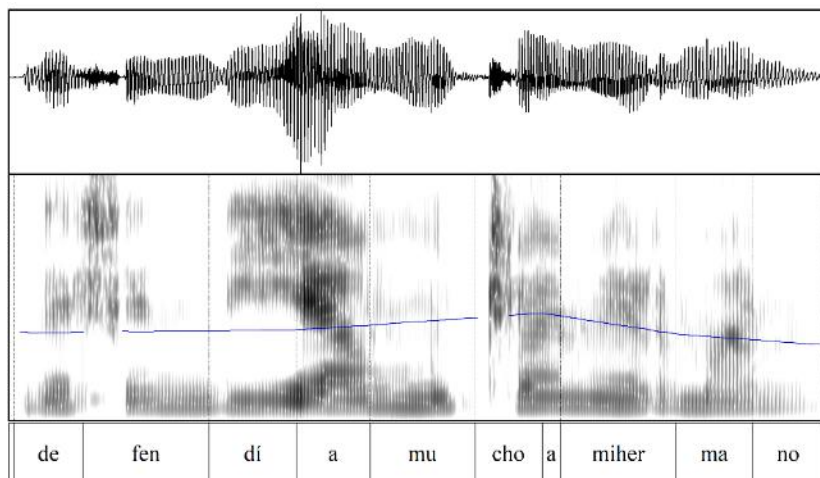


Figura 5.14. Oscilograma, espectrograma y F0 del enunciado [(La mezcla cultural de ese lugar)<sub>ip</sub> (lo hacía muy divertido)<sub>ip</sub>]<sub>IP</sub> grabado por el voluntario venezolano

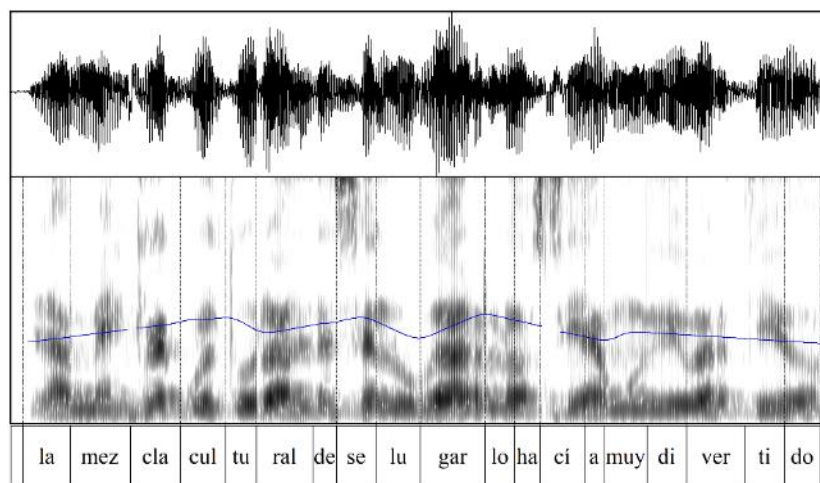
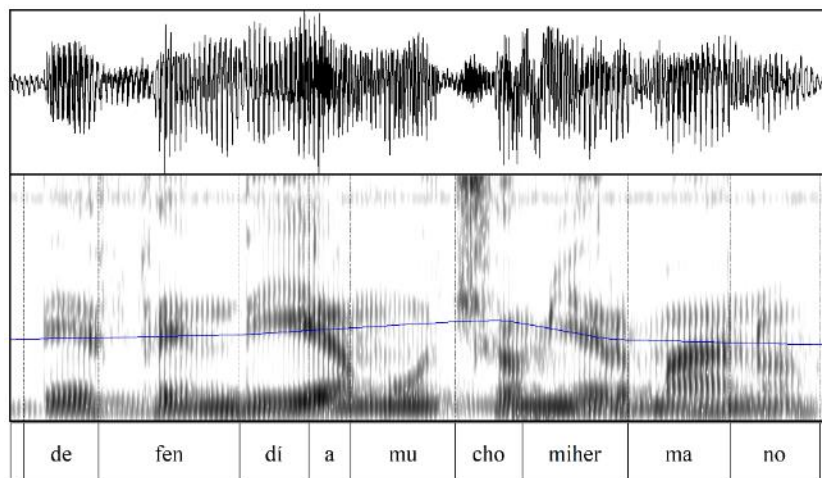


Figura 5.15. Oscilograma, espectrograma y F0 del enunciado [(Defendía mucho a mi hermano)<sub>ip</sub>] grabado por el voluntario venezolano



La prueba de validación inició con un mensaje de bienvenida (mostrado en el apéndice 15), seguido de las instrucciones para la prueba (véase apéndice 16). Posteriormente, se les preguntó a los jueces si habían nacido en la Ciudad de México o habían llegado a ella antes de los 5 años de edad, a lo cual podían responder *sí* o *no*. En la parte central de la prueba, tras reproducir cada una de las 18 señales de audio en orden aleatorio, se les pidió a los jueces que indicaran, por medio de una escala de Likert idéntica a la mostrada en §5.2.4, qué tan de acuerdo estaban con las siguientes afirmaciones (presentadas en este orden):<sup>32</sup> i) *Esta frase fue dicha por alguien de la Ciudad de México*; ii) *La frase suena artificial*; y iii) *La frase suena espontánea*. Tras haber completado todas las secciones, se presentó un mensaje que indicaba la culminación de la prueba (véase apéndice 17). Es importante señalar que la ejecución de esta prueba no les tomó a los participantes más de 8 minutos (en promedio) y que fue realizada por 9 jueces voluntarios. Las respuestas obtenidas se tomaron en cuenta para determinar cuál de las dos baterías de estímulos se incluiría en la versión final de la prueba.

Tras analizar los resultados de la prueba de validación, se decidió utilizar la primera batería de estímulos, que es aquella en la que se controlaron únicamente la F0 y la duración

<sup>32</sup> Como se explicó previamente, no se puede establecer un orden aleatorio en la presentación de las preguntas en la extensión OSWeb. Adicionalmente, el valor en la escala de Likert podía ser seleccionado únicamente mediante el uso de botones del teclado.

en las sílabas correspondientes al tonema interno y al tonema. En ambos grupos de estímulos, los promedios de la evaluación de la artificialidad se encontraron ligeramente por encima del punto medio de la escala de Likert, con una diferencia bastante reducida (4.28 para la primera batería y 4.20 para la segunda), lo cual concuerda con que hayan sido evaluadas por debajo del punto medio en términos de la espontaneidad, aunque con una diferencia un poco mayor en esta ocasión (3.48 para la primera batería y 3.98 para la segunda). Estos datos quieren decir que ambos grupos de estímulos fueron considerados artificiales y no espontáneos.

No obstante, en ninguna de estas dos evaluaciones se registró una diferencia significativa entre las puntuaciones promedio otorgadas a los distintos estímulos. Caso contrario a lo ocurrido con el grado de asociación de la muestra con el habla de la Ciudad de México, donde la diferencia de medias resultó ser significativa ( $p = 0$ , según prueba ANOVA de un factor). Para dicho criterio, la media de la primera batería de estímulos (4.43) no solo es mayor que la de la segunda (3.83), sino que su promedio se encuentra por encima del punto medio de la escala de evaluación. En resumen, lo que inclinó la balanza a favor de la selección del primer grupo de estímulos fueron los siguientes aspectos: la similitud en los resultados relacionados con la artificialidad y espontaneidad en ambas baterías, el hecho de que las medias estadísticas no fueron significativas en dichos criterios y la diferencia significativa a favor de una asociación más favorable de los estímulos menos controlados con el habla de la Ciudad de México.

La prueba piloto, cuyo objetivo fue determinar si el planteamiento de la prueba definitiva era viable o si necesitaba algún cambio, tuvo exactamente la misma estructura descrita en §5.2.4 y, en vista de lo expuesto en el párrafo anterior, se realizó con la primera batería de estímulos. Dicha prueba fue realizada por 3 voluntarios, distintos de los que participaron en la prueba de validación, cada uno de los cuales contó con características sociales correspondientes a un modo de vida distinto. Una vez que se contó con los resultados de los tres jueces, se consideró que no era necesario llevar a cabo modificaciones adicionales antes de proceder a la prueba definitiva.

### 5.3. Resultados de la primera parte de la prueba: índices sociolingüísticos

En este apartado, se presenta el análisis de la primera sección de la prueba de percepción (descrita en §5.2.4), relacionado con la asociación de los distintos índices prosódicos

propuestos en el capítulo anterior con los tres modos de vida considerados en este trabajo.<sup>33</sup> Debido a que se considera oportuno enfocarse en los intertonemas y en los tonemas de forma independiente, a continuación, se presenta el análisis de la percepción de los intertonemas oxítonos y de los tonemas por separado. En esta ocasión, a diferencia de los análisis presentados en el capítulo anterior, no se considera necesario hacer un contraste estadístico entre las respuestas de cada grupo de jueces, sino que la prioridad es ahondar en el juicio perceptivo de cada uno de ellos. Por tal motivo, las pruebas estadísticas se le realizaron a cada grupo de forma individual, es decir, se analizaron las respuestas de cada grupo de jueces por separado. Cabe señalar que, con excepción de las respuestas de los J2 y J3 en cuanto a la relación de la muestra con el habla de la Ciudad de México, todos los datos intragrupalos obtenidos y presentados en esta sección son estadísticamente significativos (con una significación menor o igual a 0.01) según las pruebas ANOVA de un factor y el estadístico Exacto de Fisher.

#### 5.3.1. Percepción de los rasgos prosódicos en los intertonemas oxítonos

Al analizar las respuestas de los J1, mostradas en los cuadros 5.4<sup>34</sup> y 5.5,<sup>35</sup> salta a la vista que todos los estímulos fueron evaluados positivamente en cuanto a que todas las medias superan el valor medio de la escala de Likert (el punto 4). Esto quiere decir que estos jueces consideraron que la prosodia en las señales de audio era compatible tanto con el habla de la Ciudad de México, como con la de los miembros de sus círculos sociales más cercanos y la de su propia habla. Adicionalmente, se observa que la mayoría de los patrones prosódicos fueron relacionados con el habla del MdV2 (55.3 %), con independencia de que la voz fuese masculina o femenina, mientras que solo la minoría de ellos (12.5 %) fue considerado

---

<sup>33</sup> Debido a que en lo sucesivo se hará referencia constante al término “modo de vida” en relación tanto con los jueces que participaron en la prueba de percepción como con la variable social considerada en el análisis, a partir de este punto, para simplificar la redacción y evitar posibles confusiones, se usará la abreviación “MdV” para hacer referencia a la variable extralingüística del modo de vida y la letra “J” seguida por un número (correspondiente a su modo de vida) para designar a los jueces. Por ejemplo, MdV1 hace referencia a la variable social modo de vida 1 y J2 a los jueces con modo de vida 2.

<sup>34</sup> En las tres primeras preguntas de la prueba de percepción, los jueces tenían que indicar qué tan de acuerdo estaban con las siguientes afirmaciones (presentadas en la prueba en el mismo orden): 1) esta persona es de la Ciudad de México; 2) esta persona habla igual o de forma similar a tus familiares y amigos más cercanos; 3) esta persona habla igual o de forma similar que tú.

<sup>35</sup> En caso de considerar necesario recordar el significado y lectura de los distintos patrones de duración presentados en los cuadros y a los que se hará referencia en el texto en lo sucesivo, se puede revisar esta información en §5.2.2.1.

compatible con su propio modo de vida (MdV1). Ahora bien, al observar de cerca los datos, queda en evidencia que ciertos rasgos prosódicos son más asociados con algunas de las variables que con otras.

Cuadro 5.4. Medias de las respuestas de los J1 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos

Patrones		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
D^D	H* H-	4.38	4.63	4.13
D^D	H* !H-	6.75	6.63	6.38
D^D	L+H* H-	4.63	4.88	4.5
D^D	L+ <sub>j</sub> H* HH-	4.25	4.75	5.13
D^(D+103)	H* H-	6.63	6.25	6.75
D^(D+103)	H* !H-	6.25	6.5	6
D^(D+103)	L+H* H-	5.88	5.88	5.62
D^(D+103)	L+ <sub>j</sub> H* HH-	6.63	6.5	6.38
D^(D+165)	H* H-	5.75	5.63	5.75
D^(D+165)	H* !H-	6.75	6.62	6
D^(D+165)	L+H* H-	6.25	5.88	5.75
D^(D+165)	L+ <sub>j</sub> H* HH-	5.38	5.88	4.75

Cuadro 5.5. Distribución de las respuestas de los J1 a la cuarta pregunta de la fase 1<sup>36</sup> en relación con los patrones en los intertonemas oxítonos

Patrón		Modos de vida (MdV)			Total
		1	2	3	
D^D	H* H-	1	4	3	8
D^D	H* !H-	0	4	4	8
D^D	L+H* H-	1	6	1	8
D^D	L+ <sub>i</sub> H* HH-	2	5	1	8
D^(D+103)	H* H-	0	5	3	8
D^(D+103)	H* !H-	2	6	0	8
D^(D+103)	L+H* H-	1	1	6	8
D^(D+103)	L+ <sub>i</sub> H* HH-	0	6	2	8
D^(D+165)	H* H-	1	6	1	8
D^(D+165)	H* !H-	0	4	4	8
D^(D+165)	L+H* H-	1	1	6	8
D^(D+165)	L+ <sub>i</sub> H* HH-	1	7	0	8

El patrón de duración que se había supuesto como índice sociolingüístico del MdV1 (D^(D+103)), propio de la red social vecinal, es reconocido, por los J1, con un alto grado de afinidad (superior a 6, siendo 7 el límite superior de calificación) tanto con su propia habla como con la de sus allegados. Sin embargo, dicha característica es principalmente asociada con el MdV2, independientemente de la configuración tonal que la acompañe. Esto se podría interpretar como que los J1 consideran que el patrón de duración D^(D+103) es un rasgo prosódico que indexa un modo de vida con mayor grado de instrucción e ingresos que el suyo, lo que le podría atribuir mayor prestigio y explicaría su prominencia en los miembros de la red vecinal en el contexto de la entrevista sociolingüística. Esta hipótesis cobra fuerza al tener en cuenta que ocurre lo mismo con la duración D^(D+165), presunto índice de la red social laboral (MdV3), aunque únicamente cuando está acompañada por el intertonema H\* !H-. En este caso, los J1 se relacionan a sí mismos y a sus círculos sociales cercanos con este patrón, pero lo reconocen como un rasgo más atribuible al MdV2.

Los patrones tonales no parecen tener algún tipo de incidencia en el grado de identificación que les dan los jueces en relación consigo mismos y con sus allegados. De hecho, el único patrón recurrente se observa en torno al intertonema H\* !H-, con el que los

<sup>36</sup> La cuarta pregunta de la prueba de percepción fue “¿A cuál de las siguientes actividades crees que podría dedicarse esta persona?”



jueces se relacionan positivamente (con una media de puntuación entre 6 y 7), excepto cuando ocurre con el patrón de duración  $D^{\wedge}D$ . Sin embargo, una vez más, esta configuración tonal es más asociada con el MdV2 que con el MdV1 (propio de estos jueces), incluso cuando está acompañada por la duración  $D^{\wedge}(D+103)$ , presunto índice de la red vecinal, de acuerdo con los resultados de la fase de producción. Por otra parte, llama la atención que el intertonema  $L+H^* H-$  es asociado con el MdV3 (excepto cuando está acompañado por la duración  $D^{\wedge}D$ ), lo cual resulta interesante debido a que se esperaba que el intertonema más relacionado con este modo de vida fuese  $L+;H^* HH-$ , cuya mayor elevación tonal había sido identificada como un posible índice sociolingüístico de la red laboral (MdV3). Esto hace pensar que, si bien los intertonemas oxítonos con movimientos tonales ascendentes pueden estar relacionados con el MdV3, al menos para los J1, cuando dicho ascenso es mayor a 3 st la elevación tonal adquiere otro significado social y pasa a representar otro MdV.

Los J2, en cambio, muestran una tendencia diferente. Como se puede ver en los datos presentados en los cuadros 5.6 y 5.7, este grupo de jueces identifica todos los patrones prosódicos como propios del Ciudad de México, al igual que con el habla de sus allegados y con la suya propia en buena medida, pues las calificaciones medias en la escala de Likert se encuentran entre 5 y 6. Así mismo, en consonancia con lo anterior, asocian todos los patrones principalmente con el habla de aquellos con su mismo MdV, con excepción de algunos que también son relacionados con el MdV3. Cabe señalar que, al igual que en los datos de los J1, dicha asociación se establece sin que exista una aparente relación entre el sexo reconocido en la voz (masculino o femenino) y los patrones prosódicos estudiados.

Cuadro 5.6. Medias de las respuestas de los J2 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos

Patrones prosódicos		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
D^D	H* H-	5.88	5.63	5.5
D^D	H* !H-	6.13	6	5.88
D^D	L+H* H-	5.5	5.25	5
D^D	L+ <sub>j</sub> H* HH-	5.63	5.75	5.5
D^(D+103)	H* H-	5.63	5.63	5.5
D^(D+103)	H* !H-	6	5.63	5.38
D^(D+103)	L+H* H-	6	5.88	5.88
D^(D+103)	L+ <sub>j</sub> H* HH-	5.88	5.75	5.5
D^(D+165)	H* H-	6	5.63	5.25
D^(D+165)	H* !H-	6	5.88	5.63
D^(D+165)	L+H* H-	5.88	5.75	5.63
D^(D+165)	L+ <sub>j</sub> H* HH-	5.88	5.75	5.63

Cuadro 5.7. Distribución de las respuestas de los J2 a la cuarta pregunta de la fase 1 en relación con los patrones en los intertonemas oxítonos

Patrón		Modos de vida (MdV)			Total
		1	2	3	
D^D	H* H-	0	6	2	8
D^D	H* !H-	0	4	4	8
D^D	L+H* H-	0	5	3	8
D^D	L+ <sub>j</sub> H* HH-	0	5	3	8
D^(D+103)	H* H-	0	6	2	8
D^(D+103)	H* !H-	0	5	3	8
D^(D+103)	L+H* H-	0	4	4	8
D^(D+103)	L+ <sub>j</sub> H* HH-	1	5	2	8
D^(D+165)	H* H-	0	5	3	8
D^(D+165)	H* !H-	0	3	5	8
D^(D+165)	L+H* H-	0	5	3	8
D^(D+165)	L+ <sub>j</sub> H* HH-	0	6	2	8

A diferencia de lo observado en los J1, la duración de los patrones prosódicos no juega un papel determinante en los J2, al momento de establecer el grado de relación entre su propia habla, su modo de vida y los distintos patrones en los estímulos. No obstante, se observan

dos similitudes bastante llamativas entre las respuestas del grupo J1 y J2. La primera de ellas es que, en ambos casos, el contorno tonal H\* H- es asociado con el MdV3 cuando está acompañado por los patrones de duración  $D^{\wedge}D$  y  $D^{\wedge}(D+165)$ . A pesar de que los J1 presentan dudas acerca de la relación plena de este intertonema con el MdV3 (pues también lo asocian con el MdV2 en proporciones iguales) y los J2 solo presentan esta duda cuando la configuración es producida con la duración  $D^{\wedge}D$ , no se puede pasar por alto la posibilidad de que exista un vínculo entre este intertonema oxítono (producido tanto con duración igual como con una considerablemente mayor que la de la sílaba átona previa) y el MdV3. Estos vínculos aparentes entre los dos intertonemas mencionados cobran mayor relevancia al tener en cuenta que no son muchos los patrones prosódicos asociados con el MdV3. Las pocas asociaciones con el MdV3 son especialmente significativas debido a que se puede identificar un elemento recurrente en común en la percepción de distintos grupos de jueces.

La segunda similitud es que los J1 y J2 relacionan el intertonema L+H\* H- con el MdV3 cuando es producido con la duración  $D^{\wedge}(D+103)$ . En este caso, los J1 atribuyen este patrón prosódico únicamente al MdV3, mientras que los J2 tienen opiniones divididas entre su relación con el MdV3 y el MdV propio. Adicionalmente, a diferencia de lo observado en el primer grupo de jueces, los J2 no relacionan esta configuración con MdV3 en presencia de ningún otro tipo de duración ni siquiera con la que se presumía índice sociolingüístico de la red laboral ( $D^{\wedge}(D+165)$ ), sino exclusivamente con  $D^{\wedge}(D+103)$ . Aunque en estos jueces no se observa la misma tendencia a asociar una mayor longitud de la sílaba tónica final de ip con el MdV3, resulta interesante que al menos consideren la posibilidad de que dicha duración no represente al MdV1 en ningún caso. Esto permite inferir que una elevación tonal (menor a 3 st en el caso de los J2) en la sílaba final (tónica) de ip, cuya duración es considerablemente superior a la de la sílaba átona que la precede, es un rasgo atribuido a miembros de la comunidad de habla que, al menos, cuentan con un grado de instrucción universitario.

Por su parte, los J3 siguen la tendencia general, relacionada con la asociación de los patrones prosódicos con el habla de la Ciudad de México, con la de sus allegados y con la suya propia (esto último se ve en la gran mayoría de los patrones estudiados), del mismo modo que relacionan casi todos los patrones con el MdV2, como se puede ver en los cuadros 5.8 y 5.9. Sin embargo, salta a la vista que existe un alto grado de similitud entre las

respuestas de los J3 y las de los J1, en relación con patrones prosódicos que asocian con el habla de sus círculos sociales más cercanos y la suya propia.

Cuadro 5.8. Medias de las respuestas de los J3 a las tres primeras preguntas de la fase 1 en relación con los patrones en los intertonemas oxítonos

Patrones prosódicos		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
D^D	H* H-	5.13	4.88	4.13
D^D	H* !H-	5.63	5.38	4.75
D^D	L+H* H-	4.88	4.63	3.88
D^D	L+ <sub>j</sub> H* HH-	5.5	5.25	4.75
D^(D+103)	H* H-	6	6	5.5
D^(D+103)	H* !H-	6.25	6	5.75
D^(D+103)	L+H* H-	5.88	5.88	5.13
D^(D+103)	L+ <sub>j</sub> H* HH-	5.38	5.25	4.75
D^(D+165)	H* H-	6.25	6.25	6
D^(D+165)	H* !H-	6	5.63	5.5
D^(D+165)	L+H* H-	5.88	5.63	5.25
D^(D+165)	L+ <sub>j</sub> H* HH-	5.88	5.5	5.38

Cuadro 5.9. Distribución de las respuestas de los J3 a la cuarta pregunta de la fase 1 en relación con los patrones en los intertonemas oxítonos

Patrón		Modos de vida (MdV)			Total
		1	2	3	
D^D	H* H-	0	4	4	8
D^D	H* !H-	1	5	2	8
D^D	L+H* H-	1	4	3	8
D^D	L+ <sub>j</sub> H* HH-	1	5	2	8
D^(D+103)	H* H-	0	4	4	8
D^(D+103)	H* !H-	0	4	4	8
D^(D+103)	L+H* H-	0	4	4	8
D^(D+103)	L+ <sub>j</sub> H* HH-	0	6	2	8
D^(D+165)	H* H-	0	5	3	8
D^(D+165)	H* !H-	0	4	4	8
D^(D+165)	L+H* H-	0	5	3	8
D^(D+165)	L+ <sub>j</sub> H* HH-	0	5	3	8

Mientras que los J1 reflejan un grado positivo de asociación (entre 4 y 5 en la mayoría de los casos) de los patrones prosódicos que poseen la duración  $D^{\wedge}D$  con el habla de sus allegados y con la propia, los J3 se distancian considerablemente de esta duración al otorgarle un grado de relación bajo (entre 2 y 4) con su propia habla, aunque la identifican más con el habla de sus allegados. Esto hace pensar que dichos patrones prosódicos podrían tener algún tipo de carga social negativa (poco prestigio) para este grupo. Sin embargo, tanto un grupo de jueces como el otro relacionan estos patrones principalmente con el MdV2, lo que descarta considerablemente la posibilidad de que la duración  $D^{\wedge}D$  esté ligada con el significado social mencionado. En todo caso, esta desvinculación por parte de los dos grupos de jueces y la asociación con el MdV2 parece corresponder con que este rasgo, de hecho, sea característico e, incluso, posiblemente un índice sociolingüístico de los hablantes con este MdV. Esta idea es especialmente llamativa debido a que los J2 relacionan los patrones que tienen esta duración con su MdV (de forma casi exclusiva), su habla y la de sus allegados.

Otro aspecto que llama la atención en las respuestas de los J3 es que muestran una tendencia a asociar los patrones de duración  $D^{\wedge}(D+103)$  con el MdV2 y con el suyo propio (MdV3), lo cual ocurre en el 75 % de los patrones que tienen esta duración. Lo interesante es que este tipo de duración había sido identificada como un potencial índice sociolingüístico de la red social vecinal (MdV1). El hecho de que este grupo de jueces opine de esta manera, aunado a que ninguno de los otros grupos relaciona el patrón de duración  $D^{\wedge}(D+103)$  con el MdV1, lleva a inferir que la condición de índice sociolingüístico, previamente relacionada con dicho rasgo, no existe. Esto trae a colación la idea de que, si bien un rasgo prosódico puede ser más frecuente o característico de un grupo social que otro, no necesariamente implica que en él se perciba un significado sociolingüístico relacionado con algún grupo social en particular. Por ello, se puede plantear la hipótesis de que la cualidad prosódica en cuestión se trata de un elemento extendido en la comunidad de habla que, aunque es producido con mayor frecuencia por los miembros de la red vecinal (MdV1), no es percibido como un rasgo distintivo o fácilmente asociable con esta, por lo que no representa un índice sociolingüístico.

Finalmente, se observa que el grupo J3, al igual que los jueces de los otros dos grupos, relacionan el intertonema  $H^* !H-$  con su propio modo de vida, cuando ocurre en concordancia con la duración  $D^{\wedge}(D+165)$ . Lo mismo ocurre con la configuración  $L+H^* H-$  en compañía

de la duración  $D^{(D+103)}$ , aunque no de forma consistente, pues la distribución de asociaciones es equitativa entre los MdV2 y MdV3. Dicha tendencia también es parcialmente observable en los jueces de los otros grupos, pues todos presentan la misma distribución equitativa en uno de los dos patrones prosódicos (los J1 en el segundo y los J2 en el primero) y una asociación predominante con el MdV3 en el otro. Si bien esta opinión generalizada en los jueces podría apuntar hacia la identificación de tales patrones prosódicos como un índice sociolingüístico en conexión con el MdV3, el hecho de que exista en los otros grupos de jueces la misma vacilación observada en los J3 (sin una asignación notablemente mayoritaria hacia el MdV2 o 3) descarta esta posibilidad. En efecto, lo único que se puede interpretar en este sentido es que ambos patrones representan índices sociolingüísticos de personas con estudios universitarios (concluidos o cercanos a concluir), pues en ningún caso se observó una atribución mayoritaria de este rasgo al MdV1.

### 5.3.2. Percepción de los rasgos prosódicos en los tonemas

A diferencia de lo observado en el análisis de los patrones prosódicos en los intertonemas oxítonos, en el caso de los tonemas se observa una tendencia bastante clara que lleva a la confirmación de uno de los índices sociolingüísticos identificados durante el análisis de los datos de producción. Como se puede ver en el cuadro 5.10, los J1 no solo rechazan la idea de que los patrones prosódicos que incluyen la duración  $D^{((DX2)+20)}$  puedan representar el habla de sus allegados y la suya propia (en ambos casos con independencia de que la voz sea masculina o femenina), sino que, incluso, consideran que estos patrones no fueron producidos por alguien de la Ciudad de México (especialmente cuando son producidos por una voz masculina). Las medias de calificación de acuerdo con la escala de Likert en las opiniones relacionadas con la duración  $D^{((DX2)+20)}$  no superan los 3 puntos en cada caso, independientemente del sexo percibido en las señales de audio, lo que los ubica por debajo del punto medio de la escala. Adicionalmente, como se muestra en el cuadro 5.11, los J1 relacionan todos los patrones con esta duración, especialmente cuando la voz es femenina, principalmente con el MdV1, correspondiente con la red social vecinal de este estudio y con su propio modo de vida. Ahora bien, se entiende que el hecho de que los J1 no consideren que dichos patrones puedan pertenecer a la Ciudad de México podría invalidar la idea de que la duración  $D^{((DX2)+20)}$  represente un índice sociolingüístico en la Ciudad de México. Sin

embargo, como se presentará en párrafos posteriores, esta idea no es compartida por los demás jueces (J2 y J3), lo que añade un matiz relevante a la baja puntuación en la asociación con el habla de la Ciudad de México.

Cuadro 5.10. Medias de las respuestas de los J1 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas

Patrones prosódicos		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
212 -157	!H* =%	6.25	5.88	5.75
212 -157	L+H* =%	4.75	4.75	4.38
212 -157	L+H* L%	6	6.5	5.38
212 -157	H* L%	5.13	4.75	4.75
212 -157	L* L%	5.25	5.38	5.25
212 -157	H+L* L%	5.38	5.5	5.13
D^D	!H* =%	6.25	6.5	5.62
D^D	L+H* =%	4.63	4.5	4.87
D^D	L+H* L%	5.12	5.25	5.37
D^D	H* L%	3.88	3.63	3.63
D^D	L* L%	6	5.75	5.38
D^D	H+L* L%	6	6.25	6
D^((DX2)+20)	!H* =%	2.75	2.88	2.75
D^((DX2)+20)	L+H* =%	3	2.63	2.75
D^((DX2)+20)	L+H* L%	2.62	3	2.88
D^((DX2)+20)	H* L%	1.63	2	1.75
D^((DX2)+20)	L* L%	3.13	3.25	2.63
D^((DX2)+20)	H+L* L%	3	3.25	2.75

Cuadro 5.11. Distribución de las respuestas de los J1 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas

Patrón		Modos de vida (MdV)			Total
		1	2	3	
212 -157	!H* =%	2	4	2	8
212 -157	L+H* =%	1	6	1	8
212 -157	L+H* L%	2	3	3	8
212 -157	H* L%	3	4	1	8
212 -157	L* L%	1	4	3	8
212 -157	H+L* L%	0	5	3	8
D^D	!H* =%	2	4	2	8
D^D	L+H* =%	3	4	1	8
D^D	L+H* L%	2	4	2	8
D^D	H* L%	3	3	2	8
D^D	L* L%	1	4	3	8
D^D	H+L* L%	1	5	2	8
D^((DX2)+20)	!H* =%	6	1	1	8
D^((DX2)+20)	L+H* =%	5	2	1	8
D^((DX2)+20)	L+H* L%	6	1	1	8
D^((DX2)+20)	H* L%	6	1	1	8
D^((DX2)+20)	L* L%	6	2	0	8
D^((DX2)+20)	H+L* L%	4	2	2	8

En líneas generales, los integrantes del grupo J1 relacionan la mayoría de los patrones con duraciones distintas a  $D^((DX2)+20)$  con el MdV2. No obstante, llama la atención que el otro patrón que obtuvo una calificación inferior a 4 puntos en las tres primeras preguntas ( $H^* L\%$  con duración  $D^D$ ) es el único que también se relaciona con el MdV1 (aunque se observa vacilación, pues fue atribuido al MdV2 equitativamente). Esto demuestra una tendencia interesante en la percepción (o posiblemente en la actitud lingüística) del grupo J1, pues los rasgos que relacionan con su propio MdV no son considerados por ellos como pertenecientes a la Ciudad de México. Dicha tendencia se discutirá con la comparación de los datos aportados por los otros grupos de jueces.

Por otra parte, las medias de puntuación obtenidas en torno a la relación establecida por los J1 con el resto de los patrones prosódicos superan el punto medio de la escala de Likert, lo cual indica que estos jueces los consideran suficientemente representativos de la comunidad de habla, el habla de sus círculos cercanos y la suya propia. De entre las



configuraciones mencionadas, se puede destacar lo observado en los tonemas H\* L% y L+H\* =%, pues ambos reciben una puntuación positiva, pero reducida (entre los puntos 4 y 5) y son asociadas principalmente con el MdV2. Lo más relevante de esto es que la segunda configuración se había identificado como un posible índice sociolingüístico de la red vecinal (MdV1). Sin embargo, debido a que estos jueces no tienden a relacionarse con los rasgos identificados como índices potenciales de su MdV, es necesario ver lo que ocurre en los otros grupos de jueces, para poder determinar si este rasgo puede representar un índice sociolingüístico de este MdV o si tal idea se debe descartar.

En relación con la percepción de los J2, de acuerdo con los datos reflejados en los cuadros 5.12 y 2.13, este grupo de jueces presenta respuestas bastantes similares a las registradas en los datos de los J1 en torno a los patrones prosódicos que incluyen la duración  $D^{\wedge}((DX2)+20)$ , aunque en este caso sí se relacionan favorablemente con el habla de la Ciudad de México (sin que el sexo percibido en la señal de audio influyera en las respuestas). Fuera de esta diferencia, los J2 niegan la posibilidad de que estos patrones puedan representar el habla de sus allegados (calificaciones entre 2 y 3 puntos) y la suya propia (con un rechazo más acentuado, expresado con puntuaciones entre 1 y 2 en la escala). Adicionalmente, los J2 también asocian todos estos patrones con el MdV1 de forma contundente, pues lo hacen en el 91.7 % de los casos.

Cuadro 5.12. Medias de las respuestas de los J2 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas

Patrones prosódicos		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
212 -157	!H* =%	6	5	4.88
212 -157	L+H* =%	5.75	5.38	5.13
212 -157	L+H* L%	5.5	5.5	5.5
212 -157	H* L%	6.13	5.75	5.63
212 -157	L* L%	5.63	5.38	5.25
212 -157	H+L* L%	6.13	5.88	6
D^D	!H* =%	5.88	5.38	5.63
D^D	L+H* =%	5.75	5.25	5
D^D	L+H* L%	5.63	5.38	5.38
D^D	H* L%	6.13	5.75	5.38
D^D	L* L%	6.13	6	6.13
D^D	H+L* L%	6.38	6.25	6.13
D^((DX2)+20)	!H* =%	5.5	2.5	2.13
D^((DX2)+20)	L+H* =%	5.38	2.63	1.5
D^((DX2)+20)	L+H* L%	5.37	2.5	1.5
D^((DX2)+20)	H* L%	5.5	2.13	1.38
D^((DX2)+20)	L* L%	4.5	2.38	1.63
D^((DX2)+20)	H+L* L%	4.38	2	1.63

Cuadro 5.13. Distribución de las respuestas de los J2 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas

Patrón		Modos de vida (MdV)			Total
		1	2	3	
212 -157	!H* =%	1	6	1	8
212 -157	L+H* =%	0	6	2	8
212 -157	L+H* L%	0	8	0	8
212 -157	H* L%	0	6	2	8
212 -157	L* L%	0	8	0	8
212 -157	H+L* L%	1	5	2	8
D^D	!H* =%	2	3	3	8
D^D	L+H* =%	1	6	1	8
D^D	L+H* L%	1	5	2	8
D^D	H* L%	0	6	2	8
D^D	L* L%	1	4	3	8
D^D	H+L* L%	0	6	2	8
D^((DX2)+20)	!H* =%	6	1	1	8
D^((DX2)+20)	L+H* =%	8	0	0	8
D^((DX2)+20)	L+H* L%	8	0	0	8
D^((DX2)+20)	H* L%	7	1	0	8
D^((DX2)+20)	L* L%	7	1	0	8
D^((DX2)+20)	H+L* L%	8	0	0	8

En cuanto al resto de los patrones prosódicos, los J2 son bastante consistentes en su opinión de que representan el habla de la Ciudad de México, la de sus allegados y la suya. Adicionalmente, asocian todos los patrones de forma casi exclusiva con su propio MdV. El único patrón que relacionan con el MdV3 consiste en el tonema !H =% en conjunto con la duración D^D, pero esta asociación no es absoluta, ya que se observa vacilación en la atribución de dicho patrón a los MdV2 y 3. Fuera de esto, el único dato adicional que sobresale de la tendencia predominante de los J2 es que asocian estrechamente (con una puntuación igual o superior a 6 en la escala de Likert) su propia habla con el tonema H+L\* L%, previamente identificado como posible índice de la red laboral (MdV3). Sin embargo, al tener en cuenta que los J2 vinculan estas configuraciones principalmente con su propio MdV, se puede contemplar la idea de que este tonema no posee el significado social enlazado con el MdV3 que se había planteado en la fase del análisis sociolingüístico (§4.4.1).

Por su parte, el grupo J3 refleja la misma tendencia que los J2 en cuanto a sus respuestas a las tres primeras preguntas de la prueba, como se puede ver en el cuadro 5.14. Desde la perspectiva de los J3, todos los estímulos son compatibles con el habla de la Ciudad de México (con puntuaciones superiores al punto 4 de la escala), pero no necesariamente con la de sus allegados y la suya propia, pues rechazan (con puntuaciones promedio inferiores a 2) la idea de que los patrones con duración  $D^{\wedge}((DX2)+20)$  puedan ser representativos tanto de su forma de hablar como de aquellos en sus círculos sociales cercanos (sin que el timbre masculino o femenino desempeñe un papel relevante en las tres instancias). De igual forma, como se puede observar en el cuadro 5.15, los J3 asocian todos los patrones prosódicos que contienen este tipo de duración con el MdV1. Dichos datos, junto con las respuestas de los otros grupos de jueces, confirma que la duración  $D^{\wedge}((DX2)+20)$  es un rasgo prosódico que indexa significado social relacionado con el MdV1 y que, con base en la distancia que marcan todos los jueces entre este rasgo y su propia habla y la de su entorno cercano, posiblemente tiene un prestigio social reducido.

Cuadro 5.14. Medias de las respuestas de los J3 a las tres primeras preguntas de la fase 1 en relación con los patrones en los tonemas

Patrones prosódicos		Pertenencia a la CDMX	Relación con habla de allegados	Relación con habla propia
212 -157	!H* =%	5.5	5.13	4.75
212 -157	L+H* =%	5.38	4.75	4.5
212 -157	L+H* L%	6	5.75	5.38
212 -157	H* L%	5.5	5.13	5.13
212 -157	L* L%	5.75	5.38	4.88
212 -157	H+L* L%	5.25	4.75	4.75
D^D	!H* =%	5.63	5.5	4.75
D^D	L+H* =%	5.13	5	4.5
D^D	L+H* L%	5.38	5.13	4.88
D^D	H* L%	5.63	5.13	4.75
D^D	L* L%	6.13	5.75	5.13
D^D	H+L* L%	6	5.75	5.5
$D^{\wedge}((DX2)+20)$	!H* =%	5	3	2
$D^{\wedge}((DX2)+20)$	L+H* =%	4.12	2	1.75
$D^{\wedge}((DX2)+20)$	L+H* L%	4.63	2.13	1.63
$D^{\wedge}((DX2)+20)$	H* L%	4.5	2.38	2
$D^{\wedge}((DX2)+20)$	L* L%	4.25	2.38	2
$D^{\wedge}((DX2)+20)$	H+L* L%	4	1.88	1.63

Cuadro 5.15. Distribución de las respuestas de los J3 a la cuarta pregunta de la fase 1 en relación con los patrones en los tonemas

Patrón		Modos de vida (MdV)			Total
		1	2	3	
212 -157	!H* =%	1	6	1	8
212 -157	L+H* =%	1	5	2	8
212 -157	L+H* L%	1	3	4	8
212 -157	H* L%	1	3	4	8
212 -157	L* L%	0	4	4	8
212 -157	H+L* L%	2	4	2	8
D^D	!H* =%	2	2	4	8
D^D	L+H* =%	2	6	0	8
D^D	L+H* L%	3	1	4	8
D^D	H* L%	1	5	2	8
D^D	L* L%	1	3	4	8
D^D	H+L* L%	0	6	2	8
D^((DX2)+20)	!H* =%	7	1	0	8
D^((DX2)+20)	L+H* =%	6	2	0	8
D^((DX2)+20)	L+H* L%	7	0	1	8
D^((DX2)+20)	H* L%	7	1	0	8
D^((DX2)+20)	L* L%	7	0	1	8
D^((DX2)+20)	H+L* L%	8	0	0	8

El resto de los patrones prosódicos no parecen presentar una tendencia concreta de los J3 en la asociación de los patrones de duración 212-157 y D^D con su propia habla o con los distintos MdV. Sin embargo, se puede observar una tendencia en las calificaciones de los tonemas. Las configuraciones nucleares no descendentes, identificadas como posibles índices sociolingüísticos del MdV1 en la fase de análisis sociolingüístico, son vinculadas con su propia habla (entre 4 y 5 puntos en la escala). Sin embargo, estos tonemas son principalmente asociados con el MdV2, excepto cuando se trata del tonema !H\* =% en compañía de la duración D^D, el cual relacionan con su propio modo de vida (MdV3). Por ende, y en vista de que ningún otro grupo de jueces asocia estas configuraciones con el MdV1, se puede descartar la idea de que los tonemas no descendentes sean un índice sociolingüístico del MdV1.

Por otra parte, los J3 vinculan el tonema circunflejo (L+H\* L%), principalmente, con su propio modo de vida, en todos los casos con duración distinta a D^((DX2)+20). Si bien

esta tendencia también se observa en los J2, no ocurre lo mismo con los J1, quienes lo relacionan principalmente con el MdV2. El que la mayoría de los jueces asocien su propia habla y MdV con este tonema no resulta sorprendente, pues dicha configuración tonal ya ha sido identificada como un elemento prosódico característico del habla del centro de México (Martín Butragueño, 2017). No obstante, resulta llamativo el hecho de que no haya sido enlazado con el MdV1 en un mayor número de ocasiones, pues Martín Butragueño (2004, 2011 y 2021) señala que este rasgo es bastante frecuente en personas con bajo nivel de instrucción. Ahora bien, en este sentido no se puede omitir que el autor realiza un análisis sociolingüístico bastante detallado de la extensión social del tonema circunflejo y que dichas precisiones no fueron tomadas en consideración en la elaboración de los estímulos de esta prueba, ya que las propiedades y necesidades metodológicas de estos estudios difieren sustancialmente. Por lo tanto, hay que considerar la posibilidad de que, si se replicara este estudio incluyendo todos los detalles explicados por Martín Butragueño (2004, 2011 y 2021), posiblemente determinadas realizaciones del tonema circunflejo (cada una con características específicas) pudiesen ser más relacionadas con un MdV que con otros. Adicional a esto, las discrepancias entre los resultados obtenidos en la fase de percepción y en el análisis sociolingüístico, en términos de la identificación de índices sociolingüísticos, abren la puerta a consideraciones teóricas que se consideran relevantes y que serán presentadas detalladamente en la sección final de este capítulo.

#### 5.4. Resultados de la segunda parte de la prueba: asociación de patrones prosódicos con el habla de la Ciudad de México.

Antes de iniciar con el análisis de los hallazgos en esta sección de la prueba (descrita en §5.2.4), es necesario señalar que los resultados que se muestran en esta sección no fueron significativos, de acuerdo con el estadístico Exacto de Fisher, en ninguno de los grupos de jueces.<sup>37</sup> Sin embargo, se considera que esto no impide llevar a cabo un acercamiento cualitativo a los datos de acuerdo con las tendencias observadas.

En las respuestas relacionadas con los intertonemas, dos grupos de jueces presentan respuestas que siguen una misma tendencia, mientras que el otro difiere, como se puede ver

---

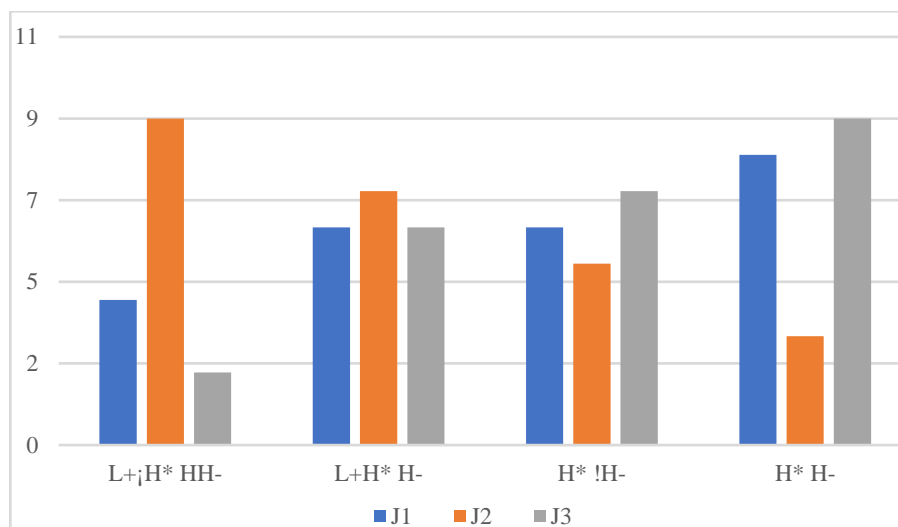
<sup>37</sup> El grado de significación del estadístico Exacto de Fisher en las pruebas realizadas con los datos obtenidos de los J1 fue de 0.58, en las de los J2 fue de 0.476 y en las de los J3 de 0.556.

en los datos mostrados en el cuadro 5.16. La discrepancia de este grupo de jueces se puede percibir más claramente en la representación gráfica de los datos, presentada en la figura 5.16, donde se observa que dos grupos opinan de forma similar, mientras que el otro considera que el grado de asociación es exactamente el inverso. Primeramente, los grupos J1 y J3 están de acuerdo en que la configuración más representativa de la Ciudad de México es H\* H-. En segundo lugar, de acuerdo con el grado de vinculación de los J1, se encuentran H\* !H- y L+H\* H- con igualdad de opiniones. No obstante, en el estímulo en el que se contrastaba justamente este par de configuraciones, estos jueces no reflejan un consenso o una opinión mayoritaria hacia una de las dos, pues ambos fueron calificados de forma equitativa. Los J3, en cambio, asocian más el intertonema H\* !H- con el habla de la Ciudad de México, aunque la diferencia con respecto a L+H\* H- fue de tan solo un punto. Sin embargo, al observar las respuestas ante el estímulo que contiene ambas configuraciones, se puede ver que este grupo relaciona más H\* !H- con la Ciudad de México, de forma unánime. En cuanto a la configuración menos vinculada con el habla de la capital del país, los J1 y J3 coinciden en que esta corresponde con el ascenso mayor a 3st en la sílaba nuclear de las ip oxítonas (L+;H\* HH-), rasgo que había sido identificado previamente como índice sociolingüístico de la red laboral (MdV3). En vista de todo lo expuesto, se puede plantear la siguiente relación de asociación de los intertonemas oxítonos con el habla de la Ciudad de México (sin olvidar que, para los J1, las dos configuraciones en los lugares intermedios tienen el mismo grado de representatividad): tonos altos sostenidos (H\* H-) > tonos moderadamente altos sostenidos (H\* !H-) > ascensos tonales inferiores a 3 st (L+H\* H-) > ascensos tonales superiores a 3 st (L+;H\* HH-). Por su parte, las respuestas de los J2 siguen la tendencia contraria a la presentada, ya que el intertonema que más vinculan con el habla de la Ciudad de México es L+;H\* HH-, seguido (en orden de vinculación descendente) por L+H\* H-, H\* !H- y H\* H-.

Cuadro 5.16. Distribución de las respuestas relacionadas con el intertonema más asociado con la Ciudad de México en contraste de pares<sup>38</sup> en cada estímulo

Combinaciones de estímulos		Modos de vida de los jueces					
		1		2		3	
Primero	Segundo	Primero	Segundo	Primero	Segundo	Primero	Segundo
L+;H* HH-	L+H* H-	1	3	4	0	0	4
L+;H* HH-	H* !H-	2	2	2	2	2	2
L+;H* HH-	H* H-	1	3	3	1	0	4
L+H* H-	H* !H-	2	2	4	0	0	4
L+H* H-	H* H-	1	3	3	1	2	2
H* !H-	H* H-	2	2	3	1	1	3

Figura 5.16. Grado de asociación de cada intertonema con la Ciudad de México en cada grupo de jueces



Al contrastar las asociaciones establecidas por los grupos de jueces, no se puede ver una tendencia especial en torno a las configuraciones que cada uno relaciona más con el habla de la Ciudad de México. Sin embargo, llama la atención que el intertonema H\* !H- sea el segundo más asociado con esta comunidad de habla por los J1 y J3 y el tercero por los J2. Si bien esta configuración es principalmente asociada con el MdV2 (al igual que todos los demás), también, es vinculada con el MdV3 de forma equitativa con el MdV2 en poco más

<sup>38</sup> Los términos “primero” y “segundo” utilizados en este cuadro y en el 5.17 hacen referencia al orden de presentación de cada estímulo en la señal de audio reproducida en esta fase de la prueba. Por ejemplo, en el audio correspondiente a la primera fila de datos del cuadro 5.16, los jueces escucharon inicialmente el estímulo que contenía el intertonema L+;H\* HH- y, posteriormente, el que contenía la configuración tonal L+H\* H-.



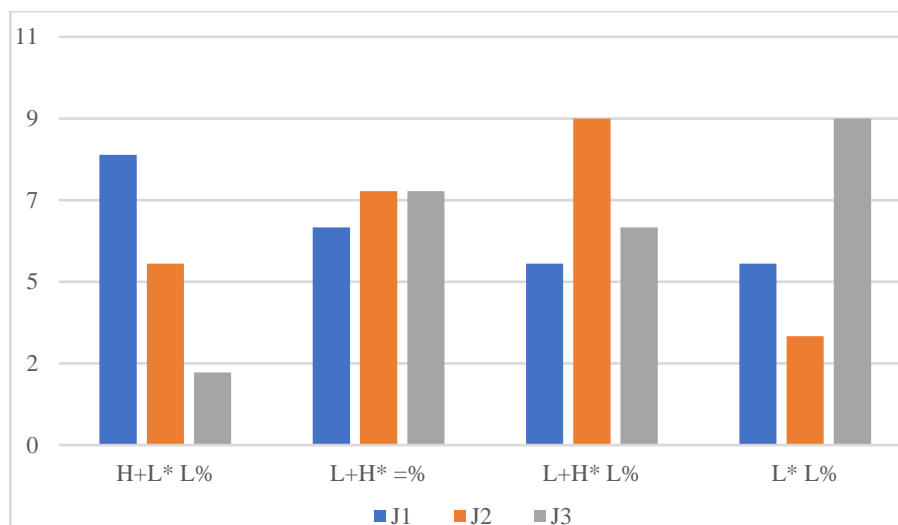
de la mitad de los casos y en una sola ocasión de manera exclusiva (situación que no ocurre con menor frecuencia con H\* H-). No obstante, el hecho de que todos los intertonemas sean principalmente relacionados con el MdV2, ya sea de forma exclusiva o equitativa con el MdV3, dificulta una profundización teórica en las implicaciones cognitivas que puede tener el grado de representatividad de la configuración H\* !H- con el habla de la Ciudad de México.

Los resultados obtenidos en el contraste de pares de tonemas, mostrados en el cuadro 5.17 y en la figura 5.17, no presentan consistencia en la selección de la configuración más asociada con el habla de la Ciudad de México. Para los J1, esta configuración es H+L\* L%, con la que se identificaron muy poco y asociaron principalmente con el MdV2. Entre los J2, el circunflejo L+H\* L% e el tonema con el que estos jueces se identificaron favorablemente y asociaron con su propio modo de vida (MdV2). Para los J3, la configuración más relacionada con el habla de la capital fue L\* L%, relacionada por ellos mismos con su propia forma de hablar, aunque no con una puntuación particularmente elevada y que asocian, principalmente, con el MdV2. Esta falta de unanimidad en la selección del tonema más característico del habla de la ciudad haría pensar, en una primera instancia, que cada grupo de jueces estaría vinculando con el habla capitalina un rasgo prosódico más asociado con su modo de vida, pero, en vista de las asociaciones registradas en la primera parte de la prueba, este solo es el caso de los J2. Por lo tanto, una posible conclusión es que estos tonemas representen configuraciones tonales que se encuentran de forma extendida en la comunidad de habla de la Ciudad de México, especialmente entre aquellas personas que se pudieran incluir en el MdV2.

Cuadro 5.17. Distribución de las respuestas relacionadas con el intertonema más asociado con la Ciudad de México en contraste de pares en cada estímulo

Combinaciones de estímulos		Modos de vida de los jueces					
		1		2		3	
Primero	Segundo	Primero	Segundo	Primero	Segundo	Primero	Segundo
H+L* L%	L+H* =%	1	3	2	2	1	3
H+L* L%	L+H* L%	4	0	1	3	1	3
H+L* L%	L* L%	3	1	2	2	0	4
L+H* =%	L+H* L%	2	2	2	2	2	2
L+H* =%	L* L%	1	3	3	1	2	2
L+H* L%	L* L%	3	1	4	0	1	3

Figura 5.17. Grado de asociación de cada tonema con la Ciudad de México en cada grupo de jueces



Si bien los tres grupos de jueces no están de acuerdo en cuál de estos tonemas es el más representativo de la Ciudad de México, sí están de acuerdo en los que ocupan el segundo y tercer lugar. De forma unánime, todos los jueces señalaron que la segunda configuración más representativa de esta comunidad de habla es la no descendente L+H\* =%. La selección no resulta del todo inesperada, teniendo en cuenta que, en la primera parte de la prueba, este tonema fue principalmente asociado con el MdV2; pero sí resulta llamativo el hecho de que la ausencia de una juntura descendente haya sido más asociada con el habla de esta comunidad que los dos tonemas descendentes restantes. Tal situación no solo resalta la ausencia de significado social enlazado con características del MdV1 (como se mencionó en

§5.3), sino que hace surgir la posibilidad de que esta configuración esté más presente en los integrantes de la comunidad de habla de la Ciudad de México de lo que se pudo vislumbrar en el análisis sociolingüístico realizado en el capítulo 5.

El tercer tonema más asociado con la comunidad de habla en cuestión fue el circunflejo (L+H\* L%). Este estatus se mantiene claramente en la opinión de los J2 y J3, pero no en la de los J1. Si bien el número de veces que dicho tonema fue seleccionado por los J1 coincide con el de L\* L%, al observar las respuestas dadas por los J1 en la señal de audio que contiene estas dos configuraciones, el 75 % de ellos consideraron que el tonema circunflejo era más característico de la Ciudad de México que el descendente L\* L%. Por tal motivo, L+H\* L% se puede considerar el tercero en grado de asociación con dicha comunidad de habla. Este nivel de vinculación por parte de todos los grupos de jueces no resulta especialmente llamativo al recordar que este mismo tonema fue el tercero más frecuente en la totalidad del corpus. Esto quiere decir que, aunque el contorno tonal circunflejo es un rasgo característico del habla del centro de México, de acuerdo con los estudios sociolingüísticos cuantitativos realizados con datos de producción de habla (Martín Butragueño, 2004, 2011 y 2021), según la percepción de los jueces dicha configuración tonal parece no tener el mismo grado de asociación con el habla de la Ciudad de México. Según su perspectiva, existen otros rasgos que son más característicos del habla de esta comunidad en particular. Sin embargo, aunque estos datos pueden generar un debate interesante en cuanto a cuál configuración tonal puede ser más característica del habla de la Ciudad de México desde las perspectivas productiva y perceptiva, no hay que perder de vista que los datos presentados consisten solo en una primera aproximación en este sentido, pues hay que recordar que la distribución de casos no fue estadísticamente significativa en las respuestas de esta parte de la prueba.

Finalmente, los tonemas con un menor grado de asociación con el habla de la Ciudad de México son L\* L%, según los grupos J1 y J2, y H+L\* L%, de acuerdo con los J3. Si bien cada una de estas configuraciones fue seleccionada como la más representativa por otros grupos de jueces (H+L\* L% por los J1 y L\* L% por los J3), no es particularmente sorprendente que también hayan sido considerados los menos vinculados con esta comunidad. Teniendo en cuenta que ambos pueden ser considerados como un descenso uniforme en la sección nuclear del enunciado, que es lo esperable de acuerdo con las

caracterizaciones prosódicas de enunciados aseverativos en español (ver §1.1.2.1), es posible que cada grupo de jueces les haya dado su respectiva calificación debido a que los consideraron parte del español general y no necesariamente representativos de la Ciudad de México en particular. Planteado en otros términos, estos tonemas descendentes podrían haber sido reconocidos por algunos jueces como característicos del habla de la Ciudad de México debido a que siguen los patrones del español general, lo que no excluye a esta comunidad de habla (hecho resaltado por la alta frecuencia de L\* L% y H+L\* L% en los datos del corpus, como se mostró previamente en §4.4.1), pero al mismo tiempo pudieron haber sido considerados poco representativos de esta comunidad por otros jueces, debido a que posiblemente consideren que, al ser rasgos típicos del español, existen otras configuraciones más peculiares (menos generales) que representan el habla de esta ciudad de forma más acertada.

En vista de lo expuesto en relación con la asociación de los tonemas con la comunidad de habla de la Ciudad de México (en términos del análisis cualitativo y el resultado del estadístico Exacto de Fisher), no se puede afirmar que exista algún tipo de consistencia en cuanto a las configuraciones con mayor y menor grado de vinculación. Sin embargo, teniendo en cuenta que todos los jueces coinciden en los contornos que ocuparían las posiciones intermedias en la escala de representatividad de la Ciudad de México, se puede decir que dichos contornos son, en general, más característicos del habla de esta ciudad. Esto no quiere decir que sean los más comunes, pues se puede constatar en los resultados de los capítulos 3 y 4 que, de hecho, no lo fueron. Lo que los datos permiten dilucidar es que, desde el punto de vista de los jueces, las configuraciones más alejadas de la declinación tonal típica del español general (L+H\* =% y L+H\* L%), aunque no necesariamente las más periféricas desde el enfoque basado en el uso, son más sobresalientes o prominentes perceptivamente, por lo que se pueden relacionar con la comunidad de habla capitalina con mayor facilidad o seguridad.

##### 5.5. Recapitulación de resultados y consideraciones finales de la fase de percepción

El análisis de los resultados relacionados con los intertonemas oxítonos reveló que, si bien todos los grupos de jueces reconocen estas combinaciones de duración y entonación como rasgos pertenecientes a su propia habla, tanto los J1 como los J3 los asocian principalmente

con el MdV2. Por tal motivo, no se puede decir que exista necesariamente un distanciamiento de los dos grupos de jueces en relación con los patrones estudiados, sino que estos representan rasgos de un grupo social con grado de instrucción y nivel de ingresos ubicados en un punto intermedio, entre aquellos del MdV1 y del MdV3. Dicho de otra forma, la mayoría de los patrones prosódicos estudiados son relacionados casi siempre con personas que tienen, al menos, la preparatoria concluida y que cuentan con un poder adquisitivo inferior al de un empleado con cargo ejecutivo en una empresa.<sup>39</sup> Desafortunadamente, no se puede afirmar que las personas con estas características representen al grupo de habitantes más predominante en la Ciudad de México ni a los que se podrían considerar como clase media. No obstante, debido a que tanto los J1 como los J3 se relacionaron con ellos, sí se puede manejar la idea de que dichos patrones representan el habla de personas cuyos rasgos sociales no se encuentran demasiado lejos de los de estos jueces en el *continuum* del espectro social de la Ciudad de México.

Existen casos que dejan menos dudas en cuanto a la relación con el MdV2, como ocurre con el patrón de duración D^D. Aunque los J1 y J3 se identifican de forma positiva con los patrones prosódicos que incluyen esta duración, lo hacen con una calificación relativamente baja y los relacionan mayoritariamente con el MdV2. Esto indicaría que este tipo de duración, en la que la sílaba tónica final de frase entonativa intermedia (ip) tiene la misma duración que la sílaba átona prenuclear, puede poseer una mayor carga de significado social que, en este caso, está conectado con el MdV2. Tal noción se ve sustentada por el hecho de que, a diferencia de los otros grupos de jueces, los J2 muestran un alto grado de autoidentificación con esta duración y, al igual que los otros jueces, la relacionan con su propio MdV. Por lo tanto, si bien la diferencia, en términos de grado de vinculación con el habla propia, no es del todo negativa en los J1 y J3, el contraste sutil entre los niveles de identificación entre esta duración y el habla de los grupos de jueces puede ser una señal de que la asociación entre el patrón D^D y el MdV2 debe ser estudiada con mayor profundidad.

Por otra parte, los patrones tonales no parecen ser los que determinan la conexión perceptiva con los distintos modos de vida. El único grupo de jueces que presenta indicios de generar asociaciones basadas en las configuraciones tonales son los J2.; coincidiendo con

---

<sup>39</sup> Cabe recordar que las ocupaciones relacionadas con el MdV2 en la prueba eran estudiante del último período de su carrera y docente de secundaria.

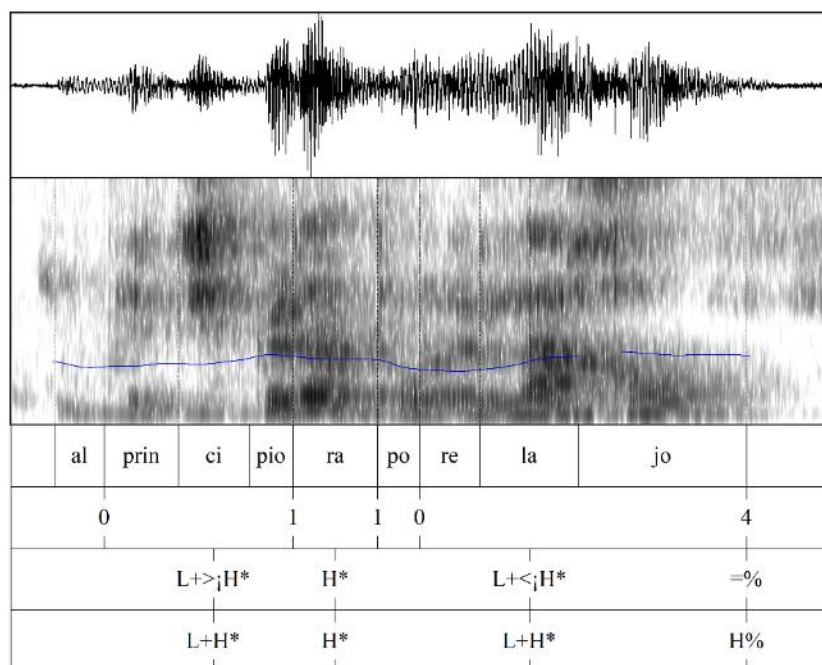
el grupo J1 en algunas relaciones establecidas con determinados modos de vida. No obstante, dichas relaciones no salen de los patrones generales de asociación con el MdV2, por lo que, aunque llamativos, estas similitudes perceptivas entre J1 y J2 en torno a las conexiones entre intertonemas y modos de vida no representan una desviación general de la orientación principal. Por ende, y teniendo en cuenta las respuestas de los J3, con una tendencia parecida a relacionar todos los patrones principalmente con el MdV2, se puede concluir que ninguno de los rasgos prosódicos identificados como potenciales índices sociolingüísticos, durante el análisis correspondiente, tiene tal carga de significado social.

Al contrario, como se mencionó en líneas anteriores, en todo caso, estos patrones están más relacionados con el habla de personas con grado de instrucción superior a la preparatoria y con un margen de ingresos menor a los hablantes de la red laboral del presente estudio. Lo más cercano a un índice sociolingüístico que se pudo identificar en la sección del enunciado referente a los intertonemas oxítonos fueron los patrones  $H^* !H-$  con duración  $D^{(D+165)}$  y  $L+H^* H-$  con duración  $D^{(D+103)}$ . En ambos casos, estos patrones fueron relacionados con los MdV2 y MdV3, aunque con cierta vacilación en algunas ocasiones. Ahora bien, el hecho de que en ningún caso se estableciera una conexión mayoritaria o parcial de estos rasgos prosódicos con el MdV1 lleva a considerar que están asociados socialmente con personas que tienen cierto nivel de estudios universitarios (concluidos o cercanos a concluir). Tal consideración se debe a que las ocupaciones propuestas para los MdV2 y 3 en la primera sección de la prueba (estudia último año en la universidad, enseña en una secundaria, diseña sistemas informáticos en un banco, y gerencia una empresa internacional) requieren necesariamente estudios universitarios concluidos o cercanos a concluir, mientras que los propuestos para el MdV1 (vende en la calle y limpia coches) no requieren este tipo de formación. Por lo tanto, se considera que este rasgo social (que puede ser vinculado con otras características, como el nivel de ingresos, por ejemplo) es el que más notoriamente pudo haber influido en las respuestas de los jueces y que, en consecuencia, es el que se consideraría como índice sociolingüístico para los patrones prosódicos mencionados previamente.

En el estudio de los tonemas, en cambio, se observó un claro patrón de respuestas que permite confirmar el estatus de índice sociolingüístico de solo uno de los rasgos fonéticos identificados como tal durante el análisis de los datos de producción. La duración  $D^{((DX2)+20)}$  (visible en las dos últimas sílabas del enunciado mostrado en la figura 5.18),

que se presumía portadora de significado social relacionada con la red vecinal (MdV1), no solo fue desvinculada del habla propia por todos los grupos de jueces, sino que se asoció en todo momento con el MdV1. Con este dato no se puede comprobar la hipótesis planteada en un inicio (§A.7), según la cual los índices sociolingüísticos están conformados por la combinación de más de un elemento prosódico, pues se pudo constatar que el significado social puede estar relacionado con solo uno de ellos, como ocurre en este caso con el patrón de duración  $D^{\wedge}((DX2)+20)$ . Ahora bien, el hecho de que la previa postulación de este rasgo como índice potencial no se haya realizado con base en su frecuencia de uso (pues el que haga referencia a un valor absoluto concreto hace que la cuantificación de la frecuencia sea muy complicada) dificulta precisar las cualidades abstractas (cognitivas) del rasgo prosódico y, debido a que se trata de un valor absoluto, no resulta razonable barajar la idea de que dicho valor concreto pueda considerarse el portador de significado social. Partiendo de las premisas de que los conceptos cognitivos son los elementos propensos a poseer cualquier tipo de significado (§1.4.3), que ellos (al ser parte del sistema cognitivo) consisten en abstracciones y que la mente humana no puede cronometrar con precisión los milisegundos que dura un sonido, no se puede considerar que un valor de duración absoluto (concreto) constituya un índice sociolingüístico (es decir, un elemento portador de significado social). Sin embargo, al plantear el rasgo de duración en términos relativos (es decir, la longitud percibida de un segmento en comparación con la de otro) se puede obtener un grado de abstracción lo suficientemente alto como para generar un concepto susceptible a poseer significado social. Por ello, se puede estimar que el índice sociolingüístico asociado con el MdV1 radica en una duración de la sílaba final que sea, al menos, el doble de larga que la sílaba nuclear (cuya representación puede ser  $D^{\wedge}(DX2)$ ). Tal estimación, basada en el planteamiento de una duración relativa, es avalada por los resultados de la prueba de percepción, ya que el patrón que representaba dos sílabas con igual duración ( $D^{\wedge}D$ ) no fue juzgado como índice sociolingüístico, mientras que aquel en el que la sílaba final fue al menos el doble de larga que la anterior ( $D^{\wedge}((DX2)+20)$ ), sí se consideró portador de significado social.

Figura 5.18. Oscilograma, espectrograma y F0 del enunciado Rs1H2-14: [(Al principio era por relajo)<sub>ip</sub>]IP



Por otra parte, no se debe pasar por alto que los integrantes del grupo J1 (cuyo MdV corresponde con el de los miembros de la red vecinal) rechaza la idea de que el patrón de duración  $D^{\wedge}((DX2)+20)$  pudiese representar el habla de la Ciudad de México. La apreciación da cabida a dos posibles interpretaciones: i) los J1 consideran que esta duración es característica de personas con MdV1, que provienen de otras comunidades de habla de México; ii) estos jueces piensan que esta duración tiene muy poco prestigio social y por ello se desvinculan de ella, incluso en términos comunitarios. En cualquiera de los dos casos, las posibles interpretaciones se limitan a los J1, pues tanto los J2 como los J3 afirman, con un grado de asociación positiva promedio superior al punto 4 en la escala (5.1 en el caso de los J2 y 4.5 en los J3), que la duración  $D^{\wedge}((DX2)+20)$  sí puede ser considerada como un elemento prosódico de la comunidad de habla de la Ciudad de México, perteneciente a personas con MdV1. Además, debido al rechazo y desvinculación unánime de todos los jueces ante este patrón de duración, junto con la asociación de dicho patrón con el MdV1, puede que lo consideren como una característica poco prestigiosa socialmente, de la que ellos se desvinculan categóricamente. No obstante, quedaría pendiente someter este rasgo a una



prueba de actitudes lingüísticas para determinar si es valorado negativamente y si es, en efecto, estigmatizado en la comunidad de habla estudiada.

Al contrario de lo observado en el caso del patrón de duración  $D^{((DX2)+20)}$ , cuyo valor de índice sociolingüístico quedó confirmado, no se encontró evidencia que permita identificar los tonemas no descendentes  $!H^* =\%$  y  $L+H^* =\%$  como configuraciones tonales con significado social relacionado con el MdV1. Esto se debe a que todos los jueces muestran un grado de asociación positivo (superior a 4, aunque variable en los promedios), entre estas configuraciones nucleares y el habla de sus allegados y la suya propia. Adicionalmente, al igual que la mayoría de patrones prosódicos, tales tonemas fueron asociados, mayoritariamente, con el MdV2. Tanto estos resultados como los obtenidos en la segunda parte de la prueba, donde todos los jueces consideran que el contorno no descendente  $L+H^* =\%$  es el segundo más representativo del habla de la Ciudad de México, reducen la posibilidad de que pueda ser asociado únicamente con un MdV en particular. Asimismo, los resultados hacen que se plantee la posibilidad de que este tipo de tonemas esté más extendido entre los habitantes de la capital del país de lo que los resultados del análisis sociolingüístico permiten inferir en este primer acercamiento.

La posibilidad de que los tonemas no descendentes con junturas sostenidas estén más presentes en el habla de la Ciudad de México de lo que se sospecha cobra mayor relevancia al observar los resultados obtenidos por Martín Butragueño (2017). El autor presenta una hipótesis geoprosódica, según la cual se plantean puntos de partida para determinar el grado de mantenimiento de los rasgos prosódicos originarios de hablantes de distintas zonas del país. Incluso, el autor propone una hipótesis espacial, la cual sugiere que la prosodia de las hablas del norte y las del centro del país se podrían clasificar como hablas centrales, caracterizadas por un acento nuclear ascendente y una juntura que varía entre un tono bajo en el habla del centro del país y un sostenimiento en la del norte. Ahora bien, Martín Butragueño (2017) expone el caso de Rosario L., una hablante cuya habla fue altamente considerada como propia de la Ciudad de México, a pesar de ser originaria del norte del país, haber vivido solo 2 años en la ciudad y preservar (aunque no en gran medida) la juntura sostenida, señalada por el autor como un rasgo prominente del habla del norte del país. El hecho de que esta hablante haya sido percibida como originaria de la Ciudad de México abre la puerta a la posibilidad de que las junturas sostenidas desde un tono alto en la sílaba nuclear

esté ganando más terreno en la comunidad de habla capitalina, lo que avalaría la hipótesis espacial de Martín Butragueño (2017) en torno a las hablas que se pueden considerar centrales.

Los resultados obtenidos en la segunda parte, en la prueba de contraste, revelaron que no hay un consenso general en relación con las configuraciones tonales más asociadas con el habla de la Ciudad de México. Por una parte, los J1 y J3 consideraron que las configuraciones tonales al final de ip oxítone más representativas de la Ciudad de México, en orden de asociación decreciente, son las siguientes: tonos altos sostenidos ( $H^* H^-$ ) > tonos moderadamente altos sostenidos ( $H^* !H^-$ ) > ascensos tonales inferiores a 3 st ( $L+H^* H^-$ ) > ascensos tonales superiores a 3 st ( $L+;H^* HH^-$ ). Los J2, en cambio, consideraron que el orden es el inverso. Por otra parte, en cuanto a los tonemas, cada grupo de jueces seleccionó como el más y el menos representativo de la comunidad de habla un tonema distinto. Sin embargo, las coincidencias entre todos los grupos de jueces se dieron en los que ocupaban uno de los lugares intermedios, pues calificaron el tonema no descendente  $L+H^* =\%$  como el segundo más característico del habla capitalina.

En cualquier caso, los resultados de la segunda prueba permiten concluir que las configuraciones tonales más asociadas con el habla de la Ciudad de México no necesariamente corresponden con las más frecuentes en el corpus. En relación con los intertonemas oxítonos, el grupo de jueces que más se acercó a una asociación con base en el uso (de acuerdo con los resultados del análisis de datos de producción) fueron los J2. Este grupo consideró que los más asociables con la Ciudad de México fueron  $L+;H^* HH^-$  y  $L+H^* H^-$  (en ese mismo orden de representatividad), lo cual corresponde con las dos configuraciones más frecuentes del corpus en orden decreciente.

En cuanto a los tonemas, si bien todos los jueces estiman que el contorno que más se relaciona con el habla de la Ciudad de México es descendente (uno distinto para cada grupo de jueces), todos están de acuerdo en que el segundo más característico es  $L+H^* =\%$ , el cual está lejos de los más frecuentes en las muestras de habla recolectadas del corpus. El más registrado en el corpus es  $L^* L\%$  y fue, de hecho, el más relacionado con el habla de la capital del país por los J3, pero tuvo el menor grado de asociación según las opiniones de los J1 y J2. Todo esto puede hacer pensar que las configuraciones tonales que se consideran más relacionadas con la Ciudad de México son las que, en cierta forma, distan más de los rasgos

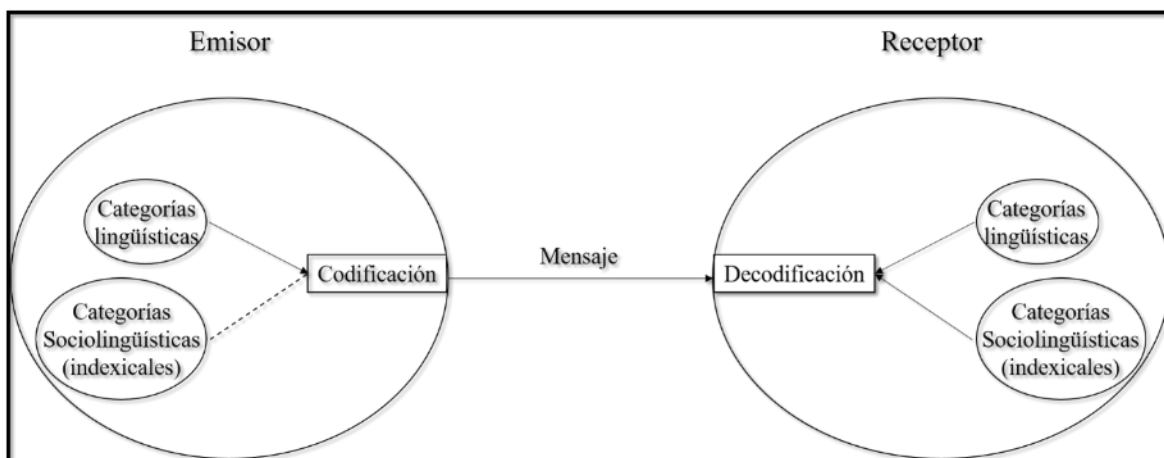
más predominantes del español general y el del centro de México o de la totalidad del país. No obstante, la falta de consenso entre los grupos de jueces dificulta mucho llegar a cualquier conclusión categórica en este sentido. Por lo tanto, es necesario profundizar en este aspecto a través de un estudio perceptivo con una metodología mucho más enfocada en la resolución de este planteamiento.

Todos los datos obtenidos a través de esta prueba, observados desde una perspectiva sociolingüística cognitiva, abren un debate en relación con la concordancia entre el grado de centralidad de un rasgo prosódico en las estructuras cognitivas de los grupos sociales de una comunidad de habla (determinado a través de la frecuencia de uso de dicho rasgo) y el valor indexical de esa característica en términos sociolingüísticos. El hecho de que un elemento prosódico se pueda relacionar con un grupo de hablantes de un corpus, debido a la diferencia de frecuencia (o nivel de prototipicidad) que este elemento presenta entre los grupos estudiados, no necesariamente implica que este rasgo prosódico será asociado directamente con el habla de personas que comparten rasgos sociales con dicho grupo, al ser escuchado por miembros de la misma comunidad de habla. Es posible que los hablantes recurran a diferentes esquemas dentro de distintas categorías cognitivas al momento de producir y percibir material lingüístico, pero hacen uso de categorías paralelas, diferentes en naturaleza, para identificar el significado social indexado en los rasgos prosódicos del habla de otras personas. Por ejemplo, durante un proceso comunicativo entre dos personas de la misma comunidad de habla que interactúan por primera vez, ambos recurren al uso de esquemas dentro de categorías cognitivas lingüísticas, para producir y comprender los mensajes producidos durante la conversación, al mismo tiempo que recurren a categorías cognitivas relacionadas con (o incluso tal vez especializadas en) la identificación de índices sociolingüísticos en el habla de su interlocutor. Esto implicaría que la constitución interna del segundo tipo de categorías es distinta de la de aquellas empleadas con una función lingüística primaria, en las que los elementos no tienen función indexical, sino lingüística.

El modelo cognitivo básico que se propone en el párrafo anterior, representado en la figura 5.19, plantea la existencia de categorías sociolingüísticas que operan de forma paralela a las lingüísticas en la producción de un mensaje. Al momento de producir un mensaje, el emisor hace uso de esquemas cognitivos extraídos de categorías lingüísticas, al mismo tiempo que incluye elementos sociolingüísticos característicos de la comunidad de habla y

grupo social al que pertenece, tomados de categorías sociolingüísticas. Ahora bien, la inclusión de los rasgos sociolingüísticos en el proceso de codificación del mensaje se presenta en la figura 5.19 con una línea punteada, porque se considera que este proceso puede llevarse a cabo de forma variable, dependiendo del grado de espontaneidad o control con el que se emite el acto de habla, pues el emisor puede ejercer mayor grado de injerencia en el contenido sociolingüístico del mensaje en caso de que desee ocultar o camuflar sus rasgos sociales reales, entre otras posibles causas. Por su parte, al momento de decodificar el mensaje, el receptor recurre a sus esquemas en las categorías lingüísticas para comprender el mensaje, al mismo tiempo que recurre a aquellos en las categorías sociolingüísticas (posiblemente constituidas únicamente por esquemas de esta naturaleza), para identificar los índices sociolingüísticos que le permitan obtener mayor información de los rasgos sociales de su interlocutor.

Figura 5.19. Modelo cognitivo básico de la intervención de categorías cognitivas lingüísticas y sociolingüísticas (indexicales) en el proceso de comunicación.



Partiendo de lo antes expuesto, se puede considerar la idea de que las categorías indexicales (nombradas así para diferenciarlas de las netamente lingüísticas) pueden estar construidas sobre nociones sociales, como los estereotipos, y no sobre el grado de centralidad que tiene un elemento lingüístico en las estructuras cognitivas de un determinado grupo social. Para Kristiansen (2010) los estereotipos lingüísticos son imágenes centrales de categorías lectales, es decir, conjuntos de esquemas cognitivos abstractos que capturan la esencia de la forma en

la que habla un grupo social.<sup>40</sup> Sin embargo, en este estudio, se observó que los rasgos centrales son los mismos en las dos redes sociales, lo que hace sospechar que dichos rasgos pueden estar más relacionados con características propias de la comunidad de habla que con los grupos sociales a los que pertenecen las redes.

El que ambas redes compartieran los esquemas entonativos prototípicos de cada categoría hizo surgir la idea de que los índices sociolingüísticos podían ser aquellos rasgos más frecuentes en un grupo que en otro, es decir, que tienen un mayor grado de centralidad en un grupo que en otros (lo que no necesariamente implica que sean los más prototípicos o centrales de una determinada categoría en un grupo social en particular). No obstante, esta idea no fue confirmada a través de la prueba de percepción, pues la mayoría de los rasgos prosódicos considerados índices sociolingüísticos, en un primer momento, resultaron no serlo. Es por ello que partir de una noción de estereotipo como la planteada por Amodio (2014) (según quien los estereotipos consisten en cuerpos de información semántica sobre las características y atributos genéricos o más abstractos de grupos sociales) podría resultar más provechoso, pues los estereotipos lingüísticos no estarían determinados por la frecuencia de uso de una serie de rasgos en particular, sino que serían construidos sobre la base de experiencias sociales de los miembros de una comunidad. Adicionalmente, esta perspectiva implicaría partir de las bases sociales, para identificar cuáles son los elementos lingüísticos que los usuarios consideran más generalizados en un grupo social, los cuales pueden no coincidir (como fue el caso en este trabajo) con los más centrales de dicho grupo (sea prototípicos o más centrales en relación con otros grupos).

Lo expuesto en el párrafo anterior explicaría la diferencia observada entre los rasgos identificados como índices sociolingüísticos en el análisis sociolingüístico cuantitativo y la percepción de los hablantes. Sin embargo, ahondar en esta idea implica profundizar en aspectos teóricos de la psicología social, que van más allá del alcance de esta investigación. No obstante, se considera que sería conveniente realizar un estudio en el que se busque determinar los rasgos prosódicos asociados con distintos estereotipos sociales para, luego, establecer el grado de centralidad que tienen dichos rasgos en las estructuras cognitivas de grupos sociales estrechamente relacionados con tales estereotipos y, finalmente, realizar una

---

<sup>40</sup> La definición de estereotipo presentada es una interpretación propia de las palabras de Kristiansen (2010: 226), quien define el concepto de la siguiente manera: “sets of fairly abstract linguistic schemata capturing the essence of what an out group speaks like, or more technically speaking: central images of lectal categories.”

prueba de percepción que incluya los rasgos prototípicos de dichos grupos, aquellos más centrales en un grupo social que en otros (como en el caso de este estudio) y los rasgos estereotípicos. Se considera que las aportaciones de un estudio desde esta perspectiva pueden ser muy valiosas, para obtener una mejor comprensión de la naturaleza cognitiva de los elementos lingüísticos que permiten relacionar el habla de un individuo con las cualidades de un grupo social con características específicas.

Por otra parte, en relación con aspectos netamente metodológicos, se considera que la prueba de percepción llevada a cabo en este trabajo posee especial relevancia, debido a que es el primer modelo, al menos hasta donde se pudo investigar antes de su realización, diseñado con la finalidad de constatar el estatus de rasgos identificados previamente como posibles índices sociolingüísticos. Si bien se entiende que todo el procedimiento de selección de los rasgos prosódicos sometidos a análisis, el diseño del procedimiento, la presentación de la información y otros aspectos de esta prueba de percepción pueden ser optimizados, la metodología que se siguió se considera un buen punto de partida. Aunque no se pudo confirmar el estatus de índice sociolingüístico de la mayoría de los patrones prosódicos previamente identificados como potenciales portadores de significado social, el hecho de que uno de ellos sí pudo ser determinado como tal y que los datos recolectados hayan servido para presentar la discusión acerca de la naturaleza cognitiva de los índices sociolingüísticos (esquemas cognitivos sociolingüísticamente prototípicos o estereotípicos) se considera suficiente para inferir que el procedimiento planteado es suficientemente efectivo como para constituir las bases del marco metodológico inicial en estudios posteriores. Sin embargo, la elaboración de las dos etapas del desarrollo de este modelo (la de validación y la prueba de percepción) revelaron algunos aspectos importantes que se deben tomar en cuenta en el diseño y ejecución de experimentos de este estilo.

Una de las decisiones que puede resultar más difícil de tomar en un inicio está relacionada con la naturaleza de los estímulos. Al manejar un enfoque sociolingüístico, lo primero que se considera es el uso de señales originales (sin ningún tipo de manipulación) tomadas del corpus de habla con el que se esté trabajando, producidas con tanta espontaneidad como sea posible en el contexto de la entrevista sociolingüística. No obstante, al tomar en cuenta las variables fonéticas (tanto segmentales como suprasegmentales) y acústicas que entran en juego durante la grabación de un corpus de habla, se pueden dar al

menos dos situaciones potencialmente negativas (que fueron las enfrentadas en el caso de este trabajo) para el cumplimiento del objetivo de una prueba de percepción: i) la presencia de rasgos prosódicos (en secciones del enunciado distintas a las que se quieren estudiar) o de material léxico que provea información sociolingüística del informante del que se tomó la muestra; ii) la existencia de factores extralingüísticos que interfieran en la percepción óptima de las cualidades prosódicas del estímulo. Esto hace que el número de señales de audio aptas para ser tomadas en cuenta en la prueba se reduzca considerablemente, al punto de que, incluso, sea imposible contar con un estímulo para algunos de los aspectos prosódicos que se desean analizar.

Ahora bien, puede que las situaciones planteadas previamente no representen un problema para el correcto desarrollo de algunas investigaciones sociolingüísticas, de acuerdo con sus objetivos particulares. Sin embargo, debido a que la metodología de análisis de este trabajo se fundamentó en la teoría basada en el uso, los elementos prosódicos que se tomaron en cuenta para la prueba de percepción respondían a frecuencias de producción bastante concretas, por lo que se consideró que la mejor alternativa para cubrir los objetivos planteados fue la manipulación de las señales de audio, para replicar artificialmente los rasgos prosódicos identificados como índices sociolingüísticos potenciales. No obstante, esto hace que se deba poner en una balanza el grado de espontaneidad que se desea presentar en los estímulos y el nivel de control para alcanzar el resultado deseado. Los resultados de la prueba de validación de estímulos permitieron ver que cualquier tipo de manipulación en la señal de audio le resta puntos en espontaneidad y le suma en artificialidad. Por ello, un estímulo que cuente con un alto grado de control puede llegar a perder todo tipo de valor sociolingüístico, al no poder ser asociado con la comunidad de habla en cuestión (o con ninguna en concreto, dependiendo del tipo y la cantidad de alteraciones acústicas realizadas, pues el resultado puede llegar a percibirse altamente artificial). Por otra parte, un menor grado de control en el estímulo permite obtener niveles más elevados de asociación con la comunidad de habla que se desea estudiar (que en este caso fue la comunidad de la Ciudad de México).

Por lo tanto, se considera que, cuando no se puede alcanzar o mantener un nivel de espontaneidad ideal en los estímulos, el factor que se debe priorizar es la identidad comunitaria de la muestra de habla. Para ello, es recomendable evitar un elevado número de

alteraciones globales (a lo largo del contenido lingüístico) o locales (en una sola sílaba, por ejemplo) en la señal de audio original. Esto es especialmente relevante cuando las modificaciones tienen que ver con la duración, pues las alteraciones de este elemento son las que pueden elevar la percepción de artificialidad con mayor facilidad. En cualquier caso, se considera de gran importancia tener en cuenta que las manipulaciones de rasgos prosódicos pueden suponer un alto grado de exigencia a los límites de tolerancia de los jueces, en términos de espontaneidad y reconocimiento comunitario. Por tal motivo, se sugiere realizar pruebas preliminares (como la prueba de validación de estímulos llevada a cabo en este trabajo) para cerciorarse de que los estímulos cubren los parámetros mínimos de aceptabilidad, para el cumplimiento de los objetivos de la investigación que se desea realizar.



## Conclusiones

En esta sección se presentan las conclusiones generales de la investigación. Adicionalmente, se comentan ciertos aspectos considerados relevantes para alcanzar una mejor comprensión de la prosodia en términos fonológicos, fonéticos, sociolingüísticos, cognitivos y sociocognitivos. En primera instancia, se hará referencia a cada una de las preguntas de investigación (presentadas aquí en el mismo orden en que se plantearon en §A.6) y a las hipótesis que se plantearon inicialmente en torno a ellas, seguidas por algunas anotaciones basadas en los resultados obtenidos a lo largo del desarrollo de la investigación. Posteriormente, se abordan los puntos teóricos y metodológicos que guiaron el diseño de todos los aspectos de las fases de producción y percepción de este trabajo. Para culminar, se exponen las reflexiones finales relacionadas con el análisis sociolingüístico.

### 6.1. Respuestas a preguntas de investigación y evaluación de las hipótesis

La primera pregunta de esta investigación, relacionada con cuáles son los parámetros prosódicos de los enunciados aseverativos producidos por cada red social, permitió el planteamiento de la hipótesis de que los parámetros prosódicos de los enunciados aseverativos de cada red corresponden con los previamente descritos para la variedad de la Ciudad de México; aunque, al mismo tiempo, existen diferencias entre ambas redes. Tras analizar los resultados, se pudo confirmar la hipótesis, al constatar que las dos redes estudiadas presentan patrones prosódicos relacionados con el habla de la Ciudad de México, como la alta frecuencia del tonema circunflejo (L+H\* L%). Asimismo, los rasgos prosódicos más frecuentes de una red coinciden en gran medida con los de la otra: acentos tonales abstractos H\* y L+H\* en la primera sílaba tónica, intertonemas no oxítonos abstractos altos (H\* H-) o ascendentes (L+H\* H-), intertonema oxítono abstracto L+H\* H-, tonemas abstractos L\* L%, H\* L%, L+H\* L% (circunflejo) y H+L\* L% (en orden de mayor a menor ocurrencia). Sin embargo, existen diferencias notorias entre las producciones prosódicas de ambas redes, como la producción de acentos tonales H+L\* en la primera sílaba tónica y una mayor duración de la sílaba final tónica de ip, en la que recae el intertonema oxítono por parte de los hablantes de la red laboral, al igual que la emisión notoriamente superior de tonemas no descendentes en hablantes de la red vecinal.

En segundo lugar, se cuestionó la posibilidad de que se pudiera aplicar, en el plano prosódico, el mismo enfoque basado en el uso que se ha empleado para la delimitación de elementos categoriales prototípicos y periféricos en el plano segmental. Dicho planteamiento originó la hipótesis de que era posible determinar los elementos prototípicos y periféricos en el habla de dos redes sociales con base en el enfoque basado en el uso. Una vez más, esta hipótesis quedó confirmada. Sin embargo, es necesario destacar que el enfoque basado en el uso no es totalmente aplicable al análisis de valores absolutos de los rasgos fonéticos (expresados en st, dB y ms), aunque sí lo es en aquellos casos en los que los datos prosódicos se pueden plantear en términos normalizados o relativos. Por ejemplo, si mil hablantes leen un mismo enunciado produciendo un ascenso tonal en la primera sílaba tónica, es probable que los valores absolutos en semitonos varíen entre cada una de las emisiones registradas, lo que generaría mil elementos distintos y haría inviable un análisis basado en el uso. En cambio, si se transforman estos valores absolutos en variables normalizadas o relativas, como son las etiquetas del sistema de transcripción ToBI, se obtiene un número de unidades estandarizadas, susceptibles de ser cuantificables más fácilmente de acuerdo con su frecuencia de producción, lo que hace viable un análisis basado en el uso. Ahora bien, no existe, hasta este momento (al menos hasta donde se pudo investigar), un sistema que permita hacer lo mismo con los valores absolutos de la duración, por lo que un método de análisis basado en duración relativa de sílabas específicas en el enunciado (como el planteado en este trabajo en las sílabas que constituyen los intertonemas no oxítonos y los tonemas) puede servir para normalizar los valores de duración y transformarlos en un repertorio reducido de unidades susceptibles de ser cuantificables de acuerdo con su frecuencia de producción en el habla.

En relación con la tercera pregunta de investigación (¿cuáles son los rasgos prosódicos prototípicos y periféricos de cada una de las redes sociales estudiadas?), se manejó la hipótesis de que los rasgos prosódicos prototípicos corresponden con los identificados previamente como rasgos propios de la Ciudad de México (Matluck, 1952; Kvavik, 1979; Sosa, 1999; Martín Butragueño, 2004, 2006a, 2006b, 2011, 2021), mientras que los periféricos son aquellos que no se alinean con las tendencias determinadas por las referidas investigaciones previas. Esta hipótesis quedó confirmada tras el análisis sociolingüístico basado en el uso y se pudieron identificar rasgos prosódicos con distintos

grados de centralidad en cada red. A continuación, se hace un breve recuento de los elementos prosódicos que resultaron ser prototípicos en ambas redes simultáneamente (al menos en la mayoría de los casos), por lo que se debe entender que los demás rasgos prosódicos poseen una posición más periférica en la organización de la categoría cognitiva. En los intertonemas no oxítonos, se encontró que el elemento entonativo más frecuentes fue H\* H- y que la duración relativa prototípica consiste en una mayor longitud de la sílaba final de ip en relación con la nuclear. En el caso de los intertonemas oxítonos, la configuración tonal más frecuente fue L+H\* H-. Finalmente, el tonema prototípico para las dos redes fue L\* L%, al tiempo que la duración relativa más frecuente de la red social laboral fue la consistente en una mayor duración de la sílaba nuclear, en contraste con la duración más breve de la sílaba final. Resulta interesante que la red vecinal no presentó un tipo de duración relativa más frecuente, pues la sílaba nuclear es más larga que la final en el mismo número de ocasiones que en las que ocurre lo contrario. Dicha igualdad de casos en ambos tipos de duración relativa lleva a concluir que, para esta red, no existe un esquema de duración prototípico en la sección nuclear del enunciado.

La siguiente pregunta de investigación planteaba la cuestión de si los índices sociolingüísticos están conformados por un solo rasgo prosódico o por la combinación de varios. La hipótesis generada sobre la base de esta interrogante fue que los índices sociolingüísticos están constituidos por las características conjuntas de más de un rasgo prosódico. Los resultados de la prueba de percepción permitieron ver que tal afirmación no es verdadera, pues el único patrón prosódico que se identificó como índice sociolingüístico consistía exclusivamente en un rasgo de la duración. Esto permite concluir que un índice sociolingüístico puede estar integrado por solo un elemento prosódico sin que exista, necesariamente, participación de otro rasgo en la percepción de un significado social.

La quinta pregunta de investigación consistió en la siguiente interrogante: ¿Los rasgos prosódicos identificados como índices sociolingüísticos mediante el análisis estadístico también son considerados así por los jueces oyentes en la prueba de percepción? Esta pregunta surgió a la par de la hipótesis de que los jueces serían capaces de reconocer el significado sociolingüístico en los rasgos identificados como índices, durante el análisis sociolingüístico. Tras analizar los datos de la fase de producción, se pudo observar que las respuestas de los jueces no confirmaron el estatus de índice sociolingüístico, previamente

atribuido a la mayoría de los rasgos prosódicos sometidos a prueba. Solo uno de los rasgos prosódicos que se presumía como índice sociolingüístico (el patrón de duración  $D^{((DX2)+20)}$  para la red vecinal) fue constatado.

La última pregunta de investigación reflejaba la duda con respecto a la cualidad cognitiva (prototípica o periférica) de los rasgos prosódicos prototípicos que indexan significado sociolingüístico, ante lo cual se planteó la hipótesis de que los rasgos prosódicos prototípicos de cada red son los que transmiten el significado social que les permite a los miembros de la comunidad de habla (en este caso, de la Ciudad de México) identificar a un hablante como perteneciente a un grupo social o a otro. La hipótesis no pudo ser comprobada debido a que, como se planteó en el párrafo anterior, los rasgos prototípicos de una red también suelen tener el mismo estatus cognitivo en la otra. Por lo tanto, son los elementos prosódicos periféricos los que se consideran más propensos a ser portadores de significado social. Sin embargo, aunque los rasgos que se consideraron índices sociales al final del análisis sociolingüístico son periféricos en sus respectivas redes, el aspecto que se estima más relevante en la asignación de estatus de índice sociolingüístico potencial es el grado de centralidad que posee un elemento en la estructura cognitiva de una red, con respecto al grado que posee en la estructura de la otra. Tómese como ejemplo de esto el tonema abstracto no descendente L+H\* H%, el cual a pesar de considerarse periférico por encontrarse considerablemente lejos (en frecuencia de uso) del tonema prototípico, no es el más periférico en ninguna de las dos redes. No obstante, los casos de producción de esta configuración por parte de la red vecinal casi duplican los de la red laboral, lo que permite inferir que este tonema tiene un mayor grado de prototipicidad (centralidad) en la categoría cognitiva de la red vecinal que en la de la laboral, por lo que se planteó como un posible índice sociolingüístico, para la primera de estas redes.

## 6.2. Conclusiones teóricas y metodológicas

El análisis prosódico no suele ser una tarea fácil debido a su complejidad. En el caso de este trabajo, desde el inicio se abordó un enfoque pluriparamétrico que resultó altamente beneficioso. El uso de este enfoque, al analizar la intensidad, duración y F0 en sílabas específicas de los enunciados, permitió descartar la intensidad como elemento de variación sociolingüística y llevó a prestar más atención a la duración y F0, que sí presentaron

diferencias estadísticamente significativas entre las dos redes estudiadas. De hecho, el único índice sociolingüístico que se pudo confirmar consistió en una característica relacionada con la duración (y no con la entonación). Por lo tanto, se recomienda partir de una base pluriparamétrica en el estudio sociolingüístico de la prosodia, a partir del cual se puede estrechar el cerco alrededor de aquellos elementos suprasegmentales, que se estimen pertinentes para el cumplimiento del objetivo de cada investigación particular.

Otro aspecto que resultó ser bastante conveniente en el desarrollo de este trabajo fue el análisis entonativo en dos niveles, como también proponen Dorta *et al.* (2013) y Hualde y Prieto (2016). Por un lado, una transcripción tonal en el nivel concreto permite representar con mayor grado de detalle los movimientos de F0 en el fragmento de habla analizado. Por otro lado, una transcripción en el nivel abstracto no solo es de ayuda para agrupar distintos elementos concretos, de acuerdo con sus características centrales invariables, sino que también permite observar el esquema entonativo subyacente que da pie a las distintas formas concretas producidas por los hablantes. El conjunto de ambos brinda la posibilidad de registrar y presentar las múltiples formas concretas en las que se puede presentar un mismo tono abstracto y, por ende, obtener un mayor grado de detalle en el análisis que se desea realizar. Por ejemplo, en el caso de este trabajo, los datos registrados durante la fase de producción sirvieron para delimitar el cerco en torno a las configuraciones tonales abstractas que eran propensas a poseer significado social. Una vez que se identificaron los contornos tonales, se procedió a determinar si había algún tipo de relación sobresaliente entre sus realizaciones concretas y las redes sociales estudiadas. De esta forma, se logró identificar que los ascensos tonales en los intertonemas oxítonos tenían dos realizaciones que podrían incidir en la percepción del esquema abstracto como índice sociolingüístico: una elevación entre 1.5st y 3st (L+H\* H-) y otra de 3st o más (L+<sub>i</sub>H\* HH-). Al someter ambas formas concretas a la prueba de percepción, fue posible confirmar que el grado de la elevación tonal implica un cambio en la asociación con el MdV del hablante que lo produce, específicamente, en el caso de los J1; mientras que, para los demás grupos de jueces, la excursión tonal no es relevante al momento de relacionarla con un MdV en particular. Este hecho hace posible concluir que las configuraciones tonales abstractas pueden tener un significado invariable en todas sus representaciones concretas (al menos en términos sociolingüísticos), como en el caso de los J2 y J3, o pueden estar asociadas a varios significados simultáneamente, los cuales

están condicionados por las realizaciones concretas asociadas con el elemento abstracto, tal como revelan los datos obtenidos de los J1.

Adicionalmente, este tipo de análisis permite eliminar representaciones ambiguas que pudieran llevar a interpretaciones erradas. Un ejemplo evidenciado tanto en el análisis fonético-fonológico como en el sociolingüístico de este trabajo es el tonema no descendente H\* H%. Al observar esta etiqueta de transcripción, lo primero que puede venir a la mente es que el tono de juntura es más elevado que el acento nuclear. Sin embargo, asumir que este es el caso en los datos del corpus de este trabajo sería un error, pues la juntura H% no corresponde con un tono más elevado que el de la sílaba nuclear en ninguno de los enunciados analizados, aunque sí corresponde con un tono alto al final del enunciado. En esta investigación, el análisis en dos niveles hizo posible dejar de lado cualquier tipo de ambigüedad o mal entendido, al reflejar que esta juntura alta corresponde, en el nivel concreto, con un tono alto sostenido desde la sílaba nuclear (cuya representación concreta corresponde con =%) y no con una elevación en la sílaba final. Todo esto sin dejar de representar la naturaleza abstracta de dicho tono como una juntura H%. Si bien es cierto que H% representa un tono alto que puede estar al mismo nivel del acento nuclear H\* o a un nivel superior, no se puede negar que esta falta de especificidad pudiera causar confusión al momento de presentar datos sin contar con un espectrograma con F0, que haga visible y explícito el rasgo tonal al que se hace referencia. Cabe señalar que Hualde y Prieto (2016) ya habían hecho referencia a este tipo de dificultades en los estudios que emplean solo un nivel de análisis entonativo y, de igual modo, destacan otros beneficios del uso de dos niveles en el estudio de la entonación. A las ventajas señaladas por estos autores, se puede sumar, desde las bases cognitivas de este trabajo, que un análisis entonativo en dos niveles permite cuantificar las producciones de un elemento abstracto junto con las de sus realizaciones concretas. Gracias a esto, en esta investigación se pudo ver, por ejemplo, la relevancia cognitiva del intertonema no oxítono abstracto H\* H- (el prototípico en ambas redes sociales) y la frecuencia con la que se manifestaba a través de sus diferentes realizaciones concretas. Esto permitió llevar a cabo un análisis sociolingüístico más profundo y detallado al tener en cuenta el trasfondo cognitivo de las representaciones abstractas y las formas concretas a través de las cuales se manifestaron.

Por otra parte, se considera relevante hacer una serie de observaciones relacionadas con la sección del enunciado que va desde la última sílaba acentuada de la frase entonativa intermedia (ip) no final (es decir, aquella cuya última sílaba no coincide con la última del enunciado) hasta el final de ella. Aunque se ha puesto en duda la pertinencia de la frase entonativa intermedia como nivel prosódico en español (como se mencionó en §1.1.3), los datos recopilados en esta investigación no solo avalan la pertinencia de esta unidad, sino que, también, permiten plantear la existencia de un núcleo en ella (denominado a lo largo de este trabajo “tonema interno” o “intertonema”). De acuerdo con Sosa (1999: 49) “el núcleo coincide con las sílabas finales a partir de la que lleva el último acento” y dicho elemento es el único requerimiento para que exista un grupo melódico. Este autor, quien maneja un enfoque netamente fonológico, es detractor de la idea de que el español es propenso a poseer frases entonativas intermedias (ip) y no considera que un núcleo deba estar necesariamente al final de una unidad sintáctica con sentido completo, pues afirma: “desde el punto de vista de la organización prosódica, el factor necesario y suficiente que va a permitir la generación de un grupo melódico es el estatus acentual de la secuencia, no su contenido léxico o sus significados potenciales” (Sosa 1999: 50). Por tal motivo, estrictamente hablando, desde la perspectiva de este autor, lo que se ha denominado “intertonema” en este trabajo debería ser considerado un tonema, pues se encuentra al final de una unidad melódica, que puede estar seguida (o no) por una pausa. Ahora bien, si por una parte se comparte la perspectiva del autor con respecto a las cualidades de un núcleo prosódico, pero no ocurre lo mismo con la noción de que los núcleos de las frases entonativas intermedias deban ser considerados tonemas. La razón de esto es que no se considera que las ip sean grupos melódicos autónomos. Partiendo de las aportaciones de Hualde (2003), Prieto (2006), Nespor y Vogel (2007) y Von Heusinger (2007), se puede plantear que los tonemas encabezan el contorno melódico de todo el material prosódico previo y le dan un significado semántico al enunciado completo, mientras que los tonemas internos solo tienen este efecto sobre el material lingüístico dentro de los límites de una ip no final, sin tener injerencia en la ip previa (en caso de haberla) o en la siguiente, y sin que el contenido de la ip en cuestión constituya un enunciado con sentido completo en sí misma. Por ende, se considera que los tonemas internos tienen un significado semántico de no conclusión del enunciado y la función de indicar que

el contorno entonativo emitido hasta el momento está subordinada a o es una fracción de uno mayor encabezado por el tonema.

En concordancia con lo anterior, en los datos obtenidos del corpus de esta investigación, las configuraciones tonales más frecuentes al final de una ip (H\* H- y L+H\* H-) concuerdan con los planteamientos del código de producción propuesto por Gussenhoven (2004), según el cual el uso de junturas altas (H-) implica la transmisión de un significado de continuidad o no conclusión del enunciado. Si bien este significado no es el primario de un enunciado, pues este recae en el tonema de la frase entonativa mayor (IP), no se puede pasar por alto que estos ascensos representan una codificación fonética y fonológica para un significado secundario contenido dentro de los límites del enunciado. Dicha codificación se inicia desde la producción de un tono alto o ascendente en la última sílaba tónica de la ip y finaliza con un tono de juntura igualmente alto, ya sea al mismo nivel tonal alcanzado en la última tónica o en un nivel mayor. De igual forma, la duración de las sílabas que constituyen los intertonemas se puede equiparar con la del tonema en cuanto a que las medias de ambas se oponen a las del resto de las sílabas del enunciado, ya que son considerablemente (y perceptivamente) más largas (ver §3.1.1). Esta coincidencia en la duración del intertonema y el tonema demuestra que los hablantes reproducen el mismo de comportamiento lingüístico (mayor duración silábica) antes de culminar lo que consideran una frase entonativa (sea intermedia o mayor), posiblemente para que los movimientos tonales nucleares (al final de las frases entonativas) puedan ser producidos y percibidos con mayor facilidad.

Por lo antes expuesto, se puede concluir que la frase entonativa intermedia no final es un nivel pertinente en la jerarquía prosódica del español de la Ciudad de México (y posiblemente del español general) debido a que su sección final posee características suprasegmentales únicas que sirven para diferenciarla prosódicamente del resto del enunciado: se encuentra delimitada por configuraciones tonales ascendentes o altas (asociadas con un significado de no conclusión del enunciado) y posee sílabas finales con duración semejante a las de núcleo de la frase entonativa mayor (IP), es decir, con mayor duración que las del resto del enunciado. Además de esto, la evidencia obtenida en la presente investigación permite constatar la presencia de un núcleo en las frases entonativas intermedias no finales. Si bien en un inicio se planteó el término “tonema interno”, sobre la base de que transmitía un significado secundario relacionado con la continuidad del



enunciado luego de la frase intermedia no final, el hecho de que la duración de las sílabas en esta sección del enunciado se asemeje a la de los tonemas cambia el panorama. Este hecho revela la misma marca o codificación de duración para producir los núcleos y lindes de las frases entonativas (no finales y finales), lo que pone el intertonema al mismo nivel que el tonema como elemento delimitador de un nivel prosódico. La entonación, sin embargo, es la que cumple un papel diferenciador en el nivel de significado (y de orden en la jerarquía) de cada una de las frases entonativas, pues este es el elemento empleado para distinguir las frases entonativas con un significado semántico primario (las IP) de aquellas con significado secundario (las ip).

### 6.3. Conclusiones sociolingüísticas

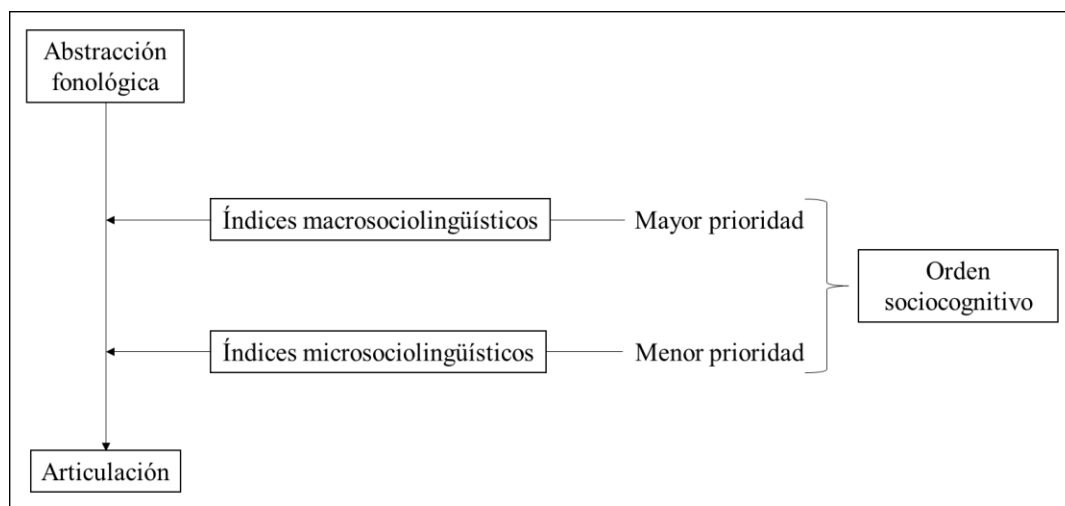
Tras llevar a cabo el análisis fonético-fonológico, se pudo constatar que las tendencias principales de una red casi siempre resultaron ser las principales de la otra: acento tonal H\* en la primera sílaba tónica, sílabas finales de ip más largas que las nucleares en los intertonemas no oxítonos, intertonema no oxítono H\* H-, intertonema oxítono L+H\* H-, tonema L\* L%. De este análisis, se concluyó que los rasgos prosódicos prototípicos de ambas redes representan características lingüísticas de la comunidad de habla (la Ciudad de México, en este caso), mientras que los periféricos son los que constituyen cualidades lingüísticas específicas de cada grupo de hablantes y que, por ende, tienen el potencial de indexar significado social. Por otra parte, la prueba de percepción confirmó el estatus de índice sociolingüístico de solo uno de los rasgos prosódicos previamente caracterizado de esta forma: tonemas de IP en los que la combinación de los valores medios de las sílabas que los constituyen se encuentren alrededor de los 376 ms como índice sociolingüístico de la red social 1 (vecinal).

Una de las implicaciones que tienen estos hechos, en términos sociocognitivos, es que dan pie a la creación de una hipótesis de prioridad indexical. Esta hipótesis plantea que los rasgos prosódicos que indexan significado sociolingüístico relacionado con la comunidad de habla se asignan cognitivamente primero que los que portan significado asociado con grupos sociales específicos. El hecho de que los esquemas cognitivos prototípicos fuesen compartidos por todos los informantes del corpus, mientras que los esquemas periféricos pueden llegar a tener más frecuencia en una red que en otra, hace surgir la idea de la

existencia de un orden sociocognitivo que responde a un sistema de priorización del tipo de significado social incluido en un acto de habla.

Según la hipótesis propuesta de prioridad indexical, la atribución de significado relacionado con la comunidad de habla tiene mayor prioridad cognitiva que la relacionada con la membresía a un grupo social determinado. Planteado en otros términos, la inclusión de índices macrosociolingüísticos (relacionados con la comunidad de habla de la Ciudad de México, en este caso) se llevaría a cabo primero, antes que la adición de los índices microsociolingüísticos (pertinentes a los de la red o grupo social al que pertenece el hablante),<sup>41</sup> como se ejemplifica en la figura B.1. Esto conlleva la posibilidad de que el orden cognitivo cambie total o parcialmente (con mayor inclusión de índices microsociolingüísticos), cuando el acto comunicativo se realice entre dos personas pertenecientes a la misma red o grupo social, ya que en estos casos la prioridad puede ser confirmar o afianzar la pertenencia a dicha red o grupo para mantener o estrechar vínculos sociales, demostrar solidaridad de grupo o llevar a cabo algún otro acto social. Sin embargo, la constatación de esta hipótesis queda pendiente para una investigación posterior.

Figura 6.1. Modelo de orden sociocognitivo de introducción de significado social en la prosodia



<sup>41</sup> Los términos macro y microsociolingüísticos planteados en la hipótesis de prioridad indexical están relacionados con el tercer y primer orden indexical (respectivamente) planteados por Johnstone *et al.* (2006).

El contraste entre los resultados obtenidos en la fase de análisis sociolingüístico y en la de percepción permite ver que el hecho de que un rasgo prosódico sea producido con una frecuencia significativamente mayor (en términos estadísticos) por un grupo social *X* que por uno *Y* no necesariamente implica que este rasgo pueda ser identificado como característico del grupo *X*. Dicho de otro modo, la percepción de los miembros de una comunidad de habla no necesariamente va de la mano con los datos cuantitativos objetivos, obtenidos a través de medios estadísticos en un análisis sociolingüístico. Esto se debe a que los estudios sociolingüísticos se enfocan (principalmente) en el estudio de las variaciones que potencialmente generan cambios lingüísticos, para lo cual los datos de producción representan obligatoriamente el eje central del análisis. Sin embargo, cuando el foco de la investigación radica en la identificación de índices sociolingüísticos (como es el caso en el presente trabajo), los datos de percepción cobran mayor relevancia.

Si se entiende este concepto como el significado social incluido dentro de un rasgo lingüístico (Irvine, 2001; Eckert, 2008), entonces no se puede afirmar la condición indexical de un elemento lingüístico sin haberlo sometido primero al juicio de miembros de la comunidad de habla. Si dichos miembros no pueden identificar significado sociolingüístico en el elemento estudiado (como fue el caso de la gran mayoría de los patrones prosódicos sometidos a la prueba de percepción), entonces no se puede decir que sea un índice sociolingüístico. A fin de cuentas, solo se puede constatar que un elemento lingüístico posee un significado cuando los usuarios de la lengua (en términos locales o generales) son capaces de reconocerlo. Por lo tanto, se sugiere tratar el uso del término “índices sociolingüísticos potenciales” u otro similar durante la fase de análisis cuantitativo o cualitativo; y únicamente otorgarles el estatus de “índices sociolingüísticos” una vez que hayan sido confirmados mediante una prueba de percepción o algún otro método que tome en cuenta la valoración de una muestra de miembros de la comunidad de habla bajo estudio.

La comparación entre los resultados de las dos fases de la presente investigación también puede implicar que el enfoque utilizado para explorar categorías cognitivas en el análisis lingüístico sociocognitivo de datos de habla difiere, o debe diferir, del utilizado en el perceptivo; pues cada uno puede hacer referencia a elementos cognitivos con naturaleza diferente en términos indexicales. Es posible que se tenga que refinar la metodología del estudio sociolingüístico y la de percepción empleada en este trabajo (desde la identificación

de índices potenciales en el análisis sociolingüístico hasta la naturaleza de los estímulos), para aumentar la identificación exitosa de índices sociolingüísticos potenciales y la tasa de confirmación de estos por parte de los jueces. Si esto también fallara (en el sentido de que las coincidencias siguieran siendo pocas o mínimas) tras múltiples intentos, tendría que considerarse la posibilidad teórica de que las estructuras cognitivas de las que se toma la información durante la producción lingüística sean diferentes a aquellas a las que el oyente recurre durante la percepción. La hipótesis planteada en este caso sería que, durante un evento comunicativo llevado a cabo con habla espontánea, el emisor recurre a categorías cognitivas especializadas en el almacenamiento de esquemas prosódicos (con significado semántico y pragmático), para producir un mensaje; y el receptor apela a categorías cognitivas especializadas en la agrupación de esquemas prosódicos necesarios para decodificar un mensaje en términos lingüísticos y sociolingüísticos (los cuales no necesariamente corresponden con aquellas categorías a las que recurre cuando toma su turno de habla).

En principio, esta hipótesis relacionada con la diferenciación entre categorías cognitivas prosódicas productivas y perceptivas radicaría en el hecho de que, durante la emisión de un enunciado aseverativo espontáneo, la producción de índices es un proceso automático (espontáneo), mientras que su identificación en la percepción es controlada. Inclusive, se considera viable que la categoría cognitiva empleada durante la percepción de significado social pueda estar más relacionada con el almacenamiento de información lingüística asociada con estereotipos sociales que con los modos de vida (que fue la variable social empleada en el análisis perceptivo de este estudio). Dicha posibilidad se plantea debido a que los patrones prosódicos analizados con base en los modos de vida pueden ser relativamente uniformes en distintos grupos sociales con modos de vida diferentes, lo que causaría que algunos patrones que se consideren índices sociales en el análisis cuantitativo y cualitativo de datos de habla no sean considerados así, por un grupo de jueces (como sucedió en este trabajo). En cambio, los estereotipos sociales están estrechamente relacionados con las características de determinados grupos sociales, por lo que es posible que los esquemas incluidos en categorías cognitivas especializadas en la percepción social consistan en aquellos patrones más prominentes de dichos grupos (y no en un rango más amplio de esquemas prosódicos como el observado en el análisis basado en modos de vida).

Sea cual sea el caso, el punto de inicio en la búsqueda de índices sociolingüísticos debe ser necesariamente el análisis sociolingüístico. El uso de la metodología sociolingüística es un recurso invaluable como guía para señalar hechos cuantificables concretos y para la interpretación de los datos. Sin embargo, no se considera que la sociolingüística tradicional (efectiva para el estudio de la variación y el cambio lingüístico) necesariamente provea las herramientas requeridas para analizar aspectos cognitivos. La naturaleza de un hecho lingüístico, en términos empíricos, marca los patrones de una realidad lingüística que puede ser ajena a la conciencia lingüística colectiva, por lo que también estarían alejados de la realidad cognitiva de los hablantes. Esto hace surgir la siguiente pregunta: ¿Qué debe tener más peso para los sociolingüistas, la naturaleza social del lenguaje o la cognitiva? La respuesta yace en medio de las dos corrientes, pues estas facetas del ser humano no son mutuamente excluyentes, sino que ambas (la formación de grupos sociales y el empleo de un lenguaje articulado complejo) forman parte de la esencia humana que nos diferencia del resto de las especies del planeta. Por lo tanto, se considera que el mejor enfoque para comprender la naturaleza del lenguaje humano debe tomar en cuenta tanto la fuente del mismo lenguaje (el sistema cognitivo, únicamente analizable por medio de pruebas con fundamentos psicológicos, como las de percepción) y el contexto de uso del lenguaje (el entorno social).

## Referencias bibliográficas

- Abdurrahman, Muhammad Ibn Abdullah. 2014. *Sociophonetic perception of African American English in Minnesota*. Tesis doctoral. University of Minnesota, Estados Unidos de América.
- Adobe Systems Incorporated. 2013. *Adobe Audition* [Programa informático]. Versión CS6. Delaware, Estados Unidos de América: Adobe Systems Incorporated.
- Amodio, David. 2014. The neuroscience of prejudice and stereotyping. *Nature Reviews: Neuroscience*, vol. 15, nº. 10. 670-682.
- Areiza Londoño, Rafael, Mireya Cisneros Estupiñán y Luís Tabares Idárraga. 2012. *Sociolingüística: enfoques pragmático y variacionista* (2ª ed.). Bogotá, Colombia: Ecoe.
- Baqué, Lorraine y Mónica Estruch. 2003. Modelo Aix-en-Provence. En Pilar Prieto (ed.), *Teorías de la entonación*, 123-153. Barcelona, España: Ariel.
- Beckman, Mary, Manuel Diaz-Campos, Julia Tevis McGory y Terrell Morgan. 2002. Intonation across Spanish, in the tones and break índices framework. *Probus*, nº. 14. 9-36.
- Bello, Andrés. 1847. *Gramática de la lengua castellana destinada al uso de los americanos*. Chile: Imprenta del Progreso.
- Boersma, Paul y David Weenink. 2019. *Praat: doing phonetics by computer* [Programa informático]. Versión 6.1.06. Disponible en: <http://www.fon.hum.uva.nl/praat/>. Consulta [08 de noviembre de 2019].
- Blanco Canales, Ana. 2000. Redes sociales y variación sociolingüística. *REIS: Revista Española de Investigaciones Sociológicas*, nº. 91. 115-136.
- Blas Arroyo, José Luis. 2004. *Sociolingüística del español: Desarrollos y perspectivas en el estudio de la lengua española en contexto social*. Madrid, España: Cátedra.
- Broś, Karolina. 2013. La aspiración y la pérdida de /s/ en el español de Chile como ejemplo de opacidad. *Onomáein*, nº.28. 56-71.

- Burgoon, Erin, Marlene Henderson y Arthur Markman. 2013. There are many ways to see the forest for the trees: a tour guide for abstraction. *Perspectives on psychological science*, vol. 8, n°. 5. 501-520.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American speech*, vol. 82, n°. 1. 32-64.
- Carrasquero, Viktor. 2010. Un caso de variación sociofonética: /-s/ posnuclear en el español actual de Caracas. *Letras*, vol. 52, n°. 81. 63-86.
- Chevrot, Jean-Pierre, Celine Dugua y Michel Fayol. 2009. Liaison acquisition, word segmentation and construction in French: A usage-based account. *Journal of Child Language*, vol. 36, n°. 3. 557-596.
- Clopper, Cynthia, Brianna Conrey y David Pisoni. 2005. Effects of talker gender on dialect categorization. *Journal of Language and Social Psychology*, vol. 24, n° 2. 182-206.
- Croft, William y D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge, Reino Unido: Cambridge University Press.
- Cuenca, Maria Josep y Joseph Hilferty. 1999. *Introducción a la lingüística cognitiva*. Barcelona, España: Ariel.
- Dąbrowska, Ewa. 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflexions: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language*, vol. 58. 931-951.
- De-la-Mota, Carmen, Pedro Martín Butragueño y Pilar Prieto. 2010. Mexican Spanish intonation. En Pilar Prieto y Paolo Roseano (eds.), *Transcription of the intonation of the Spanish language*, 319-350. München, Alemania: Lincom.
- Díaz, Chaxiraxi y Josefa Dorta. 2015. Acentos tonales y variantes: declarativas en habla formal y espontánea de La Gomera (Islas Canarias). *Estudios de Lingüística*, N° 29. 53-80.
- Díaz-Campos, Manuel y Jason Killam. 2012. Assessing language attitudes through a matched-guise experiment: the case of consonantal deletion in Venezuelan Spanish. *Hispania*, vol. 95, n°. 1. 83-102.
- Díaz-Campos, Manuel y Rebecca Ronquest. 2007. La percepción de acentos tonales en enunciados afirmativos. *Estudios de Fonética Experimental*, vol. XVI. 81-98.

- D'Introno, Francesco; Enrique Del Teso y Rosemary Weston. 2010. *Fonética y fonología del español*. Madrid, España: Cátedra.
- Dorta, Josefa, Beatriz Hernández y José Antonio Martín Gómez. 2013a. Comparación de la entonación canario-cubana. En, Josefa Dorta (ed.), *Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela*, 87-168. Santa Cruz de Tenerife, España: La página.
- Dorta, Josefa, Chaxiraxi Díaz, Beatriz Hernández, Carolina Jorge Trujillo, José Antonio Martín Gómez. 2013b. El marco de la investigación: aspectos metodológicos. En, Josefa Dorta (ed.), *Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela*, 53-83. Santa Cruz de Tenerife, España: La página.
- Dorta, Josefa, Chaxiraxi Díaz, Elsa Mora, Carolina Jorge Trujillo y Nelson Rojas. 2013c. Comparación de la entonación canario-venezolana. En, Josefa Dorta (ed.), *Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela*, 171-250. Santa Cruz de Tenerife, España: La página.
- Dorta, Josefa, José Antonio Martín y Carolina Jorge Trujillo. 2017. Intensity threshold: beyond pure tones. *Estudios de Fonética Experimental*, vol. XXVI. 133-163.
- Dugua, Céline, Elsa Spinelli, Jean-Pierre Chevrot y Michel Fayol. 2009. Usage-based account of the acquisition of liaison: Evidence from sensitivity to the singular/plural orientation of nouns. *Journal of Experimental Child Psychology*, vol. 102, nº. 3. 342-350.
- Eckert, Penelope. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society*, vol. 17, nº. 2. 183-207.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics*, vol.12, nº. 4. 453-476.
- Eckert, Penelope. 2012. Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, vol. 41. 87-100.
- Eckert, Penelope. 2018. *Meaning and linguistic variation: the third wave in sociolinguistics*. Cambridge, Reino Unido: Cambridge University Press.
- Eguren, Luis y Olga Fernández Soriano. 2006. *Terminología gramatical*. Madrid, España: Gredos.



- El DeFe. 2021. *Mapa de colonias de la Alcaldía Miguel Hidalgo*. Recuperado 08 de febrero, 2021, de <http://eldefe.com/mapa-colonias-delegacion-miguel-hidalgo/>
- Ellis, Nick. 2019. Essentials of a theory of language cognition. *The Modern Language Journal*, vol. 103. 39-60.
- Escudero, David, Lourdes Aguilar, César González, Valentín Cardeñoso y Yurena Gutiérrez. 2014. Applying a fuzzy classifier to generate Sp ToBI annotation: preliminar results. En Campbell, Nick, Dafydd Gibbon y Daniel Hirst (eds), *Social and Linguistic Speech Prosody : Proceedings of the 7th international conference on Speech Prosody*. Vol. 7. 457-461.
- Estebas Vilaplana, Eva y Pilar Prieto. 2008. La notación prosódica del español: una revisión del Sp\_ToBI. *Estudios de Fonética Experimental*, vol. XVII. 263 – 283.
- Evans, Vyvyan. 2007. *A glossary of cognitive linguistics*. Edinburgo, Reino Unido: Edinburgh University Press.
- Evans, Vyvyan. 2012. Cognitive linguistics. *WIREs Cognitive Science*, vol. 3. 129-141.
- Face, Timothy. 2006. Cognitive factors in the perception of Spanish stress placement: implications for a model of speech perception. *Linguistics* vol. 44, nº. 6. 1237-1267.
- Face, Timothy. 2007. The role of intonational cues in the perception of declaratives and absolute interrogatives in Castillian Spanish. *Estudios de fonética experimental*, vol. XVI. 185-225.
- García-Lecumberri, María Luisa. 2003. Análisis por configuraciones: la escuela británica. En Pilar Prieto (ed.), *Teorías de la entonación*, 35-61. Barcelona, España: Ariel.
- Geeraerts, Dirk. 1989. Prospects and problems of prototype theory. *Linguistics*, vol. 27, nº. 4. 587-612.
- Geeraerts, Dirk y Hubert Cuyckens. 2007. Introducing cognitive linguistics. En Geeraerts, Dirk y Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 3 - 21. Nueva York, Estados Unidos de América: Oxford University Press.
- Gili Gaya, Samuel. 1975. *Elementos de fonética general* (5ta. ed.). Madrid, España: Gredos.
- Glaser, Barney y Anselm Strauss. 1967. *The discovery of grounded theory: strategies for qualitative research*. Nueva York, Estados Unidos de América: Aldine Publishing Company.

- Goldsmith, John. 1976. *Autosegmental phonology*. Tesis doctoral. Massachusetts Institute of Technology, Estados Unidos de América.
- Grady, Joseph. 2008. Image schemas and perception: refining a definition. En Hampe, Beate (ed.), *From perception to meaning: image schemas in cognitive linguistics*. Berlin, Alemania: De Gruyter Mouton.
- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Nueva York, Estados Unidos de América: Cambridge University Press.
- Harder, Peter. 2012. Variation, structure and norms. *Review of Cognitive Linguistics*, vol. 10, nº.2. 294-314.
- Hernández Campoy, Juan Manuel y Manuel Almeida. 2005. *Metodología de la investigación sociolingüística*. Granada, España: Comares.
- Hirst, Daniel. 2013. UK declarative rises and the frequency code. En Sylvie Hancil y Daniel Hirst (eds.), *Prosody and iconicity*, 149-160. Amsterdam, Países Bajos: John Benjamins publishing company.
- Hirst, Daniel y Albert Di Cristo. 1998. A survey of intonation systems. En Hirst, Daniel y Albert Di Cristo (eds.), *Intonation systems: a survey of twenty languages*, 1-44. Cambridge, Reino Unido: Cambridge University Press.
- Hirst, Daniel, Alberto Di Cristo y Robert Espesser. 2000. Levels of representation and levels of analysis for the description of intonation systems. En Horne, Merle (ed.), *Prosody: theory and experiment*, 51-87. Dordrecht, Países Bajos: Kluwer Academic Press.
- Højrup, Thomas. 1983. The concept of life-mode: A form-specifying mode of analysis applied to contemporary Western Europe. *Ethnologia Scandinavica. A journal for Nordic Ethnology*, vol. 13. 15-50.
- Hualde, José Ignacio. 2002. Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestionis. En Wiltshire, Caroline y Joaquim Camps (eds.), *Romance phonology and variation: selected papers from the 30<sup>th</sup> Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000*, 101-116. Amsterdam, Países Bajos: Benjamins.
- Hualde, José Ignacio. 2003. El modelo métrico autosegmental. En Prieto, Pilar (ed.), *Teorías de la entonación*, 155-184. Barcelona, España: Ariel.

- Hualde, José Ignacio y Pilar Prieto. 2015. Intonational variation in Spanish: European and American varieties. En Frota, Sonia y Pilar Prieto (eds.), *Intonation in Romance*, 350-391. Oxford, Reino Unido: Oxford University Press.
- Hualde, José Ignacio y Pilar Prieto. 2016. Towards an international prosodic alphabet (IPrA). *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 7, nº. 1. 1-25.
- Hudson, Richard Anthony. 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.
- IBM Corp. 2011. *IBM SPSS Statistics for Windows* [Programa informático]. Versión 20.0. Nueva York, Estados Unidos de América: IBM Corp.
- Irvine, Judith y Susan Gal. 2000. Language ideology and linguistic differentiation. En Kroskrity, Paul (ed.), *Regimes of language: ideologies, politics and identities*, 35-84. Santa Fe, Estados Unidos de América: School of American research press.
- Irvine, Judith. 2001. "Style" as distinctiveness: the culture and ideology of linguistic differentiation. En Eckert, Penelope y John Rickford (eds.), *Style and sociolinguistic variation*, 21-43. Cambridge, Reino Unido: Cambridge University Press.
- Johnson, Keith. 1997. Speech perception without speaker normalization: an exemplar model. En Johnson, Keith y John Mullennix (eds.), *Talker variability in speech processing*, 145-165. San Diego, Estados Unidos de América: Academic Press.
- Johnstone, Barbara, Jennifer Andrus y Andrew Danielson. 2006. Mobility, indexicality, and the enregisterment of "Pittsburghese". *Journal of English Linguistics*, vol. 34, nº. 2. 77-104.
- Jorge Trujillo, Carolina, Josefa Dorta y Moisés Betancort. 2017. Estudio de la intensidad en el marco de la prosodia de Canarias. *Onomázein*, nº. 35. 145-172.
- Jun, Sun-Ah. 2005. Prosodic typology. En Jun, Sun-Ah, *Prosodic typology: the phonology of intonation and phrasing*, 430-458. Nueva York, Estados Unidos de América: Oxford University Press
- Kristiansen, Gitte. 2006. Towards a usage-based cognitive phonology. *International journal of English studies*, vol. 6, nº. 2. 107-140
- Kristiansen, Gitte. 2008. Style-shifting and shifting styles: a socio-cognitive approach to lectal variation. En Kristiansen, Gitte y René Dirven (eds.), *Cognitive*

- Sociolinguistics: language variation, cultural models and social systems*, 45-90. Berlin, Alemania: De Gruyter Mouton.
- Kristiansen, Gitte. 2010. Lectal acquisition and linguistic stereotype formation: an empirical study. En Geeraerts, Dirk, Gitte Kristiansen e Yves Peirsman (eds.), *Advances in cognitive sociolinguistics*, 225-263. Berlin, Alemania: De Gruyter Mouton.
- Kristiansen, Gitte y René Dirven. 2008. Introduction. Cognitive sociolinguistics: rationale, methods and scope. En Kristiansen, Gitte y René Dirven (eds.), *Cognitive Sociolinguistics: language variation, cultural models and social systems*, 1-20. Berlin, Alemania: De Gruyter Mouton.
- Kvavik, Karen. 1979. An interpretation of cadences in Mexican Spanish. En Lantolf, James P., Francine Wattman Frank y Jorge M. Guitart (eds.), *Colloquium on Spanish and Luso-Brazilian linguistics*, 37-47. Washington, Estados Unidos de América: Georgetown University Press.
- Labov, William. 1963. The social motivation of a sound change. *Word*, vol. 19, nº. 2. 1-42.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington D.C, Estados Unidos de América.: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, Estados Unidos de América: University of Pennsylvania Press.
- Labov, William. 1981. *Field methods of the project on linguistic change and variation. Sociolinguistic working paper number 81*. Austin, Estados Unidos de América: Southwest Educational Development.
- Ladd, Robert. 2008. *Intonational phonology* (2ª ed.). Cambridge, Reino Unido: Cambridge University press.
- Lakoff, George. 1987. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago, Estados Unidos de América: The University of Chicago Press.
- Lakoff, George y Mark Johnson. 1980. *Metaphors we live by*. Chicago, Estados Unidos de América: The University of Chicago Press.
- Langacker, Ronald. 1986. An introduction to cognitive grammar. *Cognitive science*, vol. 10. 1-40.
- Langacker, Ronald. 1987. *Foundations of cognitive grammar, vol I: theoretical prerequisites*. Stanford, Estados Unidos de América: Stanford University Press.

- Langacker, Ronald. 2008. *Cognitive grammar: a basic introduction*. Nueva York, Estados Unidos de América: Oxford University Press.
- Lange, Kristian, Simone Khün y Elisa Filevich. 2015. “Just another tool for online studies” (JATOS): an easy solution for setup and management of web servers supporting online studies. *PLOS ONE*, vol. 10, n° 7. e0134073.
- Lastra, Yolanda y Pedro Martín Butragueño. 2000. El modo de vida como factor sociolingüístico en la Ciudad de México. En Martín Butragueño, Pedro (ed.), *Estructuras en contexto: Estudios de variación y cambio*, 13-43. Ciudad de México, México: El Colegio de México.
- Lavandera, Beatriz. 1988. The study of language in its socio-cultural context. En Newmeyer, Frederick, *Linguistics: the Cambridge survey* (ed.), 1-13. Cambridge, Reino Unido: Cambridge University Press.
- LeBlanc, David. 2004. *Statistics: concepts and applications for science*. Londres, Reino Unido: Johns and Barlett Publishers.
- Lee, Kaitlyn. 2016. *The perception of creaky voice: does speaker gender affect our judgments?*. Tesis de maestría. University of Kentucky, Estados Unidos de América.
- Llisterri, Joaquim. 2018. *La melodía y la entonación*. Barcelona, España: Departament de Filologia Espanyola, Universitat Autònoma de Barcelona. Recuperado 3 de junio, 2019 de [http://liceu.uab.es/~joaquim/phonetics/fon\\_prosod/suprasegmentales\\_melodia\\_entonacion.html](http://liceu.uab.es/~joaquim/phonetics/fon_prosod/suprasegmentales_melodia_entonacion.html)
- Lipski, John. 2011. Socio-phonological variation in Latin American Spanish. En Díaz-Campos, Manuel (ed.), *The handbook of Spanish sociolinguistics*, 72-97. Massachusetts, Estados Unidos de América: Wiley-Blackwell.
- Macaulay, Ronald. 1977. *Language, social class and education: A Glasgow study*. Edinburgh, Reino Unido: Edinburgh University Press.
- Malmberg, Bertil. 1977. *La fonética*. Buenos Aires, Argentina: Editorial Universitaria de Buenos Aires.
- Martín Butragueño, Pedro. 2004. Configuraciones circunflejas en la entonación del español mexicano. *Revista de filología española*, vol. 84. 347 – 373.

- Martín Butragueño, Pedro. 2006a. El estudio de la entonación en el español de México. En Sedano, Mercedes, Adriana Bolívar y Martha Shiro (comps.), *Haciendo lingüística: homenaje a Paola Bentivoglio*, 105 - 125. Caracas, Venezuela: Universidad Central de Venezuela.
- Martín Butragueño, Pedro. 2006b. Proyección sintáctico-discursiva de la entonación circunfleja mexicana. En Company Company, Concepción, *El español en América. Diatopía, diacronía e historiografía. Homenaje a José G. Moreno de Alba en su 65 aniversario*, 35-63. Ciudad de México, México: UNAM.
- Martín Butragueño, Pedro. 2011. Estratificación sociolingüística de la entonación circunfleja mexicana. En Martín Butragueño, Pedro (ed.), *Realismo en el análisis de corpus orales: Primer coloquio de cambio y variación lingüística*, 93-121. Ciudad de México, México: El Colegio de México.
- Martín Butragueño, Pedro. 2014. Prosodia fonética de enunciados representativos e interrogativos absolutos: elementos locales y globales. *Estudios de Fonética Experimental*, vol. XXIII. 125-202.
- Martín Butragueño, Pedro. 2015a. Hacia una prosodia basada en el uso: actos de habla en el español mexicano. *Revista de filología española*, vol. 5. 97 - 115.
- Martín Butragueño, Pedro. 2015b. Acercamiento a la prosodia de los actos de habla expresivos. Datos del español de México. En Hernández, Esther y Pedro Martín Butragueño (eds.), *Variación y diversidad lingüística: hacia una teoría convergente*, 259 - 345. Ciudad de México, México: El Colegio de México.
- Martín Butragueño, Pedro. 2017. Contacto dialectal entonativo. Estudio Exploratorio. En Orozco, Leonor y Alonso Guerrero Galván (coords.), *Estudios de variación geolingüística*, 139-168. Ciudad de México, México: Secretaría de Cultura – Instituto Nacional de Antropología e Historia.
- Martín Butragueño, Pedro. 2019. *Fonología variable del español de México. Volumen II: Prosodia Enunciativa, tomo 1*. Ciudad de México, México: El Colegio de México.
- Martín Butragueño, Pedro. 2021. Historia de dos medidas: Contacto entonativo en la Ciudad de México. En Serrano, Julio y María Ángeles Soler (eds.), *Contacto de lenguas y variedades. IV coloquio de cambio y variación lingüística*, 403-460. Ciudad de México, México: Universidad Nacional Autónoma de México.

- Martín Butragueño, Pedro y Érika Mendoza. 2018. Prosodic nuclear patterns in narrow and broad focus utterances: Pragmatic and social factors in Central Mexican Spanish. En García García, Marco y Melanie Uth (eds.), *Focus realization in Romance and beyond*, 131-172. Amsterdam, Países bajos: John Benjamins.
- Martín Butragueño, Pedro, Érika Mendoza y Leonor Orozco (coords.). En preparación. *Corpus oral del español de México (COEM)*. Ciudad de México, México: El Colegio de México.
- Martín Butragueño, Pedro y Eva Velásquez Upegui. 2014. Prosodia basada en el uso: Proyecto para el estudio sociolingüístico del español de España y América. En *Reunión del proyecto Preseea*, João Pessoa, Paraíba, Brasil, Julio 14-19.
- Martín Butragueño, Pedro y Yolanda Lastra. 2011. *Corpus sociolingüístico de la Ciudad de México. vol 1. Hablantes de nivel superior*. Ciudad de México, México: El Colegio de México.
- Martínez Celadrán, Eugenio. 2003. Análisis por niveles: la escuela americana. En Pilar Prieto (ed.), *Teorías de la entonación*, 63-95. Barcelona, España: Ariel.
- Mathôt, Sebastian, Daniel Schreij y Jan Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, vol. 44, n°. 2, 314-324.
- Matluck, Joseph. 1952. La pronunciación del español en el valle de México. *Nueva revista de filología hispánica*, vol. 6, n°. 2. 109-120.
- Matluck, Joseph. 1965. *Entonación hispánica*. Austin, Estados Unidos de América: University of Texas.
- Meyerhoff, Miriam. 2006. *Introducing sociolinguistics*. Nueva York, Estados Unidos de América: Routledge.
- Milroy, Lesley y James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, vol. 21, n°. 1. 1-26.
- Mora, Elsa. 1996. Caracterización prosódica de la variación dialectal de l'espagnol parlé au Venezuela. Tesis doctoral. Université Aix-Marseille, Aix en Provence, Francia.
- Mora, Elsa. 1998. Entonación. *Español Actual*, n°. 69. 43-50.
- Mora, Elsa y Rosa Amelia Asuaje. 2009. *El canto de la palabra: una iniciación al estudio de la prosodia*. Mérida, Venezuela: Universidad de Los Andes.

- Mora, Elsa; Daniel Hirst y Albert Di Cristo. 1997. Intonation features as a form of dialectal distinction in Venezuelan Spanish. *ESCA Workshop of intonation: theory, models and applications*. Atenas, Grecia.
- Mora, Elsa; Hernán Martínez y Carmen Luisa Domínguez. 2009. Análisis audio-perceptivo y acústico de la prosodia de las cláusulas en español venezolano. *Opción* vol. 25, n°. 58. 54-69.
- Moreno Fernández, Francisco. 2009. *Principios de sociolingüística y sociología del lenguaje* (4ª ed.). Barcelona, España: Ariel.
- Moreno Fernández, Francisco. 2010. Elementos para una fonología cognitiva de la variación. En Castañer Martín, Rosa y Vicente Lagüéns García (eds), *De moneda nunca usada, estudios filológicos dedicados a José Ma. Enguita Utrilla*, 471 – 490. Zaragoza, España: Institución “Fernando el Católico”.
- Moreno Fernández, Francisco. 2012. *Sociolingüística cognitiva: proposiciones, escolios y debates*. Madrid, España: Iberoamericana.
- Murrieta, Laura. 2016. *Análisis experimental del umbral de percepción entonativa en el español del centro de México*. Tesis de licenciatura. Universidad Nacional Autónoma de México, México.
- Navarro Tomás, Tomás. 1935. *El acento castellano*. Madrid, España: Tipografía de archivos.
- Navarro Tomás, Tomás. 1957. *Manual de pronunciación española*. Nueva York, Estados Unidos de América: Hafner Publishing Co.
- Navarro Tomás, Tomás. 1966. *Manual de entonación española* (3ra ed.). Ciudad de México, México: Colección Málaga.
- Nespor, Marina e Irene Vogel. 2007. *Prosodic phonology: with a new foreword*. Berlin, Alemania: De Gruyter Mouton.
- Nesset, Tore. 2008. *Abstract phonology in a concrete model: cognitive linguistics and the morphology-phonology interface*. Berlin, Alemania: De Gruyter Mouton
- Nuyts, Jan. 2007. Cognitive linguistics and functional linguistics. En Geeraerts, Dirk y Hubert Cuyckens (eds), *The Oxford handbook of cognitive linguistics*, 543 - 565. Nueva York, Estados Unidos de América: Oxford University Press.
- Obediente, Enrique. 2005. *Fonética y fonología*. Mérida, Venezuela: Universidad de Los Andes.



- Orozco, Leonor. 2016. Aproximación a la entonación de enunciados declarativos en Guadalajara. *Revista de Lingüística Aplicada*, año 34, n°. 63. 13 – 35.
- Olivar Espinosa, Stefany. En prensa. Percepción entonativa acerca del español de contacto de San Miguel de Canoa, México. En Congosto Martín, Yolanda y Pedro Martín Butragueño (eds.), *La entonación del español americano septentrional: estudios geoprosódicos (para Peter Lang)*.
- Ortega, Alex. 2019. *Patrones entonativos en el español de Riohacha: un estudio sobre bilingües y monolingües*. Tesis doctoral. Colegio de México, México.
- Pamies Bertrán, Antonio, Ana Fernández Planas, Eugenio Martínez Celdrán, Alicia Ortega Escandell y María Cruz Amorós Céspedes. 2002. Umbrales tonales en español peninsular. En Díaz García, Jesús, *Actas del II congreso de fonética experimental*, 272-278. Sevilla, España: Universidad de Sevilla.
- Pease, Alison, Alan Smaill y Markus Guhe. 2009. Abstract or not abstract? Well, it depends.... *Behavioral and Brain Science*, vol. 32, n°. 3-4. 345-346.
- Pedersen, Ted. 1996. Fishing for exactness. *Proceedings of the South Central SAS user's group*. 188-200.
- Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Tesis doctoral. Massachusetts Institute of Technology, Estados Unidos de América.
- Pinillos, Paloma. 2016. Comparación dialectal del español limeño y madrileño: Contraste acústico y perceptivo de las vocales. *Anuario de Letras. Lingüística y Filología*, vol. IV, n°. 1. 127-157.
- Prieto, Pilar. 2006. Phonological phrasing in Spanish. En Colina, Sonia y Fernando Martínez-Gil (eds.), *Optimality-theoretic advances in Spanish phonology*, 39-61. Amsterdam, Países bajos: John Benjamins.
- Prieto, Pilar y Paolo Roseano. 2010a. *Transcription of the intonation of the Spanish language*. München, Alemania: Lincom.
- Prieto, Pilar y Paolo Roseano. 2010b. Transcription of the intonation of the Spanish language: Introduction. En Pilar Prieto y Paolo Roseano (eds.), *Transcription of the intonation of the Spanish language*, 1-15. München, Alemania: Lincom.
- Quilis, Antonio. 1981. *Fonética acústica de la lengua española*. Madrid, España: Gredos.

- Real Academia Española. 2011. *Nueva gramática de la lengua española. Fonética y fonología*. Barcelona, España: Espasa libros.
- Sagastuy, Paola y Ana Fernández Planas. 2014. La prosodia del español del Centro de México en el marco del proyecto AMPER. *Estudios de Fonética Experimental*, N° XXIII. 47-93.
- Samper Padilla, José Antonio. 2011. Socio-phonological variation and change in Spain. En Díaz-Campos, Manuel (ed.), *The handbook of Spanish sociolinguistics*, 98-120. Massachusetts, Estados Unidos de América: Wiley-Blackwell.
- Sankoff, David y Suzanne Laberge. 1978. The linguistic market and the statistical explanation of variability. En Sankoff, David (ed.), *Linguistic variation: models and methods*, 239-250. Nueva York, Estados Unidos de América: Academic Press.
- Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook*, vol. 3. 371-405.
- Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Washington D.C., Estados Unidos de América: Georgetown University Press.
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, vol. 23, n°. 3-4. 193-229.
- Shapira, Oren, Nira Liberman, Yaacov Trope y SoYon Rim. 2012. Levels of mental construal. En Fiske, Susan y Neil Macrae (eds.), *The SAGE handbook of social cognition*, 229-250. Thousand Oaks, Estados Unidos de América: Sage publications.
- Sosa, Juan Manuel. 1999. *La entonación del español: su estructura fónica, variabilidad y dialectología*. Madrid, España: Cátedra.
- Sosa, Juan Manuel. 2003. La notación tonal del español en el modelo Sp-ToBI. En Pilar Prieto (ed.), *Teorías de la entonación*, 185-208. Barcelona, España: Ariel.
- Stevens, Kenneth. 200. *Acoustic phonetics*. Cambridge, Estados Unidos de América: The MIT Press.
- Strauss, Anselm y Juliet Corbin. 2002. *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*. Antioquia, Colombia: Editorial Universidad de Antioquia.
- Tagliamonte, Sali. 2006. *Analyzing sociolinguistic variation*. Nueva York, Estados Unidos de América: Cambridge University Press.

- Terrell, Tracy. 1978. Sobre la aspiración y elisión de /s/ implosiva y final en el español de Puerto Rico. *Nueva Revista de Filología Hispánica*, vol. 27, n.º. 1. 24-38.
- Toledo, Guillermo. 1998. *El ritmo en el español. Estudio fonético con base computacional*. Madrid, España: Gredos.
- Toledo, Guillermo. 2000. Acentos en español: habla espontánea. *Estudios Filológicos*, n.º. 35. 125-137.
- Toledo, Guillermo. 2007. Fraseo en español peninsular y modelo autosegmental y métrico. *Estudios Filológicos*, n.º. 42. 227-243.
- Toledo, Guillermo. 2008. Frases fonológicas ( $\phi$ ). *Ianua. Revista philological romanica*, n.º. 8. 1-18.
- Torreblanca, Máximo. 1978. Sobre la estructura fonosintáctica de la lengua española. *Nueva Revista de Filología Hispánica*, vol. 27, n.º. 1. 1-23.
- Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, vol. 1, n.º. 2. 179-195.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge, Reino Unido: Cambridge University Press.
- Trudgill, Peter. 2000. *Sociolinguistics: an introduction to language and society* (4ta ed.). Londres, Reino Unido: Penguin Books.
- Tuggy, David. 2007. Schematicity. En Geeraerts, Dirk y Hubert Cuyckens (eds), *The Oxford handbook of cognitive linguistics*, 82 - 116. Nueva York, Estados Unidos de América: Oxford University Press.
- Van der Hulst, Harry. 2003. Cognitive phonology. En Koster, Jan y Henk van Riemsdijk (eds.), *Germania et alia: A linguistic webschrift for Hans den Besten on occasion of his 55<sup>th</sup> birthday*, 1-24. Groningen, Países Bajos: University of Groningen.
- Vida Castro, Matilde. 2004. *Estudio sociofonológico del español hablado en la ciudad de Málaga: Condicionamientos sobre la variación de /-s/ en la distensión silábica*. Alicante, España: Universidad de Alicante.
- Villena Ponsoda, Juan Andrés, Juan Antonio Moya Corral, Antonio Ávila Muñoz y Matilde Vida Castro. 2003. Proyecto de investigación de la formación de dialectos (FORDIAL). *Estudios de Lingüística Universidad de Alicante*, n.º. 17. 607-636.

- Von Heusinger, Klaus. 2007. Discourse structure and intonational phrasing. En Lee, Chungmin, Matthew Gordon y Daniel Büring (eds.), *Topic and focus: cross-linguistic perspectives on meaning and intonation*, 265-290. Dordrecht, Países Bajos: Springer.
- Watkins, Ed, Nicholas Moberly y Michelle Moulds. 2008. Processing mode casually influences emotional reactivity: distinct effects of abstract versus concrete construal on emotional response. *Emotion*, vol. 8, n°. 3. 364-378.
- Wolf, Hans-Georg y Frank Polzenhagen. 2009. *World Englishes: a cognitive sociolinguistic approach*. Berlin, Alemania: De Gruyter Mouton.
- Wolfram Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, D.C., Estados Unidos de América: Center for Applied Linguistics.
- Yi, So Young. 2015. *Sociophonetic variations in Korean constituent final –KO and –TO*. Tesis doctoral. University of Hawai‘i at Mānoa, Estados Unidos de América.
- Zeschel, Arne y Nadine Proske. 2015. Usage-based linguistics and conversational interaction: A case study of German motion verbs. *Yearbook of the German Cognitive Linguistics Association*, vol. 3. 123-144.
- Zhang, Qing. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, vol. 34. 431-466.

## Apéndices

Apéndice 1. Repertorio de temas y guion de preguntas que sirvieron de guía (no condicional) en la fase 1 de la entrevista (habla espontánea).

Para la elaboración del guion de preguntas, se siguió la siguiente línea temática: aspectos culturales de México > infancia > familia > trabajo.

El guion de preguntas de la primera fase de las entrevistas se presenta a continuación:

- ¿Qué podrías decir de la ciudad de México?
- ¿Qué recomendaciones le darías a alguien que es nuevo en la ciudad?
- ¿Qué tradiciones especiales tienen en Ciudad de México?
- ¿Cómo se celebran algunas fiestas típicas de México?
- ¿Celebraba estas tradiciones cuando era niño(a)? / ¿Cómo las celebraba?
- ¿Cuáles eran las tradiciones más importantes en su familia cuando era niño(a)?
- ¿Diría que fue muy travieso(a) cuando era niño (a)? / ¿Por qué?
- ¿Tiene algún recuerdo de su niñez que sea especial para usted? / ¿Cuál es?
- ¿Cómo fue para usted el cambio de la infancia a la adolescencia? / ¿Por qué?
- ¿Qué actividades solía hacer con sus amigos cuando era adolescente?
- ¿Se metió alguna vez en algún problema cuando era adolescente? / ¿Cómo ocurrió?
- ¿A qué edad empezó a trabajar?
- ¿Cómo fueron sus inicios en el trabajo?
- ¿Cómo ha cambiado su oficio desde que empezó hasta ahora?
- ¿Considera que su trabajo actual es fácil? / ¿Por qué?

Apéndice 2. Cuestionario sociolingüístico aplicado en la entrevista (adaptación de Villena Ponsoda *et al.*, 2003; Vida Castro, 2004; Martín Butragueño, Mendoza y Orozco, en preparación).

Confidencial

❖ Datos generales del informante:

- 1 Nombre: \_\_\_\_\_ / Código: \_\_\_\_\_
- 2 Sexo: \_\_\_\_\_ / Edad: \_\_\_\_\_ / Últimos estudios: \_\_\_\_\_
- 3 Lugar de nacimiento: (Est.Col./Del.): \_\_\_\_\_
- 4 Residencia (Col./Del.): \_\_\_\_\_
- 5 Carrera/Oficio: \_\_\_\_\_
- 6 Localidad en la que trabaja (Col./Del.): \_\_\_\_\_
- 7 Ingresos mensuales: (  ) 0 \$ – 2.700\$ / (  ) 2.700\$ - 5.400\$ / (  ) 5.400\$ - 10.000\$ / (  ) 10.000\$ - 14.000\$ / (  ) 14.000\$ - 18.000\$ / (  ) 18.000\$ - 22.000\$ / (  ) 22.000\$ - 26.000\$ / (  ) 26.000\$ - 30.000\$ / (  ) 30.000\$ en adelante.
- 8 ¿A qué clase social cree usted que pertenece?  
(  ) Baja / (  ) Media-baja / (  ) Media / (  ) Media-alta / (  ) Alta.

❖ Datos generales de padres del informante:

- 9 Origen de la madre (estado): \_\_\_\_\_
- 10 Edad de llegada de la madre a Ciudad de México: \_\_\_\_\_
- 11 Ocupación de la madre: \_\_\_\_\_
- 12 Últimos estudios de la madre: \_\_\_\_\_
- 13 ¿A qué clase socioeconómica cree usted que pertenece su madre?  
(  ) Baja / (  ) Media-baja / (  ) Media / (  ) Media-alta / (  ) Alta.
- 14 Origen del padre (estado): \_\_\_\_\_
- 15 Edad de llegada del padre a Ciudad de México: \_\_\_\_\_
- 16 Ocupación del padre: \_\_\_\_\_
- 17 Últimos estudios del padre: \_\_\_\_\_
- 18 ¿A qué clase socioeconómica cree usted que pertenece su padre?  
(  ) Baja / (  ) Media-baja / (  ) Media / (  ) Media-alta / (  ) Alta.

❖ Ubicación de contactos

- 19 ¿Cuántas amistades tiene que sean de la misma zona en la que vive?  
(  ) Todas / (  ) La mayoría / (  ) Algunas / (  ) Pocas / (  ) Ninguna

- 20 ¿Cuántas amistades tiene que sean de la misma zona en la que trabaja?  
() Todas / () La mayoría / () Algunas / () Pocas / () Ninguna
- 21 ¿Cuántas amistades tiene que sean de zonas diferentes a las que vive?  
() Todas / () La mayoría / () Algunas / () Pocas / () Ninguna
- 22 ¿Cuántas amistades tiene que sean de zonas diferentes a las que trabaja?  
() Todas / () La mayoría / () Algunas / () Pocas / () Ninguna
- 23 ¿A qué personas diría que frecuenta más? () Personas de la colonia donde vivo / () Personas de la colonia donde trabajo. ¿Por qué?

---

- 24 ¿En qué lugares frecuenta a estas personas?

---

❖ Contacto con otras comunidades

- 25 ¿Ha vivido en otra ciudad? () Sí / () No. ¿Durante cuánto tiempo?

---

- 26 Si vivió en otra ciudad, ¿por qué motivo fue?

---

- 27 ¿Ha realizado viajes dentro de México? () Sí / () No. ¿Qué lugares ha visitado?

---

---

- 28 ¿Ha realizado viajes al extranjero? () Sí / () No. ¿Qué lugares ha visitado?

---

---

❖ Actitud e integración en la cultura local

- 47 ¿Participa (o celebra) fiestas tradicionales de su comunidad?

() Sí / () No / () En ocasiones, pero no con regularidad.

- 48 ¿Qué opinión tiene acerca de la cultura de su localidad?

() Soy parte de ella / () Me agrada / () Me es indiferente / () Hay varios aspectos que no me agradan /

() No me siento parte de ella

❖ Actitud e integración a la red

- 31 Asigne un número del 1 al 4 de acuerdo al grado de importancia que tienen para usted los siguientes vínculos, teniendo en cuenta que 1 representa la importancia más baja y 4 la más alta.  
() Familia / () Amistades fuera del trabajo / () Amistades dentro del trabajo / ()  
Compañeros de trabajo

❖ Densidad de la red

- 32 ¿Las personas a las que trata con frecuencia se conocen entre sí? () Sí / () No.  
33 ¿Las personas a las que trata con frecuencia se tratan entre sí? () Sí / () No.  
34 ¿Se preocupa usted por lo que piensan esas otras personas antes de tomar decisiones importantes?  
() Sí / () No.  
35 ¿Les consulta antes de tomar una decisión? () Sí / () No.  
36 ¿Podría reunir usted a todas esas personas en caso de necesidad? () Sí / () No.  
37 ¿Se reúne usted con todos ellos alguna vez? () Sí / () No.  
38 ¿Se reúne usted con todos ellos periódicamente? () Sí / () No.

❖ Fuerza de los vínculos

- 39 ¿A cuáles de las siguientes personas les pediría usted un favor muy importante?  
A familiares inmediatos (cónyuge, padres, hijos, etc.) () Sí / () No.  
A amistades () Sí / () No.  
A compañeros de trabajo () Sí / () No.  
A conocidos () Sí / () No.  
A todas las personas a las que conoce () Sí / () No.

❖ Uso de la red

- 40 ¿Cómo se enteró acerca del trabajo que tiene actualmente? (Puede seleccionar más de una opción de ser necesario).  
() Información oficial o pública (periódicos, redes sociales, páginas de empleo) / () Por medio de la empresa directamente / () A través de familiares, conocidos, amigos, etc. / ()  
Otros.  
41 ¿Qué contacto utilizó para encontrar este trabajo? (Puede seleccionar más de una opción de ser necesario).



Por medio de familiares /  Por medio de vecinos o compañeros de trabajo /  Por medio de un amigo /  Por medio del amigo de un amigo /  Por medio de un conocido.

- 42 ¿Cómo consigue contactos para solucionar problemas con distintas entidades o instituciones (por ejemplo, hospitales, seguridad social, escuelas, bancos, oficinas o departamentos gubernamentales, etc.) (Puede seleccionar más de una opción de ser necesario).

Nunca busco conocidos para esos asuntos /  Por medio de familiares /  Por medio de amigos /  Por medio de vecinos /  Por medio del amigo de un amigo /  Por medio de un conocido /  Otros.

❖ Contacto con medios de comunicación

- 43 ¿Cuántas horas a la semana, aproximadamente, ve la televisión?

Más de 30 horas /  Entre 20 y 30 horas /  Entre 10 y 20 horas /  Menos de 10 horas /

No ve televisión

- 44 ¿Cuántas horas a la semana, aproximadamente, escucha la radio?

Más de 30 horas /  Entre 20 y 30 horas /  Entre 10 y 20 horas /  Menos de 10 horas /

No escucha la radio

- 45 ¿Cuántas horas a la semana, aproximadamente, ve videos (películas, series, documentales, noticias, etc.) en internet?

Más de 30 horas /  Entre 20 y 30 horas /  Entre 10 y 20 horas /  Menos de 10 horas /

No uso internet

❖ Contacto con la norma.

46 En su vida diaria, profesional y laboral, ¿qué importancia da usted al uso "correcto" o "bueno" de la lengua? ¿Significa algo para usted y cree que influye en alguna medida en el correcto desarrollo y buenos resultados en su trabajo o negocio y, en general, en su vida social? Indique la opción que prefiera y que más se adapte a sus circunstancias.

() Nada. Creo que la lengua vale si los demás te entienden. Da igual cómo se use. Yo la utilizo y nada más. Me da igual lo que los demás piensen de cómo hablo; hay cosas más importantes.

() Poco. No presto una especial atención a esas cosas. Son secundarias para mi trabajo o profesión. Utilizo la lengua de modo corriente, sin caer en vulgarismos excesivos, pero no valoro personalmente el uso correcto de la lengua.

() Mucho. Trato de utilizar correctamente la lengua; para mi trabajo es conveniente y seguramente es un factor que contribuye a que todo salga mejor.

() Bastante. Es lo más importante o de los más importante; si yo no empleara correctamente la lengua perdería oportunidades o descendería en posición en mi profesión y en mi relación con los demás; incluso no podría desempeñarla adecuadamente. Presto una gran atención al uso correcto de la lengua.

❖ Creencias lingüísticas

47 ¿Considera usted que habla de forma diferente en algunas situaciones o con algunas personas en la Ciudad de México? (por ejemplo, cuando visita al doctor) () Sí / () No. ¿En qué situaciones o con qué personas habla de forma diferente? \_\_\_\_\_  
\_\_\_\_\_

48 ¿Cuántas formas de hablar considera usted que hay en Ciudad de México?

() Muchas / () Algunas / () Pocas / () Solo una

49 ¿Cuáles son esas formas?

\_\_\_\_\_  
\_\_\_\_\_

50 ¿Con cuál de esas formas de hablar se identifica usted?

\_\_\_\_\_

51 ¿En qué cree que se basa esa diferencia en las distintas formas de hablar? ¿Qué elementos del lenguaje le indican que una persona de Ciudad de México habla diferente?

La edad /  La delegación o colonia /  El nivel de ingresos /  La profesión u oficio /  El nivel de estudios /  El género (hombre/mujeres) Explique:

---

---

---

52 En su opinión, ¿quiénes hablan mejor en la Ciudad de México?

jóvenes /  adultos /  adultos mayores ¿Por qué?

---

---

hombres /  mujeres ¿Por qué?

---

---

personas sin estudios /  personas con primaria completa /  personas con bachillerato completo /  personas con carreras universitarias /  personas que trabajan en medios de información ¿Por qué?

---

---

personas con bajos ingresos /  personas con ingresos medios /  personas con altos ingresos ¿Por qué?

---

---

personas de la zona en la que vivo o trabajo /  personas de otras zonas de la ciudad ¿Por qué?

---

---

❖ Escala de intensidad reticular general

- Todos

53 ¿Viven por lo regular sus parientes en el vecindario (al menos dos núcleos familiares además del propio)?

Sí /  No.

54 ¿Algunos de sus amigos más íntimos (más de dos) son parientes suyos?  Sí /  No.

55 ¿Sale con frecuencia con sus familiares en momentos de ocio (con al menos dos)? ( ) Sí / ( ) No.

56 ¿Se ve o sale con vecinos en momentos de ocio? ( ) Sí / ( ) No.

- Solo trabajadores

57 ¿Viven en su misma colonia al menos dos compañeros de trabajo? ( ) Sí / ( ) No.

58 ¿Viven en su misma colonia al menos dos compañeros de trabajo del mismo sexo? ( ) Sí / ( ) No.

59 ¿Se reúne con sus compañeros de trabajo en momentos de ocio? ( ) Sí / ( ) No.

60 ¿Algunos de sus vecinos (más de dos) trabaja en el mismo sitio que usted? ( ) Sí / ( ) No.

61 ¿Trabaja en el mismo lugar con algunos de sus parientes? ( ) Sí / ( ) No.

- Para ser respondidas por el encuestador

62 ¿Está el sujeto integrado a una piña? ( ) Sí / ( ) No.

63 ¿El individuo es miembro de un grupo de alta densidad que forme parte de la red? ( ) Sí / ( ) No.

64 ¿El individuo mantiene vínculos con dos tercios de la red? ( ) Sí / ( ) No.

65 ¿El individuo mantiene relación de parentesco con al menos tres miembros de la red? ( ) Sí / ( ) No.

66 ¿El individuo vive en el mismo barrio o zona que el individuo anclaje de la red? ( ) Sí / ( ) No.

Apéndice 3. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de duración de las distintas sílabas pertinentes en la totalidad del corpus.

Comparaciones múltiples						
(I)	(J)	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
Sílabas del enunciado	Sílabas del enunciado				Límite inferior	Límite superior
Cuerpo	Primera sílaba	4.77	4.45	1.00	-9.79	19.32
	Primera sílaba y primera tónica	-5.75	5.48	1.00	-23.67	12.17
	Primera tónica	-1.66	4.62	1.00	-16.76	13.44
	Núcleo de IP	-38.637*	3.50	0.00	-50.09	-27.18
	Linde de IP	-35.187*	3.50	0.00	-46.64	-23.73
	Linde de ip	-101.595*	4.69	0.00	-116.91	-86.27
	Primera tónica, núcleo y linde de ip	-153.928*	22.58	0.00	-227.74	-80.12
	Primera sílaba, primera tónica y núcleo de ip	26.95	45.12	1.00	-120.57	174.47
	Núcleo de ip	-48.070*	4.83	0.00	-63.86	-32.27
	Núcleo y linde de ip	-131.659*	7.91	0.00	-157.54	-105.78
	Primera tónica y núcleo de ip	-44.15	20.20	0.56	-110.19	21.88
Primera sílaba	Cuerpo	-4.77	4.45	1.00	-19.32	9.79
	Primera sílaba y primera tónica	-10.52	6.92	0.94	-33.15	12.11
	Primera tónica	-6.42	6.26	1.00	-26.89	14.04
	Núcleo de IP	-43.403*	5.49	0.00	-61.35	-25.45
	Linde de IP	-39.953*	5.49	0.00	-57.90	-22.00
	Linde de ip	-106.361*	6.31	0.00	-126.99	-85.73
	Primera tónica, núcleo y linde de ip	-158.694*	22.97	0.00	-233.79	-83.60
	Primera sílaba, primera tónica y núcleo de ip	22.18	45.32	1.00	-125.98	170.35
	Núcleo de ip	-52.836*	6.42	0.00	-73.82	-31.85
	Núcleo y linde de ip	-136.426*	8.97	0.00	-165.76	-107.09
	Primera tónica y núcleo de ip	-48.92	20.64	0.43	-116.39	18.55
Primera sílaba y primera tónica	Cuerpo	5.75	5.48	1.00	-12.17	23.67
	Primera sílaba	10.52	6.92	0.94	-12.11	33.15
	Primera tónica	4.10	7.03	1.00	-18.89	27.08

	Núcleo de IP	-32.883*	6.35	0.00	-53.66	-12.11
	Linde de IP	-29.433*	6.35	0.00	-50.21	-8.66
	Linde de ip	-95.841*	7.07	0.00	-118.97	-72.71
	Primera tónica, núcleo y linde de ip	-148.175*	23.19	0.00	-223.99	-72.36
	Primera sílaba, primera tónica y núcleo de ip	32.70	45.43	1.00	-115.83	181.23
	Núcleo de ip	-42.316*	7.17	0.00	-65.76	-18.87
	Núcleo y linde de ip	-125.906*	9.53	0.00	-157.05	-94.76
	Primera tónica y núcleo de ip	-38.40	20.88	0.80	-106.67	29.87
Primera tónica	Cuerpo	1.66	4.62	1.00	-13.44	16.76
	Primera sílaba	6.42	6.26	1.00	-14.04	26.89
	Primera sílaba y primera tónica	-4.10	7.03	1.00	-27.08	18.89
	Núcleo de IP	-36.978*	5.63	0.00	-55.37	-18.58
	Linde de IP	-33.528*	5.63	0.00	-51.92	-15.13
	Linde de ip	-99.936*	6.43	0.00	-120.95	-78.92
	Primera tónica, núcleo y linde de ip	-152.270*	23.00	0.00	-227.47	-77.07
	Primera sílaba, primera tónica y núcleo de ip	28.61	45.34	1.00	-119.61	176.82
	Núcleo de ip	-46.411*	6.54	0.00	-67.78	-25.04
	Núcleo y linde de ip	-130.001*	9.06	0.00	-159.61	-100.39
	Primera tónica y núcleo de ip	-42.50	20.67	0.65	-110.08	25.09
Núcleo de IP	Cuerpo	38.637*	3.50	0.00	27.18	50.09
	Primera sílaba	43.403*	5.49	0.00	25.45	61.35
	Primera sílaba y primera tónica	32.883*	6.35	0.00	12.11	53.66
	Primera tónica	36.978*	5.63	0.00	18.58	55.37
	Linde de IP	3.45	4.76	1.00	-12.10	19.00
	Linde de ip	-62.958*	5.68	0.00	-81.53	-44.38
	Primera tónica, núcleo y linde de ip	-115.292*	22.80	0.00	-189.85	-40.74
	Primera sílaba, primera tónica y núcleo de ip	65.58	45.24	0.95	-82.31	213.48
	Núcleo de ip	-9.43	5.80	0.90	-28.40	9.54
	Núcleo y linde de ip	-93.023*	8.54	0.00	-120.95	-65.09
	Primera tónica y núcleo de ip	-5.52	20.45	1.00	-72.38	61.35

Linde de IP	Cuerpo	35.187*	3.50	0.00	23.73	46.64
	Primera sílaba	39.953*	5.49	0.00	22.00	57.90
	Primera sílaba y primera tónica	29.433*	6.35	0.00	8.66	50.21
	Primera tónica	33.528*	5.63	0.00	15.13	51.92
	Núcleo de IP	-3.45	4.76	1.00	-19.00	12.10
	Linde de ip	-66.408*	5.68	0.00	-84.98	-47.83
	Primera tónica, núcleo y linde de ip	-118.742*	22.80	0.00	-193.30	-44.19
	Primera sílaba, primera tónica y núcleo de ip	62.13	45.24	0.97	-85.76	210.03
	Núcleo de ip	-12.88	5.80	0.53	-31.85	6.09
	Núcleo y linde de ip	-96.473*	8.54	0.00	-124.40	-68.54
	Primera tónica y núcleo de ip	-8.97	20.45	1.00	-75.83	57.90
Linde de ip	Cuerpo	101.595*	4.69	0.00	86.27	116.91
	Primera sílaba	106.361*	6.31	0.00	85.73	126.99
	Primera sílaba y primera tónica	95.841*	7.07	0.00	72.71	118.97
	Primera tónica	99.936*	6.43	0.00	78.92	120.95
	Núcleo de IP	62.958*	5.68	0.00	44.38	81.53
	Linde de IP	66.408*	5.68	0.00	47.83	84.98
	Primera tónica, núcleo y linde de ip	-52.33	23.02	0.50	-127.58	22.91
	Primera sílaba, primera tónica y núcleo de ip	128.54	45.34	0.17	-19.70	276.78
	Núcleo de ip	53.525*	6.58	0.00	32.00	75.05
	Núcleo y linde de ip	-30.065*	9.09	0.04	-59.79	-0.34
	Primera tónica y núcleo de ip	57.44	20.69	0.19	-10.19	125.08
Primera tónica, núcleo y linde de ip	Cuerpo	153.928*	22.58	0.00	80.12	227.74
	Primera sílaba	158.694*	22.97	0.00	83.60	233.79
	Primera sílaba y primera tónica	148.175*	23.19	0.00	72.36	223.99
	Primera tónica	152.270*	23.00	0.00	77.07	227.47
	Núcleo de IP	115.292*	22.80	0.00	40.74	189.85
	Linde de IP	118.742*	22.80	0.00	44.19	193.30
	Linde de ip	52.33	23.02	0.50	-22.91	127.58
	Primera sílaba, primera tónica y núcleo de ip	180.875*	50.44	0.02	15.98	345.77
	Núcleo de ip	105.859*	23.05	0.00	30.51	181.20

	Núcleo y linde de ip	22.27	23.88	1.00	-55.81	100.35
	Primera tónica y núcleo de ip	109.775*	30.26	0.02	10.84	208.71
Primera sílaba, primera tónica y núcleo de ip	Cuerpo	-26.95	45.12	1.00	-174.47	120.57
	Primera sílaba	-22.18	45.32	1.00	-170.35	125.98
	Primera sílaba y primera tónica	-32.70	45.43	1.00	-181.23	115.83
	Primera tónica	-28.61	45.34	1.00	-176.82	119.61
	Núcleo de IP	-65.58	45.24	0.95	-213.48	82.31
	Linde de IP	-62.13	45.24	0.97	-210.03	85.76
	Linde de ip	-128.54	45.34	0.17	-276.78	19.70
	Primera tónica, núcleo y linde de ip	-180.875*	50.44	0.02	-345.77	-15.98
	Núcleo de ip	-75.02	45.36	0.89	-223.31	73.27
	Núcleo y linde de ip	-158.606*	45.79	0.03	-308.31	-8.90
	Primera tónica y núcleo de ip	-71.10	49.42	0.96	-232.66	90.46
Núcleo de ip	Cuerpo	48.070*	4.83	0.00	32.27	63.86
	Primera sílaba	52.836*	6.42	0.00	31.85	73.82
	Primera sílaba y primera tónica	42.316*	7.17	0.00	18.87	65.76
	Primera tónica	46.411*	6.54	0.00	25.04	67.78
	Núcleo de IP	9.43	5.80	0.90	-9.54	28.40
	Linde de IP	12.88	5.80	0.53	-6.09	31.85
	Linde de ip	-53.525*	6.58	0.00	-75.05	-32.00
	Primera tónica, núcleo y linde de ip	-105.859*	23.05	0.00	-181.20	-30.51
	Primera sílaba, primera tónica y núcleo de ip	75.02	45.36	0.89	-73.27	223.31
	Núcleo y linde de ip	-83.590*	9.17	0.00	-113.56	-53.62
	Primera tónica y núcleo de ip	3.92	20.72	1.00	-63.83	71.66
Núcleo y linde de ip	Cuerpo	131.659*	7.91	0.00	105.78	157.54
	Primera sílaba	136.426*	8.97	0.00	107.09	165.76
	Primera sílaba y primera tónica	125.906*	9.53	0.00	94.76	157.05
	Primera tónica	130.001*	9.06	0.00	100.39	159.61
	Núcleo de IP	93.023*	8.54	0.00	65.09	120.95
	Linde de IP	96.473*	8.54	0.00	68.54	124.40
	Linde de ip	30.065*	9.09	0.04	0.34	59.79
	Primera tónica, núcleo y linde de ip	-22.27	23.88	1.00	-100.35	55.81



	Primera sílaba, primera tónica y núcleo de ip	158.606*	45.79	0.03	8.90	308.31
	Núcleo de ip	83.590*	9.17	0.00	53.62	113.56
	Primera tónica y núcleo de ip	87.506*	21.65	0.00	16.73	158.28
Primera tónica y núcleo de ip	Cuerpo	44.15	20.20	0.56	-21.88	110.19
	Primera sílaba	48.92	20.64	0.43	-18.55	116.39
	Primera sílaba y primera tónica	38.40	20.88	0.80	-29.87	106.67
	Primera tónica	42.50	20.67	0.65	-25.09	110.08
	Núcleo de IP	5.52	20.45	1.00	-61.35	72.38
	Linde de IP	8.97	20.45	1.00	-57.90	75.83
	Linde de ip	-57.44	20.69	0.19	-125.08	10.19
	Primera tónica, núcleo y linde de ip	-109.775*	30.26	0.02	-208.71	-10.84
	Primera sílaba, primera tónica y núcleo de ip	71.10	49.42	0.96	-90.46	232.66
	Núcleo de ip	-3.92	20.72	1.00	-71.66	63.83
	Núcleo y linde de ip	-87.506*	21.65	0.00	-158.28	-16.73
*. La diferencia de medias es significativa al nivel 0.05.						

Apéndice 4. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de intensidad de las distintas sílabas pertinentes en la totalidad del corpus.

Comparaciones múltiples						
(I)	(J)	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
Sílabas del enunciado	Sílabas del enunciado				Límite inferior	Límite superior
Cuerpo	Primera sílaba	-.75493	.73897	.997	-3.1709	1.6610
	Primera sílaba y primera tónica	-3.02016*	.90989	.043	-5.9949	-.0454
	Primera tónica	-2.40970	.76656	.073	-4.9158	.0964
	Núcleo de IP	1.36781	.58174	.440	-.5341	3.2697
	Linde de IP	10.46036*	.58174	.000	8.5585	12.3623
	Linde de ip	.41454	.77778	1.000	-2.1283	2.9574
	Primera tónica, núcleo y linde de ip	-5.25030	3.74748	.964	-17.5020	7.0014
	Primera sílaba, primera tónica y núcleo de ip	-5.52530	7.48956	1.000	-30.0111	18.9605
	Núcleo de ip	-.33063	.80191	1.000	-2.9523	2.2911
	Núcleo y linde de ip	2.30030	1.31375	.844	-1.9948	6.5954
Primera tónica y núcleo de ip	-3.87730	3.35265	.992	-14.8382	7.0836	
Primera sílaba	Cuerpo	.75493	.73897	.997	-1.6610	3.1709
	Primera sílaba y primera tónica	-2.26523	1.14895	.713	-6.0215	1.4911
	Primera tónica	-1.65477	1.03913	.912	-5.0520	1.7425
	Núcleo de IP	2.12274	.91138	.456	-.8569	5.1023
	Linde de IP	11.21530*	.91138	.000	8.2357	14.1949
	Linde de ip	1.16947	1.04744	.994	-2.2550	4.5939
	Primera tónica, núcleo y linde de ip	-4.49537	3.81258	.991	-16.9600	7.9692

	Primera sílaba, primera tónica y núcleo de ip	-4.77037	7.52234	1.000	-29.3634	19.8227
	Núcleo de ip	.42430	1.06548	1.000	-3.0591	3.9077
	Núcleo y linde de ip	3.05524	1.48934	.658	-1.8139	7.9244
	Primera tónica y núcleo de ip	-3.12237	3.42526	.999	-14.3207	8.0760
Primera sílaba y primera tónica	Cuerpo	3.02016*	.90989	.043	.0454	5.9949
	Primera sílaba	2.26523	1.14895	.713	-1.4911	6.0215
	Primera tónica	.61046	1.16688	1.000	-3.2045	4.4254
	Núcleo de IP	4.38797*	1.05472	.002	.9397	7.8362
	Linde de IP	13.48052*	1.05472	.000	10.0323	16.9287
	Linde de ip	3.43470	1.17429	.132	-.4044	7.2738
	Primera tónica, núcleo y linde de ip	-2.23014	3.84936	1.000	-14.8150	10.3547
	Primera sílaba, primera tónica y núcleo de ip	-2.50514	7.54105	1.000	-27.1593	22.1491
	Núcleo de ip	2.68953	1.19041	.506	-1.2023	6.5814
	Núcleo y linde de ip	5.32046*	1.58112	.037	.1513	10.4897
Primera tónica y núcleo de ip	-.85714	3.46616	1.000	-12.1892	10.4749	
Primera tónica	Cuerpo	2.40970	.76656	.073	-.0964	4.9158
	Primera sílaba	1.65477	1.03913	.912	-1.7425	5.0520
	Primera sílaba y primera tónica	-.61046	1.16688	1.000	-4.4254	3.2045
	Núcleo de IP	3.77751*	.93389	.003	.7243	6.8307
	Linde de IP	12.87007*	.93389	.000	9.8169	15.9233
	Linde de ip	2.82425	1.06709	.254	-.6644	6.3129
	Primera tónica, núcleo y linde de ip	-2.84060	3.81802	1.000	-15.3230	9.6418

	Primera sílaba, primera tónica y núcleo de ip	-3.11560	7.52510	1.000	-27.7177	21.4865
	Núcleo de ip	2.07907	1.08480	.749	-1.4675	5.6256
	Núcleo y linde de ip	4.71001	1.50321	.075	-.2045	9.6245
	Primera tónica y núcleo de ip	-1.46760	3.43132	1.000	-12.6857	9.7505
Núcleo de IP	Cuerpo	-1.36781	.58174	.440	-3.2697	.5341
	Primera sílaba	-2.12274	.91138	.456	-5.1023	.8569
	Primera sílaba y primera tónica	-4.38797*	1.05472	.002	-7.8362	-.9397
	Primera tónica	-3.77751*	.93389	.003	-6.8307	-.7243
	Linde de IP	9.09256*	.78928	.000	6.5121	11.6730
	Linde de ip	-.95327	.94313	.997	-4.0367	2.1301
	Primera tónica, núcleo y linde de ip	-6.61811	3.78525	.845	-18.9933	5.7571
	Primera sílaba, primera tónica y núcleo de ip	-6.89311	7.50853	.999	-31.4410	17.6548
	Núcleo de ip	-1.69844	.96312	.838	-4.8472	1.4503
	Núcleo y linde de ip	.93249	1.41791	1.000	-3.7031	5.5681
	Primera tónica y núcleo de ip	-5.24511	3.39482	.928	-16.3439	5.8537
Linde de IP	Cuerpo	-10.46036*	.58174	.000	-12.3623	-8.5585
	Primera sílaba	-11.21530*	.91138	.000	-14.1949	-8.2357
	Primera sílaba y primera tónica	-13.48052*	1.05472	.000	-16.9287	-10.0323
	Primera tónica	-12.87007*	.93389	.000	-15.9233	-9.8169
	Núcleo de IP	-9.09256*	.78928	.000	-11.6730	-6.5121
	Linde de ip	-10.04582*	.94313	.000	-13.1292	-6.9624

	Primera tónica, núcleo y linde de ip	-15.71067*	3.78525	.002	-28.0859	-3.3354
	Primera sílaba, primera tónica y núcleo de ip	-15.98567	7.50853	.602	-40.5335	8.5622
	Núcleo de ip	-10.79100*	.96312	.000	-13.9397	-7.6422
	Núcleo y linde de ip	-8.16006*	1.41791	.000	-12.7957	-3.5244
	Primera tónica y núcleo de ip	-14.33767*	3.39482	.001	-25.4364	-3.2389
Linde de ip	Cuerpo	-.41454	.77778	1.000	-2.9574	2.1283
	Primera sílaba	-1.16947	1.04744	.994	-4.5939	2.2550
	Primera sílaba y primera tónica	-3.43470	1.17429	.132	-7.2738	.4044
	Primera tónica	-2.82425	1.06709	.254	-6.3129	.6644
	Núcleo de IP	.95327	.94313	.997	-2.1301	4.0367
	Linde de IP	10.04582*	.94313	.000	6.9624	13.1292
	Primera tónica, núcleo y linde de ip	-5.66485	3.82029	.945	-18.1546	6.8250
	Primera sílaba, primera tónica y núcleo de ip	-5.93985	7.52626	1.000	-30.5457	18.6660
	Núcleo de ip	-.74518	1.09276	1.000	-4.3178	2.8274
	Núcleo y linde de ip	1.88576	1.50897	.985	-3.0476	6.8191
	Primera tónica y núcleo de ip	-4.29185	3.43385	.985	-15.5182	6.9345
Primera tónica, núcleo y linde de ip	Cuerpo	5.25030	3.74748	.964	-7.0014	17.5020
	Primera sílaba	4.49537	3.81258	.991	-7.9692	16.9600
	Primera sílaba y primera tónica	2.23014	3.84936	1.000	-10.3547	14.8150
	Primera tónica	2.84060	3.81802	1.000	-9.6418	15.3230
	Núcleo de IP	6.61811	3.78525	.845	-5.7571	18.9933

	Linde de IP	15.71067*	3.78525	.002	3.3354	28.0859
	Linde de ip	5.66485	3.82029	.945	-6.8250	18.1546
	Primera sílaba, primera tónica y núcleo de ip	-.27500	8.37157	1.000	-27.6444	27.0944
	Núcleo de ip	4.91967	3.82528	.981	-7.5864	17.4258
	Núcleo y linde de ip	7.55061	3.96429	.757	-5.4100	20.5112
	Primera tónica y núcleo de ip	1.37300	5.02294	1.000	-15.0487	17.7947
Primera sílaba, primera tónica y núcleo de ip	Cuerpo	5.52530	7.48956	1.000	-18.9605	30.0111
	Primera sílaba	4.77037	7.52234	1.000	-19.8227	29.3634
	Primera sílaba y primera tónica	2.50514	7.54105	1.000	-22.1491	27.1593
	Primera tónica	3.11560	7.52510	1.000	-21.4865	27.7177
	Núcleo de IP	6.89311	7.50853	.999	-17.6548	31.4410
	Linde de IP	15.98567	7.50853	.602	-8.5622	40.5335
	Linde de ip	5.93985	7.52626	1.000	-18.6660	30.5457
	Primera tónica, núcleo y linde de ip	.27500	8.37157	1.000	-27.0944	27.6444
	Núcleo de ip	5.19467	7.52879	1.000	-19.4194	29.8088
	Núcleo y linde de ip	7.82561	7.60036	.997	-17.0225	32.6737
	Primera tónica y núcleo de ip	1.64800	8.20243	1.000	-25.1685	28.4645
Núcleo de ip	Cuerpo	.33063	.80191	1.000	-2.2911	2.9523
	Primera sílaba	-.42430	1.06548	1.000	-3.9077	3.0591
	Primera sílaba y primera tónica	-2.68953	1.19041	.506	-6.5814	1.2023
	Primera tónica	-2.07907	1.08480	.749	-5.6256	1.4675
	Núcleo de IP	1.69844	.96312	.838	-1.4503	4.8472
	Linde de IP	10.79100*	.96312	.000	7.6422	13.9397
	Linde de ip	.74518	1.09276	1.000	-2.8274	4.3178

	Primera tónica, núcleo y linde de ip	-4.91967	3.82528	.981	-17.4258	7.5864
	Primera sílaba, primera tónica y núcleo de ip	-5.19467	7.52879	1.000	-29.8088	19.4194
	Núcleo y linde de ip	2.63094	1.52155	.855	-2.3435	7.6054
	Primera tónica y núcleo de ip	-3.54667	3.43939	.997	-14.7912	7.6978
Núcleo y linde de ip	Cuerpo	-2.30030	1.31375	.844	-6.5954	1.9948
	Primera sílaba	-3.05524	1.48934	.658	-7.9244	1.8139
	Primera sílaba y primera tónica	-5.32046*	1.58112	.037	-10.4897	-.1513
	Primera tónica	-4.71001	1.50321	.075	-9.6245	.2045
	Núcleo de IP	-.93249	1.41791	1.000	-5.5681	3.7031
	Linde de IP	8.16006*	1.41791	.000	3.5244	12.7957
	Linde de ip	-1.88576	1.50897	.985	-6.8191	3.0476
	Primera tónica, núcleo y linde de ip	-7.55061	3.96429	.757	-20.5112	5.4100
	Primera sílaba, primera tónica y núcleo de ip	-7.82561	7.60036	.997	-32.6737	17.0225
	Núcleo de ip	-2.63094	1.52155	.855	-7.6054	2.3435
Primera tónica y núcleo de ip	-6.17761	3.59337	.860	-17.9255	5.5703	
Primera tónica y núcleo de ip	Cuerpo	3.87730	3.35265	.992	-7.0836	14.8382
	Primera sílaba	3.12237	3.42526	.999	-8.0760	14.3207
	Primera sílaba y primera tónica	.85714	3.46616	1.000	-10.4749	12.1892
	Primera tónica	1.46760	3.43132	1.000	-9.7505	12.6857
	Núcleo de IP	5.24511	3.39482	.928	-5.8537	16.3439
	Linde de IP	14.33767*	3.39482	.001	3.2389	25.4364

Linde de ip	4.29185	3.43385	.985	-6.9345	15.5182
Primera tónica, núcleo y linde de ip	-1.37300	5.02294	1.000	-17.7947	15.0487
Primera sílaba, primera tónica y núcleo de ip	-1.64800	8.20243	1.000	-28.4645	25.1685
Núcleo de ip	3.54667	3.43939	.997	-7.6978	14.7912
Núcleo y linde de ip	6.17761	3.59337	.860	-5.5703	17.9255
*. La diferencia de medias es significativa al nivel 0.05.					



Apéndice 5. Cuadro de comparaciones múltiples resultante de la aplicación de la prueba HSD de Tukey al comparar las medias de la frecuencia fundamental (F0) de las distintas sílabas pertinentes en la totalidad del corpus.

Comparaciones múltiples						
(I)	(J)	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
Sílabas del enunciado	Sílabas del enunciado				Límite inferior	Límite superior
Cuerpo	Primera sílaba	-.71069	1.53241	1.000	-5.7207	4.2993
	Primera sílaba y primera tónica	-5.00743	1.88684	.250	-11.1761	1.1613
	Primera tónica	-3.98956	1.58962	.334	-9.1865	1.2074
	Núcleo de IP	3.77942	1.20637	.075	-.1646	7.7234
	Linde de IP	28.81286*	1.20637	.000	24.8689	32.7569
	Linde de ip	-1.34318	1.61290	1.000	-6.6163	3.9299
	Primera tónica, núcleo y linde de ip	-9.97236	7.77119	.981	-35.3790	15.4342
	Primera sílaba, primera tónica y núcleo de ip	-6.40736	15.53120	1.000	-57.1840	44.3693
	Núcleo de ip	.11495	1.66293	1.000	-5.3217	5.5516
	Núcleo y linde de ip	1.19764	2.72433	1.000	-7.7091	10.1044
Primera tónica y núcleo de ip	-6.11836	6.95243	.999	-28.8482	16.6115	
Primera sílaba	Cuerpo	.71069	1.53241	1.000	-4.2993	5.7207
	Primera sílaba y primera tónica	-4.29674	2.38259	.817	-12.0862	3.4927
	Primera tónica	-3.27887	2.15487	.935	-10.3238	3.7661
	Núcleo de IP	4.49011	1.88994	.423	-1.6887	10.6690
	Linde de IP	29.52356*	1.88994	.000	23.3447	35.7024
	Linde de ip	-.63249	2.17210	1.000	-7.7338	6.4688

	Primera tónica, núcleo y linde de ip	-9.26167	7.90620	.991	-35.1097	16.5863
	Primera sílaba, primera tónica y núcleo de ip	-5.69667	15.59919	1.000	-56.6956	45.3023
	Núcleo de ip	.82564	2.20950	1.000	-6.3980	8.0492
	Núcleo y linde de ip	1.90833	3.08846	1.000	-8.1889	12.0055
	Primera tónica y núcleo de ip	-5.40767	7.10302	1.000	-28.6298	17.8145
Primera sílaba y primera tónica	Cuerpo	5.00743	1.88684	.250	-1.1613	11.1761
	Primera sílaba	4.29674	2.38259	.817	-3.4927	12.0862
	Primera tónica	1.01787	2.41978	1.000	-6.8932	8.9289
	Núcleo de IP	8.78685*	2.18718	.003	1.6362	15.9375
	Linde de IP	33.82029*	2.18718	.000	26.6697	40.9709
	Linde de ip	3.66425	2.43514	.940	-4.2970	11.6255
	Primera tónica, núcleo y linde de ip	-4.96493	7.98248	1.000	-31.0623	21.1324
	Primera sílaba, primera tónica y núcleo de ip	-1.39993	15.63799	1.000	-52.5257	49.7258
	Núcleo de ip	5.12238	2.46856	.641	-2.9482	13.1929
	Núcleo y linde de ip	6.20507	3.27879	.764	-4.5144	16.9245
	Primera tónica y núcleo de ip	-1.11093	7.18782	1.000	-24.6103	22.3884
Primera tónica	Cuerpo	3.98956	1.58962	.334	-1.2074	9.1865
	Primera sílaba	3.27887	2.15487	.935	-3.7661	10.3238
	Primera sílaba y primera tónica	-1.01787	2.41978	1.000	-8.9289	6.8932
	Núcleo de IP	7.76898*	1.93662	.004	1.4375	14.1004
	Linde de IP	32.80242*	1.93662	.000	26.4710	39.1339
	Linde de ip	2.64638	2.21283	.989	-4.5881	9.8808

	Primera tónica, núcleo y linde de ip	-5.98280	7.91749	1.000	-31.8677	19.9021
	Primera sílaba, primera tónica y núcleo de ip	-2.41780	15.60492	1.000	-53.4354	48.5998
	Núcleo de ip	4.10451	2.24956	.805	-3.2500	11.4591
	Núcleo y linde de ip	5.18720	3.11724	.884	-5.0041	15.3785
	Primera tónica y núcleo de ip	-2.12880	7.11558	1.000	-25.3920	21.1344
Núcleo de IP	Cuerpo	-3.77942	1.20637	.075	-7.7234	.1646
	Primera sílaba	-4.49011	1.88994	.423	-10.6690	1.6887
	Primera sílaba y primera tónica	-8.78685*	2.18718	.003	-15.9375	-1.6362
	Primera tónica	-7.76898*	1.93662	.004	-14.1004	-1.4375
	Linde de IP	25.03344*	1.63674	.000	19.6824	30.3845
	Linde de ip	-5.12260	1.95577	.269	-11.5167	1.2715
	Primera tónica, núcleo y linde de ip	-13.75178	7.84953	.844	-39.4145	11.9109
	Primera sílaba, primera tónica y núcleo de ip	-10.18678	15.57054	1.000	-61.0920	40.7185
	Núcleo de ip	-3.66447	1.99723	.799	-10.1941	2.8651
	Núcleo y linde de ip	-2.58178	2.94034	.999	-12.1947	7.0312
Linde de IP	Primera tónica y núcleo de ip	-9.89778	7.03988	.963	-32.9135	13.1179
	Cuerpo	-28.81286*	1.20637	.000	-32.7569	-24.8689
	Primera sílaba	-29.52356*	1.88994	.000	-35.7024	-23.3447
	Primera sílaba y primera tónica	-33.82029*	2.18718	.000	-40.9709	-26.6697
	Primera tónica	-32.80242*	1.93662	.000	-39.1339	-26.4710

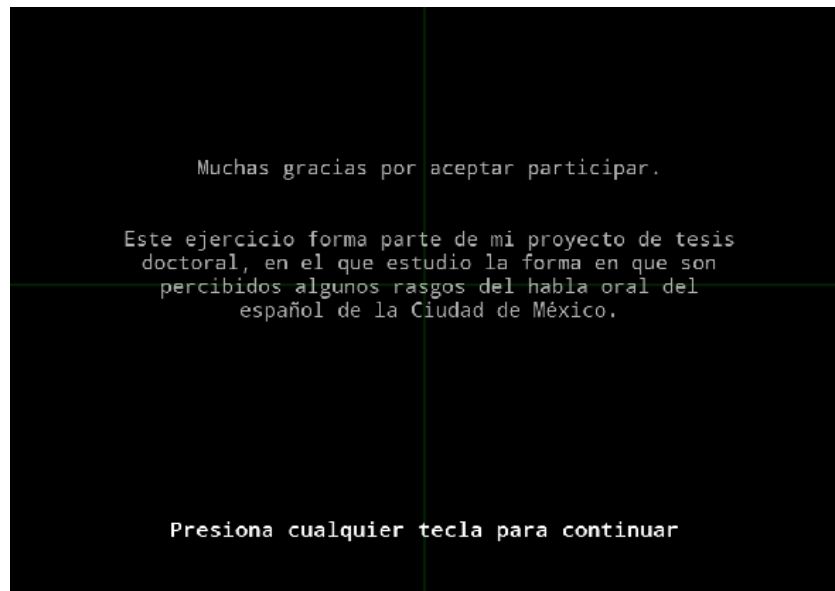
	Núcleo de IP	-25.03344*	1.63674	.000	-30.3845	-19.6824
	Linde de ip	-30.15605*	1.95577	.000	-36.5501	-23.7620
	Primera tónica, núcleo y linde de ip	-38.78522*	7.84953	.000	-64.4479	-13.1225
	Primera sílaba, primera tónica y núcleo de ip	-35.22022	15.57054	.504	-86.1255	15.6850
	Núcleo de ip	-28.69791*	1.99723	.000	-35.2275	-22.1683
	Núcleo y linde de ip	-27.61522*	2.94034	.000	-37.2282	-18.0023
	Primera tónica y núcleo de ip	-34.93122*	7.03988	.000	-57.9469	-11.9155
Linde de ip	Cuerpo	1.34318	1.61290	1.000	-3.9299	6.6163
	Primera sílaba	.63249	2.17210	1.000	-6.4688	7.7338
	Primera sílaba y primera tónica	-3.66425	2.43514	.940	-11.6255	4.2970
	Primera tónica	-2.64638	2.21283	.989	-9.8808	4.5881
	Núcleo de IP	5.12260	1.95577	.269	-1.2715	11.5167
	Linde de IP	30.15605*	1.95577	.000	23.7620	36.5501
	Primera tónica, núcleo y linde de ip	-8.62918	7.92220	.995	-34.5295	17.2711
	Primera sílaba, primera tónica y núcleo de ip	-5.06418	15.60731	1.000	-56.0896	45.9613
	Núcleo de ip	1.45813	2.26607	1.000	-5.9504	8.8667
	Núcleo y linde de ip	2.54082	3.12917	1.000	-7.6895	12.7711
Primera tónica y núcleo de ip	-4.77518	7.12082	1.000	-28.0555	18.5051	
Primera tónica, núcleo y linde de ip	Cuerpo	9.97236	7.77119	.981	-15.4342	35.3790
	Primera sílaba	9.26167	7.90620	.991	-16.5863	35.1097
	Primera sílaba y primera tónica	4.96493	7.98248	1.000	-21.1324	31.0623

	Primera tónica	5.98280	7.91749	1.000	-19.9021	31.8677
	Núcleo de IP	13.75178	7.84953	.844	-11.9109	39.4145
	Linde de IP	38.78522*	7.84953	.000	13.1225	64.4479
	Linde de ip	8.62918	7.92220	.995	-17.2711	34.5295
	Primera sílaba, primera tónica y núcleo de ip	3.56500	17.36024	1.000	-53.1914	60.3214
	Núcleo de ip	10.08731	7.93253	.983	-15.8468	36.0214
	Núcleo y linde de ip	11.17000	8.22081	.971	-15.7066	38.0466
	Primera tónica y núcleo de ip	3.85400	10.41615	1.000	-30.1998	37.9078
Primera sílaba, primera tónica y núcleo de ip	Cuerpo	6.40736	15.53120	1.000	-44.3693	57.1840
	Primera sílaba	5.69667	15.59919	1.000	-45.3023	56.6956
	Primera sílaba y primera tónica	1.39993	15.63799	1.000	-49.7258	52.5257
	Primera tónica	2.41780	15.60492	1.000	-48.5998	53.4354
	Núcleo de IP	10.18678	15.57054	1.000	-40.7185	61.0920
	Linde de IP	35.22022	15.57054	.504	-15.6850	86.1255
	Linde de ip	5.06418	15.60731	1.000	-45.9613	56.0896
	Primera tónica, núcleo y linde de ip	-3.56500	17.36024	1.000	-60.3214	53.1914
	Núcleo de ip	6.52231	15.61256	1.000	-44.5203	57.5649
	Núcleo y linde de ip	7.60500	15.76098	1.000	-43.9229	59.1329
	Primera tónica y núcleo de ip	.28900	17.00949	1.000	-55.3207	55.8987
Núcleo de ip	Cuerpo	-.11495	1.66293	1.000	-5.5516	5.3217
	Primera sílaba	-.82564	2.20950	1.000	-8.0492	6.3980
	Primera sílaba y primera tónica	-5.12238	2.46856	.641	-13.1929	2.9482
	Primera tónica	-4.10451	2.24956	.805	-11.4591	3.2500
	Núcleo de IP	3.66447	1.99723	.799	-2.8651	10.1941
	Linde de IP	28.69791*	1.99723	.000	22.1683	35.2275

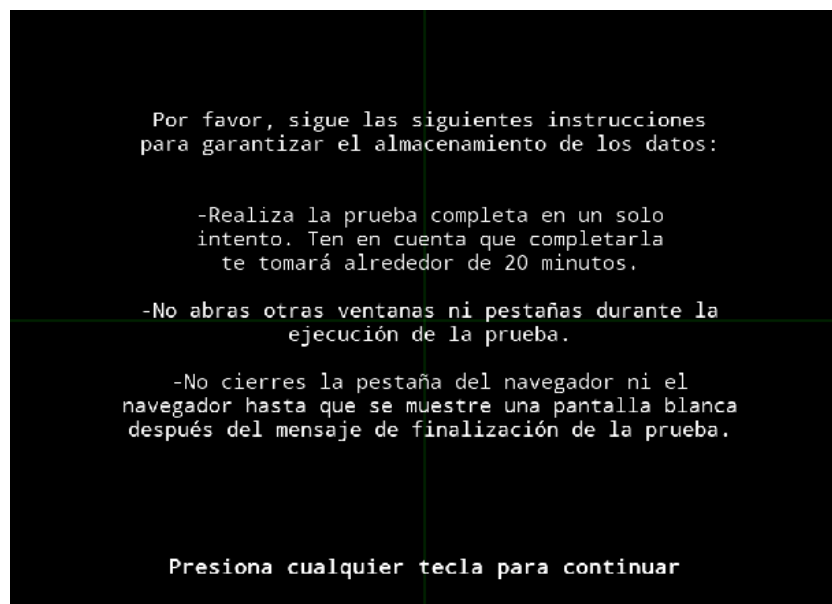
	Linde de ip	-1.45813	2.26607	1.000	-8.8667	5.9504
	Primera tónica, núcleo y linde de ip	-10.08731	7.93253	.983	-36.0214	15.8468
	Primera sílaba, primera tónica y núcleo de ip	-6.52231	15.61256	1.000	-57.5649	44.5203
	Núcleo y linde de ip	1.08269	3.15525	1.000	-9.2329	11.3983
	Primera tónica y núcleo de ip	-6.23331	7.13232	.999	-29.5512	17.0846
Núcleo y linde de ip	Cuerpo	-1.19764	2.72433	1.000	-10.1044	7.7091
	Primera sílaba	-1.90833	3.08846	1.000	-12.0055	8.1889
	Primera sílaba y primera tónica	-6.20507	3.27879	.764	-16.9245	4.5144
	Primera tónica	-5.18720	3.11724	.884	-15.3785	5.0041
	Núcleo de IP	2.58178	2.94034	.999	-7.0312	12.1947
	Linde de IP	27.61522*	2.94034	.000	18.0023	37.2282
	Linde de ip	-2.54082	3.12917	1.000	-12.7711	7.6895
	Primera tónica, núcleo y linde de ip	-11.17000	8.22081	.971	-38.0466	15.7066
	Primera sílaba, primera tónica y núcleo de ip	-7.60500	15.76098	1.000	-59.1329	43.9229
	Núcleo de ip	-1.08269	3.15525	1.000	-11.3983	9.2329
	Primera tónica y núcleo de ip	-7.31600	7.45162	.998	-31.6778	17.0458
Primera tónica y núcleo de ip	Cuerpo	6.11836	6.95243	.999	-16.6115	28.8482
	Primera sílaba	5.40767	7.10302	1.000	-17.8145	28.6298
	Primera sílaba y primera tónica	1.11093	7.18782	1.000	-22.3884	24.6103
	Primera tónica	2.12880	7.11558	1.000	-21.1344	25.3920
	Núcleo de IP	9.89778	7.03988	.963	-13.1179	32.9135

Linde de IP	34.93122*	7.03988	.000	11.9155	57.9469
Linde de ip	4.77518	7.12082	1.000	-18.5051	28.0555
Primera tónica, núcleo y linde de ip	-3.85400	10.41615	1.000	-37.9078	30.1998
Primera sílaba, primera tónica y núcleo de ip	-.28900	17.00949	1.000	-55.8987	55.3207
Núcleo de ip	6.23331	7.13232	.999	-17.0846	29.5512
Núcleo y linde de ip	7.31600	7.45162	.998	-17.0458	31.6778
*. La diferencia de medias es significativa al nivel 0.05.					

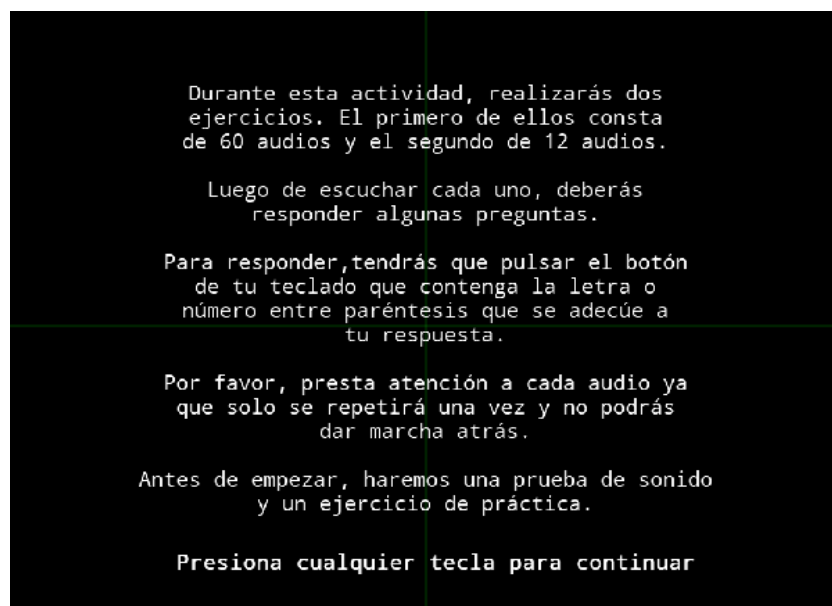
Apéndice 6. Agradecimiento inicial y breve explicación del proyecto.



## Apéndice 7. Instrucciones generales de la prueba de percepción

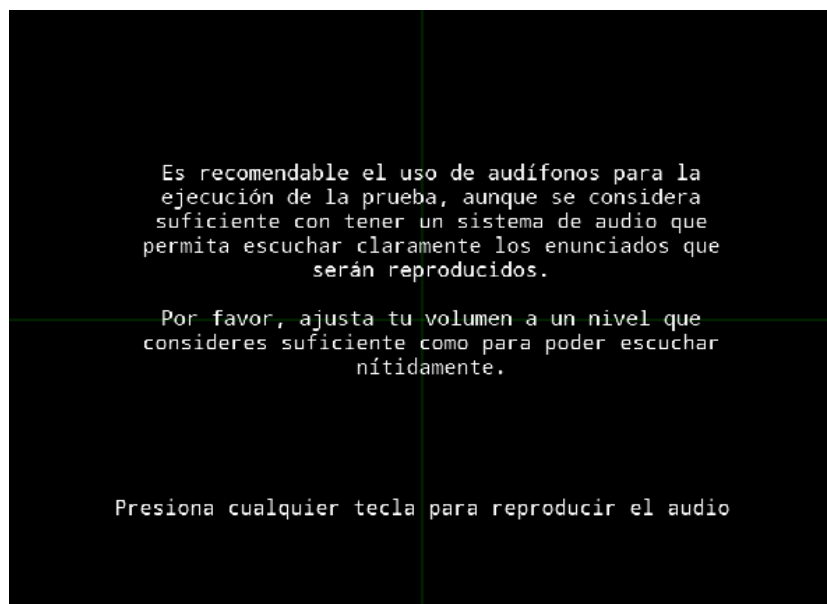


## Apéndice 8. Dinámica de la prueba de percepción

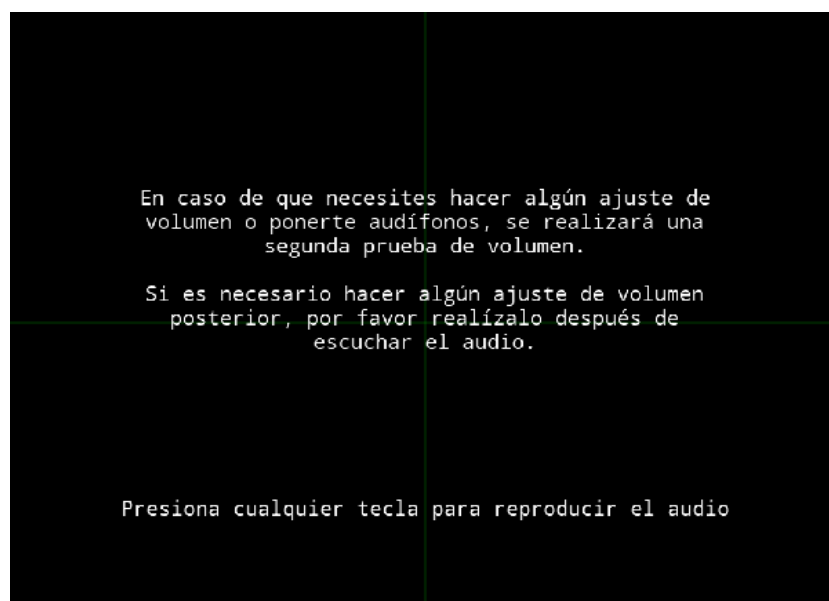




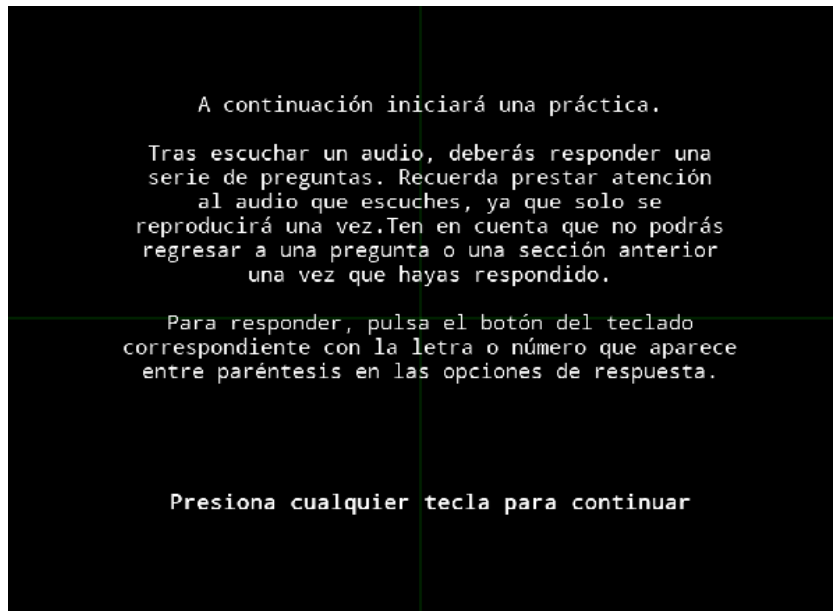
## Apéndice 9. Instrucciones de la prueba de audio



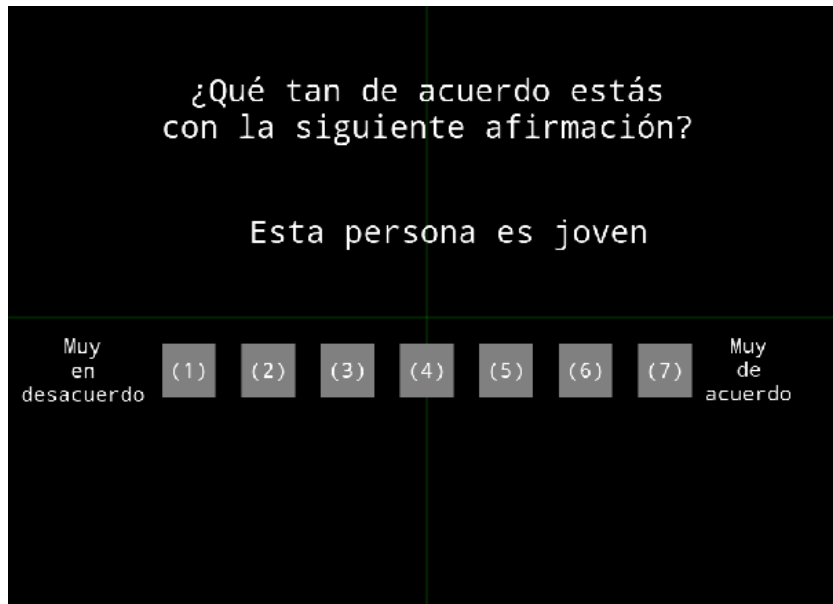
(Luego de la primera reproducción)



Apéndice 10. Instrucciones para la sección de práctica de la prueba de percepción



Apéndice 11. Preguntas de la sección de práctica de la prueba de percepción



¿Qué tan de acuerdo estás  
con la siguiente afirmación?

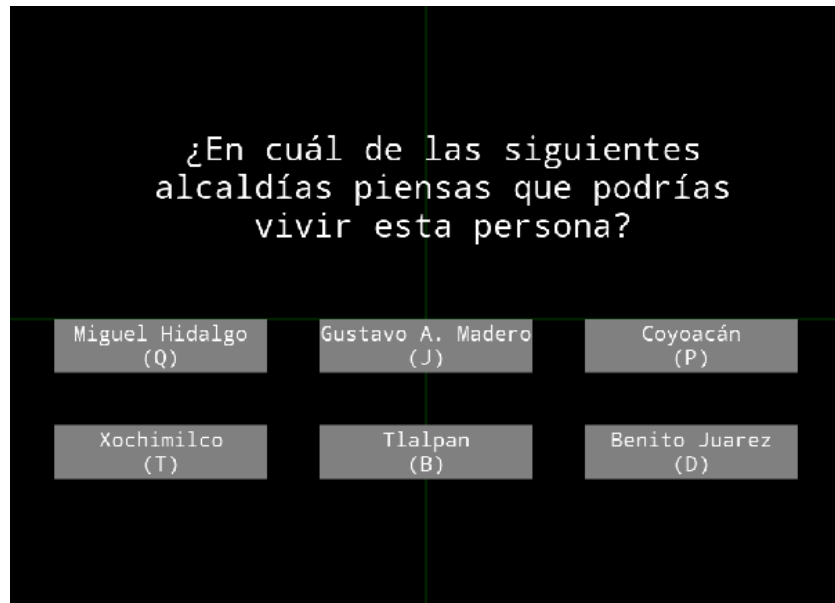
El enunciado suena natural

Muy en desacuerdo (1) (2) (3) (4) (5) (6) (7) Muy de acuerdo

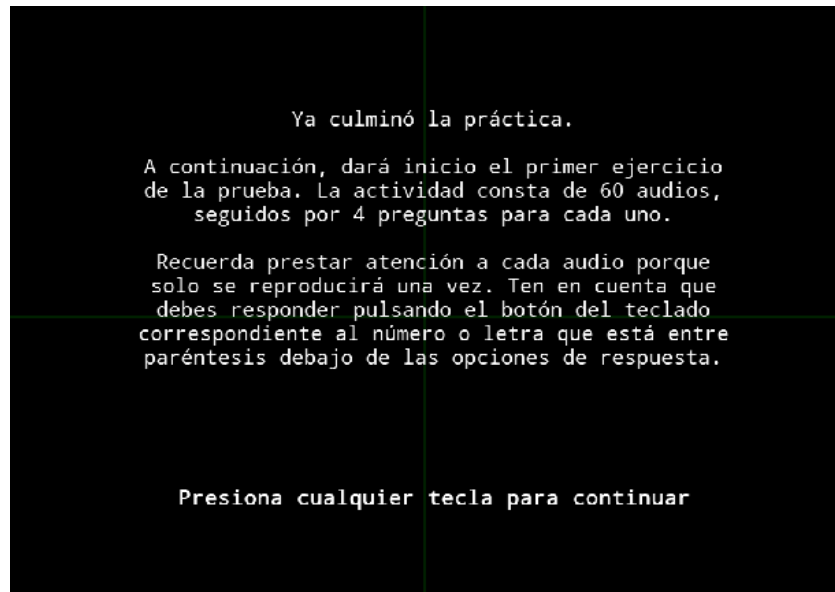
¿Qué tan de acuerdo estás  
con la siguiente afirmación?

Esta persona tiene estudios  
universitarios

Muy en desacuerdo (1) (2) (3) (4) (5) (6) (7) Muy de acuerdo



Apéndice 12. Mensaje de culminación de la práctica e inicio de la prueba de percepción



### Apéndice 13. Anuncio de culminación del primer ejercicio e instrucciones del segundo

Ya culminó el primer ejercicio.

A continuación, dará inicio el segundo ejercicio de la prueba. La actividad consta de 12 audios, seguidos por una pregunta para cada uno.

Cada audio se reproducirá solo una vez, por lo que es necesario que le prestes atención a cada uno. Ten en cuenta que debes responder pulsando el botón del teclado correspondiente al número o letra que está entre paréntesis debajo de las opciones de respuesta.

Presiona cualquier tecla para continuar

### Apéndice 14. Mensaje de culminación de la prueba de percepción.

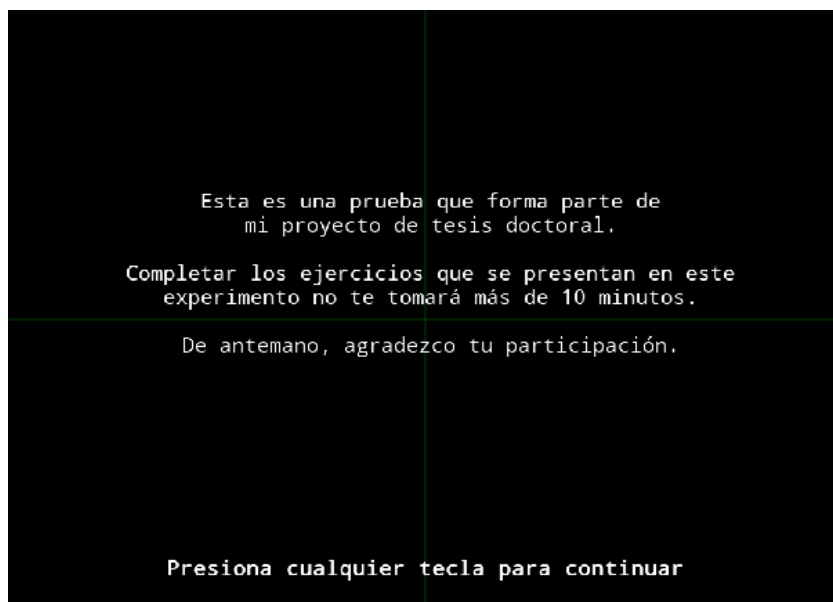
Ya solo queda un paso para culminar la prueba.

Pulsa cualquier botón para guardar los datos y por favor, no cierres la pestaña o el navegador hasta que aparezca una pantalla blanca con el siguiente mensaje:

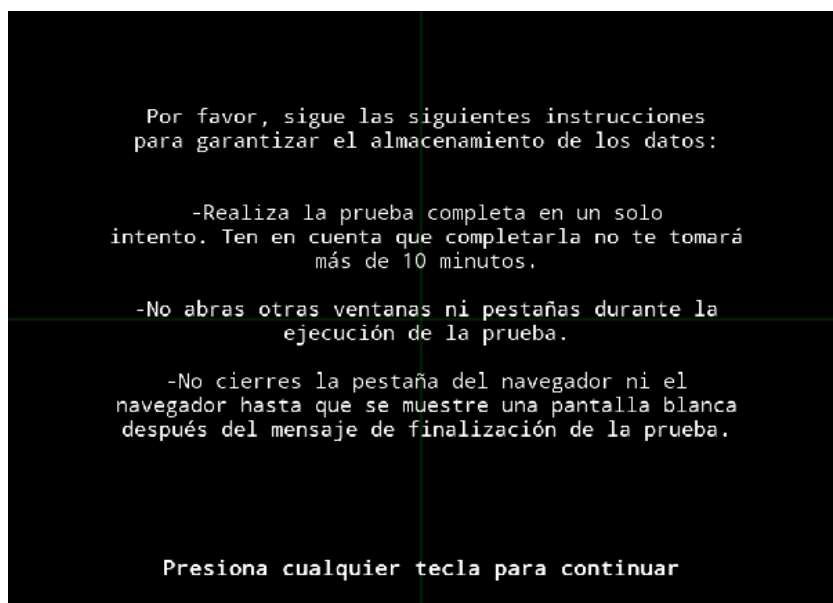
'This study is finished.  
Thank you for your participation'.

Muchas gracias por tu participación.

#### Apéndice 15: Mensaje de bienvenida de la prueba de validación



#### Apéndice 16: Instrucciones de la prueba de validación



## Apéndice 17: Mensaje de culminación de la prueba de validación

