

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE FILOLOGÍA



TESIS DOCTORAL

**Fonética y fonología de los fenómenos de refuerzo consonántico en el
seno de unidades léxicas en español**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

José María Lahoz-Bengoechea

Directora

Juana Gil Fernández

Madrid, 2015



U N I V E R S I D A D
COMPLUTENSE
M A D R I D



INSTITUTO UNIVERSITARIO DE INVESTIGACIÓN
ORTEGA Y GASSET

Tesis Doctoral

Doctorado en Lingüística teórica
y sus aplicaciones

Fonética y fonología
de los fenómenos
de refuerzo consonántico
en el seno de unidades léxicas
en español

José María Lahoz-Bengoechea

Dirigida por Juana Gil Fernández

Madrid, 2015

Agradecimientos

Quiero expresar mi más profundo agradecimiento a mi directora de tesis, la Dra. Juana Gil Fernández, por su sabiduría, por dar siempre en el clavo con sus recomendaciones; por su calidad humana y su empatía; por su capacidad de formar y alentar a un estupendo equipo de colaboradores; por todas las oportunidades que me ha brindado; por las risas y llantos compartidos; y por su dedicación incluso en los momentos más difíciles. A ella debo la supervisión de este trabajo, aunque cualquier posible defecto es responsabilidad mía.

A los doctores Pier Marco Bertinetto y Bernhard Hurch, por todas las sugerencias vertidas en sus informes de evaluación.

A mis supervisores durante mis estancias de investigación en el extranjero (Jacqueline Vaissière, Keith Johnson y Pat Keating), por todo lo que aprendí de ellos y por su apoyo con mis experimentos.

A la Dra. Alicia Puigvert Ocal, por animarme siempre, por su disponibilidad, por su confianza y su complicidad.

Al Dr. Eugenio Bustos Gisbert, por toda la ayuda prestada y por sus consejos.

A la Dra. Sandra Schwab, por su orientación con los análisis estadísticos.

A Marianela Fernández, por «cubrirme» en tantas ocasiones en que mi dedicación a esta tesis no me permitía dar más abasto; y por aportar calma y sensatez en los momentos de pánico.

A un sinfín de colegas de profesión que, conversación tras conversación, han ido ayudando a dar forma a mis ideas. Sería imposible nombrarlos a todos pero, igualmente, gracias.

A los sujetos que participaron en mi experimento. Sin ellos, esta tesis no sería posible.

A mis padres, por su apoyo incondicional en todo y por animarme a estudiar lo que realmente me apasiona. A mi hermana y a mi hermano, por sufrir las ausencias y, otras veces, soportar las presencias; y por ser, además, grandes amigos.

A mis amigos... A Jorge Sosa, por no haber pedido todavía que le convaliden créditos de libre configuración en Lingüística con cargo a horas de amistad; por sus consejos, por saber escuchar, por todo lo que aprendo de él. A Anna, por su amistad a primera vista, desde el día de la cola de la matrícula en primero de carrera; por su cariño inquebrantable. A Marta, por acudir a rescatarme siempre que lo necesito, por nuestras conversaciones, por ser como es. A Fátima, por su espontaneidad y su teatralidad, por hacerme reír, por comprenderme tan bien, por el camino de lucha compartido. A Nacho y Olga, por su paciencia, por su coherencia y su ejemplo de vida. A Toñi, por estar ahí, por su cercanía. A Begoña, por su dulzura. A Jorge Navarro, a Güisi, a Ale... por seguir estando próximos en la distancia. Y, afortunadamente, un largo etcétera.

A Dios, por todos los anteriores, y por acompañarme siempre en el camino.

Finalmente, debo dejar constancia de que la investigación conducente a esta tesis ha sido parcialmente financiada por una beca predoctoral de la Universidad Complutense de Madrid y por una beca del Programa Nacional de Formación del Profesorado Universitario (FPU) del Ministerio de Educación y Ciencia (referencia AP2005-0761).

Índice de contenidos

Agradecimientos.....	3
Índice de tablas.....	9
Índice de figuras.....	13
Lista de abreviaturas.....	15
Nota sobre las transcripciones.....	17
Resumen.....	19
Abstract.....	23
1. Introducción.....	27
2. La estructura prosódica en la arquitectura de la lengua.....	31
2.1. La jerarquía de constituyentes prosódicos.....	34
2.1.1. La sílaba (σ).....	37
2.1.1.1. La sílaba como constituyente.....	37
2.1.1.2. La estructura interna de la sílaba.....	38
2.1.1.3. El concepto de sonoridad.....	44
2.1.1.4. Ambisilabicidad.....	49
2.1.1.5. Resilabeo.....	50
2.1.1.6. La sílaba como unidad de coordinación de gestos articulatorios.....	52
2.1.1.7. La mora (μ): ¿constituyente, o propiedad de la sílaba?.....	57
2.1.2. El pie (Σ).....	61
2.1.2.1. El pie como constituyente.....	61
2.1.2.2. La estructura interna del pie.....	61
2.1.2.3. Posibles infracciones de la estratificación rigurosa.....	62
2.1.3. La palabra fonológica (ω).....	63
2.1.3.1. La constitución de la palabra fonológica.....	63
2.1.3.2. Posibles infracciones de la estratificación rigurosa.....	64
2.1.4. La frase fonológica (φ).....	68
2.1.4.1. La constitución de la frase fonológica.....	68
2.1.4.2. Posibles infracciones de la estratificación rigurosa.....	69
2.1.5. La frase entonativa (I).....	72
2.1.5.1. La constitución de la frase entonativa.....	72
2.1.6. El enunciado fonológico (U).....	72
2.1.6.1. El enunciado fonológico como constituyente.....	72
2.2. La prominencia prosódica y su jerarquía.....	74
2.2.1. Asignación de prominencias correspondientes al acento léxico.....	78
2.2.2. Asignación de prominencias correspondientes al acento de frase.....	81
2.3. Prosodia y segmentación del continuo sonoro como apoyo al acceso al léxico.....	83
2.3.1. Probabilidades transicionales.....	84
2.3.2. La frase fonológica, dominio del acceso al léxico.....	86
2.3.3. Ritmo lingüístico y segmentación preléxica.....	87
2.3.3.1. Segmentación basada en el ritmo silábico.....	89
2.3.3.2. Segmentación basada en el ritmo acentual.....	91
2.3.3.3. Segmentación basada en el ritmo moraico.....	92
2.3.4. Variaciones alofónicas.....	93
3. El refuerzo fonético como marca de la estructura prosódica.....	95
3.1. Refuerzo por la posición prosódica.....	99
3.2. Refuerzo por la prominencia prosódica.....	103
3.3. Gestos glotales.....	106

3.3.1. Gestos glotales: refuerzo por posición.....	106
3.3.2. Gestos glotales: refuerzo por prominencia	112
3.4. Gesto velar	114
3.4.1. Gesto velar: refuerzo por posición.....	114
3.4.2. Gesto velar: refuerzo por prominencia	119
3.5. Movimientos de la mandíbula	121
3.5.1. Movimientos de la mandíbula: refuerzo por posición	121
3.5.2. Movimientos de la mandíbula: refuerzo por prominencia.....	122
3.6. Gestos linguales	123
3.6.1. Gestos linguales: refuerzo por posición.....	123
3.6.2. Gestos linguales: refuerzo por prominencia	126
3.7. Gestos labiales.....	127
3.7.1. Gestos labiales: refuerzo por posición.....	127
3.7.2. Gestos labiales: refuerzo por prominencia	128
4. Fonología de la fortición y la lenición en español	129
4.1. Coda silábica.....	131
4.1.1. Debucalización (I)	137
4.1.2. Elisión (I).....	140
4.1.3. Rotacismo	141
4.1.4. Ensondecimiento pasivo (I).....	141
4.1.5. Vocalización	146
4.2. Ataque silábico en contexto intersonorántico	147
4.2.1. Sonorización pasiva.....	148
4.2.2. Degeminación	150
4.2.3. Espirantización	151
4.2.4. Elisión (II).....	153
4.2.5. Debucalización (II)	154
4.3. Otras instancias de ataque silábico	155
4.3.1. Oclusivización de aproximantes espirantes.....	155
4.3.2. Oclusivización de aproximantes laterales.....	164
4.3.3. Palatalización de sonantes interrumpidas coronales /l/ y /n/	165
4.3.4. Fricativización de la espirante palatal.....	167
4.3.5. Ensondecimiento pasivo (II)	169
5. Un corpus para el estudio del refuerzo subsegmental en español (metodología).....	173
5.1. Hipótesis	174
5.2. Sujetos.....	178
5.3. Lista de palabras	180
5.4. Desarrollo de una sesión experimental	183
5.4.1. Obtención de datos articulográficos	183
5.4.2. Obtención de datos aerodinámicos	185
5.4.3. Obtención de datos acústicos.....	186
5.5. Procesamiento preliminar de los datos	187
5.5.1. Segmentación y etiquetado del audio	187
5.5.2. Ajustes en los archivos de EMA.....	187
5.6. Extracción de los datos de interés.....	193
5.7. Análisis estadístico	197
5.7.1. Efectos fijos	197
5.7.2. Efectos aleatorios.....	199
6. Resultados.....	201
6.1. H ₁ : pico de velocidad de la constricción.....	202

6.2. H ₂ : pico de velocidad de la distensión	209
6.3. H ₃ : duración del gesto de constricción	212
6.4. H ₄ : duración del gesto de distensión	215
6.5. H ₅ : solapamiento temporal entre los movimientos articulatorios de la constricción.....	220
6.6. H ₆ : duración acústica del segmento consonántico	222
6.7. H ₇ : VOT	224
6.8. H ₈ : altura de la mandíbula.....	225
6.9. H ₉ : altura del ápice de la lengua.....	227
6.10. H ₁₀ : abertura labial	229
6.11. H ₁₁ : flujo y presión nasal.....	230
6.12. H ₁₂ : %t pico de magnitud, pico de U _n y pico de P _n	231
6.13. H ₁₃ : intensidad relativa.....	236
7. Discusión y conclusiones.....	239
7.1. Efectos de la posición prosódica	240
7.2. Efectos de la prominencia prosódica.....	244
7.3. Valoración global y aplicaciones de este estudio.....	247
8. General discussion and conclusions	251
8.1. Effects of prosodic position.....	252
8.2. Effects of prosodic prominence.....	256
8.3. Overall evaluation and applications of this study	258
Referencias bibliográficas	263

Índice de tablas

Tabla 2.1. Estructuras de coordinación de gestos consonánticos y vocálicos en el seno de una sílaba.	56
Tabla 2.2. Estructuras de coordinación intergestual y su contribución al peso silábico.	59
Tabla 2.3. Formas verbales con uno y dos pronombres enclíticos (con variación en el carácter de sílaba ligera o pesada del primero —o único— de ellos).	66
Tabla 2.4. Formas verbales de 2 Sg Imperativo con uno y dos pronombres pospuestos.	71
Tabla 4.1. Porcentaje de conservación de [s] (vs. aspiración [h]) en coda real y en coda resilabeada. Datos de Terrell (1980).	132
Tabla 5.1. Hipótesis sobre el refuerzo consonántico como consecuencia de la posición prosódica y de la prominencia.	175
Tabla 5.2. Descripción de los sujetos participantes en el estudio.	180
Tabla 5.3. Lista de palabras. Las consonantes analizadas en cada una de ellas aparecen subrayadas.	182
Tabla 6.1. Resultados de la comparación χ^2 entre los modelos con efecto aleatorio del Sujeto sobre la pendiente (y la constante) y el modelo básico con efecto solo sobre la constante.	201
Tabla 6.2. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de la mandíbula correspondiente a la fase de constricción (mm/muestra).	202
Tabla 6.3. M ₁ : pico de velocidad del movimiento de la mandíbula (fase de constricción).	202
Tabla 6.4. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento del ápice correspondiente a la fase de constricción (mm/muestra).	208
Tabla 6.5. M ₂ : Pico de velocidad del movimiento del ápice (fase de constricción).	208
Tabla 6.6. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de los labios correspondiente a la fase de constricción (mm/muestra).	208
Tabla 6.7. M ₃ : Pico de velocidad del movimiento de los labios (fase de constricción).	208
Tabla 6.8. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de la mandíbula correspondiente a la fase de distensión (mm/muestra).	210
Tabla 6.9. M ₄ : Pico de velocidad del movimiento de la mandíbula (fase de distensión).	210
Tabla 6.10. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad de la distensión correspondiente al gesto del ápice (mm/muestra).	211
Tabla 6.11. M ₅ : Pico de velocidad de la distensión correspondiente al gesto del ápice.	211
Tabla 6.12. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad de la distensión correspondiente al gesto labial (mm/muestra).	211
Tabla 6.13. M ₆ : Pico de velocidad de la distensión correspondiente al gesto labial.	211

Tabla 6.14. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de la mandíbula correspondiente a la fase de constricción.....	212
Tabla 6.15. M ₇ : Duración del movimiento de la mandíbula (fase de constricción).....	212
Tabla 6.16. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento del ápice de la lengua correspondiente a la fase de constricción.....	213
Tabla 6.17. M ₈ : Duración del movimiento del ápice de la lengua (fase de constricción).	213
Tabla 6.18. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de los labios correspondiente a la fase de constricción.....	214
Tabla 6.19. M ₉ : Duración del movimiento de los labios (fase de constricción).....	214
Tabla 6.20. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de la mandíbula correspondiente a la fase de distensión.....	216
Tabla 6.21. M ₁₀ : Duración del movimiento de la mandíbula (fase de distensión).....	216
Tabla 6.22. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento del ápice de la lengua correspondiente a la fase de distensión.	218
Tabla 6.23. M ₁₁ : Duración del movimiento del ápice de la lengua (fase de distensión).	218
Tabla 6.24. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de los labios correspondiente a la fase de distensión.	219
Tabla 6.25. M ₁₂ : Duración del movimiento de los labios (fase de distensión).....	219
Tabla 6.26. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del segmento vocálico siguiente a la consonante.....	220
Tabla 6.27. M ₁₃ : Duración acústica del segmento vocálico siguiente.....	220
Tabla 6.28. Media (y desviación estándar), por condiciones prosódicas, de la duración (s) del solapamiento temporal entre los movimientos de la mandíbula y del ápice (fase de constricción).	221
Tabla 6.29. M ₁₄ : Solapamiento temporal entre los movimientos de la mandíbula y del ápice (fase de constricción).	221
Tabla 6.30. Media (y desviación estándar), por condiciones prosódicas, de la duración (s) del solapamiento temporal entre los movimientos de la mandíbula y de los labios (fase de constricción).....	221
Tabla 6.31. M ₁₅ : Solapamiento temporal entre los movimientos de la mandíbula y de los labios (fase de constricción).	221
Tabla 6.32. Media (y desviación estándar), por condiciones prosódicas, de la duración acústica (en segundos) de las consonantes orales.....	223
Tabla 6.33. M ₁₆ : Duración acústica de las consonantes orales.....	223
Tabla 6.34. Media (y desviación estándar), por condiciones prosódicas, de la duración acústica (en segundos) de las consonantes nasales.....	223
Tabla 6.35. M ₁₇ : Duración acústica de las consonantes nasales.....	223
Tabla 6.36. Media (y desviación estándar), por condiciones prosódicas, del VOT (en segundos) de las oclusivas sordas.....	225
Tabla 6.37. M ₁₈ : VOT de las oclusivas sordas.....	225

Tabla 6.38. Media (y desviación estándar), por condiciones prosódicas, de la proporción de tiempo que ocupa la distensión de las oclusivas sobre el total de su duración acústica.	225
Tabla 6.39. M ₁₉ : Proporción de tiempo que ocupa la distensión de las oclusivas sobre el total de su duración acústica.	225
Tabla 6.40. Media (y desviación estándar), por condiciones prosódicas, de la altura (en mm) de la mandíbula.	226
Tabla 6.41. M ₂₀ : Altura de la mandíbula.	226
Tabla 6.42. Media (y desviación estándar), por condiciones prosódicas, de la altura (en mm) del ápice de la lengua.	227
Tabla 6.43. M ₂₁ : Altura del ápice de la lengua.	227
Tabla 6.44. Media (y desviación estándar), por condiciones prosódicas, de la abertura labial (distancia entre el labio superior y el labio inferior, en mm).	229
Tabla 6.45. M ₂₂ : Abertura labial.	229
Tabla 6.46. Media (y desviación estándar), por condiciones prosódicas, del flujo de aire nasal (en ml/s).	230
Tabla 6.47. M ₂₃ : Flujo de aire nasal.	230
Tabla 6.48. Media (y desviación estándar), por condiciones prosódicas, de la presión de aire nasal (en cm de H ₂ O).	230
Tabla 6.49. M ₂₄ : Presión de aire nasal.	231
Tabla 6.50. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del movimiento de ascenso de la mandíbula.	232
Tabla 6.51. M ₂₅ : %t pico de magnitud del movimiento de ascenso de la mandíbula.	232
Tabla 6.52. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del gesto de ascenso del ápice de la lengua.	232
Tabla 6.53. M ₂₆ : %t pico de magnitud del gesto de ascenso del ápice de la lengua.	232
Tabla 6.54. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del gesto de constricción labial.	233
Tabla 6.55. M ₂₇ : %t pico de magnitud del gesto de constricción labial.	233
Tabla 6.56. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del flujo nasal.	234
Tabla 6.57. M ₂₈ : %t pico de magnitud del flujo nasal.	234
Tabla 6.58. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud de la presión nasal.	234
Tabla 6.59. M ₂₉ : %t pico de magnitud de la presión nasal.	234
Tabla 6.60. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las fricativas (proporción).	236
Tabla 6.61. M ₃₀ : Intensidad relativa de las fricativas.	236
Tabla 6.62. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las espirantes (proporción).	237
Tabla 6.63. M ₃₁ : Intensidad relativa de las espirantes.	237
Tabla 6.64. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las nasales (proporción).	237
Tabla 6.65. M ₃₂ : Intensidad relativa de las nasales.	237

Tabla 7.1. Recapitulación de los resultados del refuerzo por posición y por prominencia.	241
---	-----

Table 8.1 Summary of the results for strengthening due to position and prominence.	253
---	-----

Índice de figuras

Figura 2.1. Modelo plano de la sílaba	39
Figura 2.2. Modelo plano con nivel CV	39
Figura 2.3. Modelo KLV	39
Figura 2.4. Modelo ternario	39
Figura 2.5. Modelo binario con cuerpo y coda	39
Figura 2.6. Modelo binario con ataque y rima	39
Figura 2.7. Mínima especificación dicotómica de la sonoridad en función de los rasgos distintivos (según Blevins, 1995)	47
Figura 2.8. Ambisilabicidad de una consonante intervocálica interior de pie	50
Figura 2.9. Consonante prenuclear dominada directamente por la sílaba	58
Figura 2.10. Consonante prenuclear dominada por la mora nuclear	58
Figura 2.11. Estructura de un pie ternario	63
Figura 2.12. Semisílaba en la periferia de la palabra	65
Figura 2.13. Estructura con enclíticos	67
Figura 2.14. Estructura con proclíticos	70
Figura 2.15. Estructura con falsos enclíticos	72
Figura 5.1. Colocación de los sensores de EMA	184
Figura 5.2. Colocación de la máscara y los tubos de mediciones aerodinámicas	184
Figura 5.3. Distribución de los valores del pico de P_n durante los segmentos nasales	186
Figura 5.4. Ejemplo simplificado de un fotograma	190
Figura 5.5. Traslación de las coordenadas x mediante sustacción del valor inicial $R1x$	190
Figura 5.6. Traslación de las coordenadas y mediante sustracción del valor inicial $R1y$: ahora $R1$ coincide con el centro (0,0)	190
Figura 5.7. El ángulo θ que forma el segmento $R2-R1$ sobre la horizontal se calcula como el arctan ($R2y/R2x$)	190
Figura 5.8. El ángulo $\theta_{dif} = \theta' - \theta$ es el ángulo que debe aplicarse en la rotación para hacer coincidir $R2$ (el puente de la nariz) con $R2'$ (posición de convergencia tomada del fotograma de referencia, FR)	190
Figura 5.9. Cuadrantes (I-IV) de un sistema de coordenadas en función de los valores positivo o negativo de x e y	191
Figura 5.10. Rotación de todos los puntos sobre el centro de coordenadas ($R1$) en función del ángulo θ_{dif} calculado previamente.	191
Figura 5.11. Representación idealizada de una señal articulográfica y su derivada. ...	194
Figura 6.1. M_1 : Interacción Unidad x Acento en el pico de velocidad del movimiento de la mandíbula (fase de constricción)	204
Figura 6.2. M_1 : Interacción Consonante x Acento en el pico de velocidad del movimiento de la mandíbula (fase de constricción)	206
Figura 6.3. M_1 : Efecto aleatorio del Sujeto sobre la pendiente de Unidad en el pico de velocidad de la constricción correspondiente al movimiento de la mandíbula. ...	207
Figura 6.4. M_3 : Interacción Unidad x Acento en el pico de velocidad del movimiento de los labios (fase de constricción)	209
Figura 6.5. M_9 : Interacción Unidad x Acento en la duración del movimiento labial de constricción	215

Figura 6.6. M ₉ : Interacción Consonante x Unidad en la duración del movimiento labial (fase de constricción).....	215
Figura 6.7. M ₁₀ : Interacción Consonante x Acento en la duración del movimiento de distensión de la mandíbula.	217
Figura 6.8. M ₁₁ : Interacción Unidad x Acento en la duración del movimiento del ápice (fase de distensión).....	218
Figura 6.9. M ₁₅ : Interacción Consonante x Unidad en la duración del solapamiento entre el movimiento de la mandíbula y el movimiento de los labios (fase de constricción).	222
Figura 6.10. M ₁₇ : Interacción Consonante x Unidad en la duración acústica de las consonantes nasales.....	224
Figura 6.11. M ₂₀ : Interacción Consonante x Unidad en la altura de la mandíbula.	228
Figura 6.12. M ₂₈ : Efecto aleatorio del Sujeto sobre la pendiente de Acento en el %t pico de magnitud del flujo nasal.....	235
Figura 6.13. M ₂₉ : Efecto aleatorio del Sujeto sobre la pendiente de Unidad en el %t pico de magnitud de la presión nasal.....	235
Figura 6.14. M ₃₂ : Interacción Unidad x Acento en la intensidad relativa de las nasales.	238
Figura 7.1. Solapamiento relativo de los movimientos articulatorios integrantes de un gesto.....	242
Figure 8.1. Relative overlap of articulatory movements comprising a gesture.....	254

Lista de abreviaturas

EFC	escala de fuerza de las consonantes
EMA	articulografía electromagnética
EMG	electromiografía
EPG	electropalatografía
f_0	frecuencia fundamental
L1	lengua primera o lengua materna
LA	abertura de labios
P_n	presión nasal
P_o	presión oral (y, en general, presión supraglotal)
P_s	presión subglotal
P_t	presión transglotal
Q	palabra morfosintáctica
RMS	<i>root mean square</i> (raíz cuadrada de la media de los cuadrados)
SPE	<i>Sound Pattern of English</i> (Chomsky y Halle, 1968)
%t	punto proporcional de la duración acústica en que se produce un evento
U_n	flujo de aire nasal
v-f	velofaríngeo
VOT	<i>Voice Onset Time</i>

Símbolos referidos a la estructura prosódica

U	enunciado fonológico
I	frase entonativa
φ	frase fonológica
ω	palabra fonológica
Σ	pie
σ	sílaba
$^l\sigma$	sílaba acentuada
σ	sílaba no acentuada
μ	mora

Nota sobre las transcripciones

A lo largo de esta tesis, las transcripciones fonéticas y —quizá para mayor sorpresa del lector— también las fonológicas se presentan con un elevado nivel de detalle fonético (lo que habitualmente se conoce como transcripción estrecha). Por ejemplo, para hacer alusión al fonema oclusivo dental sordo del español¹, se emplea el símbolo diacrítico correspondiente al carácter dental previsto por el Alfabeto Fonético Internacional (AFI): /t̪/. Como se verá en las próximas páginas, en esta obra se adopta la idea de que el detalle fonético, cuando refleja un comportamiento sistemático de una lengua (y no necesariamente de otras) forma parte del conocimiento lingüístico que tienen los hablantes. A menudo, por comodidad, en las publicaciones sobre fonética y fonología las variaciones predecibles se omiten de la transcripción, para aligerarla. Sin embargo, su predictibilidad para una determinada lengua es, precisamente, lo que confirma que son propias del sistema fonológico de ese idioma. El grado de detalle aconsejable para una transcripción depende de la finalidad de la misma. Dadas las características de este trabajo, y mi propia concepción de la fonética y la fonología, me ha parecido apropiado afinar aquí el detalle de la transcripción, especialmente en el caso de las consonantes, objeto de estudio de esta tesis. Además, mis investigaciones en el campo de la fonética contrastiva, aplicada al desarrollo de estrategias de enseñanza del componente fónico de segundas lenguas, me han hecho ver que en ocasiones una obra con información potencialmente muy valiosa debe tomarse con precaución porque la ausencia de detalle fonético en las explicaciones o en las transcripciones no permite una comparación (con la otra lengua de interés) sobre bases fiables.

Por otro lado, frente a la práctica habitual, se incluye la marca de sílaba acentuada incluso en las palabras monosílabas, cuando estas tienen acento léxico, p. ej. *sí* ['si], por contraposición a *si* [si]. Si bien es cierto que en los monosílabos no cabe el contraste con la asignación de la prominencia acentual a otra sílaba (porque no la hay), tanto el conocimiento fonológico como la realización fonética sí manifiestan una distinción entre palabras léxicamente acentuadas y las que no lo son, y creo que esta información debe quedar recogida en las transcripciones.

¹ Más precisamente, se trata de una articulación dentoalveolar laminal (Martínez Celdrán y Fernández Planas, 2013, p. 47, nota 3).

Resumen

Fonética y fonología de los fenómenos de refuerzo consonántico en el seno de unidades léxicas en español

Esta tesis es un estudio de los fenómenos de refuerzo consonántico, entendido como un cambio en el desarrollo espaciotemporal de los gestos articulatorios —*i. e.* como una modificación (normalmente un aumento) de su duración, de su magnitud, o de ambas— como consecuencia de la estructura prosódica en la que se insertan; en concreto, por su contacto con una frontera de constituyente prosódico o por su aparición en una unidad relativamente prominente con respecto a otras del mismo nivel.

Todas las lenguas poseen una estructura prosódica, que consiste en una organización de los elementos de contenido fonológico en una serie de unidades situadas en un nivel jerárquico progresivamente más alto. Así, los gestos articulatorios, de cuya combinación y abstracción resultan los distintos segmentos fonológicos, se organizan en sílabas, que se organizan pies, y estos a su vez en palabras, las cuales se agrupan en constituyentes cada vez mayores hasta formar un enunciado. Además, cada unidad tiene un núcleo, que es su constituyente más prominente y que sirve para señalar la presencia de dicha unidad.

La estructura prosódica es relevante no solo por la organización del contenido fonológico, sino también porque sirve de interfaz con otros niveles lingüísticos, como el morfosintáctico o el semántico. Por ello es importante que los hablantes (y los oyentes) dispongan de claves que les permitan codificar (o decodificar) la jerarquía de los elementos que componen los enunciados. El refuerzo espaciotemporal de los gestos articulatorios situados en los extremos de los distintos constituyentes supone una marca fonética que pone de relieve la frontera entre unidades, del mismo modo que el refuerzo asociado a las posiciones prominentes ayuda a indicar el número de constituyentes que hay en un determinado nivel jerárquico de esa estructura. La otra cara inevitable de la moneda, el debilitamiento gestual en las posiciones centrales de un constituyente, señala la continuidad del mismo y, por ello, también proporciona información —a su manera— acerca de la estructura prosódica.

Los fenómenos de refuerzo y debilitamiento están, pues, ligados a determinadas posiciones prosódicas. Su desencadenamiento no depende de los gestos o segmentos que haya alrededor, como sucede, por ejemplo, con otro tipo de fenómenos, los de coarticulación y asimilación (o también los de disimilación). Por ello es importante distinguir las nociones de *posición* (prosódica) y *contexto* (segmental), porque, si bien en ocasiones ambos tipos de fenómeno pueden interactuar, conviene reconocer la contribución particular de uno y otro al resultado final. Por ejemplo, los cambios por asimilación suelen ser más frecuentes en posiciones prosódicas propensas al debilitamiento, ya que en ellas se produce un mayor solapamiento entre gestos articulatorios contiguos, pero, en última instancia, el devenir de la asimilación depende de la identidad de los gestos que entran en contacto.

Las modificaciones en la magnitud y en la duración gestuales asociadas a los fenómenos de refuerzo (y debilitamiento) pueden manifestarse como pequeñas alteraciones en el detalle de la realización fonética, sin que estas modifiquen la identidad del segmento implicado. Se trata de variaciones subsegmentales que solo pueden analizarse mediante las técnicas instrumentales apropiadas. Las lenguas estudiadas hasta ahora con estos medios indican una serie de tendencias comunes, como características generales de los procesos de refuerzo, pero cada idioma presenta, igualmente, ciertas particularidades. No en vano, aunque se trate de modificaciones fonéticas, estas actúan sobre la base de los elementos de contenido fonológico y dependen de la manera en que estos se organizan prosódicamente, aspectos que sí son peculiares de cada sistema.

Por otro lado, en ocasiones los efectos de la posición prosódica sobrepasan la mera variación subsegmental y se ponen de manifiesto mediante alternancias alofónicas en un proceso de variación sincrónica, o como cambios fonológicos consolidados diacrónicamente. En estos casos, con mayor motivo, la regulación de las alternancias se integra en el sistema fonológico de la lengua y, si se produce un cambio de un fonema por otro, esta modificación se refleja en la forma léxica subyacente. Estos cambios alofónicos y fonémicos condicionan el inventario de segmentos que habitualmente suele aparecer en una posición u otra, en el seno de una sílaba o de una palabra.

Por ejemplo, la posición de coda silábica es una posición débil por excelencia, y resulta propicia a fenómenos de debucalización, elisión, rotacismo y ensordecimiento pasivo, entre otros posibles, lo que puede llegar a traducirse en una neutralización de determinadas oposiciones fonológicas. Por el contrario, la posición de ataque silábico es una posición relativamente fuerte. En el caso de los constituyentes superiores a la sílaba, son fuertes tanto la posición inicial como la final, mientras que resulta débil el interior, estableciendo un patrón en forma de V. Existe una gradación tal que el refuerzo es tanto mayor cuanto más alto en la jerarquía prosódica se sitúe el constituyente en cuestión. Es decir, los efectos del refuerzo apuntan en la misma dirección en los distintos niveles jerárquicos, y se manifiestan de manera acumulativa, aunque no siempre es posible detectar cambios discretos entre un nivel y el siguiente. También se producen interacciones, de modo que, cuando una coda (débil) coincide con la posición final de una palabra (u otro constituyente superior), ya no resulta tan débil, como consecuencia del refuerzo propio de la posición de frontera prosódica.

Como se ha dicho, el ataque silábico es relativamente fuerte (con respecto a la coda), pero es más débil que otros inicios de constituyente de mayor rango. Esto le otorga al ataque unas propiedades especiales, que lo hacen relativamente vulnerable a los efectos de la coarticulación, especialmente en contexto intersonántico. Bajo estas circunstancias, se producen fenómenos de sonorización pasiva, degeminación, espirantización, elisión o debucalización, entre otros. En cambio, en un contexto diferente, o si el ataque coincide además con la posición inicial de palabra (o de otros constituyentes superiores), eso lo convierte en una posición más fuerte, propicia a otros tipos de fenómenos, que pueden incluir la oclusivización o la fricativización de aproximantes, o el ensordecimiento pasivo.

Esta tesis describe los fenómenos de lenición y fortición manifestados como cambios segmentales en el seno de unidades léxicas en español. Por otro lado, se plantea el objetivo de estudiar experimentalmente los cambios subsegmentales (de detalle fonético) derivados de la particular dinámica gestual que caracteriza a la posición inicial de palabra (frente a la inicial de sílaba) y a las sílabas acentuadas (frente a las no acentuadas). Para ello, se ha grabado un corpus de habla con seis informantes mexicanos, que incluye datos articulográficos, aerodinámicos (flujo y presión del aire nasal) y acústicos.

El refuerzo por posición (inicial de palabra) se manifiesta como una menor velocidad de ejecución de los movimientos articulatorios, especialmente durante la fase de constricción de la consonante y, en menor medida, durante la fase de distensión. Además, se produce un menor grado de solapamiento temporal entre los distintos movimientos articulatorios que componen los gestos orales (mandíbula-ápice y mandíbula-labios). Este resultado coincide con la idea de que las posiciones fuertes son menos propicias a la coarticulación, y permite constatar que también es válida en lo que se refiere a la coordinación intragestual. En cuanto a la magnitud de los gestos, esta sufre un incremento cuando, a continuación, se realiza otro gesto que modifique la misma variable de tracto en una dirección diferente. Por ejemplo, la altura de la mandíbula y la altura del ápice alcanzan valores más elevados en posición inicial de palabra, como consecuencia del refuerzo y gracias a que la vocal siguiente también presenta una especificación de gestos que implican a la mandíbula y a la lengua. Además, el refuerzo espacial en el caso del movimiento del ápice provoca, en el caso de las fricativas, un estrechamiento más acusado y un aumento de la intensidad relativa del ruido de fricción. Por otro lado —y al contrario de lo que ocurre con la mandíbula y el ápice—, la abertura labial en el caso de las consonantes labiales o el gesto de descenso del velo del paladar en el caso de las nasales experimentan una menor magnitud en posiciones fuertes, debido al hecho de que las vocales que las seguían ([a], [i]) no estaban especificadas para gestos labiales ni velares.

Por lo que respecta a las medidas acústicas de las consonantes en posición inicial de palabra, no se observa ninguna modificación significativa de la duración segmental. Tampoco hay variación en el VOT (*Voice Onset Time*) de las oclusivas sordas, resultado contrario al de otras lenguas estudiadas previamente. La divergencia se justifica por el hecho de que el español es una lengua que basa sus oposiciones laríngeas en la presencia o ausencia de sonorización (fonación), y no en la aspiración. Por tanto, las sordas son el elemento no marcado y el refuerzo desencadenado por la posición prosódica no puede afectar al gesto glotal, ya que no hay ninguno especificado en la forma subyacente.

En cuanto al refuerzo por prominencia (*i. e.* el que afecta a las sílabas acentuadas), se constata una disminución de la velocidad articulatoria durante la fase de constricción y un aumento durante la fase de distensión. Además, la distensión se caracteriza por una mayor duración articulatoria, lo que tiene su reflejo en el aumento de las duraciones acústicas de la consonante y de la vocal. La prolongación del movimiento distensivo se interpreta como un retraso en el comienzo del subsiguiente gesto de constricción

consonántica, lo que no solo permite que la vocal de la sílaba acentuada dure más, sino también que alcance una mayor abertura y esto se manifieste en la radiación de una mayor energía acústica. Efectivamente, se comprueba que la intensidad de las consonantes, en relación a la de la vocal, disminuye como consecuencia de la mayor intensidad vocálica. Por otro lado, y al igual que en el caso del refuerzo posicional, no se observa ninguna modificación significativa del VOT.

El estudio concluye con una valoración global de los resultados y se plantea la relevancia de los mismos para estudiar los condicionantes de las restricciones fonotácticas que se observan en español en el seno de las palabras, así como de cara a fenómenos de tipo psicolingüístico, como el acceso al léxico. Por otro lado, se proponen aplicaciones para la mejora de los sistemas de reconocimiento y síntesis de habla, mediante la incorporación del detalle fonético propio de las consonantes situadas en posiciones prosódicamente fuertes. Finalmente, se avanza posibles líneas de investigación futura que tienen que ver con la interacción entre los efectos posicionales y los contextuales, y también con la conveniencia de establecer el grado en el que cada gesto articulatorio contribuye a determinados parámetros aerodinámicos, como la presión oral (parámetros que enriquecerían el modelo de la dinámica de tareas).

Abstract

Phonetics and phonology of consonant strengthening phenomena within lexical units in Spanish

This dissertation is a study of consonant strengthening phenomena, which are understood as a change in the spatiotemporal development of articulatory gestures—i.e. a modification (usually an increase) in their duration, magnitude, or both—as a consequence of the prosodic structure into which they are inserted; specifically, due to their contact with a prosodic boundary or their appearance in a relatively prominent unit as compared to others on the same level.

All languages have a prosodic structure, which consists of organizing elements of phonological content into a series of units situated in a progressively higher hierarchical level. Thus, articulatory gestures, whose combination and abstraction result in different phonological segments, are organized into syllables, which are organized into feet which, in turn, are organized into words, which are grouped into increasingly larger components until they form an utterance. Furthermore, each unit has a nucleus, which is its most prominent component and which serves to signal the presence of said unit.

Prosodic structure is relevant not only in the organization of phonological content, but also because it serves as an interface with other linguistic levels, such as morphosyntax or semantics. Therefore, it is important for speakers (and listeners) to have keys which allow them to code (or decode) the hierarchy of the elements that comprise the utterances. The spatiotemporal strengthening of articulatory gestures located at the ends of the different components provides a phonetic cue that makes the boundary between units apparent, in the same way that strengthening associated with prominent positions helps to indicate the number of components in a specific hierarchical level of that structure. On the other hand, gestural weakening in the central positions of a component signals the continuity of said component and, therefore, also provides information (in its own way) regarding the prosodic structure.

The phenomena of strengthening and weakening are, therefore, linked to certain prosodic positions. Their triggering does not depend on the gestures or segments around them, as is the case with other types of phenomena, such as coarticulation or assimilation (and also dissimilation). Therefore it is important to distinguish between the notions of (prosodic) *position* and (segmental) *context* because, while on occasion both types of phenomena may interact, we must remember the particular contribution of each one to the final result. For example, changes due to assimilation are usually more frequent in prosodic positions prone to weakening, as in these positions there is greater overlap between contiguous articulatory gestures; however, ultimately, the outcome of assimilation depends on the identity of the gestures that come into contact.

Modifications in gestural magnitude and duration associated with strengthening (and weakening) phenomena may manifest themselves as small alterations in details of phonetic performance, without these alterations modifying the identity of the segment

involved. These are subsegmental variations that can only be analyzed using the appropriate instrumental techniques. The languages studied with these means up to now indicate a series of common trends, as general characteristics of strengthening processes, but, likewise, each language presents its own special features. With good reason: although these modifications are phonetic, they act upon the basis of elements of phonological content and depend upon how these elements are organized prosodically; these aspects are unique to each system.

Moreover, sometimes the effects of prosodic position exceed mere subsegmental variation and manifest themselves through allophonic alternations in a process of synchronic variation, or as diachronically consolidated phonological changes. In these cases, the regulation of the alternations becomes integrated in the language's phonological system and if one phoneme is changed for another, this modification is reflected in the underlying lexical form. These allophonic and phonemic changes condition the inventory of segments that usually appear in one position or another in a syllable or a word.

For example, the position of the syllabic coda is the weak position par excellence and prone to phenomena of debuccalization, deletion, rhotacism and passive devoicing, among others, which may result in the neutralization of certain phonological oppositions. On the other hand, the onset of a syllable is a relatively strong position. For components on a higher level than the syllable, both the initial and final position are strong, while the medial position is weak, establishing a V-shaped pattern. There is a progression such that strengthening is greater the higher the component in question is located in the prosodic hierarchy. That is, the effects of strengthening point in the same direction at the different hierarchical levels, and manifest themselves accumulatively, although it is not always possible to detect discrete changes between one level and the next. Interactions also occur so that, when a coda (weak) coincides with the final position of a word (or another higher component), it becomes stronger, as a result of the strengthening from the prosodic boundary position.

As has been said, the syllable onset is relatively strong (compared to the coda), but it is weaker than other higher-ranking component-initial positions. This gives the onset some special properties, which make it relatively vulnerable to the effects of coarticulation, especially in an intersonorant context. Under these circumstances, passive voicing, degemination, spirantization, deletion or debuccalization phenomena occur. Conversely, in a different context, or if the onset coincides with the word-initial position (or that of other higher components), this makes it a stronger position, prone to other types of phenomena, which may include occlusivization or fricativization of approximants, or passive devoicing.

This dissertation describes the phenomena of lenition and fortition expressed as segmental changes in lexical units in Spanish. Furthermore, this thesis aims to experimentally study subsegmental changes (namely, changes in phonetic detail) arising from the specific gestural dynamics that characterize the word-initial position (compared to the syllable-initial position) and stressed syllables (compared to

unstressed). To do so, a corpus has been recorded with six Mexican participants, which includes articulographic, aerodynamic (nasal air flow and pressure), and acoustic data.

Strengthening due to position (word-initial) expresses itself as a slower speed in executing articulatory movements, especially during the consonant constriction phase and, to a lesser degree, during the release phase. Additionally, there is a lower degree of time overlap between the different articulatory movements that comprise oral gestures (jaw-tongue tip and jaw-lips). These results support the idea that strong positions are less prone to coarticulation, and also allow us to verify that this idea is also valid in terms of intra-gestural coordination. Furthermore, the magnitude of gestures has been seen to increase when, immediately after, another gesture is made that modifies the same tract variable in a different direction. For example, the height of the jaw and the height of the tongue tip have higher values in the word-initial position as a result of strengthening and due to the fact that the following vowel also presents gesture specification that involves the jaw and the tongue. Additionally, in the case of fricatives, spatial strengthening in the case of tongue tip movements causes more pronounced narrowing and an increase in the relative intensity of friction noise. On the other hand, and on the contrary to what happens with the jaw and tongue tip, lip aperture in the case of labial consonants, and the velum lowering gesture for nasals have a lesser magnitude in strong positions, due to the fact that the vowels that followed them ([a], [i]) were not specified for labial or velar gestures.

With regard to the acoustic measurements of consonants in the word-initial position, no significant modification in segmental duration was observed. There is also no variation in the VOT (Voice Onset Time) of voiceless stops, unlike in results obtained for other languages studied previously. This divergence is justified by the fact that Spanish is a language that bases its laryngeal oppositions on the presence or absence of voicing, and not on aspiration. Therefore, voiceless consonants are the unmarked element and strengthening due to prosodic position cannot affect the glottal gesture, as there is none specified in the underlying form.

In terms of strengthening due to prominence (i.e. that affects stressed syllables), a decrease in the articulatory speed during the constriction phase and an increase in the release phase were observed. Furthermore, the release is characterized by longer articulatory duration, which is reflected in an increase of the acoustic durations of the consonant and vowel. The prolongation of the release movement is interpreted as a delay in beginning the subsequent consonant constriction gesture, which not only allows the vowel of the stressed syllable to last longer, but also to achieve greater opening, which results in emitting greater acoustic energy. Indeed, it is observed that the intensity of the consonants, in relation to that of the vowel, decreases as a result of the greater vowel intensity. On the other hand, and as with positional strengthening, no significant modification of the VOT was observed.

The study concludes with a comprehensive evaluation of the results and discusses their relevance to the study of what conditions the phonotactic constraints seen in Spanish words. There is also a discussion of psycholinguistic phenomena, such as lexical access. Additionally, applications are proposed to improve speech recognition

and synthesis systems by introducing the phonetic detail specific to consonants located in strong prosodic positions. Finally, possible future areas of research are set out, which involve the interaction between positional and contextual effects, and also establishing the degree to which each articulatory gesture contributes to certain aerodynamic parameters, such as oral pressure (parameters which would enrich the task dynamics model).

1. Introducción

Esta tesis es un estudio de los fenómenos de refuerzo consonántico, entendido como un cambio en el desarrollo espaciotemporal de los gestos articulatorios —*i. e.* como una modificación (normalmente un aumento) de su duración, de su magnitud, o de ambas— como consecuencia de la estructura prosódica en la que se insertan; en concreto, por su contacto con una frontera de constituyente prosódico o por su aparición en una unidad relativamente prominente con respecto a otras del mismo nivel.

Todas las lenguas poseen una estructura prosódica, que consiste en una organización de los elementos de contenido fonológico en una serie de unidades situadas en un nivel jerárquico progresivamente más alto. Así, los gestos articulatorios, de cuya combinación y abstracción resultan los distintos segmentos fonológicos, se organizan en sílabas, que se organizan en pies, y estos a su vez en palabras, las cuales se agrupan en constituyentes cada vez mayores hasta formar un enunciado. Además, cada unidad tiene un núcleo, que es su constituyente más prominente y que sirve para señalar la presencia de dicha unidad.

La estructura prosódica es relevante no solo por la organización del contenido fonológico, sino también porque sirve de interfaz con otros niveles lingüísticos, como el morfosintáctico o el semántico. Por ello es importante que los hablantes (y los oyentes) dispongan de claves que les permitan codificar (o decodificar) la jerarquía de los elementos que componen los enunciados. El refuerzo espaciotemporal de los gestos articulatorios situados en los extremos de los distintos constituyentes supone una marca fonética que pone de relieve la frontera entre unidades, del mismo modo que el refuerzo asociado a las posiciones prominentes ayuda a indicar el número de constituyentes que hay en un determinado nivel jerárquico de esa estructura. La otra cara inevitable de la moneda, el debilitamiento gestual en las posiciones centrales de un constituyente, señala la continuidad del mismo y, por ello, también proporciona información —a su manera— acerca de la estructura prosódica.

Los fenómenos de refuerzo y debilitamiento están, pues, ligados a determinadas posiciones prosódicas. Su desencadenamiento no depende de los gestos o segmentos que haya alrededor, como sucede, por ejemplo, con otro tipo de fenómenos, los de coarticulación y asimilación (o también los de disimilación). Por ello es importante distinguir las nociones de *posición* (prosódica) y *contexto* (segmental), porque, si bien en ocasiones ambos tipos de fenómeno pueden interactuar, conviene reconocer la contribución particular de uno y otro al resultado final. Por ejemplo, los cambios por asimilación suelen ser más frecuentes en posiciones prosódicas propensas al debilitamiento, ya que en ellas se produce un mayor solapamiento entre gestos articulatorios contiguos, pero, en última instancia, el devenir de la asimilación depende de la identidad de los gestos que entran en contacto.

Aunque a veces un fenómeno de asimilación puede producir un resultado semejante a otro de refuerzo o debilitamiento, si discernimos apropiadamente los desencadenantes del cambio, podemos entender que se trata de procesos diferentes. Lo mismo se puede

decir de los casos en que un resultado similar puede ser el fruto de un debilitamiento o de un refuerzo. La determinación de las circunstancias que provocan el cambio es fundamental para clasificar dicho cambio adecuadamente.

Los cambios en la magnitud y en la duración gestuales asociados a los fenómenos de refuerzo (y debilitamiento) pueden manifestarse como pequeñas alteraciones en el detalle de la realización fonética, sin que estas modifiquen la identidad del segmento implicado. Se trata de variaciones subsegmentales que solo pueden analizarse mediante las técnicas instrumentales apropiadas. Las lenguas estudiadas hasta ahora con estos medios indican una serie de tendencias comunes, como características generales de los procesos de refuerzo, pero cada idioma presenta, igualmente, ciertas particularidades. No en vano, aunque se trate de modificaciones fonéticas, estas actúan sobre la base de los elementos de contenido fonológico y dependen de la manera en que estos se organizan prosódicamente, aspectos que sí son peculiares de cada sistema.

Por otro lado, en ocasiones los efectos de la posición prosódica sobrepasan la mera variación subsegmental y se ponen de manifiesto mediante alternancias alofónicas en un proceso de variación sincrónica, o como cambios fonológicos consolidados diacrónicamente. En estos casos, con mayor motivo, la regulación de las alternancias se integra en el sistema fonológico de la lengua y, si se produce un cambio de un fonema por otro, esta modificación se refleja en la forma léxica subyacente. Estos cambios alofónicos y fonémicos condicionan el inventario de segmentos que habitualmente suele aparecer en una posición u otra, en el seno de una sílaba o de una palabra.

En el transcurso de la historia de la fonología, los cambios segmentales se describieron antes que los subsegmentales, ya que los primeros son más fácilmente observables que estos últimos. Con el desarrollo de la lingüística histórica, los investigadores prestaron atención, especialmente, a los cambios diacrónicos que sufrían los segmentos en determinadas posiciones. Se encontró que ciertas posiciones son más propicias a rutas evolutivas cuyo último paso consiste en el borrado de un segmento ($X > \emptyset$), lo que les valió el calificativo de posiciones débiles². Por extensión, recibe la etiqueta de debilitamiento o *lenición* todo proceso típico de una posición débil, aunque no necesariamente conduzca a la eliminación de un segmento.

Por contraste, se reconoce la existencia de ciertas posiciones fuertes, más resistentes a los tipos de procesos que hemos llamado de lenición. No quiere decir que no puedan sufrir dichos procesos pero, en un determinado corte sincrónico de una lengua, la aparición de un fenómeno de lenición en una posición fuerte implica necesariamente que las posiciones débiles también se ven afectadas. En este sentido, la fuerza (o la debilidad) es una propiedad relativa; es decir, hay posiciones *más débiles* y posiciones *más fuertes* que otras. Además, dichas posiciones fuertes pueden desencadenar otros tipos de fenómenos que les son propios, y que se engloban bajo la denominación común de refuerzo o *fortición*.

² Esta definición de debilitamiento con apelación a las rutas evolutivas y al borrado fue consagrada por Hyman (1975, p. 165), quien, a partir de una comunicación personal de Theo Vennemann, declara lo siguiente: «a segment X is said to be weaker than a segment Y if Y goes through an X stage on its way to zero».

En esta tesis se reserva el uso de los términos *lenición* y *fortición* para hacer referencia a los fenómenos de debilitamiento y refuerzo, respectivamente, que van más allá de lo meramente subsegmental y se manifiestan como cambios segmentales en las palabras.

Sin embargo, son muchas las definiciones y matices que estos vocablos han recibido a lo largo de la historia de las ideas lingüísticas³. Merece especial atención el concepto que se tiene de estos procesos en el marco de la fonología natural, una escuela en la que tales nociones constituyen recurrentemente un foco de investigación (Donegan, 1978; Donegan y Stampe, 1979; Dressler, 1984; Goman, 1979; Stampe, 1973). La lenición se ve como una reducción articulatoria encaminada a facilitar la pronunciación, desde la perspectiva del hablante, mientras que la fortición conduce a un aumento de la perceptibilidad y resulta, por ello, beneficiosa, sobre todo, para el oyente. Como se ha explicado antes, en esta tesis se adopta la idea de que tanto la fortición como la lenición, en tanto que marcadoras de la estructura prosódica, resultan informativas para el oyente, cada una a su manera: el refuerzo se asocia con las posiciones de frontera de los constituyentes prosódicos, mientras que el debilitamiento denota la continuidad de estos.

Por otro lado, para la escuela natural, las leniciones son procesos sintagmáticos, sensibles al contexto; su nivel de actuación es la cadena de habla, y tienen carácter asimilatorio. Por su parte, las forticiones son siempre procesos paradigmáticos de naturaleza disimilatoria que contribuyen a la consolidación de un segmento frente a los demás del inventario de una lengua. En cambio, en esta tesis se defiende que ambos tipos de procesos están desencadenados por la posición prosódica que ocupa el segmento afectado, no por las relaciones que este contrae con su contexto. Aunque aquí se admite que, en ocasiones, una lenición puede ir acompañada por una asimilación, se insiste en la conveniencia de reconocer la diferente naturaleza de estos fenómenos. En cuanto a las características del refuerzo, se considera que sus efectos son parcialmente diferentes según dicho refuerzo esté desencadenado por la posición o por la prominencia prosódicas. Tan solo en este segundo caso (*vid.* § 3.2) se observan efectos de hiperarticulación que tengan como consecuencia el aumento de la distancia paradigmática entre los distintos segmentos del sistema fonológico, en consonancia con el concepto naturalista de fortición.

Por extensión metonímica, se suele hablar de la debilidad o de la fuerza no solo como propiedades de ciertas posiciones, sino también como atributos de los segmentos que resultan de una lenición o de una fortición, respectivamente. Sin embargo, esta idea será matizada a lo largo de este trabajo (*vid.* p. ej. § 3.3.1).

Ahora bien, como se ha dicho antes, los estudios sobre lenición y fortición han ido cambiando notablemente de enfoque y criterios en el transcurso de la historia de la fonología. Hasta época relativamente reciente se ha considerado que los resultados de dichos procesos eran segmentos diferenciables de los segmentos originales, y que tal diferenciación se daba necesariamente de un modo discreto. Se diferenciaban de otros fonemas en el caso de los cambios diacrónicos (en los estudios de corte historicista) o

³ Para un panorama de la evolución del significado de estos términos a través de los distintos modelos fonológicos, *vid.* Honeybone (2008).

de otros alófonos en el caso de las alternancias sincrónicas. En cambio, en las últimas décadas, numerosas corrientes fonológicas han ido abriéndose a la idea de que hay casos en los que el detalle fonético puede estar sujeto a una variación interlingüística sistemática y que, por tanto, también forma parte de la competencia lingüística de los hablantes y merece ser estudiado por la fonología. Esto, unido al desarrollo de la experimentación instrumental en el ámbito de los estudios fónicos, ha permitido investigar el debilitamiento y el refuerzo desde una perspectiva de variación continua (en lugar de discreta), de carácter subsegmental (con matices tan leves que no pueden reflejarse de manera apropiada en los sistemas de transcripción fonética, ni siquiera con una transcripción estrecha). Tales diferencias, no obstante, aun siendo pequeñas, pueden tener un efecto sobre la percepción y constituirse en clave para identificar la fuerza relativa de la posición que ocupan esos segmentos, y así permitir al oyente un más fácil reconocimiento de la estructura prosódica del enunciado.

Por otro lado, los estudios tradicionales han vinculado claramente los fenómenos de lenición y fortición a la fonología de la palabra, mientras que el desarrollo de la llamada fonología prosódica ha permitido comprobar que hay fenómenos similares de refuerzo tanto por debajo como por encima de dicho nivel.

El descubrimiento del refuerzo posléxico ha inclinado la balanza considerablemente en las últimas décadas hacia el estudio de dicho fenómeno en los constituyentes prosódicos de nivel superior, algo comprensible ya que estos habían sido ignorados en todas las investigaciones tradicionales sobre este tema. En cambio, hay un número relativamente escaso de estudios que hayan aplicado los nuevos enfoques, preocupados por el detalle fonético y basados en la experimentación instrumental, a los casos de refuerzo que ocurren en el seno de la palabra. Desde luego, en español hace falta una investigación profunda al respecto, y ese es el objetivo principal de esta tesis.

En el capítulo 2 se describen los distintos constituyentes que integran la jerarquía prosódica, sus condiciones de formación y sus patrones de prominencia. Además, se explica la importancia de la estructura prosódica en el diseño global de una lengua, su relaciones con otros niveles lingüísticos y su repercusión sobre fenómenos como el acceso al léxico.

El capítulo 3 revisa los efectos de la posición y la prominencia prosódicas sobre la dinámica espaciotemporal de los distintos gestos articulatorios, según se han descrito en diversos trabajos sobre distintas lenguas del mundo.

La fortición y la lenición segmental en español es el tema que centra el capítulo 4.

A partir del capítulo 5, se presenta el estudio del refuerzo subsegmental en español. En él se explica la metodología de recolección de un corpus de habla con datos articulográficos, aerodinámicos y acústicos, y se describen los pasos que se han dado para el procesamiento de los datos. El capítulo 6 se corresponde con la presentación de los resultados.

Finalmente, el capítulo 7 aborda la discusión de los resultados y expone las conclusiones de este trabajo.

2. La estructura prosódica en la arquitectura de la lengua

La prosodia se puede definir como el agrupamiento y la prominencia relativa de los elementos fonológicos (Pierrehumbert, 1999) y supone, por tanto, una estructuración de dichos elementos en distintos niveles, cada uno de los cuales puede ser el dominio de aplicación de distintos fenómenos o rasgos (*vid.* p. ej. Fox, 2000)⁴.

Si bien los niveles prosódicos más bajos (sílabas, pie) se constituyen en términos puramente fonológicos, las unidades de niveles más altos se forman atendiendo a criterios que tienen en cuenta la información provista por otros niveles lingüísticos, a saber, la morfología, la sintaxis, la semántica y la pragmática. Dicho esto, la mayoría de los autores coinciden en señalar que, aunque la prosodia sirve como interfaz entre la fonología y otros niveles de análisis, sin embargo, las unidades prosódicas no son necesariamente isomórficas con las unidades morfosintácticas en cuya información se basan, puesto que sus reglas de formación siguen siendo fonológicas (Nespor y Vogel, 1994 [1986], y Shattuck-Hufnagel y Turk, 1996, entre otros, aportan abundantes ejemplos al respecto)⁵. En cualquier caso, establecen una especie de puente que permite al componente fonológico tener acceso a cierta información de otros niveles (volveremos a hablar de esto en la introducción del capítulo 3).

Por otro lado, la relación entre sintaxis y prosodia no solo no es isomórfica, sino tampoco biunívoca. Fougeron (1999b) señala que una misma estructura sintáctica se puede corresponder con más de una estructura prosódica, pues esta última está sujeta a variación en función de determinados factores, tales como el tempo —que puede modificarse, entre otras cosas, dependiendo del estilo de habla o de lectura, o del estado de ánimo del locutor— o la tendencia rítmica a mantener duraciones similares entre constituyentes del mismo nivel. También Selkirk (2000) alerta sobre el conflicto rítmico generado, por ejemplo, por una secuencia de varias frases fonológicas constituidas por dos frases acentuales y después una con solo una frase acentual⁶. Existe la posibilidad de que se produzca una reestructuración y la frase más pequeña se integre en la frase fonológica contigua, que pasaría a tener tres frases acentuales. Como cualquiera de las dos opciones altera el ritmo de la secuencia (ninguna es menos mala que la otra), se produce una tensión que genera variabilidad en la solución prosódica alcanzada. En cualquier caso, la estructura sintáctica no cambia. No obstante, concluye Fougeron, no es una cuestión de actuación, sino que la estructura prosódica se relaciona directamente con la competencia lingüística (en términos chomskianos).

⁴ Tradicionalmente, muchos autores han definido la prosodia más bien en función de esos rasgos: sería prosodia toda variación en la frecuencia fundamental (f_0), en la duración, en la amplitud o en la cualidad espectral siempre que no pueda predecirse por la identidad de los segmentos o de sus contextos; *vid.* p. ej. Vaissière (2006) como ilustración de esta visión.

⁵ Sin embargo, *vid.* Kaisse (1985) para una visión diferente.

⁶ Aquí, «frase» se corresponde con la idea inglesa de *phrase*, que equivale a nuestro sintagma (no al sentido tradicional de frase como oración, en español). Sin embargo, las unidades prosódicas se rigen fonológica y no sintácticamente. Además, a diferencia de ellas, los «sintagmas» de la sintaxis son estrictamente binarios y endocéntricos. Por este motivo, y porque es una traducción consolidada por muchos autores (*vid.* p. ej. la edición de Nespor y Vogel, 1994 [1986], que manejo), he decidido mantener esta denominación aun a riesgo de que parezca un anglicismo.

La descripción de la estructura prosódica de un determinado enunciado incluye la delimitación de los constituyentes de que consta en cada uno de los distintos niveles de la jerarquía (*vid.* § 2.1) y el establecimiento de las relaciones de prominencia entre constituyentes del mismo nivel (*vid.* § 2.2). Además, como se analizará con detalle en el capítulo 3, existen índices fonéticos que ayudan a identificar los constituyentes prosódicos, por medio del marcado de sus fronteras (función demarcativa) o de sus núcleos o elementos prominentes (función culminativa). Dado que los distintos constituyentes pueden funcionar como ámbito de aplicación de una serie de reglas y fenómenos, tales fenómenos se convierten también en marcas indirectas de las fronteras prosódicas, pues más bien lo que señalan es la extensión de los constituyentes.

Los constituyentes como ámbito de aplicación de fenómenos fonológicos

Delais-Roussarie y Fougeron (2004) advierten que, al igual que se hace en sintaxis, es conveniente que los estudios sobre fonología distingan dos tipos de unidades, a saber, constituyentes y dominios o ámbitos, y debe tenerse en cuenta que, aunque en algún momento estos pueden coincidir, no necesariamente es así (suponer lo contrario llevaría a un razonamiento circular en muchos casos). Un constituyente se caracteriza por una estructura interna de carácter jerárquico con un núcleo. En fonología prosódica, por ejemplo, cada constituyente tiene un núcleo, que es, de todas las unidades que lo componen y que corresponden a un nivel inmediatamente inferior, la más prominente de todas, según las reglas paramétricas de asignación de prominencia de cada lengua (*vid.* § 2.2; y también Nespor y Vogel, 1994 [1986], para un análisis de distintas lenguas). Nótese, no obstante, que, a diferencia de los constituyentes sintácticos, los prosódicos no son endocéntricos, pues su núcleo es siempre una unidad de otra categoría (en cada caso, del nivel inmediatamente inferior).

En cambio, un dominio se caracteriza por ser el ámbito de aplicación de una determinada regla, de modo que tal aplicación aporta una propiedad que dota de cohesión al dominio, indicando su extensión (o la ubicación de sus fronteras).

La mayoría de los autores que trabajan en fonología prosódica (muchos de los cuales se irán mencionando en los próximos apartados) hipotetizan que los constituyentes prosódicos, además de seguir unas reglas de formación basadas en prominencias relativas, que es lo que los caracteriza como tales constituyentes, son, por otro lado, el dominio de aplicación de diversos fenómenos fonológicos. Tal visión se adopta también en esta tesis y, aun recogiendo la advertencia de distinguir los conceptos de constituyente y dominio (como tales conceptos), se considera que los constituyentes prosódicos suelen ser dominio de ciertos fenómenos⁷. Además, siguiendo a Nespor y

⁷ Delais-Roussarie y Fougeron (2004) sostienen que el dominio de aplicación de algunos fenómenos del francés, como la *liaison* y el *enchaînement*, no coincide con ningún constituyente prosódico —Selkirk (1986) propone la *frase fonológica menor*, *vid.* § 2.1.4—, pues hay casos en que, aparentemente, la aplicación rebasa las fronteras de los constituyentes candidatos, y otros casos, en cambio, en los que no se produce el fenómeno, ni siquiera en el interior de los mismos constituyentes. Sin querer entrar aquí a valorar en profundidad dichos fenómenos, ni los datos manejados por estas autoras, sí cabe apuntar que hay circunstancias bajo las cuales se pueden producir reestructuraciones de los constituyentes prosódicos, modificando sus fronteras (Nespor y Vogel, 1994 [1986]), y esto podría dar la falsa impresión de que un fenómeno ha rebasado el ámbito de un constituyente. Por otro lado, es ampliamente sabido que la

Vogel (1994 [1986]), la especificación de dichos dominios puede venir indicada de tres formas diferentes:

(1) Modos de actuación de los dominios prosódicos

- a) *Ámbito de intervalo*. Las condiciones que permiten la aparición de un fenómeno deben darse en el seno de un determinado tipo de constituyente prosódico. Por ejemplo, en la variedad toscana del italiano, una /p/, /t/, /k/ en contexto intervocálico se pronuncian [ɸ], [θ], [h] solo si la secuencia V_V se da dentro de la misma frase entonativa.
- b) *Ámbito de linde*. Las condiciones deben darse en el extremo inicial o en el extremo final del constituyente prosódico. Por ejemplo, en inglés, /p/, /t/, /k/ se pronuncian aspiradas ([p^h], [t^h], [k^h]) cuando aparecen en el extremo inicial de un pie.
- c) *Ámbito de juntura*. Las condiciones deben darse a través de la frontera entre dos constituyentes de un determinado nivel, siempre que ambos estén incluidos en un mismo constituyente de otro nivel superior. Por ejemplo, en español, una nasal en final de sílaba se asimila al punto de articulación de la consonante inicial de la sílaba siguiente (la secuencia de sílabas puede darse en cualquier posición dentro de un mismo enunciado fonológico). En cambio, en inglés, según Nespor y Vogel (1994 [1986]), esta asimilación ocurre solo si ambas sílabas se encuentran dentro del mismo pie.

Así, toda una serie de fenómenos hacen las veces de marcas indirectas de la estructura prosódica, al poner en evidencia la frontera o la extensión de los constituyentes que sirven para definir su ámbito de aplicación. Por ejemplo, el ámbito de las alternancias alofónicas suele estar determinado por la palabra prosódica o un constituyente de menor nivel que esta (*vid.* capítulo 4), aunque no siempre es el caso, como en el ejemplo de la *gorgia toscana* que se acaba de describir en (1):a. Si se conoce el ámbito regulador de dichas alternancias, la aparición de un alófono u otro puede ser un índice de la estructura prosódica. De manera análoga, los fenómenos de armonía vocálica suelen afectar a toda una palabra prosódica, y sirven para marcar su extensión, si bien es cierto que hay casos en que el patrón armónico puede verse interrumpido por diversas causas (Van der Hulst y Van de Weijer, 1995). En el capítulo 3 y siguientes

aplicación de algunos fenómenos puede ser variable (por ejemplo, condicionada por el estilo, aunque no solo), sin que por ello se deba cuestionar que su ámbito de aplicación, cuando se aplica, es uno concreto. Por ejemplo, en algunas variedades del español se puede producir la debucalización de las fricativas /f/, /θ/, /s/, /x/, entre otros segmentos, siempre que se encuentren en posición de coda. Así, *paz* se puede pronunciar *pa[h]*, o *asma*, *a[h]ma*. Los mismos individuos pueden presentar variantes sin debucalizar, pero indudablemente las condiciones de aplicación —en caso de aplicarse el fenómeno— implican la aparición en coda silábica. Con respecto a la *liaison*, Laks, Calderone y Celata (2014) hacen un estudio de corpus y concluyen, mediante un análisis inspirado en la fonología de ejemplares y en la fonología basada en el uso, que no se trata de un fenómeno homogéneo, sino que se presta a variación. Las secuencias más frecuentes muestran *liaison* independientemente de los distintos parámetros estilísticos y sociolingüísticos, lo que se interpreta como prueba de que los hablantes memorizan cada una de tales secuencias como una sola unidad. En cambio, la generalización de la *liaison* a las secuencias menos frecuentes varía en función de características de los hablantes como la edad o el nivel educativo.

volveremos a esta cuestión de las marcas fonéticas de los distintos constituyentes, mediante variaciones segmentales y subsegmentales basadas en el refuerzo (o a veces el debilitamiento) articulatorio.

Dicho esto, pasemos a ver cuáles son esos constituyentes, qué principios sigue su formación, y cómo se organizan en el conjunto de la jerarquía prosódica.

2.1. La jerarquía de constituyentes prosódicos

Todas las lenguas presentan algún tipo de estructura prosódica, es decir, una organización sintagmática y de carácter jerárquico que agrupa los elementos de contenido fonológico en distintos constituyentes de niveles sucesivamente mayores.

La idea de una jerarquía prosódica está ya esbozada de forma precursora en Hockett (1955), quien habla de macrosegmento (equivalente a un enunciado desprovisto de entonación), microsegmento (más o menos coincidente con la palabra fonológica, *vid.* § 2.1.3), sílaba, núcleo y márgenes, y segmentos. Una propuesta parecida es la de Haugen (1956), y algo más elaborada ya la de McMahon (1967), con discurso, oración, cláusula, frase, pie, sílaba y segmento. La base de las actuales propuestas de jerarquía prosódica se encuentra, sin embargo, en los trabajos de Selkirk (1978, 1980, 1981, 1986) y en posteriores reformulaciones de la misma autora. A partir de esos trabajos, otros investigadores han defendido sus propias jerarquías, con más o menos diferencias con respecto a las de Selkirk (Beckman y Pierrehumbert, 1986; Nespor y Vogel, 1994 [1986]); *vid.* Shattuck-Hufnagel y Turk (1996) para una panorámica comparativa de los principales tipos de propuestas de jerarquía prosódica.

Existen varias dificultades a la hora de presentar un inventario definitivo de constituyentes prosódicos. Por un lado, el nombre que recibe un mismo constituyente puede cambiar de unos autores a otros, pero las explicaciones proporcionadas no siempre permiten garantizar la equivalencia entre denominaciones. Por otro lado, la propia diversidad de intereses de los investigadores por unos asuntos u otros, o por unas determinadas lenguas, hace que los procedimientos analíticos desarrollados no siempre sean comparables.

En esta tesis se asume una jerarquía prosódica compuesta por seis tipos de constituyentes (cuyas características se detallan en los próximos subapartados).

(2) Jerarquía de unidades prosódicas (de mayor a menor)

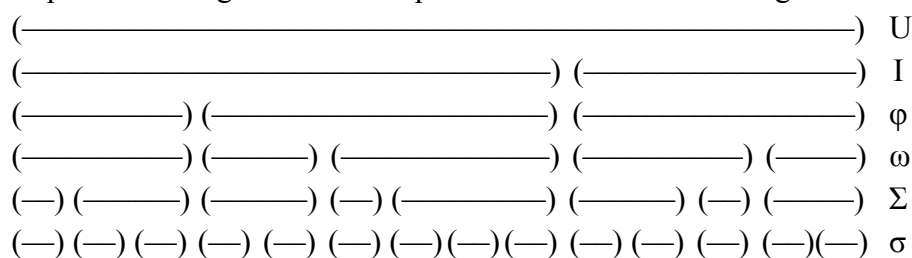
enunciado fonológico	U
frase entonativa	I
frase fonológica	φ
palabra fonológica	ω
pie	Σ
sílaba	σ

Con frecuencia, se aludirá de forma colectiva a los tres niveles más bajos de la jerarquía expuesta en (2) como *niveles inferiores* y a los tres más altos, como *niveles superiores* o también como *niveles frásticos* o *frasales*. Junto a cada nivel, se añade el

símbolo que se usará como abreviatura del mismo a lo largo de este trabajo. Estos símbolos coinciden con la propuesta original —i. e. de la versión en inglés— de Nespor y Vogel (1994 [1986]). En la discusión específica de cada nivel se aportarán otras abreviaturas utilizadas por diferentes autores, para facilitar la comparación.

Según la Hipótesis de Estratificación Rigurosa (*Strict Layer Hypothesis*, en inglés), esbozada en Selkirk (1981), aunque no recibió ese nombre hasta Selkirk (1984b), los distintos constituyentes prosódicos se encuentran anidados unos en otros, de modo que un constituyente consta única y exclusivamente de uno o más constituyentes del nivel inmediatamente inferior. Por ejemplo, replicando el diagrama de Selkirk (1986, p. 384), obtenemos un esquema como el de (3):

(3) Representación gráfica de la Hipótesis de Estratificación Rigurosa



En versiones posteriores (Selkirk, 1996), relaja un poco esta condición⁸, que desglosa en cuatro propiedades diferentes, expuestas en (4). Dos de estas propiedades (*a*, *b*) se deben respetar siempre, mientras que las otras dos (*c*, *d*) pueden verse infringidas en determinadas ocasiones. Veremos ejemplos de algunas de estas infracciones al hilo de la explicación de los distintos constituyentes.

(4) Propiedades de la jerarquía prosódica

- Estratificación*: Ninguna unidad prosódica debe dominar a otra de nivel superior.
- Nuclearidad*: Toda unidad prosódica para la que exista un nivel más bajo en la jerarquía debe dominar al menos una unidad de su nivel inmediatamente inferior.
- Exhaustividad*: Ninguna unidad prosódica debe dominar de forma directa a otra de nivel inferior si esta no es inmediatamente inferior.
- No recursividad*: Ninguna unidad prosódica debe dominar a otra de su mismo nivel.

Las unidades de cada nivel se construyen siguiendo las reglas fonológicas apropiadas según cada lengua. Esas reglas prescriben cuál es el dominio en el que se forman los distintos constituyentes. Por ejemplo, la σ y el Σ , como niveles subléxicos que son, suelen constituirse en el seno de las unidades léxicas⁹. (La palabra fonológica, ω , suele

⁸ Vid. De Pijper y Sanderman (1994) para algunos indicios experimentales que también aconsejan una interpretación no demasiado estricta de la Hipótesis de Estratificación Rigurosa.

⁹ En algunas lenguas el silabeo puede producirse de manera individual en cada morfema y permanecer inalterado con la adición de otros morfemas. En cambio, en otras lenguas se constata un silabeo cíclico,

coincidir con la palabra morfosintáctica, aunque a veces su extensión es mayor o menor que la de esta, *vid.* § 2.1.3.) Ahora bien, existen *procesos de reestructuración* que tienen el efecto de reorganizar la constitución de las unidades de un determinado nivel, en el seno de un dominio definido por un nivel superior. Uno de los casos más conocidos lo representan los procesos de resilabeo, que modifican la constitución de las sílabas de modo que estas pueden traspasar las fronteras léxicas. Estas reestructuraciones se producen bajo condiciones particulares de cada lengua, en las lenguas que las permiten, de modo que no se realiza de la misma manera el resilabeo en español que en francés, por ejemplo (*vid.* § 2.1.1.5). Por otro lado, aunque la estructura silábica definitiva en el enunciado es la resultante del resilabeo, es necesario seguir suponiendo el proceso de silabeo original, ya que hay procesos cuya aplicación solo se justifica dada la estructura silábica previa al resilabeo (p. ej. *los amigos* [lo.ha.'mi.ʝoh], con resilabeo a través de frontera de palabra morfosintáctica, pero con el sonido [h] propio de la debucalización de fricativas en final de sílaba —*vid.* § 4.1.1—, que delata un silabeo previo [loh.a.'mi.ʝoh]).

Dado que hay toda una suerte de fenómenos que requieren aludir a las distintas posiciones dentro de un constituyente (como, p. ej., el final de sílaba que acabamos de ver), resulta conveniente establecer una convención notacional para indicar dichas posiciones. Los recursos formales utilizados por distintos investigadores son diferentes y a veces pueden incluso provocar ciertas confusiones, aunque hay que reconocer que no es lo habitual. En todo caso, quede claro que, como tales convenciones, en principio, no hay ningún sistema de notación formal que sea mejor que otro; simplemente debe establecerse de forma clara desde un inicio.

En esta tesis, de manera análoga a como se hace en Nespor y Vogel (1994 [1986]), se usarán corchetes para indicar los límites inicial o final de un constituyente, y se anotará en subíndice el tipo de unidad mediante el símbolo correspondiente, según se ha expuesto en (2). Así, por ejemplo, $[\sigma_]$ representa la posición inicial de sílaba, $[_]\Sigma$ indica la posición final de pie, y $[\dots_]\omega$ denota una posición media o interior de la palabra fonológica. Además, se propone usar la letra griega π para representar, en general, una unidad prosódica de cualquier nivel, como en la notación $][\pi$ para indicar frontera de constituyente (*vid.* § 3.1 para una justificación del símbolo π). Con este mismo sentido, el general de frontera, muchos autores utilizan #. Sin embargo, a lo largo de la historia de la fonología, este símbolo se ha empleado para expresar distintos significados, como el límite de palabra morfosintáctica (Chomsky y Halle, 1976 [1968]; su modelo será referenciado como SPE de aquí en adelante), o la presencia de una pausa, y por ello se podría considerar ambiguo. Para indicar palabra morfosintáctica, Nespor y Vogel usan la letra Q. En cuanto a la pausa, dado que es una característica ligada a la frontera de determinados constituyentes (la frase entonativa y el enunciado fonológico), *vid.* § 3, parece preferible marcarla directamente como $][\iota$ o como $][\cup$. De manera alternativa a los corchetes, una práctica habitual en muchos trabajos de fonología prosódica es escribir la abreviatura del tipo de unidad y, en subíndice, una aclaración de

de modo que la constitución silábica de una base morfológica puede verse modificada al añadirle otro morfema. En este sentido, cada entrada léxica presenta su propia estructuración en sílabas.

la posición de que se trata: i(nicial), m(edial), f(inal). Por ejemplo, el inicio de palabra aquí se representa como [_ω pero otros autores escribirían W_i para indicar esa posición.

2.1.1. La sílaba (σ)

Resulta, desde luego, imposible hacer una revisión exhaustiva de todo lo que se ha dicho acerca de la sílaba. En los próximos apartados se tratarán tan solo los aspectos más relevantes para el objetivo de esta tesis, pero el lector puede remitirse —para mayor información— a algunos de los trabajos que han abordado el tema desde múltiples perspectivas (Blevins, 1995; Bosch, 2011; Colina, 2012; Féry y Van de Vijver, 2003; James W. Harris, 1991 [1983]; Van der Hulst y Ritter, 1999b; y los capítulos dedicados a la sílaba de obras como Gil Fernández, 2000a; Martínez-Gil y Colina, 2006), entre otros muchos posibles.

Símbolos y abreviaturas de la sílaba

Algunas formas de referirse a la sílaba en distintos trabajos incluyen las siguientes: Syll., Syl., Síl., S, σ . Aquí se unificarán estas notaciones bajo la forma simbólica σ .

2.1.1.1. La sílaba como constituyente

Podría parecer que la sílaba es una de las unidades fonológicas cuya existencia resulta menos controvertida. Sin embargo, no todos los autores aceptan la sílaba como un primitivo lingüístico, sino que la consideran una estructura emergente a partir de ciertas características propias de las secuencias que se establecen entre los elementos de contenido fonológico (*vid.* p. ej. Dziubalska-Kołaczyk, 1995; Vennemann, 1994; *apud* Bertinetto, Agonigi, Cioni, García Lecumberri y González Parra, 1998). En esta línea de pensamiento, Ohala y Kawasaki-Fukumori (1997) defienden que la sílaba surge con carácter epifenoménico a partir de la necesidad de modular la señal acústica de tal forma que los segmentos sucesivos se diferencien nítidamente. Como se verá con más detalle en el § 2.1.1.3, una de las principales estrategias de modulación consiste en crear diferencias de abertura entre un segmento y el siguiente. Cuando la secuencia llega a un segmento de máxima abertura, como puede ser una vocal, y especialmente una /a/, necesariamente el elemento subsiguiente presentará una abertura menor. A la inversa, cuando se alcanza un mínimo de abertura, la secuencia continuará en la dirección que tiende hacia un nuevo máximo. El reconocimiento de estos ciclos da lugar a lo que interpretamos como agrupación silábica de los segmentos en torno a máximos locales de abertura.

Ahora bien, emergente o no, parece que la sílaba cobra carta de naturaleza como unidad prosódica y hay argumentos que apuntan a su realidad psicológica. Blevins (1995) aporta cuatro posibles argumentos que apoyan que la sílaba es una verdadera unidad fonológica.

Por un lado, ejerce como ámbito de aplicación de diversos fenómenos. Por ejemplo, la asignación de prominencias acentuales siempre recae sobre una sílaba: «in all true stress languages, the syllable appears to be the stress-bearing unit, that is, there is no contrast between tautosyllabic $\acute{v}v$ and $v\acute{v}$ » (*op. cit.*, p. 208), *vid.* también el § 2.2.

También en japonés, una lengua de tono léxico —e incluso tratándose de un sistema que muchos autores analizan como moraico (*vid.* § 2.1.1.7)—, dicho tono es una propiedad de toda la sílaba, no de la mora (Fox, 2000, p. 98).

En segundo lugar, hay multitud de fenómenos y alternancias alofónicas cuyo ámbito de aplicación solo puede justificarse de manera natural haciendo referencia a las fronteras (inicial o final) de una sílaba.

Por otro lado, son numerosos los juegos lingüísticos que se basan en la división y en la estructura silábicas, como el de encadenar palabras haciendo coincidir la última sílaba de una con la primera sílaba de la siguiente; o los lenguajes inventados, como el *verlan* (que suena como la expresión francesa *l'envers*, pero invirtiendo el orden de las sílabas) o el «lenguaje de la *p*», que consiste en intercalar entre sílaba y sílaba otra nueva integrada por la *p* y el núcleo vocálico de la sílaba precedente, que se reduplica (p. ej. *abreviaturas* se convierte en a-pa-bre-pe-via-pa-tu-pu-ras-pa).

Finalmente, dice Blevins, los hablantes nativos dividen las palabras en sílabas de una manera totalmente intuitiva. Esta opinión es compartida por Quilis (1993, pp. 360-361, nota 3) y por Marrero (2001, pp. 50-51), quien atribuye la capacidad de silabeo a «niños preescolares, personas analfabetas e incluso lesionados cerebrales».

Además de los argumentos de Blevins, otros autores han aportado otras pruebas a favor de la existencia de la sílaba. Por ejemplo, Quilis (*ibid.*), entre otras cosas, menciona que «las palabras cantadas se dividen en sílabas, nunca en fonemas». Cutler, Mehler, Norris y Segui (1986) resaltan la importante implicación de la unidad silábica en los *lapses linguae*. Por último, en el § 2.3.3.1 se verán más argumentos, relacionados con el papel de la sílaba en los procesos de acceso al léxico.

2.1.1.2. La estructura interna de la sílaba

Bosch (2011), recogiendo aportaciones de Blevins (1995), plantea un panorama de los distintos tipos de modelos que se han propuesto para la estructura interna de la sílaba, que se recogen de manera esquemática en (5)¹⁰.

(5) Modelos de la sílaba

- a) Modelos planos
- b) Modelos planos con intermediación de un nivel CV
- c) Modelos con estructura pero sin sílaba (solo ataque y rima por separado)
- d) Modelos con estructura ternaria
- e) Modelos con estructura jerárquica binaria (cuerpo y coda)
- f) Modelos con estructura jerárquica binaria (ataque y rima)

En los modelos planos, (a), la sílaba domina directamente todos los segmentos, al mismo nivel, sin estructura intermedia (Figura 2.1). Algunos modelos, del tipo (b), (Clements y Keyser, 1983; *vid.* también algunos de los capítulos incluidos en Brandão de Carvalho, Scheer y Ségéral, 2008) proponen que los segmentos están dominados por un nivel (en inglés, *tier*) intermedio, antes de unirse a la sílaba, y que ese nivel está

¹⁰ Para un tipo de propuesta totalmente distinta, *vid.* § 2.1.1.6.

formado por una hilera de elementos de dos tipos: C y V (consonantes y vocales), que determinan las combinaciones posibles de segmentos en función de sus propiedades melódicas (Figura 2.2). El modelo de Kaye, Lowenstamm y Vergnaud (1985), conocido como KLV por las iniciales de sus autores, es un representante del tipo (c) (Figura 2.3). En el caso de la estructura ternaria, (d), encontramos sílabas que incluyen ataque, núcleo y coda, con estos tres elementos al mismo nivel (*vid.* p. ej. Hockett, 1955) (Figura 2.4).

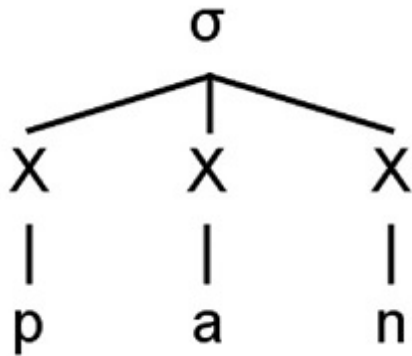


Figura 2.1. Modelo plano de la sílaba

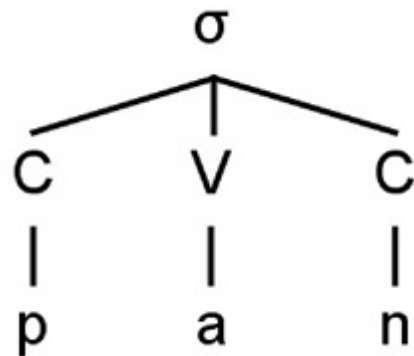


Figura 2.2. Modelo plano con nivel CV

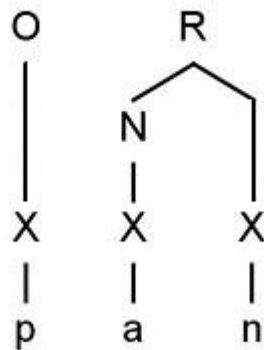


Figura 2.3. Modelo KLV

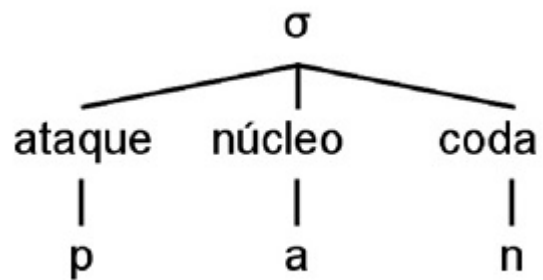


Figura 2.4. Modelo ternario

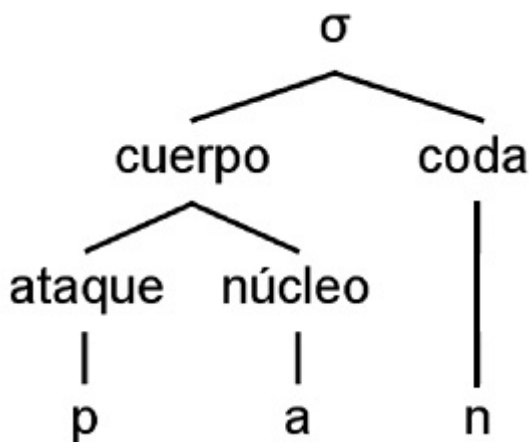


Figura 2.5. Modelo binario con cuerpo y coda

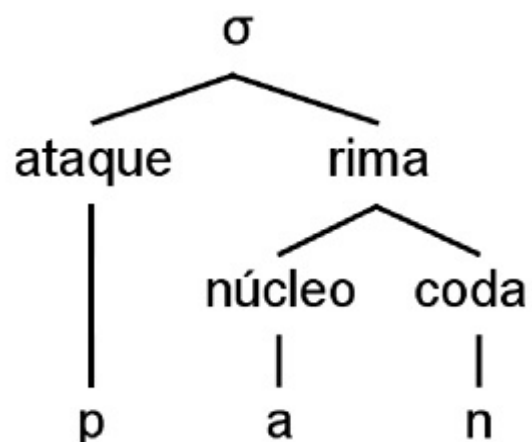


Figura 2.6. Modelo binario con ataque y rima

Lo más aceptado a día de hoy parece ser un modelo con estructura jerárquica binaria, aunque existen dos versiones de esta propuesta, (e) y (f). En la primera de ellas, el conjunto integrado por el ataque y el núcleo forma un constituyente intermedio, llamado cuerpo, que se une a la coda para dar lugar a la sílaba (Figura 2.5). En la segunda, son el núcleo y la coda los que forman un constituyente, la rima, hermano del ataque (Figura 2.6). Parece que cada uno de estos modelos resulta más adecuado para describir la estructura silábica de unas lenguas u otras: aunque el más extendido es el de ataque y rima, la opción de cuerpo y coda es la preferida para lenguas como el chino o el coreano, según defienden respectivamente Wang y Cheng (2008) y Yoon y Derwing (2001), ambos trabajos citados por Bosch (2011)¹¹.

Ahora bien, la propia Bosch se muestra escéptica ante las posibles pruebas de que exista una estructura interna de la sílaba, y hay autores que prefieren no pronunciarse acerca del carácter de constituyente del ataque, la coda, etc., aunque reconocen la utilidad descriptiva de estas etiquetas para referirse a los segmentos prevocálicos y posvocálicos, respectivamente (B. Hayes, 2009; Steriade, 1999). Además, un buen número de investigadores abogan por otro tipo de nivel subsilábico: la mora (μ), entre ellos B. Hayes (1989) y Hyman (1985), pero *vid.* la discusión acerca de la mora en el § 2.1.1.7.

Fox (2000, § 6.2.3-6.2.4) advierte que es necesario diferenciar rigurosamente lo que son verdaderas unidades de lo que son sus clases, funciones o propiedades. Por ejemplo, según él, la distinción entre C y V con respecto a las posiciones del esqueleto (esto es, de la secuencia de espacios temporales sucesivos que ocupan los segmentos) es una distinción de clases dentro de la unidad segmento. Por otro lado, sostiene que las etiquetas *ataque*, *rima*, *coda*, etc. no deben entenderse como constituyentes, esto es, unidades subsilábicas, sino que deben tratarse como *expansiones* de los segmentos que los forman, pues afirma que funcionalmente no se diferencian de dichos segmentos.

Ante esta diversidad de opiniones, Blevins (1995) sentencia: «co-occurrence restrictions between two segment positions within a syllable is evidence that the two positions form a constituent» (p. 212). Dado que en español los segmentos presentan este tipo de restricciones de coaparición¹², parece razonable apoyar la estructura más reconocida de ataque y rima, pero *vid.* también el § 2.1.1.6.

La estructura de la sílaba en español

La estructura jerárquica de los componentes subsilábicos en español puede describirse de la siguiente manera. La sílaba se divide en dos constituyentes, ataque¹³ y

¹¹ Según Bertinetto, Agonigi, Cioni, García Lecumberri y González Parra (1998), los resultados de algunos de sus experimentos psicolingüísticos podrían apuntar a ciertas características silábicas del español más bien propias de la estructura de cuerpo y coda, aunque los datos son contradictorios y los resultados de la tarea más importante de dichos experimentos apoyan la estructura de ataque y rima para esta lengua. A partir de esta constatación, y por comparación con los datos muchos más nítidos sobre el uso de ataque y rima en inglés, sugieren que la preferencia de cada idioma por un tipo de estructura silábica u otro puede prestarse a cierto nivel de gradación entre ambos extremos.

¹² La obra auspiciada por la Real Academia Española (2011) describe, en su capítulo 8 (pp. 283-354), los detalles de las combinaciones segmentales posibles en cada una de las partes internas de la sílaba.

¹³ En la tradición española el ataque también se ha llamado cabeza, *vid.* p. ej. Núñez Cedeño (2000 [1984]) y las referencias allí citadas.

rima, de los cuales el ataque es opcional. La rima, a su vez, se divide en cima y coda, de las cuales la coda es opcional. Los tres constituyentes (ataque, cima y coda) pueden incluir, cada uno, uno o dos segmentos; es decir, se asume que cada constituyente admite una ramificación binaria. La cima (en inglés, *peak*) es el único componente obligatorio de la sílaba, y contiene el segmento que ejerce de núcleo silábico, que en español es necesariamente una vocal. Aunque muchos modelos etiquetan el nudo del árbol correspondiente a la cima ya como *núcleo*, me parece conveniente distinguir estos dos conceptos, ya que la cima puede contener dos segmentos, pero solo uno de ellos es el núcleo.

Un ataque simple puede estar integrado por cualquier consonante del español, excepto la percusiva /r/ si se trata de un ataque inicial de palabra (*vid.* § 2.1.3.2 para una posible causa de esta exclusión). Un ataque complejo¹⁴ tiene como primer miembro una de las siguientes obstruyentes: /p/, /β/, /f/, /t/, /ð/, /k/, /ɣ/¹⁵; y como segundo miembro una /r/ o una /l/ (exceptuando las combinaciones *tl* y *dl*, que tan solo pueden ser tautosilábicas en determinados territorios hispanohablantes de influencia nahua).

El comportamiento de las codas consonánticas depende de si son simples o complejas, y si son o no también final de palabra. En general, una coda simple admite cualquiera de las consonantes /r/, /s/, /l/, /n/, /ð/, /θ/ (tanto en interior como en final de palabra); en interior, también tolera /p/, /β/, /f/, /t/, /k/, /ɣ/, pero estas consonantes en final de palabra —junto con la /x/— están mucho más restringidas y suelen aparecer, fundamentalmente, en préstamos. En ningún caso puede haber un fonema consonántico palatal en posición de coda. Las codas complejas presentan una mayor limitación, y suelen consistir en un primer elemento /p/, /β/, /ð/, /n/, /l/, /r/, /k/, seguido de /s/ (o de /f/ en préstamos como *surf* o *golf*). Son más habituales en interior que en final de palabra.

Mayor complejidad de análisis presentan las secuencias vocálicas, como reflejan por ejemplo varios de los trabajos compilados por Gil Fernández (2000a). Al igual que esta autora (Gil Fernández, 2000b), utilizaré aquí el término *paravocal* para referirme indistintamente a cualquier segmento vocálico no nuclear, englobando así las tradicionales denominaciones de semiconsonante (prenuclear) y semivocal (posnuclear). Mientras son auténticas paravocales, no parece que haya necesariamente más cualidades consonánticas en la posición prenuclear que en la posnuclear. Por otro lado, como vamos a ver, en ambos márgenes puede producirse una reasignación del segmento a una posición silábica propia de una consonante, para favorecer la presencia de un ataque lleno de contenido. En estos casos ya no se puede hablar de paravocal, sino que se convierten en consonantes de pleno derecho.

Las paravocales prenucleares deben ocupar la primera posición de la cima ramificada y, por tanto, son «hermanas» de la vocal nuclear. Se puede deducir que esta posición es la única que queda libre en la estructura, dado el límite de ramificación binaria de cada constituyente subsilábico, y a partir de la constatación empírica de sílabas que ya tienen ocupadas las dos posiciones de un ataque complejo y además presentan un diptongo

¹⁴ También se puede llamar compuesto, ramificado o binario.

¹⁵ En el § 4.3.1 se justificará el uso de los símbolos de las aproximantes espirantes para representar estos fonemas del español.

creciente (p. ej. *abriendo* [a'βri̯endo], *crueldad* [krue̯l'dað̃], *sobrio* ['soβri̯o]). No debe confundirse esta situación con la de aquellas palabras que aparentemente empiezan por un diptongo creciente (p. ej. *hielo*). La preferencia por sílabas con ataque obliga a analizar el segmento prenuclear como una verdadera consonante —[ʝelo]—, como demuestra el hecho de que esté sujeta a procesos típicos de consonante. Por ejemplo, no permite el resilabeo de la consonante final de la palabra anterior (*vid.* § 2.1.1.5) y se pronuncia como interrupta cuando le precede una interrupta homorgánica (*vid.* § 4.3.1): *con hielo* [kon'ʝ̩e.lo] y no *[ko.'ɲ̩e.lo]. Además, al no permitir el resilabeo, la consonante anterior sufre las variaciones alofónicas que le correspondan por su posición de coda, p. ej. *las hierbas* [laʃ.ʝ̩e.ɓas], con sonorización de la /s/ por ir seguida de consonante sonora. Por otro lado, la palatal adopta una pronunciación fricativa en las regiones que presentan esta variante (*vid.* § 4.3.4): [ʝ̟elo].

Como explica Alcoba (1999), en algunas formas flexivas de determinados verbos, una paravocal /i̯/, bien perteneciente a la raíz o a una desinencia, se consonantiza en [j̥] cuando, por algún motivo, debe pasar a ocupar el ataque de una sílaba. Esto puede ocurrir en posición inicial de palabra y seguida de vocal: [ω__V, o en interior de palabra en contexto intervocálico: [...V__V...]_ω, donde la paravocal presenta un valle de sonoridad (*vid.* § 2.1.1.3). La primera condición la cumplen los verbos de diptongación en posición inicial de palabra (p. ej. *yerro*, del verbo *errar*; *hierro*, del verbo *herrar*; *yergo*, del verbo *erguir*; *hiervo*, del verbo *hervir*). En cuanto a la segunda condición, hay dos supuestos. Por un lado, tenemos los verbos vocálicos de la 3.^a conjugación con paravocal subyacente (raíz en -u̯), p. ej. *concluye*, del verbo *concluir*; *disminuye*, del verbo *disminuir*, etc. Por otro lado, también cumplen este requisito los verbos vocálicos de la 2.^a y la 3.^a conjugación en las formas en que la vocal temática adopta el alomorfo *ie* y en la tercera persona del singular del pretérito perfecto simple, en que el morfema de tiempo, aspecto y modo (TAM) es *ió*: p. ej. formas como *cayeron*, *trayendo*, *creyera*, *sustituyó*, etc.

En el caso de las palabras que empiezan por *hue-* (p. ej. *huevo*, *hueso*, *hueco*, etc.), debemos suponer que el ataque está ocupado por la aproximante (de tipo espirante) velar /ɣ/ y que hay una paravocal /u̯/ en la cima: ['ɣ̟ue̯βo], ['ɣ̟ueso], ['ɣ̟ueko]. Por supuesto, la espirante puede pronunciarse como oclusiva en las posiciones y contextos adecuados (*vid.* § 4.3.1), p. ej. *con huevo* [kon'g̟ue̯βo]. También hay un verbo con alternancia entre vocal simple y diptongo en posición inicial de palabra: *huelo*, del verbo *oler*. El caso de la consonantización de la paravocal anterior /i̯/ era más fácil. Si suponemos que, por el mero hecho de ocupar el ataque, se cambia el carácter vocálico por el consonántico, la /j̥/ es un resultado natural, pues los gestos articulatorios implicados son muy similares a los de la /i/: tan solo se producen con una constricción un poco más estrecha (*i. e.* aumenta la magnitud del gesto de constricción)¹⁶. En el caso de la paravocal posterior /u̯/, no hay en español ninguna consonante que implique una acción simultánea del cuerpo lingual y de los labios —como sí pasa en la vocal—. Eso puede justificar el desglose en dos segmentos (/ɣ/ y /u̯/). Dado que las vocales se

¹⁶ Keating (1988) aporta datos radiográficos que muestran que la forma adoptada por la lengua es mucho más parecida entre una /i/ y una consonante palatal que entre una /u/ y una consonante velar.

producen fundamentalmente con un gesto lingual (el labial tan solo es un acompañamiento para las vocales posteriores), se entiende también la preferencia por ese gesto en la consonante escogida [ʏ], frente a una posible labial [β], si bien es cierto que esta aparece en la pronunciación de estas palabras en algunas ocasiones (con menor frecuencia que la [ʏ]).

Pasemos ahora a las paravocales posnucleares. Como se argumentará a continuación, existen pruebas de que, a diferencia de las prenucleares, no comparten la cima con la vocal nuclear, sino que se sitúan en la coda (cabe hablar, entonces, de codas vocálicas y codas consonánticas). Por un lado, dada la existencia de triptongos y el límite de ramificación binaria de los constituyentes subsilábicos, si la paravocal pre nuclear y la vocal nuclear ocupan la cima, como se ha argumentado anteriormente, entonces la posnuclear necesariamente debe ocupar la coda. Además, la presencia de una coda vocálica limita drásticamente las posibilidades de que esta vaya seguida por una consonante tautosilábica, restricción propia de las codas complejas. De hecho, prácticamente la única consonante que puede seguir a una (para)vocal en coda es la /s/, al igual que en el resto de codas complejas (p. ej. *seis*, *amáis*, *cambiéis*, etc.). La única excepción (al menos, entre el léxico patrimonial) la constituyen *veinte* y *treinta*, donde el segundo miembro de la coda compleja es una /n/.

Al igual que pasa con las paravocales prenucleares, existen circunstancias en las que una posnuclear se ve forzada a cambiar su posición estructural y a consonantizarse. Esto ocurre en los plurales en *-es* de las palabras acabadas en diptongo decreciente (o triptongo), p. ej. *reyes*, *leyes*, *bueyes*. Al quedar en contexto intervocálico y constituir un mínimo de sonoridad, la paravocal pasa a ocupar el ataque de la sílaba siguiente y se consonantiza (p. ej. ['reʝes]), y puede adoptar incluso una producción fricativa en las variedades lingüísticas que tienen este tipo de pronunciación (p. ej. ['reʒes]).

De las anteriores descripciones sobre los segmentos permitidos en cada una de las posiciones estructurales de la sílaba se pueden extraer algunas generalizaciones. Los constituyentes más «permisivos» son la cima —en concreto el núcleo—, donde puede aparecer cualquier vocal del español, y el ataque, donde puede aparecer cualquiera de las consonantes (en ataque simple). Las combinaciones de ataque y núcleo no presentan ninguna restricción. En cambio, los segmentos susceptibles de llenar la coda son más limitados, lo que encaja en la tendencia universal a que las posibilidades en coda sean más restrictivas (de hecho, hay lenguas que no permiten ningún tipo de coda silábica; otras solo admiten segmentos que compartan total o casi totalmente sus especificaciones con la consonante heterosilábica siguiente, como es el caso de las geminadas o de las nasales homorgánicas). Por otro lado, el número de combinaciones posibles siempre es menor en el interior de constituyentes binarios. Por ejemplo, si bien el ataque (simple) es un constituyente subsilábico que tolera bien cualquier segmento consonántico, el ataque complejo solo admite ciertos grupos de consonantes¹⁷. También la cima

¹⁷ Además, la posibilidad de tener rima ramificada (*i. e.* coda) es una condición previa para que una lengua admita también los ataques ramificados, al parecer, por razones de evolución histórica, *vid.* Ségéral y Scheer (2008, p. 145).

compleja y la coda compleja están sujetas a restricciones. En el § 2.1.1.6 se verá una posible causa de esta desigual distribución.

2.1.1.3. *El concepto de sonoridad*

La regulación de las posibles secuencias segmentales dentro de una sílaba se ha explicado tradicionalmente aludiendo al escurridizo concepto de sonoridad.

Antes que nada, hay que aclarar que en español se ha usado la palabra *sonoridad* (y su adjetivo correspondiente, *sonoro*) para expresar realidades totalmente distintas, lo que puede provocar cierta confusión. Por un lado, es un concepto fonético que alude a la presencia de una fuente periódica, fruto de la fonación (en inglés, *voicing*). Por otro, se ha equiparado a la idea de perceptibilidad (en inglés, *sonority*)¹⁸. Por último, y en relación con esta última idea de sonicidad (*vid.* nota 18), se ha utilizado como concepto fonológico que apela a características de los segmentos en relación con sus posibilidades distribucionales en el seno de una sílaba. Es a estas dos últimas acepciones de sonoridad a las que está dedicado este apartado.

En líneas generales, hay una idea común de que la sonoridad se refiere a la perceptibilidad relativa de cada segmento pero, como se ha dicho, es en última instancia un principio regulador de la distribución de los segmentos en la estructura interna de una sílaba. Por ejemplo, Blevins (1995) la define así:

the sonority of a sound is roughly defined as its loudness relative to other sounds produced with the same input energy (i.e., with the same length, stress, pitch, velocity of airflow, muscular tension, etc.) [...] with each sonority peak defining a unique syllable (p. 207).

El llamado Principio de Secuenciación de la Sonoridad (PSS) establece que cada pico local de sonoridad se debe asociar con el núcleo de una sílaba. Los segmentos prenucleares presentan un patrón de sonoridad creciente hasta el núcleo, mientras que los segmentos posnucleares muestran una sonoridad progresivamente menor, hasta que la sonoridad comience a crecer de nuevo en el inicio de la sílaba siguiente. De este modo se establece un ciclo de ascensos y descensos de sonoridad (Clements, 1990)¹⁹.

El PSS permite un procedimiento casi automático de silabeo, que operaría de la siguiente manera. En primer lugar, se localizan los picos de sonoridad y se asocian a los núcleos silábicos. Después, empezando desde cada núcleo hacia su izquierda, se debe afiliar el máximo número de segmentos posible a su misma sílaba, siempre que den lugar a combinaciones tautosilábicas permitidas por la fonología de esa lengua. Por

¹⁸ Con este sentido, también se habla de *sonicidad* (*vid.* p. ej. Gil Fernández y Llisterri, en prensa).

¹⁹ Hay lenguas que no respetan esta secuencia a rajatabla, y admiten, p. ej., ataques complejos en los que la sonoridad no es creciente. Aun así, estos ataques se sienten como diferentes a los formados conforme al PSS. Por ejemplo, en griego clásico, el perfecto se marca prefijando una sílaba formada mediante reduplicación de la consonante del ataque seguida de la vocal ε (λύω → λέλυκα). Si el ataque es complejo y de sonoridad creciente, se hace la reduplicación con la primera consonante (πλέω → πέπλευκα). En cambio, si el ataque era de sonoridad equilibrada o incluso decreciente, simplemente se prefijaba una ε (πταίω → ἔπταικα; σπάω → ἔσπακα).

último, se vinculan los segmentos aún sin asociar a un pico de sonoridad previo a ellos y pasan, por tanto, a ocupar una posición posnuclear²⁰.

No obstante, se sabe que una secuencia de los mismos segmentos puede silabearse de manera diferente en distintas palabras. Por ejemplo, en español tenemos el contraste entre *varias* y *varías*. En este tipo de casos con secuencias vocálicas, la forma subyacente de cada palabra debe incluir alguna información acerca de si la vocal de menor sonoridad (la /i/) debe ocupar una posición paravocálica o, por el contrario, formar el núcleo de su propia sílaba. En palabras de Blevins (1995):

in the general case, syllable structure is not present in underlying representations [...] For exceptional forms [...] we can assume that minimal structure is specified in the lexicon. In this case, it is sufficient to mark [an otherwise paranuclear vowel] as a syllable nucleus in the UR (p. 221).

En cualquier caso, esta marcación que se incluye en el léxico debe entenderse como una cuestión estructural (*i. e.* no se trata de un rasgo [$\pm$ silábico] al estilo del que postulaba el modelo SPE). Por lo demás, descontando estos casos, el silabeo debe poder deducirse directamente de los patrones de sonoridad²¹.

El interés de los investigadores por la sonoridad ha hecho que se propongan unas *escalas de sonoridad*, que describen el grado relativo de sonoridad de distintas clases de segmentos. Según Ohala (1990a, p. 319), la primera noticia que tenemos de alguien que propusiera una escala de sonoridad data del año 1765, con De Brosches, y su *Traité de la formation mécanique des langues, et de principes physiques de l'étymologie*. Posteriormente, encontramos propuestas en Whitney (1874), Sievers (1893), Jespersen (1904) o Saussure (1916)²², y más tarde la fonología natural concedió un nuevo impulso a las investigaciones en este ámbito. Para el caso del español, la propuesta más conocida es la de Hooper (1974, 1976) —en el marco de la fonología generativa natural—, quien en realidad habla de una *escala de fuerza de consonantes* (EFC), equivalente a la escala de sonoridad, pero de polaridad inversa (*i. e.* a mayor fuerza, menor sonoridad, y viceversa). En efecto, los conceptos de fuerza y sonoridad aparecen entremezclados en la bibliografía de manera habitual, si bien es cierto que deben diferenciarse por varios motivos, como se verá enseguida. Algunos otros autores que han propuesto escalas para el español son Méndez Dosuna (1987) —encuadrado puramente en la fonología natural—, y James W. Harris (2000 [1989]) —con un enfoque generativista que reproduce una estructura similar a la de la teoría de la $\bar{X}$ en sintaxis—.

Tanto Hooper como Méndez Dosuna proponen una escala muy semejante, que se recoge en (6):

²⁰ Vid. Blevins (1995) para una discusión de otras posibilidades.

²¹ De acuerdo con la idea que se presentará en el § 2.1.1.6, según la cual la sílaba se entiende como una unidad de coordinación de gestos articulatorios, cada unidad silábica debe estar especificada como tal en las formas subyacentes del léxico.

²² Para una revisión detallada de la evolución histórica del concepto de sonoridad, *vid.* p. ej. Clements (1990).

(6) Escala de fuerza de consonantes del español

(de menor a mayor fuerza o, lo que es lo mismo, de mayor a menor sonoridad)

1: semiconsonantes y aspiración²³

2: líquidas

3: nasales

4: fricativas sonoras

5: fricativas sordas y oclusivas sonoras

6: oclusivas sordas

Ahora bien, aunque arraigadas en una buena parte de la tradición fonológica, las escalas de sonoridad presentan numerosos problemas descriptivos y explicativos. Por ejemplo, no captan las diferentes posibilidades fonotácticas de los distintos tipos de líquidas (/l/, /r/, /ɾ/)²⁴. Además, queda ambigua la sonoridad relativa de las fricativas sordas /f/, /θ/, /s/, /x/ y de las oclusivas sonoras /b/, /d/, /g/, puesto que todas están en una misma posición de la escala.

Por otro lado, hay lenguas que presentan patrones reversos de sonoridad en los ataques complejos y en las codas complejas: p. ej. inicios silábicos /st/, /xt/, /vd/ y finales /ts/, /kʃ/, /dz/ (Ohala, 1990a)²⁵.

Tampoco se explica por qué no todos los grupos consonánticos de coda se admiten como ataques complejos cuando se invierte el orden, a pesar de respetar la secuencia de sonoridad. Los grupos de ataque complejo necesitan mayor distancia de sonoridad entre sus miembros, mientras que los segmentos de una coda compleja pueden tener un grado de sonoridad más parecido entre sí (Blevins, 1995).

Por estas y otras razones, las escalas de sonoridad han sido fuertemente criticadas por algunos autores, como Ohala (1990a), Ohala y Kawasaki-Fukumori (1997), o John Harris (2006). En efecto, los argumentos propuestos en su defensa son circulares, puesto que no hay ninguna demostración empírica independiente. A menudo se dice, por ejemplo, que una /a/ presenta más sonoridad porque constituye núcleo silábico, y se justifica su capacidad de ser núcleo alegando que tiene más sonoridad.

Según Gil Fernández (2000b, p. 72, nota 75),

se ha discutido mucho si la sonoridad ha de entenderse como una dimensión derivada de los valores asignados a los rasgos distintivos habituales, o si constituye ella misma un único rasgo polivalente. Esta última postura es la defendida por Selkirk (1984a), quien propone eliminar los rasgos distintivos de clase mayor, que resultarían redundantes al incorporarse al modelo el nuevo rasgo. Por el contrario, Clements (1990), además de repasar los argumentos que se han dado a favor de una y otra concepción, aboga claramente por definir la sonoridad en función de los principales rasgos binarios (en el análisis de Clements, cuantas más especificaciones positivas (+) suma un segmento con

²³ La aspiración es una aportación específica de Méndez Dosuna a la propuesta original de Hooper.

²⁴ Sobre esta cuestión, *vid.* p. ej. Pons (2008).

²⁵ *Vid.* nota 19.

respecto a los rasgos [silábico], [vocoide], [aproximante] y [sonante], mayor será su grado de sonoridad).

También Blevins (1995, p. 211) presenta un cálculo de la sonoridad basado en rasgos distintivos, estableciendo una propuesta de mínimos, es decir, que solo incluye aquellas relaciones de sonoridad que, según ella, no han sido desmentidas con ningún contraejemplo de ninguna lengua. La Figura 2.7 recoge esta propuesta; «for each node, the left branch is more sonorous than the right branch, and sonority relations for a given feature are only defined with respect to segments with the feature specification of the mother node» (*ibid.*). En las ramas terminales de abajo a la derecha, cuando se habla del rasgo [$\pm$ sonoro], se está haciendo referencia a la presencia o ausencia de vibración laríngea que opone los sonidos sonoros a los sordos.

Además de este tipo de explicaciones fonológicas, basadas en rasgos, se ha intentado buscar una definición operativa de la sonoridad buscando sus posibles manifestaciones fonéticas. Sin embargo, resulta difícil establecer un correlato exacto. La sonoridad

no es fácil de definir ni de explicar en términos fonéticos y, si bien se han barajado muchos parámetros como sus posibles condicionantes (el grado de abertura de los segmentos, la intensidad de la espiración, el volumen, la presión, la combinación de algunas de estas variables, etc.), no se ha llegado a una conclusión definitiva, incluso se duda de que este sea un objetivo alcanzable (Gil Fernández, 2000b, p. 73).

Algunos estudios experimentales han descrito altos grados de correlación entre sonoridad e intensidad (Parker, 2002, *apud* John Harris, 2006). Cabe matizar que no hay un rango de valores de intensidad (en decibelios) que corresponda a cada uno de los grupos definidos por la escala de sonoridad. El nivel de intensidad depende, en cada producción concreta, del grado de eficiencia vocal (*vid.* p. ej. Titze, 2000) y, en todo caso, la intensidad que llega al oído —o al micrófono, en un contexto experimental— decrece de manera proporcional al cuadrado de la distancia respecto a la fuente del sonido (Baken, 1996, p. 105). Ahora bien, dadas las circunstancias de una misma emisión, se espera que los sonidos con mayor sonoridad presenten también una intensidad relativa mayor.

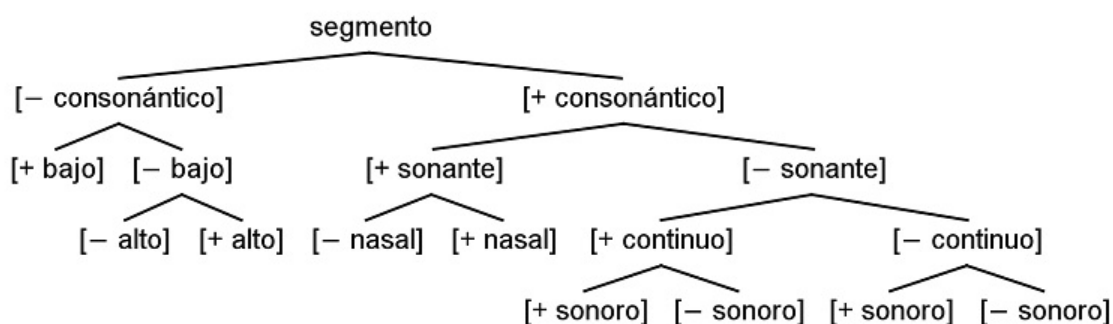


Figura 2.7. Mínima especificación dicotómica de la sonoridad en función de los rasgos distintivos (según Blevins, 1995)

La intensidad se ha interpretado normalmente en términos de perceptibilidad. Es decir, que el segmento que ocupa el núcleo silábico es el más intenso y, por ello, el más perceptible. John Harris (2006) acepta que, aunque con ligeras inadecuaciones, la intensidad es bastante paralela a la sonoridad. Sin embargo, critica la asociación entre intensidad y perceptibilidad, pues más bien lo que hace que un determinado segmento sea perceptible es que se diferencie notablemente de los segmentos contiguos. Esta diferencia puede ser en términos de intensidad, pero no solo. La perceptibilidad se vincula más bien a la *modulación* consecutiva de la onda portadora, como también opinan Ohala (1990a) y Ohala y Kawasaki-Fukumori (1997).

Por descontado, una de las maneras de modular dicha señal es variar el grado de abertura del tracto vocal con respecto al segmento precedente. Dada una consonante oclusiva, p. ej. /p/, cualquier variación en la abertura debe ir en la dirección de una abertura mayor en el siguiente segmento, p. ej. /l/, y después de este se puede seguir abriendo más todavía, y llegar p. ej. a una /a/; alcanzado este extremo, si se quiere modular la señal mediante una modificación de la abertura, ya no queda más remedio que volver a un segmento más cerrado que la vocal, p. ej. /n/. Este tipo de secuencias explicaría el carácter cíclico de la sonoridad en el seno de una sílaba, p. ej. /plan/. También se podrían vincular a los movimientos, también cíclicos, de la mandíbula (*vid.* § 3.5)²⁶. Las mayores aberturas permiten radiar más energía acústica y, por tanto, se relacionan con mayor intensidad.

Sin embargo, hay otras maneras de modular la señal acústica que no tienen que ver con la variación en la abertura o en la intensidad, y son estos mecanismos, seguramente, los que puedan explicar las aparentes contradicciones o insuficiencias descriptivas de las tradicionales escalas de sonoridad. La posibilidad de tener una determinada secuencia segmental en una sílaba se debe justificar mediante una legitimación por claves perceptivas (en inglés, *licensing by cue*), pero estas claves pueden ser fiables por muchos más motivos que la variación de intensidad. Wright (2004) aporta una excelente explicación de la solidez de las distintas claves perceptivas en diferentes contextos, y cómo esta justifica la preferencia por determinadas secuencias segmentales, sin necesidad de recurrir a conceptos abstractos como la sonoridad.

Ahora bien, la sonoridad establece una asociación preferente entre determinados tipos de segmentos y ciertas posiciones de la sílaba. Dada la caracterización de estas posiciones como fuertes o débiles (*vid.* § 3.1), muchos autores han creado una relación (inversa) entre la sonoridad y la fuerza, como ya se ha adelantado al hablar de la EFC en (6). Sin embargo, es preciso diferenciar adecuadamente ambos conceptos.

Para empezar, la posición de ataque es una posición fuerte y en ella suelen aparecer elementos de menor sonoridad. En cambio, la coda es una posición débil, pero también admite segmentos de escasa sonoridad según las escalas. Esa constatación ya sacude los cimientos de la relación entre sonoridad y fuerza.

²⁶ Algunos trabajos de MacNeilage (p. ej. los citados por Goldstein, Byrd, y Saltzman, 2006) apuntan a un posible origen de la sílaba vinculado a los movimientos de la mandíbula, pero Goldstein *et al.* señalan que esta teoría aporta explicaciones insuficientes para algunos aspectos de la adquisición de la estructura silábica.

La sonoridad es una propiedad de los segmentos relativa a los otros segmentos que aparecen en su contexto (aunque las escalas de sonoridad la presentan más bien como una propiedad de las clases naturales a las que pertenecen esos segmentos). En cambio, la fuerza es, ante todo, una propiedad de determinadas posiciones prosódicas, y solo indirectamente atribuible a los segmentos que las ocupan.

Además, el término *sonoridad* se refiere a la posibilidad de que un segmento aparezca en un orden lineal determinado, con respecto a otros segmentos, en el seno de una sílaba. Por el contrario, los fenómenos de refuerzo pueden afectar a unidades prosódicas de distintos niveles, no solo a la sílaba (de hecho, dado su efecto acumulativo, se manifiestan en mayor grado en niveles superiores a la sílaba), *vid.* § 3.1.

Por último, aunque se ha intentado relacionar las escalas de sonoridad con las rutas de lenición (o de fortición), las predicciones evolutivas que se desprenden de las escalas no se constatan empíricamente. Por ejemplo, si una fricativa se debilita no se convierte en una nasal (John Harris, 2009, p. 29; Kirchner, 2001, p. 87). Y, si bien las vocales son los elementos más sonoros posibles, no es cierto que el último paso de la lenición sea la vocalización (como bien señala Lavoie, 2001, p. 94), sino la elisión.

En definitiva, aunque en algunos casos pueda haber una relación indirecta, conviene diferenciar los conceptos de sonoridad y fuerza. Mientras que los fenómenos de refuerzo sirven para delimitar los constituyentes de la estructura prosódica (*vid.* capítulo 3), la sonoridad permite dar cuenta de los patrones fonotácticos de la lengua (*i. e.* de las secuencias de segmentos permitidas dentro de una sílaba)²⁷. No obstante, nótese que hay autores (Blevins, 2003; Steriade, 1999; entre otros) que sostienen que los patrones fonotácticos se aprenden por inducción a partir de la forma de las palabras, y solo en segunda instancia se extrapolan a la sílaba. En el § 2.1.3.2 volveremos a este asunto, con la discusión de las diferencias fonotácticas entre los extremos y el interior de las palabras.

2.1.1.4. Ambisilabicidad

El carácter ambisilábico de una consonante (intervocálica) consiste en estar afiliada simultáneamente a dos sílabas²⁸, siempre que fonotácticamente se trate de un segmento admisible en la coda de la primera sílaba y también sea admisible en el ataque de la segunda. A lo largo de las últimas décadas, algunos autores (p. ej. Selkirk, 1980, 1982) han planteado la necesidad de un análisis ambisilábico de algunos segmentos en algunas lenguas, para justificar sus características fonéticas o la particularidad de los fenómenos a que se veían sometidos.

El inglés, concretamente, parece ser una de esas lenguas, y esto se ha relacionado con el hecho de que su ritmo acentual otorga más importancia al pie que a la sílaba, de modo

²⁷ Sin embargo, *vid.* § 3.2, pues parece ser que el refuerzo provocado no por la posición sino por la prominencia prosódica sí tiene como efecto un aumento de la distancia perceptiva entre los segmentos (*i. e.* una mayor separación en términos de sonoridad). En este sentido, además, los casos de refuerzo por prominencia se asemejarían más al concepto de fortición empleado por los partidarios de la fonología natural, como se explicó en el § 1.

²⁸ En términos de coordinación de gestos articulatorios (*vid.* § 2.1.1.6), eso significa que está coordinada a la vez (en antifase) con el gesto vocálico precedente y (en fase) con el gesto vocálico siguiente (Gick, 2003).

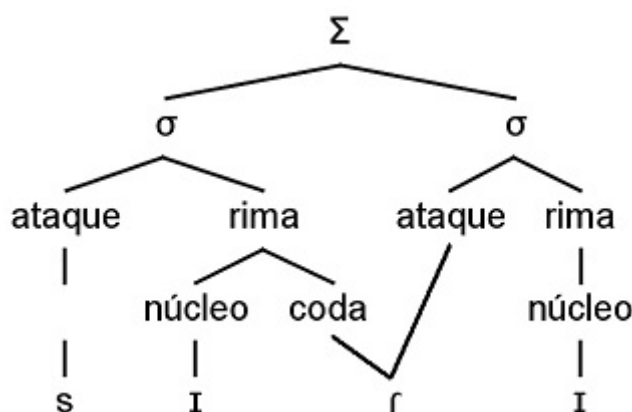


Figura 2.8. Ambisilabidad de una consonante intervocálica interior de pie

que las fronteras silábicas dentro de un mismo pie quedan desdibujadas. Se ha propuesto que el carácter ambisilábico de determinadas consonantes inglesas es la causa del fenómeno del aleteo (en inglés, *flapping*), entre otros. En una palabra como *city* /'sɪtɪ/ 'ciudad', el Principio de Maximización del Ataque²⁹ (Selkirk, 1982) exige que la /t/ sea el ataque de la segunda sílaba. Por otro lado, en general, la fonotáctica del inglés no permite una vocal laxa en sílaba abierta (con alguna excepción, como una /ɪ/ final de pie). En consecuencia, la /t/ debe trabar a la primera vocal para dar lugar a una sílaba bien formada /sɪt/. Según Gussenhoven y Jacobs (2011), es esa dualidad —la /t/ es ataque pero también es coda— la que desencadena un debilitamiento y da pie a una pronunciación percusiva en algunas variedades dialectales: [ˈsɪɾɪ] (Figura 2.8). La posibilidad de ser ambisilábica se extiende a cualquier consonante que pueda ser tolerada como coda de una sílaba, ya incluso aunque esté precedida por vocal tensa (p. ej. *writer* /ˈraɪtə/ 'escritor', pronunciado como [ˈraɪɾə]³⁰). Eso sí, la reasignación silábica solo se puede producir en el seno de un mismo pie: *agitation* /ædʒɪˈteɪʃən/ 'agitación' no puede pronunciarse *[ædʒɪˈɾeɪʃən], porque la /t/ es inicial de pie.

La ambisilabidad consiste, por tanto, en una especie de resilabeo en el interior de un pie. Ahora veremos los fenómenos de resilabeo a través de frontera de palabra, y las vinculaciones que se han establecido entre las manifestaciones fonéticas de este resilabeo y la ambisilabidad.

2.1.1.5. Resilabeo

En el § 2.1 se presentaron los procesos de reestructuración de los constituyentes prosódicos, que modifican la organización de las unidades de un determinado nivel cuando se insertan en otro nivel superior. La ambisilabidad, que acabamos de ver, es un ejemplo de ello (la estructura silábica se reorganiza en el seno de un pie). Ahora

²⁹ Este principio establece que, dada una secuencia de consonantes entre dos vocales, debe afiliarse el mayor número posible de consonantes al ataque de la sílaba a la que pertenece la segunda vocal, siempre que las restricciones fonotácticas de la lengua permitan que en posición de ataque aparezca la secuencia resultante.

³⁰ Esta pronunciación puede provocar una cuasi-homofonía con *rider* /ˈraɪdə/ 'jinete; ciclista', pronunciado [ˈraɪɾə], con alargamiento extra de la vocal por ir seguida de consonante sonora.

analizaremos los casos que corresponden al concepto más tradicional de resilabeo, que es el que se produce a través de frontera de palabra.

En español, por ejemplo, se produce un resilabeo en cualquier punto de un enunciado fonológico, siempre que se cumpla la siguiente condición.

(7) Regla del resilabeo en español

Si la primera sílaba de una palabra empieza directamente por su núcleo, dicha sílaba se unirá a los segmentos finales de la palabra anterior hasta conseguir formar una sílaba con el máximo ataque posible.

Como explica Lahoz-Bengoechea (2015),

mis amigos se resilabea como *mi.sa.mi.gos*³¹; *el arte antiguo* como *e.lar.tean.ti.guo*³²; *otro invierno* pasa a *o.tro.in.vier.no*; y *una hormiga*, a pesar de la “h” gráfica, también experimenta resilabeo: *u.nahor.mi.ga*. La tendencia a la sinalefa en español es muy fuerte, y une incluso vocales que no formarían diptongo en interior de palabra. Únicamente puede romperse (y no siempre) si la segunda vocal está acentuada, p. ej. *com.pra.o.ro*. Además, siguiendo la regla expuesta en [(7)], también hay resilabeo si la segunda palabra empieza por diptongo decreciente: *el aire*, *su audacia*, etc. En cambio, no se produce resilabeo si el diptongo es creciente [puesto que, como se ha visto en el § 2.1.1.2, la aparente paravocal realmente es una consonante que ocupa el ataque]. [Las notas al pie forman parte de la cita.]

El ámbito de los procesos de reestructuración puede ser diferente de una lengua a otra. En francés, son varios los fenómenos que pueden tener como efecto un resilabeo de este tipo: *elision*, *enchaînement* y *liaison* (vid. Spinelli, Cutler y McQueen, 2002, para una explicación de las diferencias entre ellos). Aunque con ciertas reticencias por parte de algunos autores (vid. nota 7 en la p. 32 de esta tesis), parece que el ámbito de aplicación de la *liaison*, por ejemplo, es la frase fonológica menor (vid. § 2.1.4), como defiende Selkirk (1986), entre otros.

Algunos autores han apuntado que las consonantes sujetas a resilabeo pueden presentar características fonéticas propias de segmentos ambisilábicos. Por ejemplo, Gick (2003) aporta ciertas pruebas, aunque limitadas, de tal comportamiento de la lateral /l/ y las deslizantes³³ /j/ y /w/ en inglés. Cuando estas consonantes aparecen a

³¹ Se señala en negrita la sílaba acentuada de las nuevas palabras fonológicas resultantes.

³² El resilabeo es el nombre genérico del proceso, tanto si la palabra anterior acaba en consonante como en vocal. Sin embargo, este último caso recibe también el nombre específico de *sinalefa*. En la sinalefa, dos sílabas originales se funden en una sola; en cambio, en el resilabeo consonántico, se produce una reasignación del último segmento de coda de una sílaba al ataque de la siguiente.

³³ Existen varios subtipos de consonantes aproximantes (aquellas en las que los órganos articuladores se aproximan pero sin llegar a formar un cierre completo ni tampoco un estrechamiento crítico como para producir ruido de fricción). Uno de estos subtipos de aproximantes son las espirantes, que pertenecen a la gran clase de las obstruyentes, ya que la constricción formada en el punto de articulación es suficiente para provocar un diferencial de presión a ambos lados de ese punto (Martínez Celadrán y Fernández Planas, 2013). Entre ellas se encuentran las consonantes españolas [β], [ð], [j], [ɣ]. Otro subtipo (no obstruyente) es el integrado por las que podemos llamar deslizantes (por la palabra inglesa *glide*), e incluye las

final de palabra y les sigue una palabra que empiece por vocal, la magnitud de algunos de los gestos linguales implicados resulta significativamente distinta tanto de las consonantes que sin duda ocupan una posición de coda como de las que sin duda ocupan una posición de ataque. La coordinación temporal de los gestos también muestra tendencias a un comportamiento de las consonantes resilabeadas que es intermedio entre los dos casos no resilabeados, aunque este resultado no obtiene significatividad estadística. Estos datos se interpretan en términos de un posible análisis ambisilábico de las consonantes resilabeadas.

Para el español, Strycharczuk y Kohlberger (2014) prueban que la /s/ en coda dura menos que en ataque, mientras que la /s/ en resilabeo muestra una duración intermedia y significativamente distinta de los otros dos casos, algo que también se analiza como prueba de su carácter de segmento ambisilábico.

En cualquier caso, está claro que las consonantes resilabeadas mantienen propiedades típicas de la posición de coda. Por ejemplo, en español, la rótica que aparece en final de sílaba es la percusiva /r/, que, aunque admite diversos alófonos en tal posición, puede pronunciarse, en efecto, como [r] (*vid.* § 4.1). Cuando el final de sílaba coincide con final de palabra y además se produce un resilabeo, la consonante se afilia al ataque de la nueva sílaba creada a través de la frontera léxica. Sin embargo, y a pesar de la prohibición de percusivas en inicio de palabra, en estos casos es la pronunciación [r] la que se mantiene obligatoriamente. Como se verá en el § 2.3.3.1, los oyentes nativos de lenguas de ritmo silábico como el español se apoyan en la sílaba como estrategia de segmentación para el acceso al léxico. Podría pensarse que los procesos de resilabeo a través de frontera de palabra podrían restar eficacia a esta estrategia. Sin embargo, las peculiaridades fonéticas de los segmentos resilabeados pueden ser una marca de su pertenencia original a la coda de la palabra anterior.

2.1.1.6. *La sílaba como unidad de coordinación de gestos articulatorios*³⁴

La idea de los modelos fonológicos tradicionales de que la cadena hablada se compone de una serie de segmentos fónicos es una abstracción y, como tal, aunque pueda tener una base psicológica fundada, no siempre se corresponde con la realidad fonética de la producción del sonido articulado. Las fronteras entre los segmentos distan mucho de ser nítidas. Existen casos de coarticulación y asimilación (*vid.* p. ej. Hardcastle y Hewlett, 1999), que se han analizado como la cesión de un rasgo distintivo por parte de un elemento a otro presente en su contexto —en trabajos basados en el modelo SPE (Chomsky y Halle, 1976 [1968])— o como la extensión de un autosegmento de una posición del esqueleto a otra adyacente a ella —en enfoques basados en el modelo autosegmental (Goldsmith, 1976, 1979)—. En el marco de la fonología autosegmental, el Principio de Contorno Obligatorio (OCP, por sus siglas en inglés) prohíbe la aparición de dos autosegmentos del mismo valor asociados a dos

semiconsonantes [j] y [w] del inglés, por ejemplo. Nótese la diferencia entre la [j] española (obstruyente) y la [j] inglesa (sonante). Un tercer grupo lo formarían las aproximantes laterales como la [l] o la [ʎ], y un cuarto grupo las de tipo rótico, como [ɾ].

³⁴ Este apartado se basa fundamentalmente en el trabajo de Goldstein *et al.* (2006), excepto donde se indiquen otras fuentes.

posiciones contiguas y obliga también a que sea un solo autosegmento el que se extienda y afecte a ambas posiciones.

Esta noción de la extensión temporal de un autosegmento puede considerarse precursora del carácter inherentemente durativo de los primitivos fonológicos propuestos por el modelo de la fonología articulatoria (Browman y Goldstein, 1986, 1989, 1992; y subsiguientes trabajos). La fonología articulatoria pretende superar los problemas planteados por la indefinición fonética de las fronteras entre unidades segmentales que, desde el punto de vista fonológico, se consideran discretas. Para ello, sugiere que las unidades de información fonológica (esto es, las que aparecen codificadas en la forma subyacente del léxico y establecen oposiciones de significado) deben ser isomórficas con las unidades de producción (basadas en acciones que modifican transitoriamente las características del tracto vocal).

Así, los primitivos fonológicos son los llamados *gestos articulatorios*, definidos en tanto que objetivos de constricción de mayor o menor grado en determinados puntos del tracto vocal. Estos gestos son unidades de carácter intrínsecamente espaciotemporal (presentan una magnitud y una duración concretas), y son dinámicos, en el sentido de que el gesto supone la puesta en acción de una serie de movimientos de distintos órganos articulatorios para conseguir la constricción deseada. En algunos gestos (p. ej. de la glotis, o del velo del paladar) solo hay un órgano participante, mientras que en otros (p. ej. de los labios, o del ápice de la lengua) intervienen distintos órganos coordinados (mandíbula, labio inferior y labio superior —en un caso—, mandíbula, cuerpo de la lengua y ápice lingual —en el otro—). Los movimientos activados por el gesto pueden modelarse según una dinámica de tareas (en inglés, *task dynamics*) —*vid.* Saltzman y Kelso (1987); Saltzman y Munhall (1989)— de una manera similar a como se ha hecho con los movimientos de otros órganos del cuerpo, como las extremidades. Esto permite captar la invariancia en la consecución de los objetivos (constricciones del tracto vocal) al tiempo que deja libertad para que exista variación en los movimientos concretos de los órganos implicados. En otras palabras, aun asegurando que se alcanza la constricción deseada, los movimientos articulatorios pueden variar (y, de hecho, varían) de una ejecución a otra del mismo gesto. Esta variabilidad se puede explicar por condicionantes contextuales (el punto en el que se encontraban antes los órganos, o al que tienen que llegar después) o prosódicas (la cercanía de una frontera de constituyente, *vid.* § 3.1 para una explicación detallada sobre los llamados gestos prosódicos o gestos π), entre otras causas. La velocidad global de habla también afecta a la ejecución de los movimientos articulatorios. Además, los grados de libertad aumentan cuantos más órganos diferentes estén implicados en la consecución de un mismo gesto, debido a la posibilidad de que existan efectos de compensación entre los movimientos de unos y otros (*vid.* § 3.5.1).

Además de esta coordinación entre órganos articulatorios, existe una coordinación de los distintos gestos entre sí, formando la llamada partitura de gestos (en inglés, *gestural score*). Dado que los diferentes gestos regulan distintas variables del tracto vocal, dichos gestos son independientes entre sí (de una manera que recuerda también a la independencia entre los autosegmentos de los distintos niveles o *tiers*). Por ello,

podemos encontrar gestos que se suceden en el tiempo o gestos que se solapan total o parcialmente. Los tradicionales segmentos proceden de la abstracción de una serie de características fonéticas más o menos estables derivadas de un conjunto de gestos que habitualmente se producen de manera simultánea. Sin embargo, los gestos pueden abarcar una duración mayor o menor que la del segmento con el que típicamente se asocian.

Las partituras de gestos se consideran una parte integrante de la forma léxica subyacente. Entre los patrones de coordinación intergestual que estas definen, hay unos que merecen especial interés, pues dan cuenta de la estructura interna de las sílabas y predicen las propiedades de las distintas posiciones silábicas en función de los tipos de coordinación implicados, sin necesidad de aludir a estructuras de constituyentes subsilábicos como las que hemos visto en el § 2.1.1.2:

In the gestural approach, there are configurational properties intrinsic to gestural constellations that identify syllable-initial vs. syllable-final consonants without necessarily invoking any hierarchical structure. Thus, affiliated consonants both preceding and following a vowel ('syllable-initial' and 'syllable-final' consonants) are phased with respect to the vowel, but in different ways. Moreover, because the configurational properties (the phase relations) also determine the patterns of overlap among gestures, differences in configuration are automatically associated with differences in overlap, and in turn, with the articulatory and acoustic consequences of such overlap (Browman y Goldstein, 1995, p. 20).

Un patrón coordinativo entre dos gestos se define en términos de *fase* (*i. e.* en qué fase del desarrollo de un gesto da comienzo el otro gesto), y las fases se establecen como grados sexagesimales. Así, 0° representa el comienzo de un gesto; 90°, el momento de máxima velocidad en la formación de la constricción; 180°, el punto en que se alcanza la mayor magnitud; 270°, el momento de máxima velocidad de distensión; 360°, el final del gesto (Browman y Goldstein, 1986). Alternativamente, se suele hablar de coordinación *en fase* para dos gestos simultáneos y de coordinación *en antifase* para dos gestos consecutivos. Estos dos patrones de coordinación se consideran los más naturales y, por ello, los más estables (y el patrón en fase es todavía más estable que en antifase). Cualquier otro modo de coordinación no surge de manera natural, y requiere un aprendizaje específico.

Pues bien, una serie de experimentos y mediciones articulatorias (*vid.* p. ej. Goldstein, Byrd y Saltzman, 2006; Krakow, 1993) parecen indicar que una consonante se coordina en fase con una vocal siguiente (y a esto se lo llama la relación de ataque); por el contrario, una vocal y una consonante posvocálica se coordinan en antifase (lo que constituye la relación de coda). La selección de un tipo de fase u otro está encaminada fundamentalmente a asegurar la perceptibilidad de los distintos gestos (o de las consecuencias acústicas derivadas de sus constricciones³⁵).

³⁵ No pretendo aquí entrar en el debate acerca de la adecuación de la llamada teoría motora de la percepción del habla, que —simplificando mucho— propone que en la percepción del habla el oyente es capaz de representarse los gestos articulatorios que han dado lugar a un determinado producto

Una secuencia CV, aun presentando gestos simultáneos, permite la percepción de la consonante y la vocal por los siguientes motivos. Los gestos vocálicos son intrínsecamente más lentos y duran más que los consonánticos, de modo que, aunque dos gestos —uno de cada tipo— empiecen a producirse a la vez, el consonántico acabará más o menos a la mitad del desarrollo del vocálico, *i. e.* cuando este está alcanzando su pico de magnitud. Durante la primera mitad del gesto vocálico, la menor constricción representada por este no enmascara el efecto de la constricción consonántica y, por tanto, la consonante se puede percibir correctamente. Durante la segunda mitad del gesto vocálico, la consonante ya ha acabado, y así la vocal también se puede percibir. Dadas estas características, la coproducción de CV resulta eficiente en la transmisión de la información; además, la mayor coarticulación se refleja en claves acústicas durante la transición CV que resultan ventajosas para la percepción de ambos segmentos, especialmente de la consonante³⁶.

En cambio, dada una secuencia VC, para que la consonante realmente aparezca tras la vocal y además no la enmascare, es necesario que el gesto C sea consecutivo (y no simultáneo) al gesto V. Esto justifica la coordinación en antifase propia de las codas. La secuenciación es más exhaustiva aún cuando la consonante y la vocal ponen en juego distintos valores de la misma variable de tracto. En otras palabras, los gestos vocálicos son, fundamentalmente, gestos del cuerpo de la lengua; si la consonante de coda también se articula con el cuerpo lingual, lo normal es que su ejecución deba esperar a que acabe por completo la vocal, a menos que los gestos apunten en la misma dirección (p. ej. una vocal posterior y una consonante velar); en cambio, si la consonante se articula con otros órganos diferentes, a veces se adelanta un poco su inicio, que coincide con los últimos estadios de la vocal.

Por otro lado, si se deben producir dos segmentos consonánticos consecutivos o dos segmentos vocálicos consecutivos, la coordinación tampoco puede ser simultánea o, si no, uno de los dos segmentos podría no percibirse correctamente. Esto es lo que ocurre, por ejemplo, con los ataques complejos. Cada una de las consonantes se coordina en fase con la vocal (eso es lo que las define como consonantes de ataque). Sin embargo, a su vez, las consonantes deben coordinarse en antifase entre sí. Esto genera una competición en el patrón coordinativo, que resulta en el llamado efecto de centro C (en inglés, *c-center*), *vid.* Browman y Goldstein (1988); Byrd (1995). En estas estructuras, cada consonante empieza más o menos a la mitad del desarrollo de la anterior (la presión por mantenerse en fase con la vocal no permite que el inicio de la segunda

sonoro. Para un seguimiento de la polémica, el lector puede referirse a A. M. Liberman y Mattingly (1985); Ohala (1996); Fowler (1996); Fowler y Galantucci (2005); Iskarous (2010).

³⁶ Ohala (1990a) aduce que las transiciones CV son más informativas que las VC como consecuencia de la acumulación de presión oral (P_o) que se produce durante la consonante. La mayor P_o se traduce en una explosión distensiva más intensa, y esa explosión es portadora de claves acústicas fundamentales para la percepción del punto de articulación (por las características formánticas y las transiciones), el modo (por la amplitud y la duración de la explosión) y la acción glotal (puesto que el momento de la distensión es el que se toma como referencia para medir el tiempo que tarda en empezar la vibración laríngea, el *Voice Onset Time* o VOT, *vid.* nota 57 en la p. 90). En cambio, en una secuencia VC, la mayor acumulación de P_o no se produce hasta el final de la consonante, y eso hace que las claves que se producen durante la transición no sean tan perceptibles.

consonante se demore hasta que la primera haya finalizado). En cuanto a la vocal, ocurre lo siguiente. Cuando tan solo hay una consonante de ataque, los inicios de C y de V están alineados, y hay un punto X en el desarrollo de la vocal que coincide con el centro de la consonante. Pues bien, ese punto X debe coincidir con el centro del *grupo* consonántico en los casos de ataque complejo. En eso consiste el centro C (Nam, 2007).

Con las codas complejas, en principio, no se da esta competición, sino que la primera consonante se coordina en antifase con la vocal, y la segunda consonante se fija en antifase con la primera.

En los diptongos, el inicio de una vocal debería coincidir con el punto medio de la anterior, aunque es posible que haya variación de una lengua a otra (*vid.* p. ej. los datos del rumano aportados por Marin y Goldstein, 2012). En cambio, dos vocales en hiato, como pertenecientes a dos sílabas distintas, deben ser totalmente secuenciales la una con respecto a la otra.

La Tabla 2.1 resume el tipo de coordinaciones implicadas en las distintas posiciones silábicas, de ataque y coda (simples y complejos).

Estos patrones de coordinación, derivados de la necesidad de asegurar la perceptibilidad, dan cuenta de manera unificada de los distintos tipos de relaciones que establecen los segmentos de una sílaba, sin necesidad de postular constituyentes subsilábicos (aunque las etiquetas tradicionales —ataque, coda, etc.— sigan siendo útiles desde el punto de vista descriptivo). La sílaba, entendida así como una unidad de coordinación gestual, agrupa directamente las distintas unidades primitivas de contenido fonológico (los gestos). Las diferentes propiedades del ataque y de la coda, p. ej., derivan del tipo de fase en que se coordinan dichos gestos.

Como se ha dicho, la coordinación en fase es la más estable de todas. Esto explica la preferencia universal por las sílabas de tipo CV (de hecho, hay lenguas como el hawaiano que solo admiten este esquema). Además, en el § 2.1.1.2 se ha visto que las posibilidades combinatorias entre los posibles ataques y rimas son bastante libres. En cambio, la coordinación en antifase de la coda explica las mayores restricciones fonotácticas asociadas normalmente a esta posición. También hay fuertes limitaciones en los grupos consonánticos (codas y ataques complejos), por la misma razón. Por otro lado, el efecto de competición de los ataques complejos, que hace que las consonantes se solapen parcialmente aun estando en antifase, puede ser el motivo de que esos segmentos deban guardar entre sí una mayor distancia de sonoridad, para asegurar la perceptibilidad. En cambio, en las codas complejas, la mayor separación entre ambas

Tabla 2.1. Estructuras de coordinación de gestos consonánticos y vocálicos en el seno de una sílaba.
La línea continua indica coordinación en fase y la discontinua, en antifase.

	Ataque	Coda
Simple	C ____ V	V C
Complejo	C C ____ V —————	V C C

consonantes hace que estas se perciban aunque presenten una menor diferencia de sonoridad.

Las necesidades de coordinación intergestual no solo afectan a la relación entre segmentos sucesivos, sino también a los gestos de un mismo segmento. En principio, los gestos glotales y velares pueden ser simultáneos a los gestos orales porque sus consecuencias acústicas son independientes de las de estos, y no se pone en peligro la percepción. En efecto, en la posición de ataque se constata un inicio simultáneo del gesto oral y el gesto velar en las consonantes nasales. Sin embargo, sabemos que en coda el gesto oral debe ser consecutivo al gesto de la vocal. El gesto velar, en cambio, se presta a variación: puede adelantarse al gesto oral y solaparse con la vocal, que se nasalizará parcialmente (como pasa en inglés o en español) o puede permanecer en antifase con la vocal y simultáneo al gesto oral (como pasa en francés, para preservar la oposición entre vocales orales y nasales).

En inglés también se observa un efecto de disociación de gestos de un mismo segmento en el caso de la velarización de /l/ en coda. Siempre que dos constricciones se suceden temporalmente en posición de coda, lo harán de manera que una no enmascare a la otra. Eso suele significar que primero se produce la constricción más ancha y luego la más estrecha; también las constricciones más atrasadas en el tracto vocal anteceden temporalmente a las más adelantadas. La /l/ se produce con un gesto del cuerpo de la lengua y otro del ápice. En final de sílaba, el gesto del cuerpo antecede al apical y eso provoca la variante velarizada [ɫ].

De este modo, la coordinación intrasilábica de los gestos ofrece una explicación unificada a dos fenómenos superficialmente inconexos, como son la nasalización de vocales y la velarización de /l/ (Browman y Goldstein, 1995).

2.1.1.7. La mora (μ): ¿constituyente, o propiedad de la sílaba?

Los fonólogos han reconocido desde hace mucho tiempo que la diferencia fundamental entre la rima y el ataque de una sílaba radica en que es la rima la que condiciona una serie de fenómenos relacionados con el *peso silábico*. Por ejemplo, en muchas lenguas la construcción de pies (*vid.* § 2.1.2.2) —y, por tanto, los procesos de asignación de acento (*vid.* § 2.2.1)— es sensible al peso de las sílabas.

El peso depende del número de segmentos de la rima, aunque no consiste en un mero recuento. Las sílabas ligeras son aquellas cuya rima consta únicamente de una vocal. En cambio, la presencia de un diptongo o de una vocal larga (en las lenguas de cantidad distintiva) hace que la sílaba sea pesada. Además, una coda consonántica puede convertir una sílaba en pesada en algunas lenguas (*vid.* Gordon, 2004, para una explicación de por qué las codas de diferentes lenguas varían en su capacidad de otorgar peso a la sílaba, e incluso por qué pueden comportarse de diverso modo con respecto a distintos fenómenos sensibles al peso, como la asignación de acento y la alineación de tono).

Para dar cuenta de estas diferencias de peso, algunos autores (p. ej. B. Hayes, 1989; Hyman, 1985) proponen una estructura intermedia entre la sílaba y el segmento, mediante otro tipo de constituyente llamado la mora (de la familia léxica de *demorar*,

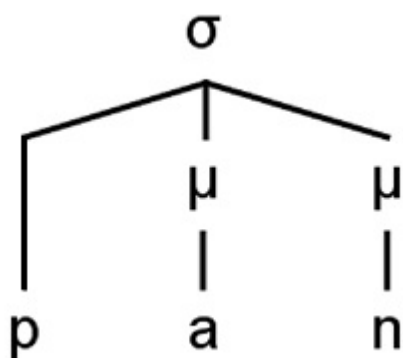


Figura 2.9. Consonante prenuclear dominada directamente por la sílaba

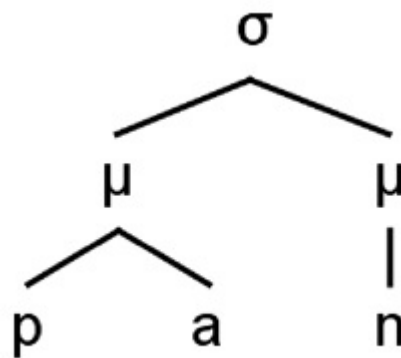


Figura 2.10. Consonante prenuclear dominada por la mora nuclear

por ser una unidad de medición de la cantidad). Así, se dice que las sílabas ligeras tienen una mora y las pesadas, dos moras. Sin embargo, es raro que una lengua haga una ulterior distinción fonológica entre sílabas de dos y de tres moras, aunque haya sílabas que contengan, p. ej., una vocal larga y además una coda consonántica. Por eso el cálculo de las moras no consiste en un mero recuento del número de segmentos de la rima, ya que, en ese caso, no haría falta el nivel moraico (bastaría referirse al nivel segmental para calcular el peso de la sílaba).

También en otro sentido el peso no es una simple cuestión de duración fonética de la sílaba, sino que parece una cuestión de estructura. En efecto, la presencia de grupos de varias consonantes no suele añadir peso si estas se encuentran en el ataque (aunque *vid.* Topintzi, 2010, para algunas excepciones). Normalmente se asume que esos segmentos del ataque se unen directamente a la sílaba, como en B. Hayes (1989) (Figura 2.9). En cambio, para Hyman (1985) la consonante del ataque queda bajo la misma mora que el núcleo (Figura 2.10), en una configuración que recuerda a la de cuerpo y coda (*vid.* § 2.1.1.2).

Además de la inadecuación del recuento de segmentos para el cálculo del peso silábico, otro de los argumentos que se suelen usar a favor de la mora es la existencia de fenómenos de alargamiento compensatorio. En este tipo de procesos, la pérdida de un elemento de la rima de una sílaba (la coda consonántica) supuestamente provoca el alargamiento de otro elemento de la rima (la vocal nuclear), de modo que se mantiene el cómputo moraico de la sílaba (*vid.* Fox, 2000, pp. 71-76, para una comparación de distintos enfoques no lineales del proceso, con un nivel moraico, con un nivel CV, o con un nivel de segmentos X, sin distinción entre subtipos consonántico y vocálico). Por otro lado, en un enfoque gestual, la elisión de un segmento puede estar provocada por su pérdida de perceptibilidad si el gesto de otro segmento se solapa con él y lo enmascara. Desde esta perspectiva, la desaparición de un segmento no necesariamente es la causa del aumento de duración del otro. A la inversa, si este se alarga por cualquier otro motivo, ello puede tener *como consecuencia* la elisión del segmento contiguo. Obviamente, si el cambio toma carta de naturaleza en la forma fonológica de la palabra,

en la partitura gestual se especificará ya la mayor duración de un gesto, y no se incluirá el otro.

A pesar de los argumentos a favor de un nivel moraico, el carácter de constituyente de la mora es controvertido, y parece que puede tener más sentido considerar que se trata de un atributo o una característica de la sílaba, más que de una unidad en sí misma (Blevins, 1995; Fox, 2000; Van der Hulst y Ritter, 1999a). En efecto, Nam (2007) argumenta que el peso es una propiedad que se puede derivar del tipo de coordinaciones intergestuales que presenta una sílaba. Para ello parte de un análisis similar al que se ha presentado en el § 2.1.1.6, con la particularidad de que subdivide cada gesto consonántico en dos momentos: constricción y distensión, separados 180° el uno del otro, con lo que el segmento entero debe completar un ciclo de 360° en condiciones normales. En posición de ataque, aunque la constricción y la distensión son secuenciales, ambos momentos se especifican como coordinados en fase con la vocal. El efecto de competición resultante hace que los distintos gestos no lleguen a separarse tanto entre sí, sin completar el ciclo de 360°, y eso puede explicar que el ataque no aporte peso a la sílaba. En cambio, en posición de coda, la formación de la constricción es secuencial a la vocal nuclear, mientras que la distensión es, a su vez, secuencial con respecto a la constricción. Los 360° se completan y eso contribuye al peso silábico.

Nam sostiene también que, en las lenguas en que las codas no son moraicas (no aportan peso), hay que suponer una relación de coordinación adicional. Como antes, la constricción es secuencial a la vocal nuclear y la distensión es secuencial a la constricción. Sin embargo, ahora además se postula una relación secuencial entre la vocal y la distensión. El efecto de la competición entre estas coordinaciones hace que se contraiga un poco el conjunto y ocupe 120° entre cada gesto y el siguiente (240° en total). Este análisis también justificaría por qué la presencia de codas no moraicas es menos habitual, ya que implica una relación de antifase adicional, y el modo de antifase es, como se sabe, un modo de coordinación menos estable.

La Tabla 2.2 muestra, en forma esquemática, la diferencia entre las coordinaciones intergestuales propias de un ataque (no moraico), de una coda moraica y de una no moraica.

Este enfoque del peso silábico lo considera, por tanto, como una propiedad de la sílaba derivada de los tipos de coordinación intergestual especificados para dicha sílaba. Las vocales siempre aportan una mora a la sílaba, dado el carácter más lento y largo de

Tabla 2.2. Estructuras de coordinación intergestual y su contribución al peso silábico.

La línea continua indica coordinación en fase y la discontinua, en antifase.

Con = constricción; Dis = distensión; V = vocal.

	Ataque	Coda
Sílaba ligera (1 μ)	Con Dis ____ V -----	V Con Dis !.....!
Sílaba pesada (2 μ)		V Con Dis

los gestos vocálicos. Las consonantes solo aportan una mora si, en virtud de la fase relativa entre los distintos gestos, desarrollan un ciclo completo de secuencialidad (360°). Esto no resta poder descriptivo a la noción de peso silábico, que sigue siendo necesaria para explicar numerosos fenómenos, como se ha mencionado (p. ej. la construcción de pies y la asignación de acento). Simplemente, el pie (Σ) se forma seleccionando unidades del nivel σ (de uno u otro tipo, es decir, ligeras o pesadas). Bien visto, tiene sentido que la formación de Σ atienda a su nivel inmediatamente inferior (la σ), y no a otro hipotético aún más bajo (la μ).

En el § 2.3.3.3 veremos pruebas psicolingüísticas, basadas en el acceso al léxico, que parecerían confirmar la existencia de la mora como unidad. Los mecanismos de acceso al léxico se apoyan, según las lenguas, en la σ (cuando el ritmo es silábico), en el Σ (cuando el ritmo es acentual) o en la μ (cuando el ritmo es moraico). Dado el carácter de constituyentes de la σ y el Σ , cabría extender esta naturaleza también a la μ . No obstante, me gustaría objetar que los experimentos que se han hecho sobre el papel de la μ en el acceso al léxico en japonés se han basado exclusivamente en la coda nasal (ninguno en las geminadas), y no faltan quienes sostienen que la coda nasal moraica del japonés es en realidad una vocal nasalizada [ũ] (Yoshida, 1996), lo que justificaría que durara tanto como cualquier otra sílaba (porque, de hecho, sería una sílaba en sí misma).

Sea como sea, en cualquier caso, ningún autor —que yo conozca— ha hablado nunca de una posible vinculación de la mora con los fenómenos de refuerzo consonántico, que es el tema que nos interesa. Por este motivo, podemos permitirnos en este punto no decantarnos firmemente por la caracterización de la mora como constituyente o como propiedad silábica, y este apartado simplemente habrá servido —es de esperar— para plantear algunas de las distintas posturas a favor y en contra de su inclusión en la jerarquía prosódica.

2.1.2. El pie (Σ)

Símbolos y abreviaturas del pie

En los estudios que aluden al pie, este se suele representar de manera abreviada como F, Ft., ϕ , Σ . Debe advertirse que en este trabajo la letra ϕ simboliza la frase fonológica (*vid.* § 2.1.4). Para el pie aquí usaremos la sigma mayúscula (Σ).

2.1.2.1. El pie como constituyente

El nivel Σ está muy vinculado a la asignación de prominencias de acento léxico (*vid.* § 2.2.1), hasta el punto de que a veces se ha prescindido de la unidad Σ como constituyente y se ha intentado justificar sus efectos por las meras prominencias. Por ejemplo, Selkirk (1984b, 1986) descarta la necesidad de postular un nivel de constituyente Σ y argumenta que todos los fenómenos vinculados con él se pueden explicar como propiedades emergentes de la parrilla métrica de prominencias (*vid.* § 2.2.1).

Sin embargo, Beckman y Edwards (1994) sostienen que las prominencias no se pueden explicar sin referencia a un constituyente. Dado un constituyente de nivel n integrado por varios otros de nivel $n - 1$, una de esas unidades del nivel inferior resulta prominente por encima de las demás unidades de su mismo nivel. Esa unidad se convierte así en núcleo de la de nivel n y marca la presencia de esta, cumpliendo una función culminativa. Además, como se explicará en el § 2.2.1, la existencia —en el seno de una palabra— de dos niveles de prominencia acentual (primaria y secundaria) exige la participación de un constituyente intermedio entre la ω y la σ , y ese es el Σ .

Por otro lado, el pie es ámbito de aplicación de otros fenómenos distintos al acento. El inglés, por ejemplo, es una lengua en la que el pie juega un papel prosódico fundamental en muchos sentidos. Es el pie el que regula el ritmo acentual de la lengua y además hay estrategias de segmentación para el acceso al léxico que se apoyan en la identificación de pies (*vid.* § 2.3.3.2). Las oclusivas tensas presentan una distensión aspirada [p^h], [t^h], [k^h] cuando se encuentran en posición inicial de pie y el fonema /h/ solo aparece en esa misma posición (cualquier «h» gráfica intervocálica en posición media de pie resulta muda), *vid.* § 3.3.1. Por último, la formación de consonantes ambisilábicas conducentes al aleteo siempre se produce en interior de pie (*vid.* § 2.1.1.4).

Todos estos fenómenos justifican la consideración del pie como constituyente.

2.1.2.2. La estructura interna del pie

Aunque en algunas lenguas los pies se construyen sin restricción alguna, en muchas otras siguen un patrón sistemático. B. Hayes (1995) propone que existe un inventario universal de tres posibles esquemas de pie y cada lengua opta por uno de esos tres tipos. Los esquemas básicos de pie se consideran binarios (*i. e.* constan de dos sílabas), pero existen construcciones unarias y ternarias, que se suelen explicar como reducciones o expansiones de aquellos. Los pies pueden ser o no sensibles al peso de las sílabas que los integran. Por otro lado, dada la estrecha relación entre el modo en que se construyen los pies y se asignan las prominencias acentuales, las descripciones de la estructura

interna de los pies suelen aludir a la distribución de dicha prominencia, que puede recaer sobre la sílaba inicial (en el caso de los troqueos) o la final (en los yambos). Aunque dejaremos el resto de la discusión sobre las prominencias para el § 2.2.1, el esquema presentado en (8) recoge la variación tipológica propuesta para la estructura interna de los pies³⁷:

(8) Los tres tipos de pie según B. Hayes (1995)

troqueo silábico: $\acute{\sigma} \sigma$

dos sílabas, con prominencia en la primera

troqueo moraico: $\acute{\cup} \cup | \acute{-}$

dos sílabas ligeras, con prominencia en la primera; o una sílaba pesada

yambo: $\cup \acute{\sigma} | \acute{-}$

dos sílabas (la primera ligera), con prominencia en la segunda (ligera o pesada); o una única sílaba pesada

σ = una sílaba cualquiera, independientemente de su carácter ligero o pesado

$\cup$ = una sílaba ligera

$-$ = una sílaba pesada

2.1.2.3. Posibles infracciones de la estratificación rigurosa

En algunas lenguas es necesario suponer la existencia de pies que incluyen tres sílabas, p. ej. en inglés la palabra *logical* /'lɒdʒɪkəl/ consta de un pie ternario. Algunos autores, como Nespor y Vogel (1994 [1986]), consideran que la constitución interna de las unidades prosódicas no necesariamente es binaria (no existe una jerarquía entre las distintas unidades del nivel inferior, más que la que otorga el efecto de prominencia: si hubiera cuatro sílabas en el pie, habría una fuerte y tres débiles, pero estas estarían al mismo nivel). Otros investigadores, en cambio, abogan por respetar el esquema binario de las clases de pies presentado en (8), y eso los obliga a adoptar un tipo de constituyente intermedio, llamado superpié o colon (*vid.* referencias en Martínez-Paricio, 2012). Sin embargo, esta autora sostiene que no es necesario suponer la existencia de un nuevo tipo de constituyente intermedio, sino que el binarismo se puede respetar si se asume la posibilidad de construir los pies de manera recursiva, infringiendo parcialmente los principios de la estratificación rigurosa presentados en (4) en la p. 35.

Si se procede así, simplemente es necesario diferenciar entre la proyección mínima y máxima de un determinado constituyente (en este caso, un nivel Σ puede ser $\Sigma^{\min}$ o $\Sigma^{\max}$). Es una proyección mínima aquella que no domina otra de su misma clase, y una proyección máxima es la que no es dominada por otra de su misma clase (Figura 2.11).

³⁷ Los pies sensibles al peso suelen definirse en términos de moras. Aquí, sin embargo, ha parecido preferible establecer las correspondencias con sílabas (aun indicando su carácter ligero o pesado), por dos motivos. En primer lugar, porque en el § 2.1.1.7 hemos expuesto serias dudas acerca de si la mora es un constituyente en sí mismo, o más bien un atributo de la sílaba. En segundo lugar, porque parece lógico que, dada la Hipótesis de Estratificación Rigurosa (*vid.* § 2.1), la formación de una unidad de un nivel se especifique haciendo alusión a unidades del nivel inmediatamente inferior (aunque sea especificando cualidades de esas unidades), pero no a las de niveles aún más bajos.

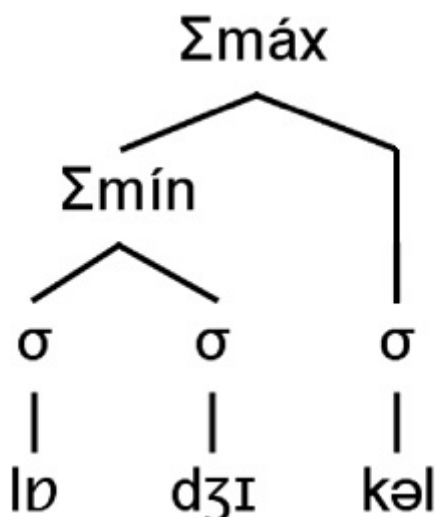


Figura 2.11. Estructura de un pie ternario

Lógicamente, cuando se respeta la restricción de No Recursividad, la proyección mínima y máxima coinciden.

2.1.3. La palabra fonológica (ω)

La palabra fonológica también se denomina palabra prosódica. A lo largo de este trabajo, se usarán ambas etiquetas indistintamente. Asimismo, se mencionará simplemente como palabra en contextos en los que quede claro a qué se hace referencia, para evitar un exceso de repetición. Dado que este es un estudio sobre prosodia, lo normal es que, cuando la etiqueta *palabra* no vaya acompañada de ninguna especificación, se trate de la palabra prosódica. Para la palabra morfosintáctica normalmente se utilizará esta precisión, a menos que resulte innecesaria y se desprenda del contexto.

Símbolos y abreviaturas de la palabra fonológica

La palabra fonológica se representa en la bibliografía como W, PhW, PW, PrWd (y otras combinaciones por el estilo) o con la letra griega ω (aquí se usará este último símbolo).

2.1.3.1. La constitución de la palabra fonológica

Los niveles ω y φ conforman puntos de interfaz entre la prosodia y la morfosintaxis, de modo que en la constitución de estas unidades se hace alusión a las proyecciones sintácticas (mínimas, $X^{\min}$, o máximas, $X^{\max}$) de las categorías léxicas. Sin embargo, no hay que olvidar que ω y φ no necesariamente son isomórficas con los constituyentes sintácticos a los que deben referirse para su formación, y su manera de construirse atiende reglas que son fonológicas.

La palabra prosódica tiende a coincidir, al menos en uno de sus dos extremos, con un límite de palabra morfosintáctica. Según el algoritmo propuesto por Selkirk (1986), el punto de alineamiento es el límite inicial o final de la proyección mínima de una

categoría léxica en la sintaxis, de ahí que la información necesaria no sea puramente morfológica sino morfosintáctica. Por ejemplo, en lenguas como el español, de recursividad sintáctica a la derecha, se puede establecer la siguiente regla de formación de ω (verbalización mía a partir de las propuestas de Selkirk).

(9) Alineación derecha de la palabra morfosintáctica y la palabra fonológica

En cada final de $X^{\min}$ coloca un límite $]_{\omega}$

Por su parte, Nespor y Vogel (1994 [1986]) aducen que el grado de correspondencia entre la palabra prosódica y la palabra morfosintáctica puede variar también según el tipo de constitución morfológica de esta. La correspondencia entre ω y Q puede ser absoluta, como en latín y en griego, por ejemplo. En otras lenguas, como el español, los compuestos sintagmáticos (p. ej. *salón comedor*, *coche cama*, *mar de fondo*, *caja fuerte*, *media naranja*, *centro-izquierda*, etc.) y otros modismos o expresiones pluriverbales lexicalizadas (p. ej. *estirar la pata*, *saltar a la vista*), aun formando una sola unidad léxica, corresponden a más de una palabra fonológica (una por cada raíz morfológica con sus afijos y clíticos), como demuestra la presencia de más de un acento primario. Lo mismo puede decirse de los adverbios en *-mente*. Hay otras lenguas en las que ese es el comportamiento de todos los compuestos (no solo los sintagmáticos, sino también los léxicos); el turco es un ejemplo de ello. Finalmente, en lenguas como el húngaro, los prefijos forman palabra fonológica por sí mismos; cada raíz, otra; y la última raíz junto con los sufijos, otra.

2.1.3.2. Posibles infracciones de la estratificación rigurosa

Peculiaridades fonotácticas de los extremos de palabra

En el § 2.1.1.3 se adelantó que, en muchas lenguas, el comportamiento fonológico de determinadas secuencias segmentales y las restricciones fonotácticas son diferentes en interior de palabra o en sus extremos (inicial o final). Por ejemplo, en las posiciones periféricas de la palabra podemos encontrar secuencias de sonoridad inversas a las esperadas. Uno de los casos más conocidos lo constituyen los grupos consonánticos $s + C$ en ataque inicial de palabra, lo que ha llevado a decir que esa $/s/$ es extramétrica, o que es parte de una semisílaba previa (*i. e.* una sílaba degenerada, sin núcleo ni ninguna mora), *vid.* Goad (2011) y varios de los capítulos incluidos en Féry y Van de Vijver (2003):

In English [...] $/s+stop/$ clusters are unique in many ways. First, these are the only clusters where the second consonant may be an obstruent. Second, these are the only cases of onsets with three members (*str* vs. $*ftr$). Third, they are immune to the restrictions on homorganicity between the members of the onset consonants (*sl*, *sn* vs. $*dl$, $*pw$). Finally $/s/$ is the only obstruent that can be combined with a nasal (*sm* vs. $*tm$, $*fn$) (Y.-M. Y. Cho y King, 2003, p. 185).

Es interesante recalcar que los grupos $s + C$ se analizan como heterosilábicos en interior de palabra.

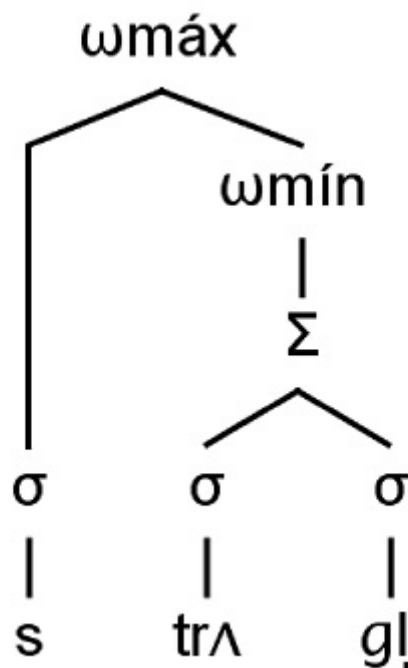


Figura 2.12. Semisílaba en la periferia de la palabra

Por otro lado, la posición final de palabra también aloja grupos consonánticos de coda diferentes a los permitidos en su interior, lo que ha motivado múltiples propuestas de análisis: como coda, como una posición silábica especial —el *apéndice*—, como semisílaba, como segmentos extramétricos (*vid.* la revisión de Côté, 2011). Sin duda, uno de los análisis más interesantes para el apartado que nos ocupa es aquel que supone la existencia de un nivel $\omega^{\text{máx}}$ que domina a la vez a una $\omega^{\text{mín}}$ y a una σ , sin intermediación de Σ en este caso (Figura 2.12). Esta configuración supone a la vez una infracción de la restricción de Exhaustividad y de la de No Recursividad, (4):c-d (p. 35).

Sin embargo, Côté (2011) no ve ninguna necesidad de recurrir a la presencia de semisílabas, sino que los segmentos en cuestión deben analizarse como parte de la coda [o del ataque, según corresponda]. Ella encuentra una explicación natural a las peculiaridades de la periferia de las palabras en los procesos de refuerzo consonántico propios de las posiciones de frontera de ω (*vid.* § 3.1). Los alargamientos característicos de estas posiciones hacen que los gestos articulatorios se solapen menos unos con otros, y eso puede hacer más perceptibles a determinados segmentos, que, en cambio, no se toleran en posición interior: «the higher the prosodic boundary a consonant is adjacent to, the more easily it surfaces without the support of an adjacent vowel; consonants not adjacent to any prosodic boundary are the weakest» (*op. cit.*, p. 863). La autora aporta más ejemplos, como el caso de las oposiciones laríngeas en *tojolabal*³⁸, que se neutralizan en coda interna de palabra pero no en final de palabra.

³⁸ En todo este trabajo sigo los glotónimos propuestos por Moreno Cabrera (2003).

Además, parece que también hay que buscar en la morfología algunas de las causas de la mayor permisividad con determinadas consonantes en posición final. Estas consonantes, precisamente, suelen coincidir con morfemas (sobre todo, flexivos), y deben preservarse para que no se pierda la información morfológica que aportan. También es más fácil mantener grupos consonánticos cuya última consonante se resilabea (y refuerza) en inicio de sílaba cuando va seguida de sufijo que empieza por vocal. Por ejemplo, la existencia de *helping* y *helper* ayuda a la estabilidad del grupo /lp/ en *help*.

Ahora bien, los límites de palabra no siempre son más permisivos que las posiciones interiores. En español, por ejemplo, la percusiva /r/ está prohibida en inicio de palabra, pero no en ataque interior. Sin embargo, propongo que hay una posible explicación siguiendo la misma línea de razonamiento que se acaba de plantear: precisamente, el alargamiento de inicio de palabra es incompatible con la /r/, pues la brevedad es característica de las percusivas. Lógicamente, esta exclusión está fonologizada en español, puesto que no entraña ningún problema pronunciar un sonido similar en inicial de palabra, por ejemplo, en japonés. En cualquier caso, es razonable suponer que ese patrón fonotáctico encuentra una base fonética en los procesos de refuerzo consonántico condicionados por la posición prosódica.

Los pronombres enclíticos

Aunque Nespor y Vogel (1994 [1986]) proponen la inclusión en la jerarquía prosódica de un nivel de constituyente llamado *grupo clítico*, aquí seguiré un análisis basado en las propuestas de Selkirk (1996). Las propiedades prosódicas de las palabras gramaticales inherentemente inacentuadas se pueden explicar asumiendo que es posible infringir algunas de las restricciones asociadas a la estratificación rigurosa, en concreto, las restricciones de Exhaustividad y de No Recursividad, *vid.* (4):c-d.

Observemos en la Tabla 2.3 el siguiente paradigma de formas verbales con pronombres enclíticos (*i. e.* no acentuados, posverbiales)³⁹.

Se han resaltado en negrita las formas que contravienen el patrón habitual de asignación del acento (*vid.* § 2.2.1). Por ejemplo, las dos formas de gerundio seguido de dos clíticos son sobreesdrújulas (anteproparoxítonas). Exceptuando el caso de los

Tabla 2.3. Formas verbales con uno y dos pronombres enclíticos (con variación en el carácter de sílaba ligera o pesada del primero —o único— de ellos).

En negrita, las formas que contravienen el patrón habitual de asignación del acento.

		2 Pl Imperativo ⁴⁰	Infinitivo	Gerundio
1 clítico	cl. ligero	contadme	contarme	contándome
	cl. pesado	contadnos	contarnos	contándonos
2 clíticos	cl. ligero	contádmelo	contármelo	contándomelo
	cl. pesado	contádnoslo	contárnoslo	contándonoslo

³⁹ En las variedades del español habladas en Galicia y Asturias la aparición de enclíticos no está limitada al imperativo, infinitivo y gerundio, sino que se admiten también otras formas, p. ej. *contábaselo*, *cerrarémosla*.

⁴⁰ La 2 Sg Imperativo sigue un comportamiento diferente y será analizada en el § 2.1.4.2.

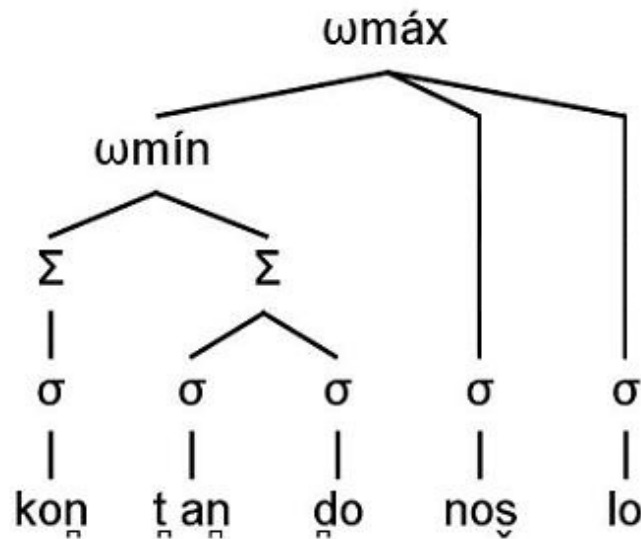


Figura 2.13. Estructura con enclíticos

gerundios con dos klítics, todas las otras formas de palabra en español tienen limitada la posición del acento a una de las tres últimas sílabas. Por otro lado, las formas *contádnoslo* y *contárnoslo* también son anómalas, pues en el resto del idioma no puede haber palabras esdrújulas (proparoxítonas) si la penúltima sílaba es pesada.

Todo esto parece indicar que los klítics se hallan fuera del ámbito de asignación del acento léxico, que es la ω . Se trata de sílabas «hermanas» al nivel ω , pero dominadas directamente por el nivel que domina a ω . Ahora bien, no podemos suponer que esas σ estén dominadas por ϕ , porque, como se verá en el § 2.1.4.2, hay una regla que alinea todos los límites $[_\phi$ con un límite $[_\omega$ y, si no existe, crea el límite $[_\omega$ para poder satisfacer ese requisito. Si ese fuera el caso, la creación de una nueva unidad ω asignaría acento léxico al pronombre; por tanto, debemos descartar esta opción. Hay que hacer que la σ klítica sea hermana de ω pero que su límite derecho ya coincida con una ω , para evitar que la regla anteriormente referida cree una nueva unidad. En definitiva, la estructura que propongo es una $\omega^{\text{máx}}$ que domina a una $\omega^{\text{mín}}$ (la forma verbal) y a una o varias σ (los klítics) (Figura 2.13). Esto supone una infracción de la restricción de No Recursividad, porque hay una capa ω que domina a otra capa ω , y también una vulneración de Exhaustividad, porque hay una ω que domina una σ sin pasar por el nivel Σ (si hubiera un Σ , habría acento secundario). Si se postula que el ámbito de asignación del acento léxico es la $\omega^{\text{mín}}$, para el momento de la adjunción de los klítics, la posición del acento ya se habrá fijado y estos no podrán modificarla⁴¹.

⁴¹ Algunos autores, p. ej. Schiering, Hildebrandt y Bickel (2007), a partir de estudios tipológicos, ponen en cuestión la universalidad de los niveles propuestos por la fonología prosódica, y es, precisamente, el nivel de la palabra prosódica (ω) uno de los que más dudas les plantean. No puedo entrar a valorar los datos que manejan, porque desconozco los detalles de las lenguas en que se apoyan. Sin embargo, una de las críticas que reiteran es que la construcción recursiva del nivel ω lleva a predicciones que se ven desmentidas empíricamente, pues —dicen— eso conllevaría aplicar de manera también recursiva las reglas propias de ese nivel a todas las proyecciones del tipo ω , pero los datos lingüísticos muestran que no es el caso. En contraste con esta idea, la propuesta que yo hago aquí es que hay reglas que pueden

2.1.4. La frase fonológica (φ)

En esta zona de la jerarquía prosódica es donde existen más discrepancias acerca del número de niveles de constituyentes necesarios y sus denominaciones. Para Selkirk (1986), por ejemplo, existen dos tipos de frase: la frase mayor y la frase menor. Aunque, en su proceso de construcción, ambas se refieren a una proyección máxima de la sintaxis, se diferencian en que la frase mayor puede incluir tanto la posición de especificador como los complementos de esa $X^{\text{máx}}$, mientras que la frase menor no puede incluir los complementos. Por su parte, Beckman y Pierrehumbert (1986) dicen que una frase entonativa completa puede estar compuesta por varias frases (entonativas) intermedias, que a su vez constan de frases acentuales. Shattuck-Hufnagel y Turk (1996) señalan una posible equivalencia entre la frase intermedia y la frase mayor, por un lado, y entre la frase acentual y la frase menor, por otro. Sin embargo, esta asociación, seguramente, no es del todo correcta, pues, si la frase acentual se caracteriza por la presencia de una palabra con un único acento primario y sus posibles clíticos, entonces en nuestro análisis equivale a la $\omega^{\text{máx}}$ cuando hay enclíticos (*vid.* § 2.1.3.2).

Símbolos y abreviaturas de la frase fonológica

En este caso, junto al símbolo φ , que usamos aquí, y otras abreviaturas como PP o PhP, conviven las representaciones de los constituyentes emanados de las distintas propuestas (MaP = frase mayor; MiP = frase menor; iP = frase intermedia; AP = frase acentual; para la frase acentual también se usa una α).

2.1.4.1. La constitución de la frase fonológica

Como es de esperar, una multiplicidad de propuestas de constituyentes —con diferentes matices, como acabamos de ver— lleva aparejada una diversidad de algoritmos de constitución, *vid.* Nespor y Vogel (1994 [1986]); Selkirk (1986), para algunos de los más conocidos; estas alternativas y otras se debaten en muchos de los trabajos recogidos en Inkelas y Zec (1990).

La frase fonológica, según Selkirk (1986), se alinea con alguno de los límites de la proyección máxima (en la sintaxis) de alguna categoría léxica. Dicho alineamiento se puede producir a la derecha o a la izquierda, según las lenguas, y se ha propuesto que coincide con el lado contrario a aquel hacia el que es recursiva la sintaxis en cada lengua (Nespor y Vogel, 1994 [1986]). Como consecuencia de este modo de constitución, el otro extremo de la frase fonológica suele estar ocupado por palabras de alguna de las categorías gramaticales. En el § 2.3.2 veremos que esto tiene una influencia positiva sobre los procesos de adquisición del léxico, y también sobre la segmentación del continuo sonoro para el reconocimiento de las palabras.

especificar como ámbito de aplicación una proyección mínima o una proyección máxima de un determinado nivel prosódico. En el ejemplo que acabamos de ver, la asignación del acento léxico se circunscribe a $\omega^{\text{mín}}$ y ya no vuelve a aplicarse en las capas superiores del tipo ω , en el caso de que las haya. Así, hay reglas que afectan a las proyecciones mínimas; otras, a las proyecciones máximas; y es posible que otras reglas afecten a todas las proyecciones de un mismo tipo (en efecto, se predice que, si una regla afecta a una proyección intermedia —ni mínima ni máxima—, entonces afectará a todas las proyecciones de ese nivel prosódico).

En las lenguas como el español, de recursividad sintáctica a la derecha, la regla de formación de φ se puede enunciar como en (10) (verbalización mía a partir de las propuestas de Selkirk).

(10) Alineación izquierda del sintagma y la frase fonológica

En cada inicio de $X^{\text{máx}}$ coloca un límite $[_\varphi$

En algunas lenguas existe la opción de que lo que en principio deberían ser dos φ separadas se constituyan en una sola φ si la segunda solo contiene una ω y además es el complemento sintáctico de la primera por el lado recursivo —en este caso, la derecha— (Nespor y Vogel, 1994 [1986]; *vid.* también Selkirk, 2000). Esto da lugar a estructuras prosódicas como la de (11), donde las fronteras de las frases «de arriba» y «goteras», dibujadas en trazo difuso, son opcionales.

(11) La vecina de arriba tiene goteras de difícil solución

$[_\varphi$ La vecina $[_\varphi$ de arriba $[_\varphi$ tiene $[_\varphi$ goteras $[_\varphi$ de difícil solución

El algoritmo se completa cerrando la frase con un límite $]_\varphi$ justo en la posición precedente a la siguiente apertura de frase (o en el final del enunciado).

2.1.4.2. Posibles infracciones de la estratificación rigurosa

Los proclíticos

Las palabras inherentemente inacentuadas —como el artículo determinado o las preposiciones propias (*vid.* Quilis, 1993, pp. 390-395)— no pueden estar dominadas por Σ ni formar su propia ω , o de lo contrario adquirirían acento léxico. Selkirk (1996) las trata como sílabas que se adjuntan directamente a la φ . Aunque ello suponga una violación de la Exhaustividad, es la opción menos mala, según su análisis. Además, eso provoca la alineación izquierda de la palabra morfosintáctica y ω , que es un resultado bienvenido. Si añadimos al enunciado de (11) la formación de ω resultante de la regla (9) y dejamos los proclíticos como σ dominadas directamente por φ , nos queda la estructura de (12). La Figura 2.14 muestra el esquema arbóreo de la última φ de este enunciado.

(12) La vecina de arriba tiene goteras de difícil solución

$[_\varphi$ La [vecina] $[_\omega$] [de [arriba] $[_\omega$]] [tiene] $[_\omega$]] [goteras] $[_\omega$]] [de [difícil] $[_\omega$] [solución] $[_\omega$]] $]_\varphi$

Alternativamente, se podría proponer una imagen especular a la que hemos planteado para los enclíticos (*vid.* § 2.1.3.2), con una $\omega^{\text{máx}}$ que domina las σ proclíticas más una $\omega^{\text{mín}}$ correspondiente a la palabra de contenido léxico. Dejo para el futuro la comparación crítica de estas dos alternativas (Lahoz-Bengoechea, en preparación), aunque a priori parece que la inclusión de los proclíticos bajo φ es preferible, puesto que

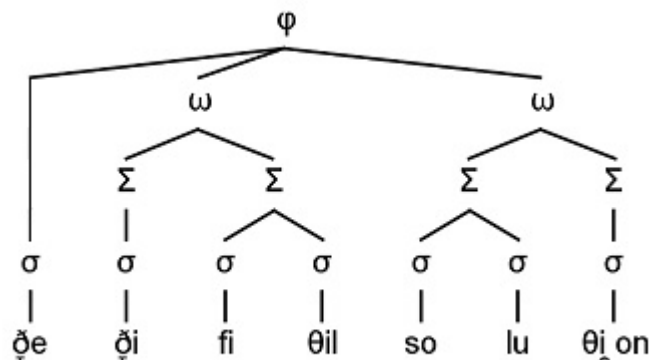


Figura 2.14. Estructura con proclíticos

solo infringe Exhaustividad, mientras que bajo $\omega^{\text{máx}}$ infringe Exhaustividad y No Recursividad.

Los falsos enclíticos

En el § 2.1.3.2 se presentó un análisis de los pronombres enclíticos como σ fuera del dominio $\omega^{\text{mín}}$, donde se asigna el acento léxico, y dominados directamente por una $\omega^{\text{máx}}$. De este modo se puede justificar formalmente el carácter no acentuado de estos pronombres cuando siguen a un infinitivo, a un gerundio, o a un imperativo en 2.^a persona del plural. Ahora bien, por alguna razón, en algunas variedades de español —aunque no en todas— esos mismos pronombres sí adquieren algún tipo de prominencia cuando van pospuestos a un imperativo de 2.^a persona del singular (*vid.* Tabla 2.4)⁴².

En apartados anteriores se han presentado las reglas de constitución de ω (9) y de ϕ (10), que se repiten aquí como (13) y (14) para facilitar la lectura. La aplicación de estas reglas hace que suele coincidir el límite final de una ϕ con el límite final de una ω . Esta coincidencia provoca que normalmente no se adviertan los efectos de otra regla, formulada por Selkirk (1996) como restricción en el marco de la teoría de la optimidad, que aquí expresamos en (15).

(13) Alineación derecha de la palabra morfosintáctica y la palabra fonológica

En cada final de $X^{\text{mín}}$ coloca un límite $]_{\omega}$

(14) Alineación izquierda del sintagma y la frase fonológica

En cada inicio de $X^{\text{máx}}$ coloca un límite $[_{\phi}$

(15) Alineación derecha de la frase fonológica y la palabra fonológica

En cada final de ϕ coloca un límite $]_{\omega}$

⁴² La misma situación se produce con las formas verbales de subjuntivo (3 Sg, 3 Pl; 1 Pl) que se usan en oraciones imperativas: p. ej. *cuénteselo* ['kuentese'lo], *cuéntenselo* ['kuentense'lo], *contémoselo* [kon'temose'lo].

Tabla 2.4. Formas verbales de 2 Sg Imperativo con uno y dos pronombres pospuestos.

2 Sg Imperativo			
1 clítico	cl. ligero	cuéntame	[¹ kuent̪a ¹ me]
	cl. pesado	cuéntanos	[¹ kuent̪a ¹ nos]
2 clíticos	cl. ligero	cuéntamelo	[¹ kuent̪ame ¹ lo]
	cl. pesado	cuéntanoslo	[¹ kuent̪anos ¹ lo]

La aplicación de esta regla (15), que hasta ahora resultaba vacía, cobra relevancia en el análisis de estos pronombres posverbiales. Por algún motivo —que aún hace falta dilucidar (Lahoz-Bengoechea, en preparación)—, los pronombres que se posponen a formas de 2 Sg Imperativo no se insertan en estructuras de $\omega^{\text{máx}}$, hermanados con $\omega^{\text{mín}}$, sino que quedan como σ bajo el dominio directo de φ . Si bien la regla general de formación de ω (13) no puede otorgar estatus de ω al pronombre, sí lo hace la regla (15), al haber quedado un límite $]_{\varphi}$ sin alinear con uno $]_{\omega}$. Esto puede explicar el carácter acentuado el pronombre. Advuértase el contraste entre (16) y (17).

(16) Pinta la roja

$$\left[\begin{array}{c} \left[\text{Pinta} \right]_{\omega} \\ \left[\varphi \right] \end{array} \right] \left[\begin{array}{c} \left[\text{la} \right]_{\sigma} \left[\text{roja} \right]_{\omega} \\ \left[\varphi \right] \end{array} \right]$$

(17) Píntala roja

$$\left[\begin{array}{c} \left[\text{Pínta} \right]_{\omega} \left[\text{la} \right]_{\omega} \\ \left[\varphi \right] \end{array} \right] \left[\begin{array}{c} \left[\text{roja} \right]_{\omega} \\ \left[\varphi \right] \end{array} \right]$$

Si esta explicación es correcta, equivale a decir que el pronombre (que ya no se puede llamar clítico con propiedad) recibe un acento primario⁴³. Según el análisis que estoy proponiendo, esta forma consta de dos palabras prosódicas, cada una con su acento primario (este enfoque resulta coherente con el tratamiento de los adverbios en *-mente*, también con dos acentos primarios, como se ha explicado en el § 2.1.3.1).

Las formas con dos pronombres se analizarían del mismo modo, p. ej. *cuéntamelo* [¹kuent̪a]_ω [me¹lo]_ω. La razón por la que el conjunto pronominal recibe el acento en *lo* y no en *me* —cuando ninguno de los dos está léxicamente especificado para ello— puede estar relacionada con el hecho de que esta prominencia se adquiere como último recurso para satisfacer los requisitos de alineamientos de frontera y, por tanto, aparece en el elemento más próximo a dicha frontera. Además, la regla de (15) solo establece que los clíticos queden dominados por ω , pero no la existencia de un nivel Σ (troqueo) intermedio, que sí podría ser responsable de la asignación de prominencia al primer pronombre. La estructura prosódica correspondiente a estos falsos enclíticos supone, por tanto, una infracción de la condición de Exhaustividad (Figura 2.15).

⁴³ Vid. Navarro Tomás (1996 [1918], § 173) para ejemplos de usos métricos en la poesía española que, efectivamente, sitúan este acento al mismo nivel que el acento primario de cualquier palabra de contenido léxico.

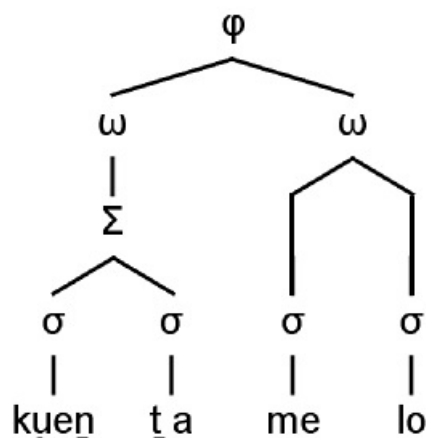


Figura 2.15. Estructura con falsos enclíticos

2.1.5. La frase entonativa (I)

Símbolos y abreviaturas de la frase entonativa

Algunas de las representaciones abreviadas de la frase entonativa incluyen IP, I, E, y la letra griega ι. Aquí usaremos la I mayúscula, como en la versión inglesa de Nespor y Vogel (1994 [1986]).

2.1.5.1. La constitución de la frase entonativa

Como indican Nespor y Vogel, por lo general, una frase entonativa coincide con una oración sintáctica, con su proposición principal y sus posibles proposiciones subordinadas. Sin embargo, hay algunas construcciones sintácticas que forman frase entonativa por sí mismas. Si se encuentran en un extremo de la oración, el resto por el extremo contrario forma, en principio, otra frase entonativa. Si se encuentran en medio, dividen a la oración por lo menos en tres frases: la del medio, y las de los dos extremos. Las construcciones que forman I por sí mismas incluyen los vocativos, interjecciones, y todo tipo de expresiones parentéticas, como proposiciones subordinadas adjetivas explicativas, las subordinadas llamadas de la enunciación, preguntas coetilla, o dislocaciones sintácticas. Según las lenguas, existen opciones de reestructuración en virtud de distintos factores sintácticos y, sobre todo, rítmicos.

2.1.6. El enunciado fonológico (U)

A lo largo de este trabajo, normalmente nos referiremos a él como enunciado, simplemente, por aligerar la expresión.

Símbolos y abreviaturas del enunciado fonológico

Las posibilidades más habituales son Utt., U, EF, v. Aquí usaremos U.

2.1.6.1. El enunciado fonológico como constituyente

Algunas propuestas prosódicas, como la de Beckman y Pierrehumbert (1986), no incluyen el enunciado como un nivel distinto de la jerarquía, y lo equiparan a la frase

entonativa. Sin embargo, el enunciado es necesario, por ejemplo, para explicar por qué en español se produce resilabeo a través de frontera][_i pero no a través de frontera][_u, como refleja la distinta pronunciación de la /r/ de coda: [r] en (18), con resilabeo, [ɾ] en (19), sin él.

(18) Pilar_i, ¿has visto a Joaquín esta mañana?

(19) Este año nos vamos de vacaciones al mar_i. A propósito, ¿tú dónde te vas?

Según Nespor y Vogel (1994 [1986]), el enunciado se corresponde con el nudo sintáctico más alto, aunque también existen posibilidades de reestructuración que fusionan dos enunciados en uno, en función de determinados factores sintácticos, semánticos y discursivos.

2.2. La prominencia prosódica y su jerarquía

Cuando dos o más constituyentes de un determinado nivel de la jerarquía se juntan para formar una unidad del nivel inmediatamente superior, se establece un patrón de prominencias de modo que uno de aquellos se considera más fuerte que el resto de sus hermanos y se convierte en núcleo del constituyente mayor. Esta afirmación implica, para empezar, dos cosas. En primer lugar, que las prominencias se definen siempre de manera relativa entre constituyentes del mismo nivel. En segundo lugar, que la presencia de un constituyente fuerte (núcleo) apunta necesariamente a la existencia de una unidad del nivel superior. El núcleo cumple así una función culminativa, entendida como en Trubetzkoy (1973 [1939]).

La alternancia métrica entre unidades fuertes y débiles se encuentra en la base del ritmo, que es la recurrencia de elementos a intervalos más o menos regulares. El ritmo, por tanto, es un fenómeno estratificado (Jun, 2005); es decir, hay alternancias rítmicas en todos los niveles de la jerarquía de constituyentes (si bien es cierto que la clasificación tipológica del ritmo de las lenguas siempre se ha basado en los niveles subléxicos)⁴⁴. Así, el hecho de que la estructura de prominencias dependa de la estructura de constituyentes permite que aquella ayude a poner a esta de manifiesto. En otras palabras, si la unidad prominente desempeña una función culminativa, entonces el ritmo es clave, en la medida en que pauta la sucesión de unidades de cada nivel. Fox (2000) también escribe a propósito de esta estratificación:

[T]he same categories, strong and weak, are used to label the whole tree, in spite of the fact that different units are involved at different levels; s(trong) and w(eak) are, of course, purely ‘relational’ labels, with no intrinsic content, and it is this which allows the use of these categories at different levels. They represent not phonetic features but *structured* categories. In fact, the s category represents nothing more than the *head* constituent at a particular level of structure (p. 347).

Este párrafo recoge otra idea fundamental, y es que, al ser una propiedad estructural, la prominencia carece de contenido fonético propio (aunque en el § 3.2 veremos que existe una serie de marcas fonéticas que suelen contribuir a señalar dicha prominencia). Desde el desarrollo de la fonología métrica (M. Liberman, 1975; M. Liberman y Prince, 1977; Prince, 1983), los fenómenos de prominencia, como p. ej. el acento léxico, no se consideran un rasgo distintivo que sirva para establecer oposiciones paradigmáticas, sino una propiedad de una relación sintagmática⁴⁵. Además, esa prominencia es,

⁴⁴ Vid. § 2.3.3 para una explicación de estos tipos rítmicos y su repercusión sobre el acceso al léxico y *vid.* también la discusión en el § 7.2 para una visión más moderna sobre la tipología rítmica de las lenguas.

⁴⁵ Esta manera de proceder supone un gran avance con respecto al tratamiento estructuralista y del primer generativismo, con el modelo SPE de Chomsky y Halle (1976 [1968]). En SPE, por ejemplo, el acento léxico era un rasgo anómalo por varios motivos. En primer lugar, no era binario, sino que se distinguían varios niveles acentuales (en vez de [$\pm$ acento] tenemos [1 acento], [2 acento], [3 acento], etc.). Además, el acento es en parte predecible contextualmente: si una sílaba es acentuada, las sílabas adyacentes (dentro de la misma palabra) no pueden serlo. De hecho, el acento no se asimila, también a diferencia de la mayoría de los rasgos. Dentro del dominio acentual de la palabra, solo una sílaba puede llevar acento primario (esto no pasa con los demás rasgos fonológicos). Por otro lado, en una época en la que la

fundamentalmente, de naturaleza abstracta: no está ligada a la aparición de una característica fonética concreta (p. ej. un determinado nivel absoluto de intensidad), sino que se basa en la comparación entre unidades hermanas. Así, una unidad solo puede interpretarse como relativamente débil por contraste con otra más fuerte, que encabeza el constituyente mayor del que forman parte. En definitiva, como dice B. Hayes (1995):

stress is parasitic, in the sense that it invokes phonetic resources that serve other phonological ends. Partly as a consequence of this, stress is phonetically realized on a language-specific basis (p. 7).

En la construcción de la jerarquía de prominencias, el constituyente sílaba tiene un carácter especial. Por un lado, aunque está claro que la sílaba tiene un núcleo —que es su elemento más prominente—, no presenta una prominencia interna en el mismo sentido que lo hacen los otros niveles (p. ej. si se adopta una estructura interna jerárquica, como en el § 2.1.1.2, los constituyentes internos de la sílaba no son «hermanos»). En consonancia con esta idea, B. Hayes (1995) afirma: «the smallest constituent in metrical structure is the foot» (p. 2). Por otro lado, la asignación de prominencias opera de manera acumulativa, de modo que la fuerza siempre recae, en última instancia, sobre una sílaba. Por ejemplo, cuando se forma un pie, una de sus sílabas será la prominente. Este primer grado de prominencia se corresponde con el acento léxico secundario. A su vez, una palabra prosódica puede estar formada por varios pies, de entre los que destacará uno. En este caso, el pie fuerte es el responsable de la asignación del acento primario pero, aunque la prominencia se establece entre las unidades del nivel pie que componen la palabra, el acento no recae sobre el pie en su totalidad, sino sobre su sílaba prominente. Si sobrepasamos el nivel de la palabra, varias de estas pueden formar una frase fonológica, pero habrá una que destaque sobre las demás. Pues bien, será la sílaba con acento primario de esa palabra la que también reciba el acento de frase correspondiente al siguiente nivel de prominencia. Por último, de entre todas las sílabas con acento de frase, habrá una con acento nuclear, que será la que corresponda a la frase fonológica más fuerte dentro de la frase entonativa. Es decir, que un determinado nivel de prominencia siempre se asigna entre las unidades ya prominentes del nivel inferior.

De la explicación anterior se desprende que podemos distinguir cuatro niveles de prominencia: dos corresponden al acento léxico (secundario y primario) y dos, al acento de frase (prenuclear y nuclear). Cabe advertir aquí que no es infrecuente encontrarse con un problema de traducción de estos términos a partir del inglés, especialmente cuando se hace equivaler indistintamente *acento* (sin más) a *stress* y *accent*. En esta lengua, *stress* suele utilizarse para significar acento léxico (o, en algunas visiones, como la de Beckman, 1986, un tipo concreto de acento léxico), aunque algunos autores lo emplean también en el sentido acumulativo de prominencias de distinto nivel. Por otro lado, *accent* suele equivaler al acento de frase, pero también hay investigadores que utilizan

fonología autosegmental había abierto la puerta a la existencia de una jerarquía entre los rasgos (o ahora autosegmentos) definitorios de los segmentos, tiene sentido que la fonología métrica postule la necesidad de una jerarquía también entre las unidades —y las prominencias— prosódicas, ya que esto dota de coherencia a la teoría fonológica.

el término para aludir al carácter sintagmático de cualquier contraste de prominencia. Para terminarlo de complicar, a veces se emplea *phrase accent* para referirse no a un tipo de prominencia, sino al tono que puede aparecer, con función demarcativa, en el final de una frase entonativa intermedia (del mismo modo que se habla de tonos de frontera en el final de una frase entonativa plena). Al acento de frase también se lo llama *pitch accent* (acento tonal), porque una de sus características es que las sílabas prominentes en los niveles frásticos son puntos de anclaje para el alineamiento entre el contenido segmental y los contornos tonales adecuados, escogidos del inventario entonativo de cada lengua en función de factores que incluyen la modalidad oracional, la estructura informativa y una serie de valores epistémicos como la obviedad, la duda, la sorpresa, etc. (siguiendo los presupuestos del modelo métrico-autosegmental: *vid.* Beckman y Pierrehumbert, 1986; Pierrehumbert, 1980; para una revisión más actual, Ladd, 1996, y su 2.^a edición, corregida y aumentada, del año 2008). Sin embargo, la etiqueta *pitch accent* también se aplica para describir un tipo de sistema de prosodia léxica de algunas lenguas como el japonés⁴⁶. El acento léxico (*stress*) es compatible con la presencia de tono léxico (Hyman, 2011; Ladd, 2008, pp. 164-168, y especialmente su tabla 4.1), pero no es esa una propiedad intrínseca del *stress*, sino que este se caracteriza por otras claves relacionadas con la duración y la intensidad, fundamentalmente, y, en algunas lenguas, como el inglés, con una diferencia de timbre en la vocal (Beckman, 1986).

En español, por otro lado, parece que la tradición ha usado de manera intercambiable las expresiones *sílabas acentuadas* y *sílabas tónicas*, frente a *sílabas no acentuadas* y *sílabas átonas*, que también serían sinónimas entre sí. Sin embargo, teniendo en cuenta la explicación anterior, la alusión a la presencia o ausencia de tono no es procedente en el nivel del acento léxico para una lengua como el español, donde los tonos son posléxicos y se alinean únicamente con los acentos de frase (para un enfoque métrico-

⁴⁶ Efectivamente, muchos autores clasifican el japonés como una lengua de acento melódico (en inglés, *pitch accent*), *vid.* p. ej. Beckman (1986). Según esta visión, el tratamiento del japonés como una lengua con acento léxico se justifica por el contraste sintagmático creado entre una sílaba —especificada léxicamente para ser portadora de un determinado contorno tonal (en concreto, un tono alto seguido de un tono bajo en la sílaba siguiente)— y las demás sílabas de la palabra, que solo adquieren su tono mediante reglas posléxicas. La supuesta sílaba acentuada no manifiesta, sin embargo, el tipo de marcas fonéticas propias del acento dinámico (en inglés, *stress*), que vienen provocadas por una modificación en la coordinación y en la dinámica gestuales (*vid.* § 3.2), y que consisten, fundamentalmente, en un aumento de la duración silábica y de la intensidad de su núcleo vocálico. En cambio, para Hyman (2011), resulta más acertado clasificar los sistemas lingüísticos según la presencia o no de acento léxico y según la presencia o no de tonos léxicos, con cuatro posibilidades combinatorias, pero no existe ningún tipo de sistema que se pueda llamar de acento melódico. En japonés existen muchas palabras en las que ninguna sílaba está especificada léxicamente para ser portadora de tono. Por otro lado, en el dialecto de Osaka, a diferencia del de Tokio, puede haber hasta dos sílabas portadoras de tono en una misma palabra: «in addition to the possibility and place of a H[igh] to L[ow] drop, lexical entries must indicate whether the form starts H or L. Thus, H-H-L contrasts with L-H-L» (*op. cit.*, p. 220). Por tanto, no se puede decir con propiedad que haya una sílaba prominente (acentuada) en cada palabra, porque a veces no hay ninguna y a veces hay dos y, si bien la presencia del tono puede poner de relieve la sílaba en cuestión, la asignación de prominencias no está métricamente regulada. El sistema del japonés se caracteriza, según Hyman, por ser un sistema que asigna léxicamente tonos (no acentos) a algunas sílabas. El hecho de que otras sílabas no tengan tono léxico —considera— no es distinto de la existencia de sílabas igualmente inespecificadas en cuanto al tono (las llamadas sílabas con tono neutro) en mandarín, una lengua que nadie duda en clasificar como tonal.

-autosegmental de la entonación española, *vid.* Prieto y Roseano, 2010; Sosa, 1999)⁴⁷. Propongo que se reserven los términos *sílaba tónica* y *sílaba átona* para referirse a la presencia o ausencia de acento de frase y especializar las denominaciones *sílaba acentuada* y *sílaba no acentuada* en la mención del acento léxico o la falta del mismo⁴⁸.

A la prominencia prosódica del acento de frase se le puede superponer una prominencia que tiene un carácter más bien semántico, relacionado con la estructura informativa. Esta prominencia informativa es lo que llamamos *foco*, y se manifiesta mediante la elección de los contornos tonales posléxicos apropiados, que en el inventario de cada lengua estén gramaticalizados para expresar los distintos tipos de foco. El ámbito informativo del *foco ancho* abarca el equivalente a la respuesta a la pregunta *¿qué sucede?*; el *foco estrecho* corresponde a las respuestas de las interrogativas parciales; el *foco contrastivo* sirve para hacer hincapié o énfasis sobre un elemento en construcciones que marcan una oposición o una corrección con respecto a otro elemento mencionado previamente. Estos tipos de foco, en este orden, pueden añadir prominencia sucesiva sobre la mera presencia de un acento de frase. Pero es importante resaltar que se manifiestan, en todo caso, como distintas realizaciones de dicho acento de frase. Por tanto, de alguna manera, son como subtipos de él, no añaden otro nivel de prominencia prosódica para completar la escala de los acentos secundario, primario, prenuclear y nuclear, porque, de hecho, no constituyen una prominencia prosódica sino informativa, aunque puedan tener un reflejo prosódico.

En el capítulo 3 veremos que algunos autores han estudiado el foco (a veces dicen el énfasis) de manera conjunta con la prominencia prosódica. Desde luego, a la hora de realizar un diseño para elicitar y recopilar un corpus de habla, una manera de asegurar la presencia de un acento de frase es inducir una estructura de foco. Pero, cuando revisemos esos trabajos, habrá que tener en mente que se trata de dos escalas de prominencia paralelas, y no de una sola. Aparte de las menciones pertinentes en esa revisión de estudios (capítulo 3), en esta tesis no se profundizará en el estudio de la prominencia focal y sus efectos.

A continuación revisaremos algunas características de la asignación de prominencias en los distintos niveles de la jerarquía prosódica.

⁴⁷ Muchos estudios han defendido el papel primordial que, en la percepción del acento léxico, juega la presencia de un pico de f_0 (para el español, *vid.* p. ej. Llisterri, Machuca, De la Mota, Riera, y Ríos, 2003; Llisterri, Machuca, Ríos, De la Mota, y Riera, 2003). No obstante, como decimos, parece que ese rasgo no es una característica general del acento léxico (primario), sino de aquellos acentos léxicos que coincidan con un acento de frase (prenuclear o nuclear). Según la constitución prosódica de la frase, podrían ser todos, pero eso no es necesario. Beckman y Edwards (1994) también señalan esta frecuente confusión de tomar los movimientos de f_0 como un correlato acústico directo del acento léxico. Además, el objetivo del tono posléxico no tiene por qué incluir «un pico», *i. e.* un tono alto (H). Hechas estas salvedades, por supuesto, estoy de acuerdo en que la presencia de un objetivo tonal asociado a la sílaba contribuye de manera decisiva a la percepción de su prominencia.

⁴⁸ Nótese que, en cualquier caso, se habla de *sílaba acentuada*, *tónica*, etc., y no de *vocal*. Aunque a menudo se encuentra en la bibliografía la expresión «vocal tónica», esto constituye un uso relajado, pues la prominencia es una propiedad que afecta a toda la sílaba en su conjunto (y no al segmento vocálico en exclusiva).

2.2.1. Asignación de prominencias correspondientes al acento léxico

En lo que sigue, me basaré fundamentalmente en el análisis propuesto por B. Hayes (1995). Recuérdese que hay una sílaba prominente por cada pie, y un pie prominente por cada palabra. La distribución métrica de ambos niveles de prominencia define las posiciones de los acentos léxicos secundarios y del acento primario, respectivamente.

En el § 2.1.2.2 se presentó el repertorio tipológico de pies defendido por B. Hayes (1995), que se repite aquí como (20) para facilitar la lectura.

(20) Los tres tipos de pie según B. Hayes (1995)

troqueo silábico: $\acute{\sigma} \sigma$

dos sílabas, con prominencia en la primera

troqueo moraico: $\acute{\cup} \cup | \acute{-}$

dos sílabas ligeras, con prominencia en la primera; o una sílaba pesada

yambo: $\cup \acute{\sigma} | \acute{-}$

dos sílabas (la primera ligera), con prominencia en la segunda (ligera o pesada); o una única sílaba pesada

σ = una sílaba cualquiera, independientemente de su carácter ligero o pesado

$\cup$ = una sílaba ligera

$-$ = una sílaba pesada, *i. e.* con vocal larga, diptongo decreciente, o coda consonántica

La construcción de pies puede operar de derecha a izquierda o de izquierda a derecha, en el seno de la palabra, en función de una opción paramétrica de cada lengua. Además, hay algunos mecanismos que regulan la inclusión exhaustiva o no de todas las sílabas de la palabra bajo un pie, como las reglas del pie degenerado, las que permiten saltar una sílaba ligera si es necesario para poder formar un pie, o las encargadas de la designación de sílabas extramétricas, que se pueden usar para justificar algunas idiosincrasias, aunque existe el claro riesgo de utilizar esta estrategia de manera estipulativa (*ad hoc*)⁴⁹.

Una vez contruidos todos los pies de la palabra, una opción paramétrica designa como más prominente el primero o el último de esos pies, y de este modo determina la ubicación del acento léxico (primario) en la palabra, que corresponde a la sílaba fuerte del pie fuerte.

Partiendo de estos presupuestos, el español se puede considerar una lengua con troqueos moraicos contruidos de derecha a izquierda y con prominencia del último pie de la palabra⁵⁰. Esta selección de un pie sensible al peso silábico encaja con el patrón

⁴⁹ Côté (2011) busca una explicación natural para una serie de factores que pueden hacer preferible que se intente evitar el acento léxico en final de palabra (y conllevar aparente extrametricidad). El alargamiento final propio de la posición de frontera (*vid.* § 3.1) disminuye la proporción relativa en la duración entre sílabas abiertas y cerradas, y eso dificulta el contraste de peso y de acento. Además, la deceleración de los gestos articulatorios en frontera parece contradictoria con la aceleración del movimiento de apertura para el acento (*vid.* § 3.2). Por último, la posible ubicación de un acento de frase en posición final de palabra hace que este quede demasiado junto al tono de frontera.

⁵⁰ Debemos suponer una licencia. Dada la construcción de derecha a izquierda, si el algoritmo tiene que empezar a construir un pie y encuentra una sílaba ligera, no lo completará, sino que saltará a la sílaba siguiente. Ahora bien, es posible que la siguiente sílaba no sea ligera (para conformarse al esquema $\acute{\cup} \cup$, y

predominante de palabras llanas (paroxítonas) cuando acaban en vocal —p. ej. *pasajero*, *tesoro*, *alarma*—, como se ilustra en (21), y agudas (oxítonas) cuando acaban en consonante o en diptongo decreciente y, por tanto, la última sílaba es pesada —p. ej. *emperatriz*, *carey*, *caracol*—, *vid.* (22). En estos ejemplos se indican los tres niveles de constituyentes relevantes, a saber, la sílaba (σ), el pie (Σ) y la palabra (ω). Mediante paréntesis se indica la extensión de cada unidad. Los asteriscos marcan la unidad prominente en el seno del nivel superior, mientras que los puntos representan las unidades no prominentes.

(21) pa.sa.je.ro

∪ ∪ ∪ ∪	σ
(* ·) (* ·)	Σ
(· *)	ω

(22) em.pe.ra.triz

— ∪ ∪ —	σ
(*) (* ·) (*)	Σ
(· · *)	ω

No obstante, hay un número considerable de palabras que no se atienen a este esquema. Desde luego, en la flexión verbal la posición del acento viene determinada morfológicamente, en función del tema verbal al que pertenezca la forma (*vid.* p. ej. Alcoba, 1999). En este subsistema, la ubicación acentual no sigue ningún principio métrico; es posible encontrar formas con cadenas de segmentos idénticas y una distinta acentuación, donde el acento tiene, precisamente, el valor morfológico que sirve para desambiguar. Pero incluso fuera del ámbito de la conjugación, existen palabras con una acentuación excepcional, para cuya descripción se suele recurrir a la consideración como extramétrica de la última sílaba (o a veces el último segmento⁵¹) de la palabra, *vid.* James W. Harris (1991 [1983], p. 117 y ss.); B. Hayes (1995, pp. 56-60). La existencia de estos casos idiosincrásicos quiere decir que la posición del acento primario es una característica ligada al conocimiento de cada pieza léxica. En los siguientes ejemplos, los elementos extramétricos se señalan entre ángulos $\langle \rangle$. Las palabras de (23) y (24) son llanas acabadas en consonante; la primera, con su último segmento extramétrico, y la segunda con su última sílaba extramétrica. En las esdrújulas, como (25), la extrametricidad siempre afecta a una sílaba entera.

cerrar el pie exitosamente), sino que sea pesada. En este caso, podemos entender que se forma un pie del tipo $\prec \cup$ o, equivalentemente, suponer que la ligera de la derecha se deja sin asignar a ningún pie y se forma un troqueo moraico con una única sílaba $\prec$. Por otro lado, si, al llegar al límite izquierdo de la palabra, solo queda una sílaba ligera, se formará un pie con ella, para satisfacer el requisito de que toda frontera entre unidades debe coincidir con una frontera de unidades de los niveles inferiores (por la Hipótesis de Estratificación Rigurosa). Además puede relacionarse con la aplicación acumulativa de fenómenos de refuerzo en posición inicial de constituyente prosódico (*vid.* capítulo 3).

⁵¹ La extrametricidad solo puede aplicarse sobre un constituyente (B. Hayes, 1995, p. 57).

(23) al.mí.ba<r>

—	∪	∪	σ
(*)	(*	·)	Σ
(·	*	)	ω

(24) al.féi.<zar>

—	—	σ
(*)	(*)	Σ
(·	*	ω

(25) mur.cié.la.<go>

—	∪	∪	σ
(*)	(*	·)	Σ
(·	*	)	ω

La limitación de la extrametricidad a una sola unidad —ya sea segmento o sílaba— del final de la palabra podría explicar por qué en español no pueden existir palabras esdrújulas con la segunda sílaba pesada. Por ejemplo, *pústula* es una palabra del español, *púntula* no lo es pero podría serlo, pero *pústurla* no podría ser una palabra española porque, aunque la sílaba final fuera extramétrica, *tur* es pesada y formaría pie por sí misma, atrayendo para sí el acento primario por ser el último pie de la palabra.

En cuanto a las agudas acabadas en vocal, la excepcionalidad está en permitir que el pie del borde derecho de la palabra esté constituido por una sílaba que, en principio, es ligera, como en (26). Al igual que con la extrametricidad, este tipo de comportamiento excepcional solo puede darse en el límite de una palabra.

(26) o.ja.lá

∪	∪	∪	σ
(*	·)	(*)	Σ
(·	*	)	ω

Ahora bien, una vez conocida la posición del acento primario, los secundarios sí siguen un patrón rítmico, y se distribuyen en sílabas alternantes desde la posición del acento primario hacia la izquierda. Quiero llamar la atención sobre el hecho de que esta alternancia ya no es sensible al peso silábico. Por ejemplo, dada la palabra *otorrinolaringología*, si hubiera un acento secundario en cada sílaba fuerte de los troqueos moraicos que se forman para decidir el acento primario, entonces tendríamos el siguiente esquema: *òtòrrinòlarìngòlogìa* (donde ` indica acento secundario). Es decir, esperaríamos un acento en la sílaba pesada *rin*, que debería formar pie por sí misma. Sin embargo, la intuición de los hablantes nativos de español parece apuntar más bien a una forma *òtòrrinòlarìngòlogìa*, donde la alternancia acentual es puramente silábica, sin atender al peso.

Esto parece reflejar un esquema de troqueos silábicos, en vez de troqueos moraicos. Podría suponerse que tal afirmación cae en contradicción con toda la descripción previa. Sin embargo, dado que el acento primario en español es una característica ligada al léxico, podríamos proponer que el acento secundario se asigna posléxicamente. En efecto, la pauta del acento secundario es mucho más regular que la del primario, característica habitual de las reglas posléxicas frente a las léxicas. En el § 2.1 hemos visto la posibilidad de que los constituyentes prosódicos sufran reestructuraciones al llegar a determinados niveles. En este caso, al llegar al nivel de la frase acentual ($\omega^{\text{máx}}$, según el análisis propuesto para las palabras clíticas en el § 2.1.3.2), una reestructuración de los pies situados a la izquierda del acento primario podría explicar el patrón de los secundarios. Dado que es un nuevo proceso de formación, nada tiene de extraño que se haya podido escoger aquí otro tipo de pie. Desde luego, esto es solo una propuesta y, en cualquier caso, los acentos secundarios no tienen un papel importante en los esquemas de prominencia del español, a diferencia de otras lenguas como el inglés, donde van acompañados de un cambio notorio en la cualidad tímbrica de la vocal.

Finalmente, como se ha explicado en el § 2.1.3.1, no siempre hay una correspondencia exacta entre palabra morfosintáctica y palabra prosódica. En español, por ejemplo, hay dos tipos de unidades léxicas que pueden equivaler a más de una palabra prosódica (típicamente dos, pero a veces más): los compuestos sintagmáticos y los adverbios en *-mente*. Dado que el acento es una propiedad (relacional) de la estructura prosódica, su ámbito de asignación son las unidades prosódicas. En concreto, el acento primario es la prominencia propia de la palabra prosódica. Por consiguiente, en estos tipos de unidades léxicas hay más de un acento primario, uno por cada palabra prosódica. La prominencia de la parte adjetival de los adverbios en *-mente*, por ejemplo, no es un acento secundario, como se ha dicho muchas veces (*vid.* p. ej. James W. Harris, 1991 [1983]), seguramente para salvaguardar la llamada ventana de tres sílabas (la limitación de la posibilidad de aparición del acento primario a las tres últimas sílabas de una palabra, dados los principios métricos explicados más arriba).

2.2.2. Asignación de prominencias correspondientes al acento de frase

Anteriormente (*vid.* § 2.1.3.1 y 2.1.4.1) hemos visto las reglas encargadas de la formación de los constituyentes ω y φ , a partir de una correspondencia parcial entre prosodia y sintaxis. Repetimos estas reglas aquí como (27) y (28) para facilitar la lectura.

(27) Alineación derecha de la palabra morfosintáctica y la palabra fonológica

En cada final de $X^{\text{mín}}$ coloca un límite $]\omega$

(28) Alineación izquierda del sintagma y la frase fonológica

En cada inicio de $X^{\text{máx}}$ coloca un límite $[\varphi$

Pues bien, dada esta manera de construir las frases fonológicas, la prominencia se asigna a la palabra prosódica del límite derecho de la frase. Por ejemplo, en el enunciado (12), repetido aquí como (29), los elementos subrayados corresponden a las

ω prominentes en el seno de cada φ . Estas son las palabras sobre las que recaerán los acentos de frase (*i. e.* los tonos posléxicos se alinearán con las sílabas portadoras del acento léxico de cada una de estas palabras).

(29) La vecina de arriba tiene goteras de difícil solución

$$\left[\underset{\varphi}{\text{La}} \left[\underset{\omega}{\text{vecina}} \right] \right] \left[\underset{\varphi}{\text{de}} \left[\underset{\omega}{\text{arriba}} \right] \right] \left[\underset{\varphi}{\text{tiene}} \right] \left[\underset{\varphi}{\text{goteras}} \right] \left[\underset{\varphi}{\text{de}} \left[\underset{\omega}{\text{difícil}} \right] \left[\underset{\omega}{\text{solución}} \right] \right]$$

En este punto conviene recordar que, en determinadas construcciones, los pronombres de objeto en posición posverbal son falsos enclíticos, pues reciben una prominencia, y ese acento ha delatado la necesidad de proponer una estructura de constituyentes especial para ellos (*vid.* § 2.1.4.2).

Por último, se suele asumir que la prominencia en el seno de una frase entonativa (y, por tanto, el acento de frase nuclear) suele recaer sobre la última φ que forme parte de dicha frase (*vid.* p. ej. Ladd, 2008; Prieto y Roseano, 2010). No obstante, según Nespor y Vogel (1994 [1986]), puede haber factores semánticos y de estructura informativa que modifiquen su ubicación habitual.

2.3. Prosodia y segmentación del continuo sonoro como apoyo al acceso al léxico

La estructura prosódica tiene una repercusión importante sobre el conocimiento de los hablantes acerca de la forma fonológica habitual de las piezas léxicas de su lengua. Además, los distintos constituyentes prosódicos ayudan a segmentar el continuo sonoro de una manera estructurada. Por las relaciones que establece la prosodia entre la fonología y otros niveles lingüísticos (*vid.* § 2), la correcta identificación de dicha estructura prosódica puede ayudar a reconocer la extensión de algunos constituyentes sintácticos, como los sintagmas y las palabras. Aquí nos centraremos en el papel de la estructura prosódica como factor coadyuvante a los procesos de acceso al léxico. A lo largo de este apartado, se revisarán algunas evidencias psicolingüísticas de la importancia de las posiciones de frontera prosódica (sobre todo de las iniciales, pero también de las finales). A partir del próximo capítulo, estudiaremos cómo los fenómenos de refuerzo articulatorio constituyen una estrategia de marcación fonética que pone de relieve la ubicación de dichas fronteras prosódicas y, por ende, la estructura de un enunciado⁵².

Así como en la lengua escrita las palabras suelen ir delimitadas por espacios, la lengua hablada, en cambio, destaca por su carácter continuo. Lo más común es que las palabras se pronuncien unas seguidas de otras, sin silencios intermedios que las demarquen (desde luego, la posible aparición de pausas no responde a esta función, sino que se explica por otros motivos, *vid.* Gil Fernández, 2007, pp. 298-305, y las referencias allí incluidas). La inexistencia de una separación evidente entre las palabras de la lengua oral plantea al oyente el problema de la segmentación del flujo sonoro para reconocer las piezas léxicas que componen un determinado enunciado.

En realidad, conviene separar los conceptos de segmentación e identificación, que, aunque relacionados, son dos procesos cognitivos diferentes (Cutler y Norris, 1988). La identificación de palabras, o acceso al léxico, depende, como es lógico, del conocimiento del vocabulario. A medida que una secuencia de sonidos va llegando al oyente, este la coteja con su diccionario mental y emite múltiples hipótesis de palabras que podrían coincidir con alguna porción de lo oído hasta el momento. De este modo se activan distintos candidatos, que entran en competición, hasta la selección de un ganador (McQueen, Norris y Cutler, 1994). Tal proceso de competición implica que el reconocimiento no es estrictamente secuencial, pues muchas veces solo se puede

⁵² Smith (2005) asume que, en el marco de la teoría de la optimidad, existe un tipo de restricciones que son generalizables a cualquier posición, mientras que hay otras restricciones —por lo demás, de idénticos efectos— cuya aplicación se limita a una determinada posición. Como es habitual en la teoría de la optimidad, un orden jerárquico diferente entre ambos tipos de restricciones establece diferentes predicciones. Pues bien, Smith defiende que en el caso de las restricciones por posición, no todas las combinaciones lógicas posibles (*i. e.* cualquier restricción en cualquier posición) forman realmente parte del conjunto universal de restricciones. Por el contrario, el tipo de restricciones que pueden aplicarse en posiciones psicolingüísticamente relevantes como los inicios de constituyente está limitado por la necesidad de maximizar las diferencias paradigmáticas entre los segmentos y favorecer el acceso al léxico indicando la identidad segmental y la ubicación de las fronteras.

resolver a favor de una palabra cuando ya se ha oído parte de la siguiente. Esta situación es especialmente notoria en el caso de las palabras incrustadas. Por ejemplo, sea la secuencia *está vacío*. Al oír esta cadena de sonidos, un individuo puede reconocer, entre otras, las siguientes palabras o formas de palabra: *es*, *está*, *estaba*, *va*, y *vacío*. Sin embargo, solo la combinación *está* + *vacío* da cuenta exhaustiva de la secuencia, pues asigna cada segmento a una palabra y solo a una (si se optara, en cambio, por *estaba*, sobraría *cío*, que no es ninguna palabra en español). Esta regla de asignación exhaustiva se une a otra que impide asignar únicamente consonantes a una supuesta palabra (a excepción de algunas lenguas, como el bereber, donde sí existen palabras que solo tienen consonantes, *vid.* El Aissati, McQueen y Cutler, 2012). Por otro lado, tienen más peso en la competición las palabras muy frecuentes, y tanto más peso tendrán cuanto menos frecuentes sean, a su vez, sus competidores. Adicionalmente, la pugna se resuelve con ayuda de los llamados procesos de arriba abajo (que incluyen la interpretación sintáctica, semántica y pragmática, y en este caso también el conocimiento fonológico de la lengua), y también con los datos puramente fonéticos que llegan de abajo arriba.

En cuanto al segundo proceso cognitivo —la segmentación del continuo sonoro—, cabe distinguir entre segmentación léxica y segmentación preléxica (Gout, Christophe y Morgan, 2004). La primera es el fruto del acceso al léxico: lógicamente, en el momento en que se identifica una palabra, se establecen sus fronteras y esto delimita dicha pieza en el flujo de habla. La segmentación preléxica, en cambio, se refiere a toda una serie de estrategias que sirven de ayuda al acceso al léxico, en tanto que detectan los puntos preferenciales para dicho acceso, y otorgan más peso en la competición a los candidatos que empiezan en esos puntos.

Estas estrategias de segmentación, como tales estrategias, son generales y comunes a todas las lenguas. Sin embargo, a la vez, son específicas, en la medida en que dependen de características fonéticas y fonológicas particulares de cada lengua (Cutler, Mehler, Norris y Segui, 1992). Su adaptación a las peculiaridades de cada sistema lingüístico hace que, cuanto más específica sea una estrategia, más eficiente resulte en el procesamiento de esa lengua por parte de sus oyentes nativos. En cambio, esto plantea un reto adicional a los aprendices de una segunda lengua, que pueden ver mermada su capacidad de comprensión auditiva si no consiguen llevar a cabo, y con la rapidez suficiente, una segmentación adecuada para el acceso al léxico (*ibid.*).

En lo que sigue, se detallarán algunas de las estrategias de segmentación utilizadas habitualmente por los hablantes (oyentes) de una lengua: aquellas basadas en probabilidades transicionales, en las unidades prosódicas de la lengua, y en las variaciones alofónicas condicionadas por la posición.

2.3.1. Probabilidades transicionales

Cuando un niño está adquiriendo su lengua, sus problemas de segmentación difieren de los del adulto en un aspecto fundamental, y es que todavía no tiene ningún vocabulario mental, con el que reconocer las palabras, sino que ese vocabulario tiene que crearlo, sin ningún conocimiento previo de la forma que pueden tener las palabras

en esa lengua. El cálculo de las probabilidades transicionales aporta una estrategia de segmentación adecuada, pues se basa en un razonamiento puramente lógico, y puede empezar a usarse sin que sea necesario ningún conocimiento previo del idioma (Saffran, Newport y Aslin, 1996).

Este proceso de segmentación se basa en la idea de que, si un elemento sigue a otro con frecuencia, es probable que ambos pertenezcan a una misma unidad. La fórmula (30) permite calcular el poder predictivo que tiene un elemento x acerca de si el elemento que lo sigue será (o no) el elemento y .

(30) Cálculo de la probabilidad transicional

$$x|y = \frac{\text{frecuencia}(xy)}{\text{frecuencia}(x)}$$

Por ejemplo, si se observa la secuencia *bo.ca#gran.de*, se comprenderá que la probabilidad de que, en español, la sílaba *gran* vaya seguida de la sílaba *de* es mucho mayor que la probabilidad de encontrar *ca* seguido de *gran*. Así, donde la probabilidad transicional es alta, se entiende que ambos elementos forman parte de la misma palabra, mientras que es razonable esperar una frontera léxica allí donde la probabilidad transicional es baja.

Saffran *et al.* (1996) argumentan que esta fórmula es aplicable a cualquier tipo de unidad, no solo a la sílaba, sino también al segmento, e incluso a cualquier otro elemento que no sea un sonido habitual del habla, como los silbidos, trinos y gorjeos del experimento de J. R. Hayes y Clark (1970). Además, estos estudios demuestran que esta estrategia no solo es utilizada por los bebés, sino que también la emplean los adultos cuando se enfrentan al aprendizaje de una lengua artificial, de la que no tienen ningún conocimiento previo. Aunque el cálculo computacional parezca complejo, responde a una habilidad innata y se comprueba que los sujetos la utilizan de manera inconsciente al elaborar los repertorios léxicos de las lenguas que adquieren.

Por otro lado, para Saffran *et al.* (1996), la aplicación de este proceso justifica que en los estadios iniciales de la adquisición de la lengua materna (L1) el final de las palabras resulte más prominente que el inicio (frente a lo que ocurre en la vida adulta). Por ejemplo, se sabe que los niños tienden a repetir más bien esos finales. Este hecho cobra sentido bajo la consideración de que la probabilidad de que en un punto determinado haya una frontera léxica se calcula sobre la frecuencia del elemento previo (la x , no la y). De este modo, una frontera es más bien un final de palabra (aunque la lógica diga que, al mismo tiempo, es un inicio de la siguiente).

En cambio, conforme avanza el proceso de adquisición y se aplican otras estrategias segmentadoras, la prominencia se desplaza al inicio de palabra. Cutler, Hawkins y Gilligan (1985) sostienen que, dado el proceso de competición, resulta ventajoso que los morfemas más informativos de una palabra aparezcan cuanto antes, para desambiguar entre candidatos. Las lenguas han ido evolucionando de manera que prefieren el orden *raíz + afijo*, frente a la posibilidad *afijo + raíz*, incluso cuando por su tipología deberían haberse inclinado más hacia la prefijación, y han consolidado esas preferencias, que

resultaban beneficiosas para el procesamiento. Esto dota de una gran relevancia informativa a los inicios de palabra, que juegan un papel fundamental en el acceso al léxico, como demuestra el fenómeno de *tener algo en la punta de la lengua*, cuando alguien solo recuerda cómo empieza la palabra que busca. Además, Nootboom (1981) demuestra que es más fácil recordar una palabra cuando se sabe el inicio que cuando se sabe el final. En neerlandés, la palabra *kannibaal* ‘caníbal’ presenta una secuencia inicial de cuatro fonemas que no aparecen en el inicio de ninguna otra palabra, y lo mismo pasa con los cuatro fonemas finales, que tampoco son compartidos por ningún otro final de palabra. A priori, por tanto, ambas secuencias tetrafonemáticas son igual de informativas. Sin embargo, cuando los sujetos oían el inicio, completaban la palabra correctamente en un 95 % de los casos, mientras que el final solamente desencadenaba un 60 % de respuestas correctas. Igualmente, la distorsión de los inicios provoca más errores en el acceso al léxico que la misma distorsión en los finales (*vid.* Cutler *et al.*, 1985, y las referencias allí incluidas).

En definitiva, en los primeros estadios de la adquisición, el final de palabra es más relevante como consecuencia del uso de las probabilidades transicionales en la segmentación. Posteriormente, el peso se desplaza hacia el inicio por motivos de procesamiento. A partir de ese momento, el principio de la palabra es lo más prominente, aunque el final conserva una relevancia parcial. En todo caso, la parte central es la menos informativa. Esta relación se ha denominado, de manera metafórica, el *efecto bañera* (Aitchison, 1987): en una bañera, lo que más sobresale es la cabeza, seguida de los pies, mientras que el resto del cuerpo permanece bajo el agua.

Finalmente, Gout *et al.* (2004) hacen notar que las palabras de clase cerrada se aprenden muy pronto debido a su elevada frecuencia de aparición, unida al hecho de que suelen aparecer alineadas con los límites de frase fonológica (*vid.* § 2.1.4.1 y 2.3.2). Teniendo en cuenta que estas palabras gramaticales desencadenan una muy baja probabilidad transicional —pueden combinarse con muchos elementos distintos—, suponen una ayuda para detectar fronteras léxicas, y apoyan la hipótesis de que el resto de la secuencia (no reconocido) debe de ser una palabra de contenido léxico.

2.3.2. La frase fonológica, dominio del acceso al léxico

Aunque, según se ha visto en el § 2.1.3.1, la palabra morfosintáctica, es decir, aquella que se almacena en el léxico mental, no siempre coincide con la palabra fonológica, este nivel prosódico es la mejor aproximación a la unidad léxica, y normalmente hay alguna coincidencia entre un límite de palabra morfosintáctica y un límite de palabra fonológica. El resto de las unidades prosódicas, superiores o inferiores al nivel de la palabra, pueden ayudar a enmarcar la búsqueda de dicho nivel, acotándolo por arriba y por abajo, como se verá en este apartado y el siguiente.

Según la Hipótesis de Estratificación Rigurosa, una frontera entre unidades de un determinado nivel siempre coincide con una frontera entre unidades del nivel inferior. Así, por ejemplo, una frontera de enunciado, de frase entonativa, o de frase fonológica es también, imperativamente, una frontera de palabra fonológica. En el caso del

enunciado y la frase entonativa, una de las marcas de frontera más habituales es la pausa y, lógicamente, toda pausa se interpreta como un límite entre dos palabras.

Christophe, Peperkamp, Pallier, Block y Mehler (2004) van un paso más allá y demuestran que la frase fonológica es el dominio del acceso al léxico. No solo se da más peso a los candidatos léxicos con un límite alineado con otro límite de frase fonológica, sino que al alcanzar la frontera de frase fonológica se resuelven todas las competiciones pendientes. Estudios como el de Christophe *et al.* (2004), con hablantes adultos de francés, o el de Gout *et al.* (2004), con niños aprendices de inglés como L1, parecen corroborar el carácter innato —e independiente del tipo rítmico de la lengua— de la frase fonológica como dominio de acceso al léxico.

Como se verá en el capítulo 3, las fronteras de la frase fonológica suelen manifestarse por medio de alguna clave fonética asociada a fenómenos de refuerzo, y parece que los oyentes son capaces de percibir dichas claves y apoyarse en ellas para reconstruir la estructura prosódica. De este modo, el acceso al léxico se puede ver facilitado por el alargamiento de los segmentos del final de la frase (Christophe *et al.*, 2004; Saffran *et al.*, 1996), o una menor coarticulación a través de frontera (Byrd, Kaun, Narayanan y Saltzman, 2000; Christophe *et al.*, 2004), entre otras claves.

Además, como se anticipaba en el § 2.3.1, Gout *et al.* (2004) consideran que la habitual aparición de palabras gramaticales, de clase cerrada, en las posiciones limítrofes de frase fonológica contribuye al aprendizaje de dicho tipo de palabras. Las palabras más frecuentes de una lengua suelen pronunciarse con mayor rapidez —dada su predictibilidad—. Para Spinelli, Grimault, Meunier y Welby (2010), esa menor duración, junto con la elevada frecuencia de aparición en una posición prosódica determinada, también ayuda a aprender el repertorio de palabras gramaticales. Identificar estas piezas en uno de los extremos de la frase fonológica, concluyen Gout *et al.* (2004), permite deducir que el resto de la secuencia, no identificada, debe interpretarse como una palabra de contenido léxico.

2.3.3. Ritmo lingüístico y segmentación preléxica

Condon y Sander (1974) demuestran que los recién nacidos son capaces de sincronizar sus movimientos con el ritmo del habla adulta (no así con otros elementos rítmicos no lingüísticos, como una cadencia de golpecitos, por ejemplo). También Cutler y Mehler (1993) afirman que los bebés experimentan un condicionamiento innato para buscar elementos periódicos en el habla. De ahí que presten especial atención a la unidad prosódica que, en cada lengua, constituye la base del ritmo⁵³.

Después de que James (1940) propusiera la conocida metáfora acerca de la distinción entre lenguas con ritmo de metralla y con ritmo de código Morse, Pike (1945)

⁵³ Jun (2005) aclara que, aunque la percepción del ritmo global de la lengua viene determinada fundamentalmente por alguna de las unidades prosódicas inferiores al nivel de la palabra, en realidad también las unidades superiores manifiestan propiedades rítmicas. Es decir, que la tendencia a la regularidad temporal puede ocurrir en varios niveles a la vez, *vid.* Saltzman, Nam, Krivokapić, y Goldstein (2008) sobre una posibilidad de modelado de las tensiones que esto genera, y su resolución.

estableció que la contraposición se basaba en la isocronía⁵⁴ silábica o en la isocronía acentual, respectivamente, y más tarde Bloch (1950) añadió un tercer tipo rítmico, basado en la mora⁵⁵. En efecto, los estudios tipológicos constatan que la percepción del ritmo de una lengua se basa en la recurrencia aproximadamente periódica de una de las siguientes unidades: la mora, la sílaba, o el pie acentual. Felizmente, todas estas unidades son inferiores al nivel de la palabra y, como se verá enseguida, esto tiene una gran trascendencia en los procesos de segmentación y, por consiguiente, en el acceso al léxico.

Dada la Hipótesis de Estratificación Rigurosa, todo inicio de palabra fonológica debe estar alineado con el inicio de su primer pie, su primera sílaba, y su primera mora. Lo contrario no es necesariamente cierto; es decir, un inicio de pie (o de sílaba o de mora) puede no coincidir con el inicio de una palabra, sino aparecer en el interior de la misma. Sin embargo, resulta evidente que las fronteras entre unidades subléxicas constituyen puntos preferenciales de segmentación, y otorgan mayor peso a los candidatos léxicos que aparecen alineados con dichos puntos. McQueen *et al.* (1994) puntualizan que se trata simplemente de eso: de dar más peso a ciertas hipótesis en el proceso de competición, pero no se descartan totalmente los otros candidatos. De otro modo, se podría incurrir en errores por omisión. Por ejemplo, en español y en francés, que son lenguas de ritmo silábico, existen procesos de resilabeo que rompen la alineación del inicio de palabra con el inicio de sílaba. En inglés, una lengua de ritmo acentual, la mayoría de las palabras de contenido léxico empiezan por sílaba fuerte (*i. e.* con vocal plena), pero también hay vocablos que comienzan por sílaba débil (con vocal reducida). Si solo se permitiera el acceso a las palabras estrictamente alineadas con inicios de sílaba o pie, sería imposible identificar estos casos excepcionales. Por ello, la posición más prudente es considerar estos puntos no como los únicos lugares de acceso al léxico, aunque sí como los preferidos.

Según se ha explicado anteriormente, los recién nacidos empiezan a aplicar la segmentación basada en probabilidades transicionales para aprender el repertorio léxico de su lengua. A medida que van aprendiendo palabras, empiezan a reconocer el alto grado de coincidencia entre los inicios de palabra y los inicios de esa otra unidad prosódica que, por su sesgo innato hacia la búsqueda de patrones rítmicos, les llama la atención. Eso les permite adquirir un conocimiento específico sobre las propiedades fonológicas de su lengua, con el que desarrollan una estrategia de segmentación más eficiente, por mejor adaptada a las condiciones particulares del idioma. La propia constatación de su eficiencia, además, consolida dicha adquisición. Al priorizar puntos de segmentación definidos por unidades superiores al segmento, se acelera el proceso, puesto que las posibilidades se restringen. En cualquier caso, las hipótesis avanzadas por la segmentación rítmica se complementan con las probabilidades transicionales para confirmar los límites de palabras (McQueen *et al.*, 1994).

⁵⁴ El concepto de isocronía como una propiedad acústica real ha sido cuestionado numerosas veces (*vid.* las referencias citadas por Jun, 2005, por ejemplo), y se habla más bien de una tendencia, dentro de un continuo, hacia un tipo rítmico u otro.

⁵⁵ Recuérdense las reticencias sobre el carácter de constituyente de la mora expresadas en el § 2.1.1.7.

Cutler y Butterfield (1992) sostienen que, como norma general, una hipótesis léxica es tanto más eficiente cuanto más restrictiva sea, especialmente en condiciones adversas, como pueden ser las siguientes: presencia de ruido ambiente, distancia entre los interlocutores, acento extranjero, *lapsus linguae*, etc. Dilley, Mattys y Vinke (2010) también comprobaron que, ante un estímulo mermado mediante un filtro pasabajo, con el consiguiente deterioro del contenido segmental (y léxico), los oyentes recurrían a los patrones rítmicos para apoyarse en sus decisiones, a la hora de elegir entre distintas opciones de lo que podía haber sonado. No en vano, la percepción del ritmo se fundamenta, a su vez, en la percepción del tiempo y, como diría Kant, el tiempo es una forma apriorística del entendimiento humano, de modo que siempre se encuentra disponible a nuestros juicios.

El hecho de afirmar que la hipótesis más restrictiva es la más eficiente no equivale a decir que la segmentación basada en pies sea preferible sobre aquella basada en sílabas o en moras. En abstracto, podría ser; pero, lógicamente, lo más adecuado y eficiente es utilizar la unidad rítmica propia de cada lengua.

Efectivamente, la estrategia rítmica de segmentación es universal, pero lo que cambia de una lengua a otra es la unidad utilizada como base del ritmo y, por tanto, como apoyo a la segmentación. Los próximos subapartados profundizan en las características de la segmentación basada en sílabas, pies o moras.

2.3.3.1. Segmentación basada en el ritmo silábico⁵⁶

La sílaba es una unidad con realidad psicológica, pues interviene de manera natural en juegos metalingüísticos por parte de hablantes no expertos, así como en *lapsus linguae* (vid. 2.1.1.1). Además, frente a la variabilidad de realizaciones que experimentan los segmentos fónicos —p. ej. debido a la coarticulación—, la sílaba presenta una mayor estabilidad perceptiva (vid. p. ej. Marrero, 2001). Si a ello se le une el uso de esta unidad como base del ritmo, su prominencia se vuelve aún mayor y la sílaba se convierte en una unidad óptima de segmentación.

Segui, Frauenfelder y Mehler (1981), en su experimento con franceses, no encontraron diferencias significativas en los tiempos de reacción en la detección de sílabas en palabras reales frente a palabras inventadas, lo que apoya la idea de que es posible un reconocimiento preléxico de la sílaba y, por tanto, el empleo de esta unidad en la segmentación del continuo sonoro, como ayuda para la posterior identificación de palabras. En cambio, según sus resultados, las sílabas se detectan más rápido que los fonemas, de modo que constituyen una unidad de segmentación más eficiente que estos.

El uso de la sílaba para estos fines se confirma por el menor tiempo de reacción en la detección de secuencias que coinciden con la estructura silábica de la palabra. Mehler, Dommergues, Frauenfelder y Segui (1981), en un estudio ya clásico, pidieron a sus informantes —franceses— que pulsaran un botón cuando oyeran una palabra, de entre una lista, que comenzara por una determinada secuencia, presentada visualmente. Los tiempos de reacción en la detección de la secuencia *pa* fueron significativamente

⁵⁶ Algunos ejemplos de lenguas de ritmo silábico son los siguientes: español, italiano, francés, coreano, o mandarín.

menores con la palabra *palace* ‘palacio’ (donde dicha secuencia coincide exactamente con la primera sílaba) que con *palmier* ‘palmera’ (donde *pa* es solo parte de la primera sílaba). Y a la inversa: la secuencia *pal* se detectó más rápido en *palmier* que en *palace* (ya que en este último la /l/ es el primer segmento de la sílaba siguiente).

Ahora bien, en lenguas como el francés o el español, que tienen ritmo silábico, también pueden existir fenómenos de resilabeo a través de frontera de palabra, lo que puede desdibujar notablemente dichas fronteras (*vid.* § 2.1.1.5).

Dada la desalineación que este fenómeno provoca entre los límites silábicos y los de palabra, cabe preguntarse cómo se realiza la segmentación y el acceso al léxico en estos casos, pues es evidente que los oyentes nativos de estas lenguas no tienen ningún problema en identificar las fronteras léxicas en contextos de resilabeo. Spinelli, McQueen y Cutler (2003) estudiaron el efecto de la *liaison* en francés. Para ello, utilizaron el paradigma experimental de la preactivación intermodal (en inglés, *cross-modal priming*). Los participantes oían unas frases y, mientras sonaba la última palabra, también veían una secuencia de letras escrita en una pantalla. Su tarea era decidir si esa cadena de caracteres correspondía a una palabra real del francés o no. Este paradigma supone que, si la secuencia era una palabra que había sido preactivada como candidato léxico por el estímulo sonoro, entonces el tiempo de reacción sería menor. Los resultados mostraron una activación léxica débil cuando la secuencia objetivo había aparecido como palabra incrustada en otra (p. ej. *oignon* ‘cebolla’ inserta en *demi rognon* ‘medio riñón’); es decir, que la desalineación entre la palabra meta y la estructura silábica tiene un coste. Los ejemplos de *liaison*, por su parte, estaban contruidos para que fueran ambiguos (p. ej. *dernier oignon* ‘última cebolla’ se pronuncia, en principio, igual que *dernier rognon* ‘último riñón’). En estos casos, la palabra con consonante inicial (palabra, por tanto, alineada con la estructura silábica) recibía un alto grado de activación. Sin embargo —y esto es lo interesante—, los oyentes también activaban sin problemas la palabra de vocal inicial cuando esta correspondía a la intención del hablante (p. ej. *oignon* en *dernier oignon*). Esto quiere decir que existe alguna pequeña diferencia acústica, subsegmental, a la que los oyentes son sensibles, que permite distinguir cuándo una consonante se encuentra en situación de resilabeo y cuándo es originalmente un ataque silábico. Los análisis acústicos de Spinelli *et al.* (2003) revelan una duración consonántica significativamente menor en contextos de resilabeo, y la equiparan a los resultados de la tesis inédita de Dejean de la Bâtie (1993), quien también constata un menor VOT⁵⁷ en estos casos.

En definitiva, el estudio de Spinelli *et al.* (2003) confirma, por un lado, la existencia de alguna diferencia perceptible entre las palabras con consonante inicial y aquellas con vocal inicial en contextos de resilabeo por *liaison*; y, por otro, una diferencia acústica relativa a la duración de la consonante fronteriza. En sentido estricto, no se puede afirmar que sea justamente esa la clave acústica a la que prestan atención los oyentes para desambiguar la selección léxica en situaciones de potencial resilabeo, y haría falta una prueba en la que se controlara específicamente esa variable. Pero, en cualquier caso,

⁵⁷ Siglas en inglés correspondientes al concepto de *Voice Onset Time* (Lisker y Abramson, 1964), o tiempo que tarda en establecerse la vibración laríngea periódica desde la distensión de una consonante.

no parece una casualidad que Strycharczuk y Kohlberger (2014) detectaran los mismos efectos duracionales en las consonantes en posición de resilabeo en español. Según estos autores, las consonantes en coda duran menos que en ataque, y las consonantes en resilabeo muestran una duración intermedia y significativamente distinta de los otros dos casos. Por otro lado, tiene sentido que sea la duración la clave que ayude a marcar la frontera de palabra en estos casos, del mismo modo en que señala las fronteras de otros niveles prosódicos (Byrd y Saltzman, 2003). Ello no obsta para que se deba seguir investigando con mayor profundidad la relevancia de dicha clave en la percepción del resilabeo, y también la posibilidad de intervención de otras claves. Spinelli *et al.* (2010), por ejemplo, apuntan a la subida de la frecuencia fundamental (f_0) en inicio de palabra de contenido léxico como clave explotada para la desambiguación en casos de elisión en francés, como *l'affiche* ‘el cartel’ frente a *la fiche* ‘la ficha’, de modo que una mayor f_0 en la [a] apoyaría la primera opción, con elisión de la vocal del artículo.

2.3.3.2. Segmentación basada en el ritmo acentual⁵⁸

Vista la eficacia de la segmentación silábica, y teniendo en cuenta que todas las lenguas tienen sílabas, podría pensarse que esta estrategia debería seguir estando disponible en lenguas con una tipología rítmica diferente. Sin embargo, Cutler *et al.* (1986) argumentan que en las lenguas de ritmo acentual (*i. e.* basado en los pies acentuales) los límites silábicos se vuelven confusos en algunas condiciones, como las llamadas consonantes ambisilábicas (*vid.* § 2.1.1.4), y eso restaría eficiencia a una segmentación basada en la sílaba.

En efecto, estas lenguas utilizan el pie como unidad rítmica de segmentación (Cutler y Butterfield, 1992; Cutler y Norris, 1988). Esto quiere decir que se identifican las sílabas fuertes como puntos preferenciales de acceso al léxico, entendiendo por fuerte toda sílaba con una vocal plena y un acento, no necesariamente primario, pero sí al menos secundario. Esta estrategia resulta muy eficiente si se tienen en cuenta las propiedades del vocabulario de estas lenguas. Por ejemplo, Cutler y Carter (1987), en su estudio sobre el léxico inglés, afirman que la mayoría de las palabras de contenido de esta lengua empiezan por sílaba fuerte. De hecho, unas tres cuartas partes de todas las sílabas fuertes de su corpus son iniciales de palabra léxica, mientras que dos tercios de las sílabas débiles son iniciales de palabra gramatical⁵⁹. Por su parte, Dilley *et al.* (2010) sostienen que las vocales plenas son más informativas que las reducidas, por ser menos predecibles y por aportar claves acústicas segmentales más claras. Por ello tienen un mayor poder restrictivo sobre la identidad de la palabra en el proceso de competición de candidatos.

La tendencia del inglés a hacer coincidir el inicio de las palabras de contenido con sílabas fuertes se explota incluso en muchos chistes, como uno de los que recogen Cutler y Butterfield (1992, p. 220):

⁵⁸ Este tipo de ritmo se encuentra, por ejemplo, en inglés, alemán, neerlandés, sueco, árabe, o griego.

⁵⁹ En estas proporciones se tienen en cuenta las frecuencias de los ejemplares y no las de los tipos; esto es, las de cada aparición real en el corpus, y no las calculadas sobre la lista de lemas de un diccionario.

(31) «Q: When is a door not a door? A: When it's ajar.»

Estas autoras también corroboran que un alto porcentaje de errores de comprensión auditiva en oyentes nativos de inglés, con material en su propio idioma, se justifica por un error en la ubicación de una frontera de palabras. Se tiende a interpretar las sílabas fuertes como inicio de palabras léxicas, y las sílabas débiles como inicio de palabras gramaticales o como sílabas no iniciales de palabra. Por ejemplo (*ibid.*: p. 222, Tabla 1):

(32) Input: «Into opposing camps» → Error: «Into a posing camp»

(33) Input: «The effective firing rate» → Error: «The effect of...»

Esta estrategia de segmentación queda confirmada por los experimentos de Cutler y Norris (1988). Los sujetos debían reconocer la presencia de palabras reales del inglés que estaban insertas en una secuencia de sonidos sin sentido. Los tiempos de reacción fueron mucho más rápidos en los casos en que la palabra iba seguida de sílaba débil (p. ej. *mint* ‘menta’ en *mintesh* [mɪntəʃ]_Σ) que cuando iba seguida de sílaba fuerte (p. ej. *mint* en *mintayve* [mɪn]_Σ [tɛɪv]_Σ). La presencia de una sílaba fuerte se interpreta como inicio de un nuevo pie, y la interposición de la frontera de pie rompe la integridad de la palabra meta (*mint*), que queda repartida a ambos lados de dicha frontera. Al intentar el acceso al léxico en el inicio de la sílaba fuerte [tɛɪv]_Σ, la porción *min* queda desgajada, dificultando notablemente la identificación del candidato *mint*. En cambio, no hubo diferencias significativas en los tiempos de reacción a la hora de identificar *thin* ‘fino, delgado’ en *thintesh* o en *thintayve* (la detección fue rápida en ambos casos), porque la separación *thintayve* en dos pies ([θɪn]_Σ [tɛɪv]_Σ) no deja dividida la palabra meta a través de la frontera prosódica.

Por su parte, Field (2005) comprobó, también para el inglés, que el desplazamiento del acento a una sílaba que debía ser átona comportaba un descenso notable en la inteligibilidad —medida como precisión en la transcripción— por parte de oyentes tanto nativos como extranjeros, de diversas L1 pero con un nivel alto de inglés. El efecto era moderado cuando el cambio de posición en el acento iba aparejado a un cambio de timbre vocálico en la nueva sílaba acentuada (ahora con vocal plena). Por el contrario, la pérdida de inteligibilidad era aún mayor si la nueva sílaba acentuada se mantenía reducida, lo que se atribuye a una pérdida de valor informativo o predictivo por parte de la vocal que debía ser más prominente de la palabra.

2.3.3.3. Segmentación basada en el ritmo moraico⁶⁰

En las lenguas de ritmo moraico, la segmentación se apoya en la mora para localizar los puntos preferenciales de acceso al léxico. Otake, Hatano, Cutler y Mehler (1993) realizaron un experimento en el que pedían a sus informantes japoneses que señalaran cuáles de las palabras que oían contenían una determinada secuencia, que habían presentado por escrito y en alfabeto romano, para no sesgar la interpretación en

⁶⁰ El japonés es la lengua contemporánea más conocida por tener ritmo moraico.

términos de moras. Por ejemplo, debían detectar la secuencia *na* en las palabras *na.no.ka* ‘el día siete del mes’ y *nan.ka* ‘algo’. Los tiempos de reacción no fueron significativamente distintos en estas dos condiciones, frente a los resultados obtenidos en condiciones semejantes en las lenguas de ritmo silábico (recuérdese el experimento de Mehler *et al.*, 1981, con las palabras *pa.la.ce* y *pal.mier*). En el ejemplo japonés resulta crucial que, aunque *na* no coincide con una sílaba entera en la palabra *nan.ka*, sí que constituye, en cambio, una mora entera. En contraste, cuando la secuencia que se debía detectar era *nan*, los resultados para ambas palabras sí eran muy diferentes. La secuencia *nan* se identificó en *nan.ka* sin problemas, aunque el tiempo de reacción fue mayor que cuando la secuencia meta era *na* (demora atribuible al hecho de que *nan* tiene dos moras, frente a *na*, con una). Por otro lado, la identificación de *nan* en *na.no.ka* se saldó con un elevado número de errores por omisión, debido a que *nan* no se correspondía con un número entero de moras en el estímulo sonoro.

En otro experimento, en el que la secuencia meta consistía en un único segmento (la vocal *o* o la consonante *n*), los tiempos de reacción y las tasas de omisión también fueron mayores cuando dichas secuencias no se correspondían exactamente con una mora (Cutler y Otake, 1994). Esto demuestra la preferencia por la mora como unidad de segmentación, no solo sobre la sílaba sino también sobre el fonema.

2.3.4. Variaciones alofónicas

Los fonemas de una lengua pueden manifestar variaciones en su realización fonética (variaciones alofónicas). Estas pueden depender del contexto segmental inmediato (por coarticulación o asimilación, a consecuencia de una anticipación o una persistencia de los gestos articulatorios propios de los sonidos circundantes). Otras veces, las variaciones en un segmento están motivadas por la posición prosódica que este ocupa (dicha posición puede inducir fenómenos de refuerzo o de debilitamiento)⁶¹, *vid.* § 4.

Cuando la aparición de cierto alófono indica el inicio de una determinada unidad prosódica, dicha variante puede ayudar indirectamente a la segmentación (*vid.* Christophe *et al.*, 2004; Gout *et al.*, 2004; y las referencias citadas en ambos trabajos). Por ejemplo, las formas aspiradas de las consonantes oclusivas sordas del inglés —[p^h], [t^h], [k^h]— siempre señalan inicio de pie y, por tanto, puntos preferenciales de acceso al léxico.

En cualquier caso, la existencia de alófonos posicionales que marquen sistemáticamente el inicio de una unidad prosódica es una cuestión que depende de la fonología particular de cada lengua.

⁶¹ Algunos casos de asimilación se restringen a posiciones débiles, pero siguen perteneciendo, fundamentalmente, al primer tipo de variación.

3. El refuerzo fonético como marca de la estructura prosódica

En el capítulo anterior se ha analizado el importante papel que juegan los constituyentes prosódicos en las fonologías de las lenguas. En el presente capítulo se analizarán los distintos rasgos fonéticos que suelen caracterizar a dichos constituyentes y que pueden ayudar, de este modo, a reconocer la estructura prosódica de una determinada cadena hablada.

Muchos estudios, algunos de los cuales se revisarán a continuación, han demostrado que, efectivamente, la estructura jerárquica de los diversos niveles prosódicos no es una mera abstracción fonológica, sino que tiene un reflejo fonético en diferentes rasgos (lo que da lugar con frecuencia a una caracterización multiparamétrica). Como se verá en los próximos apartados, la marcación fonética de los distintos constituyentes puede realizarse señalando las fronteras entre unidades (en inglés, *edge marking*) o bien señalando sus elementos más prominentes (*head marking*)⁶². Ambas estrategias se corresponden con la función demarcativa y la función culminativa de las que habló ya Trubetzkoy (1973 [1939]), respectivamente.

Entre las dimensiones fonéticas empleadas para marcar los constituyentes prosódicos se integran tanto aspectos habitualmente identificados con los rasgos suprasegmentales (f_0 , intensidad y duración) como también rasgos propios de los tradicionales segmentos⁶³.

Las marcas suprasegmentales aparecen fundamentalmente en los niveles más altos de la jerarquía. Incluyen la presencia de tonos de frontera en los límites de las frases entonativas, el reinicio de la línea de declinación tonal⁶⁴, y las pausas (de mayor duración en los niveles superiores). En cuanto a la prominencia, destaca el aumento de

⁶² La extensión de un constituyente también puede ponerse de manifiesto por la presencia de fenómenos que se propagan a lo largo de toda la unidad, que les sirve de dominio o ámbito de aplicación (*vid.* § 2). Desde luego, esto no puede entenderse como marcación de la prominencia; en cambio, en cierto modo, sí encaja en la idea de marcación de las fronteras, ya que el fenómeno fonético-fonológico en cuestión no puede traspasarlas y, de esta forma, las señala. No profundizaremos más aquí en esta manera de reflejar fonéticamente la estructura prosódica. Para más información, el lector puede referirse, entre otros trabajos, al clásico de Nespoy y Vogel (1994 [1986]).

⁶³ Cantero (2002) advierte que los segmentos son una abstracción basada en los rasgos tímbricos (donde el timbre es el correlato perceptivo que se deriva del patrón de intensidades distribuidas entre las distintas frecuencias del espectro). «La duración de un sonido, entonces, es una abstracción de segundo grado: medimos la duración de un segmento [...] definido por los rasgos de timbre que lo constituyen» (p. 16). Así vista, la duración acústica es una propiedad que no es independiente de los segmentos, aunque tradicionalmente se la caracterice como suprasegmental en el sentido de que en muchos modelos no forma parte de la matriz de rasgos que oponen paradigmáticamente unos segmentos con otros (*vid.* Fox, 2000, § 2, para una discusión sobre los posibles análisis —paradigmáticos y sintagmáticos— de la duración en los casos de lenguas con valor contrastivo de la cantidad).

Por su parte, la duración articulatoria se puede medir (con distintas técnicas, como se verá en los próximos apartados) como la extensión temporal del movimiento de un determinado articulador en su contribución a un gesto articulatorio. A diferencia de la duración acústica, no requiere una sucesión de abstracciones. También a diferencia de ella, puede entenderse como subsegmental, segmental o suprasegmental, dependiendo de los casos (téngase en cuenta que los gestos articulatorios, definidos a la manera de la fonología articulatoria de Browman y Goldstein, 1986, 1989, 1992, poseen una extensión temporal inherente y pueden abarcar periodos que coincidan con nuestra abstracción de lo que es un segmento, pero también pueden durar menos o más que un segmento).

⁶⁴ Se habla de este concepto un poco más adelante.

la duración y de la intensidad propio de las sílabas con acento léxico, y la aparición de objetivos tonales ligados a la sílabas sobre las que recaen los acentos de frase.

Las marcas segmentales se pueden encontrar tanto en los niveles bajos como en los más altos de la jerarquía, aunque son más acusadas en los altos. En general, se manifiestan como un aumento en la duración y la magnitud de los gestos articulatorios situados en la frontera de un constituyente, y también de una manera más notoria en los niveles superiores. Este refuerzo temporal y sustantivo puede afectar a los diferentes gestos articulatorios, correspondientes a la glotis, el velo, el cuerpo de la lengua, el ápice de la lengua, y los labios. También afecta a los movimientos articulatorios de la mandíbula, que actúa de manera conjunta con otros órganos articulatorios en la ejecución de buena parte de los gestos mencionados. Además, el aumento de la duración de los gestos reduce el nivel de coarticulación entre gestos consecutivos, lo cual también puede servir como clave para la detección de las fronteras prosódicas, en torno a las cuales disminuirá el grado de coarticulación.

Estas variaciones segmentales como consecuencia de la estructura prosódica son el foco de atención de esta tesis y serán analizadas con más detalle a lo largo de este capítulo y los siguientes. Sin embargo, antes de proseguir, debemos plantearnos otra cuestión. Para poder dar por buena la idea de que estos tipos de variación fonética ayudan a la reconstrucción de la estructura prosódica por parte del oyente, primero es necesario comprobar si las modificaciones producidas son perceptibles o no. Lo cierto es que todavía no hay demasiados estudios al respecto, y hace falta más investigación sobre el tema, pero a continuación se recogen algunas de las conclusiones a las que han llegado los trabajos consultados, que apuntan a una confirmación de que, en efecto, existen marcas fonéticas perceptibles que permiten identificar la estructura de constituyentes prosódicos.

*Percepción y reconstrucción de la estructura prosódica*⁶⁵

En el capítulo 2 se ha defendido la importante relación que guarda la estructura prosódica con otros aspectos de la fonología de las lenguas, así como con otros niveles lingüísticos. Por ejemplo, la correcta percepción de las marcas de frontera entre unidades prosódicas es importante para la reconstrucción de la estructura sintáctica⁶⁶ y

⁶⁵ Por supuesto, el refuerzo articulatorio no solo ayuda a la correcta interpretación de la estructura prosódica, sino también a la identificación perceptiva de los segmentos. Este efecto es especialmente importante en posición inicial de palabra, por su repercusión en el acceso al léxico (*vid.* § 2.3) y aún más si coincide con el inicio de un enunciado, donde la falta de contexto previo hace que las secuencias segmentales sean menos predecibles (Keating, 2006). Además, son conocidos los efectos de adaptación perceptiva, por los que la respuesta a un determinado estímulo auditivo pierde intensidad ante una exposición prolongada al mismo, debido a la saturación de los transmisores electroquímicos (*vid.* p. ej. Wright, 2004). En posición inicial de enunciado, la presencia de un silencio previo realza, por contraste, cualquier segmento inicial, y lo hace más perceptible (*ibid.*).

⁶⁶ Se trata del proceso que en inglés se denomina *parsing*, que podríamos traducir como *deslinearización*, basándonos en la intuición de que el proceso cognitivo de reconstruir las jerarquías lingüísticas, dada una cadena lineal de sonido, viene a ser lo contrario de lo que sucede en el proceso de producción, es decir, la adaptación de la jerarquía de constituyentes sintácticos a la linealidad impuesta como requisito por el sistema articulatorio. En el marco del Programa Minimista, se llama *linearización* al mecanismo que adapta el educto de la sintaxis *sensu stricto* (provisto de jerarquías) para que sea un aducto aceptable en la

del significado de la oración (§ 2). Asimismo, la identificación de determinadas unidades prosódicas tiene trascendencia para los procesos de segmentación del continuo sonoro y el acceso al léxico (§ 2.3). Dadas estas razones, se puede comprender el interés de los investigadores por estudiar si la estructura prosódica es una mera abstracción mental o si, por el contrario, como parece lógico, tiene un reflejo fonético en los mecanismos de producción del habla y, aún más importante, si dichas marcas fonéticas son perceptibles por los oyentes, que deben saber reconocer dicha estructura.

Sanderman y Collier (1997) comprueban, en efecto, que una agrupación prosódica adecuada repercute positivamente en la comprensión del significado de lo que oyen los participantes de su experimento. La detección de las fronteras prosódicas incluso puede ayudar a una más rápida adquisición de determinados aspectos sintácticos (*vid.* Christophe, Mehler y Sebastián-Gallés, 2001, y las referencias allí incluidas).

Christophe *et al.* (2001) comprobaron que los recién nacidos de entre dos y cuatro días ya podían discriminar entre secuencias de dos sílabas interrumpidas por una frontera de frase fonológica (según se cortaban y extraían de una grabación sonora con enunciados completos) y secuencias bisilábicas compuestas por los mismos segmentos pero sin tal frontera de frase fonológica. También Gout *et al.* (2004) verificaron experimentalmente la repercusión que tenía el reconocimiento de la frontera de frase fonológica sobre el acceso al léxico, con niños de diez y doce meses. El mismo resultado se replica con adultos en Christophe *et al.* (2004), donde se establece que, al llegar a una frontera de frase fonológica, el oyente resuelve todas las búsquedas léxicas pendientes, pues identifica tal frontera con un límite de palabra.

El estudio firmado por De Pijper y Sanderman (1994) pone en juego distintos tipos de fronteras prosódicas, desde la frontera entre palabras hasta la existente entre distintas frases fonológicas, pasando por las opciones intermedias. Sus informantes, oyentes sin ningún tipo de entrenamiento fonético, percibieron como más fuerte la separación entre las unidades que, efectivamente, por conocimiento lingüístico, se predecían como más separadas, incluso en condiciones en las que se distorsionaba el contenido léxico y, por tanto, se impedía la interpretación sintáctica y semántica del mensaje sonoro. Para ello se basaron, especialmente cuanto mayor era la percepción de la fuerza de la frontera, en claves puramente fonéticas como la discontinuidad entonativa, la presencia de pausas, o el reinicio de la línea de declinación tonal. A este respecto, los citados autores aclaran lo siguiente (*ibid.*, p. 2038):

Declination is the tendency of pitch to float down over the course of an utterance, and pitch contours can be thought of as being superimposed on a baseline that falls gradually as a result of this declination ('t Hart *et al.*, 1990). Resets of this baseline usually coincide with breath intake, but resets can also be used as a cue to the presence of a boundary ('t Hart *et al.*, 1990).

También T. Cho, McQueen y Cox (2007) estudian la percepción de las fronteras de frase entonativa y su repercusión sobre el acceso al léxico. En este caso, apuntan a que

las posibles claves fonéticas de la frontera de ese tipo de constituyente son las siguientes: en posición previa a la frontera, alargamiento de las duraciones segmentales, reducción de la coarticulación y presencia de tonos de frontera; y en posición posterior a la frontera, alargamiento, reducción de la coarticulación y refuerzo articulatorio.

El papel perceptivo que tiene el alargamiento acústico encuentra un paralelismo en el ámbito de la percepción visual, donde el alargamiento afecta a claves ópticas⁶⁷. Keating (2006, pp. 18-19) resume así los hallazgos de Auer *et al.* (2004):

[We] tested the visual perceptibility of prosodic boundaries. Talkers read minimal pairs of sentences differing in the presence/absence of boundaries, while movements of reflective markers on their faces, including on the chin, were tracked by the Qualysis infrared system. Some sentences contrasted in presence/absence of a word boundary (e.g. *He was noted for getting the right number* vs. *He just hated forgetting the right number*), while others contrasted in presence/absence of a phrase boundary (e.g. *When you sing, his songs are better* vs. *When you sing his songs, they're better*). The duration of the movement of the marker on the chin was measured, and was found to be longer around a boundary. Acoustic durations were also longer across a boundary. Perceivers then saw sentence fragments and had to decide which sentence (with or without boundary) the fragment had come from. Perceivers were generally good at doing this.

Aunque, como se ha advertido, son necesarios muchos más trabajos sobre la perceptibilidad de las claves de la estructura prosódica, parece que los estudios señalados anteriormente corroboran la idea de que los oyentes son sensibles a la presencia de las fronteras entre constituyentes prosódicos gracias a las distintas marcas acústicas (y ópticas) que las caracterizan. Esto justifica el interés de investigar cuáles son exactamente esas claves acústicas en los distintos niveles de la jerarquía prosódica.

A continuación se expondrán las posibles manifestaciones fonéticas de la estructura prosódica a través del fenómeno conocido como refuerzo articulatorio. Los próximos apartados presentan las características generales del refuerzo experimentado por las posiciones adyacentes a las fronteras prosódicas (§ 3.1) y por las sílabas caracterizadas con distintos niveles de prominencia (§ 3.2). A continuación, se revisarán los estudios que han analizado los efectos de dicho refuerzo sobre los distintos gestos articulatorios (§ 3.3-3.7). Estos trabajos ponen de manifiesto dos cosas igualmente ciertas: por un lado, existen diferencias en la marcación de los distintos niveles prosódicos entre distintas lenguas (incluso entre distintos hablantes); por otro, sin embargo, se pueden intuir algunas tendencias similares, incluso en lenguas que, a juzgar por otros aspectos prosódicos, se consideran tipológicamente muy dispares (Keating, 2006).

⁶⁷ Este paralelismo tiene mucho sentido a la luz de los estudios sobre la dinámica de tareas (en inglés, *task dynamics*), que defienden la existencia de patrones de control motor similares en el ámbito del habla y en otros ámbitos de la (loco-)motricidad (Saltzman y Kelso, 1987; Saltzman y Munhall, 1989).

3.1. Refuerzo por la posición prosódica

En este apartado se introducen los principios generales del refuerzo consonántico motivado por la posición prosódica. Desde los estudios más tradicionales sobre fortición y lenición, se ha considerado que determinadas posiciones —dentro de una sílaba y dentro de una palabra— son más fuertes que otras (para una revisión, *vid.* por ejemplo Honeybone, 2008; Szigetvári, 2008). A partir del desarrollo de la fonología prosódica (Nespor y Vogel, 1994 [1986]; Selkirk, 1978, 1980, 1981, 1986), el estudio de las posiciones fuertes y débiles se extendió a otros constituyentes prosódicos superiores, como se ha recogido en diversos estados de la cuestión (Fletcher, 2010; Fougeron, 1999b; Keating, 2006; ente otros).

La idea más extendida a día de hoy es que la proximidad a una frontera prosódica condiciona el desarrollo espaciotemporal de los gestos articulatorios y puede provocar un aumento de su duración y, a menudo, de su magnitud, entre otras cosas. Dependiendo del gesto articulatorio concreto de que se trate, el refuerzo temporal y sustantivo se puede manifestar mediante distintos tipos de marcas fonéticas, que sirven como huella o «firma» (Byrd *et al.*, 2000) que refleja al menos algunos aspectos de la estructura prosódica de un enunciado.

Como señalan los estados de la cuestión citados, parece que existe consenso acerca de que la diferencia entre el refuerzo correspondiente a los distintos niveles de la jerarquía prosódica es más cuantitativa que cualitativa. En esencia, los mecanismos que subyacen al refuerzo son los mismos en cualquier tipo de constituyente, solo que sus efectos son más intensos en los niveles superiores.

Sin embargo, me gustaría hacer notar una asimetría en la que parece que no se ha insistido lo suficiente. Los estudios tradicionales señalan el carácter fuerte del ataque silábico, frente a la debilidad de la coda (*vid.* p. ej. los citados en el capítulo 4). Así, durante mucho tiempo se ha argumentado que los inicios de constituyente son fuertes y los finales, débiles. En cambio, los estudios sobre niveles superiores a la sílaba (que veremos en este capítulo) reflejan la existencia de un refuerzo en ambos márgenes (inicial y final) de cada tipo de constituyente, mientras que las posiciones relativamente débiles serían las del medio^{68 y 69}.

⁶⁸ Algunos artículos (Fougeron y Keating, 1997; Keating, 2006) apuntan que, incluso en los niveles altos, lo apropiado es hablar únicamente de refuerzo inicial. No habría debilitamiento final (ni, por tanto, un efecto general de declinación a lo largo del constituyente) ni tampoco un patrón en forma de V con debilitamiento medial. Sin embargo, es probable que esta condición esté mediatizada por el empleo de sílabas del tipo CV, donde el segmento que queda al final de cada constituyente es una vocal y no una consonante (por este motivo, y dado que el efecto de la frontera prosódica suele limitarse a los gestos inmediatamente adyacentes, no se puede observar el comportamiento de las consonantes finales, ni para afirmar ni para negar su posible refuerzo). Por otro lado, Keating, Wright, y Zhang (1999) sí comprueban el comportamiento de algunas consonantes en posición final de palabra, y encuentran que hay un mayor refuerzo cuando el final de palabra coincide con el final de enunciado, *vid.* (34) en la p. 101. Estos datos parecen confirmar el refuerzo inicial y final, con debilitamiento medial (patrón en forma de V).

⁶⁹ Keating (2006) hace notar que el debilitamiento normalmente se asocia a una reducción de esfuerzo por parte del hablante y a una menor perceptibilidad para el oyente. Sin embargo —razona—, el debilitamiento no constituye una pérdida de información, sino que es tan informativo acerca de la estructura prosódica como lo pueda ser el refuerzo: el refuerzo marca la frontera entre unidades y el debilitamiento, la continuación de una misma unidad.

En el conjunto de los constituyentes de la jerarquía prosódica, la sílaba tiene un carácter especial. En principio, cualquier constituyente prosódico se compone de un número n de constituyentes del nivel inmediatamente inferior (Nespor y Vogel, 1994 [1986])⁷⁰. En cambio, la sílaba tiene su propia estructura interna con ataque y rima (*vid.* § 2.1.1.2)⁷¹. Resulta difícil comparar la fuerza de las consonantes de las posiciones marginales con la posición media, precisamente porque la media corresponde a la cima, con el núcleo, que suele estar ocupado por una vocal (no una consonante) y, desde luego, ese es necesariamente el caso en español⁷².

Por otro lado, como se ha visto en el § 2.1.1.6, la sílaba es una unidad de coordinación intergestual. Los gestos consonánticos del ataque están coordinados en fase (simultáneamente) con el gesto vocálico, mientras que los gestos consonánticos de la coda se coordinan en antifase (secuencialmente) con la vocal nuclear (*vid.* p. ej. Goldstein *et al.*, 2006). Además, en la coda la secuencia es tal que los gestos de mayor abertura preceden siempre a los más cerrados (Browman y Goldstein, 1995). De una parte, la secuencialidad de los gestos de la coda explica la mayor duración de los segmentos en esta posición. De otra parte, teniendo en cuenta que las coordinaciones en fase son más naturales que en antifase (*ibid.*), Nam (2007) justifica la mayor estabilidad del cierre consonántico en ataque que en coda, lo que equivale a decir que una consonante es más propensa al debilitamiento precisamente en la coda.

En definitiva, la demarcación entre sílabas sucesivas del continuo sonoro se apoya normalmente en el contraste entre vocal y consonante (cuando la sílaba es del tipo básico, CV). En cambio, cuando existen consonantes en coda, la frontera viene señalada, entre otras cosas, por el contraste entre el debilitamiento final de una sílaba y el refuerzo inicial de la siguiente.

En cambio, en los constituyentes superiores a la sílaba —como decíamos— las distintas unidades quedan demarcadas por un efecto generalizado de refuerzo en torno a la frontera (tanto antes como después de esta). Existe un efecto acumulativo, de modo que el refuerzo es tanto mayor cuanto más alto se sitúa el constituyente en la jerarquía prosódica (de hecho, a veces, no se observa refuerzo aparente en los niveles bajos). En cualquier caso, la diferenciación entre niveles no es categórica, ni se produce una distinción entre todos y cada uno de dichos niveles, sino que el grado de refuerzo varía de una lengua a otra y de un hablante a otro.

⁷⁰ En líneas generales, esta afirmación es válida, incluso aceptando la relajación de las condiciones de formación de la estructura prosódica, tal y como propone Selkirk (1996), *vid.* (4), en la p. 35.

⁷¹ Recuérdese lo dicho acerca de la mora (§ 2.1.1.7).

⁷² Hay pocos estudios sobre la fuerza relativa o sobre la dinámica gestual de las consonantes nucleares (o «silábicas») en las lenguas que las tienen. Fougeron y Ridouane (2008) investigan la producción de este tipo de consonantes en un hablante nativo de bereber tachelí y llegan —con las precauciones propias de tener un solo informante— a las siguientes conclusiones (entre otras). La duración del núcleo consonántico no es diferente a la de una consonante similar en otra posición silábica, aunque sí se constata una menor varianza en los datos de duración obtenidos. En cambio, se aprecia un aumento de la magnitud gestual, en el sentido de que son más frecuentes los ejemplares que presentan un contacto linguopalatal con cierre completo. No parece que sea apropiado, pues, hablar de debilitamiento de la posición central en el caso de la sílaba (de hecho, no tendría sentido que el núcleo fuera una posición débil).

The prosodic effects can be seen to differ across speakers and consonants within a language. It differs enough that we cannot say that any single prosodic hierarchy is exhibited by all languages and speakers, or that speakers are marking every level of a hierarchy [but what is universal is the very existence of such hierarchies] (Keating, Cho, Fougeron y Hsu, 2003, p. 159).

A veces se pueden producir interacciones entre diferentes niveles. Por ejemplo, aunque un final de sílaba es, en principio, débil (o más débil que la posición inicial), resulta menos débil cuando se alinea con una frontera (final) de otro constituyente más alto, como puede ser la palabra, donde se produce refuerzo. A la inversa, a pesar del refuerzo propio de las fronteras de enunciado, es más fuerte la frontera inicial —que coincide con un ataque silábico— que la final —cuando aparece una coda— (Byrd, Lee, Riggs y Adams, 2005). El segundo experimento de Keating, Wright y Zhang (1999) también establece una interacción entre la palabra y el enunciado, que se traduce en la siguiente escala:

(34) Orden decreciente de fuerza, a partir de Keating *et al.* (1999)

inicial de palabra en inicio de enunciado
 inicial de palabra en medio de enunciado
 final de palabra en final de enunciado
 final de palabra en medio de enunciado

Los gestos prosódicos o gestos π

Byrd y Saltzman (2003) proponen un interesante modelo que justifica de manera unificada los efectos del refuerzo posicional, tanto de duración como de magnitud, y lo mismo en posición previa que posterior a la frontera prosódica. Para ello parten del modelo de la fonología articulatoria (Browman y Goldstein, 1986, 1989, 1992) y de la dinámica de tareas (Saltzman y Kelso, 1987; Saltzman y Munhall, 1989).

La propuesta se articula sobre la idea de que existen varios tipos de gestos articulatorios. Los gestos articulatorios por antonomasia son los llamados gestos de constricción, responsables de la ejecución de los segmentos consonánticos y vocálicos. A estos hay que añadirles los gestos prosódicos (gestos π)⁷³, cuyo efecto es la ralentización local en la ejecución de los gestos de constricción con los que se solapan temporalmente. Así, la existencia independiente de los gestos π permite abstraer patrones que tendrán la misma aplicación en enunciados de diferente contenido segmental pero igual estructura prosódica.

Los gestos prosódicos tienen magnitud y duración, igual que todos los gestos articulatorios, lo que dota de coherencia al modelo. En principio, los centros de los gestos π se alinean con las fronteras prosódicas, aunque algunos trabajos han explorado la posibilidad de que estén ligeramente escorados a la derecha, *i. e.* hacia el inicio del

⁷³ Katsika, Krivokapić, Mooshammer, Tiede, y Goldstein (2014) proponen la existencia de un tercer tipo de gesto: los gestos tonales, que también estarían coordinados con los gestos de constricción, posiblemente con intermediación de los gestos prosódicos.

segundo constituyente (Byrd *et al.*, 2005; S. Lee, Parrell y Byrd, 2009). En cualquier caso, el carácter temporal de los gestos hace que los gestos π se extiendan a ambos lados de la frontera y afecten tanto a los últimos gestos de constricción de la unidad prosódica que termina como a los primeros de la que empieza.

La magnitud de un gesto π presenta una activación y una desactivación progresivas, de modo que afecta a los gestos de constricción en mayor medida cuanto más cercanos a la frontera se sitúan y de una manera más leve conforme aumenta su distancia con respecto a ella. El pico de magnitud (en realidad, una meseta) representa el grado de ralentización ejercida sobre los gestos de constricción, y es tanto mayor cuanto más alto sea el nivel jerárquico de la frontera. Así, el efecto esperado en las fronteras de distintos niveles no cambia de tipo, sino solo de grado.

El efecto primario de la ralentización en la ejecución de los gestos puede tener, dependiendo de diversos factores, otras consecuencias indirectas. En general, los gestos articulatorios se solapan menos entre sí, de modo que las zonas cercanas a las fronteras prosódicas son menos propicias a la coarticulación.

La dinámica de los gestos es tal que, en posición previa a la frontera (*i. e.* final de constituyente), sobre todo prima el aumento de la duración. En efecto, el alargamiento final suele ser una clave de la frontera de constituyente⁷⁴.

En cambio, en la posición posterior a la frontera (*i. e.* inicial de constituyente), el aumento de duración puede provocar indirectamente un aumento en la magnitud⁷⁵. Según las simulaciones hechas con este modelo, el hecho de que se produzca o no el refuerzo espacial depende de la secuencia de gestos que ocupe el inicio de constituyente.

Una posibilidad es que en posición inicial haya una secuencia de especificaciones para el mismo gesto pero con un valor diferente (p. ej. primero la glotis abierta y luego cerrada, o primero una constricción linguopalatal en un sitio y luego en otro). En ese caso, el retraso relativo en la activación de los gestos produce un «detruncamiento», es decir que, al tener más tiempo para desarrollarse, el primer gesto alcanza una magnitud mayor de la habitual, por no quedar interrumpido por una orden contradictoria. Con este

⁷⁴ Wightman, Shattuck-Hufnagel, Ostendorf, y Price (1992) afirman que, en inglés, la duración de la rima de la última sílaba de cada constituyente puede ser suficiente para distinguir entre los niveles de palabra prosódica, frase acentual, frase intermedia, y el conjunto formado por la frase entonativa y el enunciado. Sin embargo, entre estos dos, no se pudo establecer una diferencia significativa basada en el alargamiento final. Se sugiere que en las unidades mayores la duración deja de ser una clave, pero el correcto reconocimiento de las fronteras está asegurado por la existencia de otras claves, como los tonos de frontera y las pausas. *Vid.* también Katsika *et al.* (2014) sobre la posibilidad de que la aparición de dichas claves se desencadene solo cuando el gesto π alcanza cierta magnitud (y, por tanto, es algo que solo puede ocurrir en los niveles superiores).

⁷⁵ Aunque estos aspectos del desarrollo temporal y la coordinación de los gestos no coinciden exactamente con el fenómeno del ritmo (Fletcher, 2010), es interesante llamar la atención sobre una similitud y una posible relación entre estos dos ámbitos de la temporización del habla. El efecto de la frontera prosódica favorece el aumento de la duración en posición final y el aumento de la magnitud en posición inicial. Esto recuerda a los efectos de la Ley Yámbico/Trocaico (B. Hayes, 1995), que establece que, en una serie de elementos, si uno se percibe como más prominente que los demás por su mayor duración, suele interpretarse como final de grupo, mientras que, si la prominencia viene señalada por una mayor intensidad, se interpreta como inicial. En cualquier caso, nótese que aquí surge el concepto de prominencia, un aspecto de la organización prosódica que en esta tesis se ha querido tratar expresamente por separado de los efectos de posición. Haría falta mucha más investigación para profundizar en esta posible relación.

tipo de secuencias, por tanto, aumenta la duración y la magnitud (del primer gesto). Además, el pico de magnitud es más tardío, como consecuencia de la ralentización local.

En cambio, si en posición inicial hay una secuencia de especificaciones de dos gestos diferentes (relativos a «variables de tracto» diferentes), los efectos sobre el primer gesto consisten en que aumenta la duración —igual que antes—, pero disminuye la magnitud. El pico de magnitud sigue siendo tardío. En el § 3.4.1 veremos una posible aplicación de esto a la resolución de una aparente paradoja en el comportamiento del gesto velar (sirva como anticipo decir que el gesto velar es monovalente —solo se puede especificar el movimiento de apertura— y, por tanto, no es posible que se sucedan dos especificaciones de distinto valor para el mismo gesto, lo que sería el requisito para el aumento de magnitud).

3.2. Refuerzo por la prominencia prosódica

Este apartado expone los principios generales del refuerzo por influjo de los distintos niveles de prominencia prosódica: acento léxico (fundamentalmente el primario, aunque T. Cho y Keating, 2009, sugieren que algunos efectos a veces se pueden propagar a los acentos secundarios en inglés); y acento de frase (prenuclear —*i. e.* el de frase fonológica o frase intermedia— y nuclear —*i. e.* el de frase entonativa—). Como se verá en los siguientes apartados, algunos autores han combinado el estudio del refuerzo por causa prosódica con los casos de prominencia informativa o foco (estrecho y contrastivo)⁷⁶. Por lo que yo sé, no existe ninguna comparación sistemática entre todos estos niveles, al estilo de la que sí se ha hecho con los distintos constituyentes.

Como se ha explicado en el § 2.2, dada una agrupación de constituyentes prosódicos de nivel n en otro de nivel $n + 1$, siempre hay una de las unidades n que resulta más prominente que sus hermanas y se erige en núcleo de la unidad $n + 1$. La sílaba es el único constituyente que no presenta una prominencia interna en el mismo sentido que lo hacen los otros. Por supuesto, la sílaba tiene un núcleo, que es su elemento más prominente. Pero los constituyentes internos de la sílaba no son «hermanos», *i. e.* no son elementos del mismo nivel, de entre los que pueda destacar uno⁷⁷. En esto volvemos a notar una asimetría entre el nivel de la sílaba y los demás, al igual que en los efectos de la posición veíamos que la sílaba es la única que presenta debilitamiento en la posición final (en lugar de en la medial), *vid.* § 3.1.

La construcción de la jerarquía de prominencias mediante una parrilla métrica (M. Liberman, 1975; M. Liberman y Prince, 1977; Prince, 1983) supone que el elemento prominente de un nivel siempre se escoge de entre los elementos prominentes del nivel inferior. Por este motivo, cualquiera de los tipos de prominencia afecta en última instancia a una sílaba. El acento secundario corresponde a la sílaba fuerte del pie. El acento primario, a la sílaba (fuerte del pie) fuerte de la palabra. El acento de frase (prenuclear), a la sílaba (fuerte del pie fuerte de la palabra) fuerte de la frase fonológica.

⁷⁶ En general, la información nueva (más prominente) se asocia con un aumento de duración. Por ejemplo, la primera mención de una palabra en un discurso suele durar más y estar más hiperarticulada que las siguientes apariciones (*vid.* Fletcher, 2010, para una revisión).

⁷⁷ De nuevo, recuérdese lo dicho sobre la mora en el § 2.1.1.7.

Y el acento de frase nuclear, a la sílaba (fuerte del pie fuerte de la palabra fuerte de la frase fonológica) fuerte de la frase entonativa.

El patrón de prominencias se construye por acumulación sucesiva de alternancias métricas entre elementos fuertes y débiles. Esto sugiere que los efectos de la prominencia sobre el desarrollo espaciotemporal de los gestos articulatorios también deben de ser acumulativos y, de hecho, parece que es así, al menos entre los distintos grados de acento de frase y énfasis informativo (Hirst y Di Cristo, 1998; Ladd, 1996, § 5.3.2; y otras referencias citadas en T. Cho y Keating, 2009). En cambio, estos últimos autores consideran que se da un cambio cualitativo en la manera de expresar fonéticamente la prominencia dentro de la palabra frente a los niveles de frase, aunque no explican cuál podría ser la diferencia en el mecanismo subyacente del refuerzo en ambos casos. A este respecto, me gustaría sugerir lo siguiente —aun a falta de una investigación concienzuda—. En el refuerzo posicional veíamos que no hay un aumento significativo de duración en la frontera de enunciado con respecto a la frase entonativa, pero que la intervención de claves adicionales a partir de esos niveles, como los tonos de frontera y las pausas, aseguraba el mantenimiento de la distinción. Del mismo modo, con la prominencia bien podría ocurrir algo parecido, es decir que, en el paso de los niveles inferiores a los superiores, la aparición de otros rasgos prosódicos como los acentos de frase (nucleares y prenucleares) acapare la función de permitir identificar el nivel de prominencia correspondiente.

En cuanto al mecanismo causante del refuerzo, Saltzman, Nam, Krivokapić y Goldstein (2008), basándose en el modelo de gestos π de Byrd y Saltzman (2003) —*vid.* § 3.1—, proponen una categoría más general de *gestos de modulación* (gestos μ). Habría gestos de modulación temporal (gestos μ_T) y de modulación espacial (gestos μ_S). Los gestos π serían un subtipo de gesto de modulación temporal, puesto que suponen la ralentización de los gestos de constricción con los que se solapan. Del mismo modo que los gestos π se encargan del refuerzo posicional, habría otros gestos de modulación temporal responsables del refuerzo por prominencia. De esa manera, las explicaciones para ambos tipos de refuerzo se basarían en la modulación local de la ejecución de los gestos de constricción. Sin embargo, a falta de que se desarrolle más este modelo de gestos μ , la existencia de una serie de diferencias notables entre los efectos de uno y otro refuerzo parece indicar que los motivos subyacentes también deben ser distintos entre sí.

De la parrilla métrica se desprende que el ámbito del refuerzo es siempre una sílaba y, en efecto, como veremos, los efectos de dicho refuerzo se manifiestan en todos los segmentos que componen dicha sílaba. En esto se diferencia del refuerzo por posición, que actúa de un modo mucho más local (en la proximidad de la frontera) y normalmente solo afecta de manera notable al segmento —o incluso al gesto— que resulta estrictamente adyacente al límite prosódico.

Parece sensato pensar que, si el refuerzo por prominencia afecta a toda una sílaba y la sílaba es la unidad de coordinación intergestual, el refuerzo se manifestará como una modificación en dicha coordinación. Beckman, Edwards y Fletcher (1992) modelan el refuerzo por acento de frase nuclear como un retardo en el gesto de cierre (posterior a la

cima vocálica) con respecto al gesto de apertura. Las predicciones de ese modelado coinciden con los efectos constatados empíricamente, a saber, un aumento en la duración de todos los segmentos de la sílaba (y no tan localizado como el que se produce en frontera prosódica); y un mayor desplazamiento en ambos movimientos —apertura y cierre— (recuérdese que en el refuerzo posicional el aumento de magnitud solo ocurre en la posición posterior a la frontera). Además —dicen—, se produce un aumento en la velocidad del gesto de apertura (en claro contraste con la ralentización local provocada por los gestos π y, probablemente, la diferencia decisiva para que desestimemos la existencia de una causa dinámica similar, como propone la teoría de los gestos μ).

En efecto, las consecuencias de este desfase intergestual son múltiples. Al tener más tiempo para desarrollarse, el movimiento de descenso mandibular alcanza mayor magnitud, lo que, en principio, da lugar a vocales más abiertas⁷⁸. Una vocal más abierta da lugar a una mayor radiación de energía acústica al exterior. La peculiar dinámica articulatoria de las sílabas prominentes las hace, por tanto, no solo más largas sino también más intensas, rasgos acústicos tradicionalmente asociados con el acento⁷⁹.

El desfase también explica que en las sílabas prominentes haya un menor solapamiento entre sus gestos y, por ello, una menor coarticulación. Al contrario, la mayor magnitud de los desplazamientos propicia posiciones más extremas de los órganos articulatorios, lo que lleva a De Jong (1995) a proponer que uno de los efectos más característicos de la prominencia es la hiperarticulación localizada (entendiendo hiperarticulación a la manera de Lindblom, 1990). Esta exacerbación de las articulaciones contribuye a la diferenciación paradigmática de los distintos segmentos con respecto a los demás elementos del sistema fonológico⁸⁰.

De nuevo, aquí encontramos otra diferencia con respecto al refuerzo posicional, que favorece el contraste sintagmático entre unos elementos y otros y resalta la ubicación de las fronteras. Además, como se verá al hablar de las consecuencias sobre el gesto velar (§ 3.4.2), la hiperarticulación por prominencia hace a los segmentos nasales más nasales, mientras que el refuerzo por posición los hace menos nasales (§ 3.4.1).

Recapitulando, algunas de las características más relevantes del refuerzo por prominencia, provocado por el desfase entre el gesto de cierre y el de apertura, son las siguientes: incremento progresivo de la duración acústica de la sílaba, en general, y de la vocal, en particular; aumento de la magnitud; aumento de la velocidad de apertura; mayor abertura vocálica, que se traduce en un aumento del primer formante (F1) y una mayor intensidad; menor coarticulación y carácter hiperarticulado de los segmentos, con vocales más periféricas y valores formánticos más extremos (cada uno, dentro de su rango de variación posible). Por supuesto, como aclara Fletcher (2010), no todas las lenguas poseen todas estas características en el mismo grado, ni tienen por qué distinguir el mismo número de niveles de prominencia (en esto, sí ocurre igual que con el refuerzo por posición, *vid.* § 3.1).

⁷⁸ Pero *vid.* § 3.6.2 sobre posibles interacciones entre el gesto de la mandíbula y los de la lengua.

⁷⁹ *Vid.* nota 47 en la p. 77.

⁸⁰ Así, los efectos de este tipo de refuerzo por prominencia son los que mejor encajan con la idea de fortición defendida en el modelo de la fonología natural (*vid.* § 1).

En definitiva, como expresan T. Cho y Keating (2009), los efectos de la prominencia son independientes del refuerzo posicional, tienen causas distintas, y se pueden manifestar en las mismas dimensiones fonéticas, pero también en otras distintas, o de manera diferente en la misma dimensión. Además, por consiguiente, puede haber interacciones entre ambos tipos de refuerzo. En ocasiones, pueden respaldarse mutuamente, pero otras veces, por ejemplo, se halla una especie de efecto techo, de modo que ambos refuerzos no pueden acumularse. Veremos ejemplos de esto en los próximos apartados.

3.3. Gestos glotales

3.3.1. Gestos glotales: refuerzo por posición

Numerosos estudios sobre diversas lenguas han demostrado que el gesto glotal fonológicamente especificado para un determinado segmento se ve acumulativamente reforzado en la posición inicial de los constituyentes prosódicos superiores. A continuación se repasarán algunos de los datos que avalan el refuerzo posicional del gesto de apertura glotal, empezando por los análisis del VOT, el parámetro que, seguramente, más atención ha recibido, y pasando después a otros tipos de mediciones que corroboran la misma idea. En contraste con el refuerzo temporal y sustantivo en posición inicial de constituyente prosódico, se analizará también, a continuación, el debilitamiento en posición final de sílaba. Por último, se argumentará que el refuerzo también puede explicar la habitual inserción de una oclusión glotal [ʔ] ante vocal inicial de constituyente si asumimos la especificación de un gesto de aducción glotal para las vocales.

Empecemos, pues, por los efectos del refuerzo sobre el gesto de abducción glotal⁸¹. Diversos artículos han abordado la modificación del VOT de las oclusivas dentoalveolares sordas en posición inicial de distintos niveles prosódicos y, en general, encuentran un incremento progresivo del VOT conforme aumenta el nivel en la jerarquía de constituyentes. Así lo documentan T. Cho y Keating (2001) para el coreano, con diferencias significativas entre todos los tipos de unidades estudiadas (sílaba, palabra, frase acentual —*vid.* § 2.1—, frase entonativa y enunciado). Keating *et al.* (2003) comparan los resultados del coreano obtenidos en el estudio anterior con otras lenguas. Para el francés, hay un efecto mucho más limitado de la posición prosódica sobre el VOT, y solo se encuentra una diferencia significativa entre el nivel de la sílaba y el de la frase entonativa. En taiwanés no se encuentra efecto significativo de ninguno de los niveles prosódicos. En cualquier caso, cada uno de estos dos estudios solo cuenta con dos sujetos, de modo que la falta de efectos significativos bien puede deberse a la falta de potencia estadística.

⁸¹ Los trabajos que se citarán en este y los siguientes apartados sitúan a las consonantes estudiadas en distintos contextos segmentales, pero la identidad de las vocales colindantes no parece influir en los resultados, ya que todos esos artículos apuntan a las mismas tendencias. Por ese motivo, los contextos no serán mencionados en esta revisión, algo que está en consonancia con el propósito de esta tesis de circunscribir el estudio del refuerzo a los fenómenos de causa prosódica (no contextual).

En inglés, en cambio, parece haber un claro efecto de la posición inicial de frase entonativa sobre el VOT de la /t/ (Pierrehumbert y Talkin, 1992). T. Cho y Keating (2009) también observan un efecto de la posición inicial de enunciado frente a una posición que es simultáneamente interior de enunciado y de frase entonativa, correspondiente a inicio de palabra. En ese estudio encuentran, sin embargo, una interacción entre el efecto de la posición prosódica y la prominencia del acento de frase, de modo que el VOT en inicio de enunciado solo es significativamente mayor que en posición interior cuando esas posiciones no reciben el acento de frase. Además, aunque no hay un efecto principal de la prominencia, el VOT de la /t/ en posición interior de enunciado es mayor cuando dicho segmento se halla en la sílaba sobre la que recae el acento de frase.

En neerlandés, la posición inicial de frase entonativa induce una reducción del VOT en la /t/ y una reducción de la preonorización en la /d/ (T. Cho, 2003). Llama la atención este resultado, en comparación con los señalados anteriormente para otras lenguas. Según los datos comentados, la posición prosódica puede no tener efecto sobre el VOT en algunas lenguas o en algunas circunstancias pero, en caso de tenerlo, parece que siempre está orientado hacia el aumento del VOT. Cho justifica este comportamiento anómalo invocando una diferencia generalizada entre el neerlandés y otras lenguas como el inglés, tal que la /t/ neerlandesa siempre tiene un menor VOT que la inglesa, lo que él atribuye a una especificación fonológica con valor negativo del rasgo [– glotis ensanchada] en neerlandés. El refuerzo por efecto de la posición prosódica, entonces, consistiría en exacerbar la manifestación fonética de ese rasgo. Sin embargo, teniendo en cuenta que las lenguas eligen establecer sus oposiciones laríngicas en torno a uno de los dos rasgos —[± glotis ensanchada] o [± sonoro]— (Jessen y Ringen, 2002), si la /t/ se caracteriza como [– glotis ensanchada] entonces la /d/ debería adoptar el valor positivo del rasgo, algo que no tiene ningún sentido. Una posible explicación alternativa vincularía los cambios de VOT no tanto a una modificación efectiva del gesto glotal, sino de los gestos orales, puesto que el trabajo de Cho documenta una mayor duración acústica tanto de la /t/ como de la /d/ en posición inicial de frase entonativa. Una mayor duración del gesto de oclusión oral, sin ninguna modificación del gesto glotal, podría explicar, en el caso de la /t/, que el desfase entre los finales de ambos gestos se redujese, lo que se traduce, efectivamente, en un menor VOT. En cuanto a la /d/, también una mayor duración del momento oclusivo podría ser la causa de la reducción de la preonorización, por motivos aerodinámicos.

No se pretende aquí dar una interpretación definitiva a los datos del neerlandés, ni tampoco ratificar el uso de uno u otro rasgo laríngeo en la oposición entre /t/ y /d/. Tan solo parecía oportuno apuntar un posible problema de la explicación de T. Cho (2003), así como la existencia de justificaciones alternativas. En cualquier caso, los datos de esta lengua son interesantes en tanto que ponen de relieve el hecho de que, aunque existen ciertas tendencias comunes, los efectos del refuerzo por posición prosódica son distintos en cada lengua.

Dado que el VOT es una medida indirecta, que pone en relación el gesto glotal con los gestos orales, algunos autores han buscado otros parámetros que apunten de manera

más específica al refuerzo del gesto glotal. Por ejemplo, T. Cho y Keating (2001) miden la duración del intervalo sordo de la /t^h/ del coreano (incluyendo la fracción sorda posterior a la distensión) en inicio de sílaba, palabra, frase acentual, frase entonativa y enunciado⁸². La posición prosódica tiene un efecto significativo y permite distinguir todos los niveles entre sí, de modo que parece que esta medida de la duración del gesto de apertura glotal es más fiable que el VOT o que la duración del cierre.

Otra medida empleada para el estudio del refuerzo del gesto glotal es el porcentaje de duración de la porción sonora sobre la duración total del segmento. Kuzla, Cho y Ernestus (2007) detectan que, en alemán, las fricativas sonoras mantienen un mayor porcentaje de duración de la porción sonora cuando encabezan constituyentes prosódicos de bajo nivel en la jerarquía que cuando aparecen al inicio de unidades de nivel superior. En cambio, en los niveles más bajos resulta más fácil que se asimilen a una sorda precedente, dada la menor resistencia de estos niveles a la coarticulación (nótese, no obstante, que este ensordecimiento está desencadenado por el contexto, y no por la posición prosódica). En cualquier caso, aclaran, la diferencia entre constituyentes mayores y menores no afecta a la probabilidad de ensordecimiento sino al grado (medido en términos porcentuales sobre la duración).

Esta medida, además, interactúa con la duración acústica, de modo que, al parecer, en alemán, cuando la posición inicial de un constituyente superior es causante de un mayor ensordecimiento, solo se produce también un alargamiento de la duración total del segmento cuando esto no pone en peligro una oposición fonológica. Por ejemplo, en posición inicial de palabra solo puede darse una /z/ (no una /s/). Si ese inicio de palabra coincide además con un inicio de constituyente mayor, el porcentaje de duración de voz disminuirá y también podrá aumentar la duración total del segmento. En cambio, una /v/ en inicio de unidad superior verá disminuido el porcentaje de duración de voz, pero no aumentará de duración total, para preservar una clave que permita distinguirla de /f/.

El estudio de Pierrehumbert y Talkin (1992) sobre el inglés también utiliza otro parámetro para argumentar a favor del aumento de la abducción glotal en inicio de frase entonativa: se trata de la disminución de la energía (tomada como RMS, *root mean square*, de la amplitud) en el ruido de fricción de la /h/, como consecuencia del menor estrechamiento del canal.

Algunos estudios han hecho uso de medidas más directas (aunque también más invasivas), como la transiluminación en el caso de la tesis de Cooper (1991) sobre el inglés, en la que comprueba también una mayor magnitud y mayor duración del gesto de apertura glotal en inicio de constituyente prosódico superior (*apud* T. Cho y Keating, 2001), o la fibroscopia en el estudio de Jun, Beckman y Lee (1998), con conclusiones similares para el coreano.

Ahora bien, el refuerzo en posición inicial de constituyente prosódico contrasta con el debilitamiento en posición final de sílaba y su consiguiente tendencia a la neutralización de las oposiciones basadas en la especificación de un determinado gesto glotal. En las lenguas en las que el rasgo determinante de la oposición es [$\pm$ glotis

⁸² *Vid.* el trabajo mencionado para más detalles sobre las precauciones metodológicas a la hora de medir el inicio de una oclusiva sorda en posición inicial absoluta (inicial de enunciado).

ensanchada], como por ejemplo el alemán, las obstruyentes especificadas para un gesto de abducción glotal no se manifiestan, sin embargo, como aspiradas cuando aparecen en final de sílaba (John Harris, 2009). Por su parte, en las lenguas donde el gesto glotal sustenta una oposición entre obstruyentes con o sin acompañamiento de fonación, es frecuente que las sonoras subyacentes se vean sometidas a un ensordecimiento pasivo en final de sílaba. En español, por ejemplo, existen las pronunciaciones [aðmi'raɪ] e incluso [aθmi'raɪ] (*admirar*) o [iɲno'raɪ], [ixno'raɪ] (*ignorar*), *vid.* § 4.1.4 para una posible explicación del subsiguiente cambio a fricativa, como modo de respetar el sistema fonológico. En catalán pasa algo parecido tras el ensordecimiento de las obstruyentes laxas, que adoptan una pronunciación oclusiva, para ajustarse a una combinación de características fonéticas ya existente en el sistema fonológico de esa lengua (p. ej. *Madrid* [ma'ðrid]⁸³).

En definitiva, la tendencia de las codas a la reducción de la magnitud del gesto glotal correspondiente da lugar al predominio de un único resultado (la sorda no aspirada). Además de esta propensión neutralizadora, encontramos que algunas lenguas presentan variantes glotalizadas de sus obstruyentes sordas no aspiradas, es decir, que los pliegues vocales no se mantienen ligeramente separados, como correspondería a una producción sorda, sino que llegan a tocarse, aunque sin entrar en vibración. Se trata, por ejemplo, de las pronunciaciones [pʔ], [tʔ], [kʔ] de /p/, /t/, /k/ en inglés, en las que a veces pueden llegar a imponerse las claves perceptivas del elemento laríngeo y desembocando en una mera oclusión glotal [ʔ]. Este fenómeno de debucalización afecta en primer lugar a la coronal, seguida de la labial y la dorsal, orden que Blevins (2004) justifica del siguiente modo:

The fact that closure durations may differ for stops at different points of articulation means that debuccalization itself may occur first with the stops of shortest duration, and only later with those whose closure durations are longer. Coronal closure durations can be shorter than those for labial and velar stops, and tongue-tip movements show higher velocities than tongue-dorsum or lip movements (p. 121).

Por este motivo, en inglés la debucalización de /t/ en coda con resultado [ʔ] se ha generalizado en muchas variedades dialectales mientras que todavía se encuentra variación entre [pʔ] y [ʔ], por un lado, y entre [kʔ] y [ʔ], por otro (*ibid.*). En algunas lenguas, como el chino hablado en Fuzhou, la debucalización ha afectado a las tres consonantes, con la consiguiente neutralización del punto de articulación (*ibid.*).

A continuación intentaré defender la idea de que un fenómeno habitual en muchas lenguas como es la inserción de una oclusión glotal [ʔ] ante vocal inicial de constituyente prosódico también se puede explicar como un efecto de refuerzo. Para ello, suponemos que las vocales deben estar caracterizadas por un gesto de aducción glotal (para la consecución de la vibración laríngea). Al igual que hemos visto con las

⁸³ Parece preferible transcribir [d̥] en lugar de [t] para recordar el origen fonológico de ese segmento. Incluso es probable que existan claves secundarias que permitan discriminar entre [d̥] y [t] en esa posición, aunque no es este un tema en el que se quiera profundizar aquí, por motivos de espacio.

consonantes, la posición inicial provoca un refuerzo espacial del gesto glotal especificado. Si en las consonantes el incremento sustantivo de la abducción glotal conducía a un aumento del VOT, en las vocales el refuerzo de la magnitud de un gesto de aducción puede ser la causa de una oclusión glotal momentánea⁸⁴. En ambos casos se trata del mismo fenómeno de refuerzo, pero en dirección contraria, debido a la diferente especificación gestual subyacente. Felizmente, además, la inserción de este elemento oclusivo da lugar a una sílaba con ataque, resultado preferido sobre la situación de partida, en la que el ataque se encontraba vacío de contenido. Es decir, sin rechazar la explicación tradicional que justifica la inserción del momento oclusivo [ʔ] como estrategia para conseguir un tipo silábico preferido, me parece que tiene sentido proponer que la aparición de tal elemento está causada por los efectos que las fronteras de constituyente prosódico tienen sobre la ejecución motora de los gestos articulatorios.

Nespor y Vogel (1994 [1986]) proponen que algunas lenguas pueden tener reglas de reestructuración de determinados constituyentes prosódicos, que se aplican una vez formados estos, cuando se procede a la construcción de alguno de los niveles superiores. Por ejemplo, es conocido el caso de las reglas de resílabeo de lenguas como el español o el francés (*vid.* § 2.1.1.5). En estas lenguas, aunque la forma fonológica de una palabra morfosintáctica pueda incluir una vocal inicial, sin embargo, por efecto del resílabeo, dicha vocal no coincidirá normalmente con una frontera de palabra prosódica, de modo que no se producirá el refuerzo de su gesto de aducción glotal (no se insertará un elemento [ʔ]). En cambio, en las lenguas sin resílabeo, como el inglés o el alemán, es muy frecuente la aparición de la oclusiva glotal ante palabras que empiezan por vocal, y tanto más si aparecen en inicio de constituyente prosódico de más alto nivel.

Interacción entre los fenómenos de refuerzo y la base de articulación

A lo largo de este apartado hemos visto cómo la posición inicial de distintos constituyentes puede tener un efecto en el detalle fonético de la distensión aspirada de las consonantes oclusivas (medido como variaciones en el VOT). En esta subsección nos centraremos en la alternancia fonológicamente regulada entre variantes aspiradas y no aspiradas en inglés.

En esta lengua, los alófonos aspirados [p^h], [t^h], [k^h] aparecen en posición inicial de palabra o en inicio de sílaba fuerte (aunque no reciba acento primario, sino solo secundario). Es decir, lo que tienen en común estas dos condiciones es ser inicio de pie. Dada la importancia del pie en la fonología de la palabra en inglés (para establecer los

⁸⁴ Pierrehumbert y Talkin (1992) interpretan que no siempre es necesario un cierre total pero, al menos, una mayor tensión puede ser la causa de la desestabilización del patrón vibratorio de los pliegues vocales, propia de las glotalizaciones. *Vid.* también Dilley, Shattuck-Hufnagel, y Ostendorf (1996) y las referencias allí citadas. Dilley *et al.* relacionan frecuencia y magnitud de la glotalización con la posición inicial de frase intermedia y frase entonativa, la prominencia propia del acento de frase, y el refuerzo propio de las palabras menos frecuentes. Además, perfila un posible paralelismo entre la glotalización inicial y la aparición de voz crepitante (en inglés, *creaky*) en el final de enunciado, como rasgos con capacidad de demarcación de constituyentes. Este apunte me sugiere —aun a falta de una investigación más sistemática— que la desestabilización propia de la voz crepitante podría estar motivada por el refuerzo final de enunciado, y encontraría su causa material en la mayor lentitud del gesto de aducción glotal, que tardaría más en alcanzar su pico de magnitud. Esto dejaría los pliegues vocales más laxos, una de las posibles características fisiológicas de esta cualidad de voz, pero *vid.* Titze (2000) para más detalles.

acentos rítmicos), llama la atención que ninguno de los estudios revisados en este apartado incluya el nivel Σ entre los constituyentes investigados.

Por otro lado, las posiciones en las que aparecen las variantes aspiradas forman una clase natural (inicio de pie), mientras que no puede decirse lo mismo de las posiciones en las que encontramos los alófonos con distensión simple [p], [t], [k], que constituyen la condición alternativa (en inglés, *elsewhere condition*, *vid.* Kiparsky, 1973). Por ello, cabe suponer que la dirección del proceso de cambio es desde la simple hacia la aspirada. En definitiva, las aspiradas son el producto de un refuerzo desencadenado por la posición inicial de pie (y, como se ha visto anteriormente, una vez dentro de la categoría de las aspiradas, el VOT puede seguir incrementando progresivamente según estos segmentos se sitúen en el inicio de constituyentes cada vez más altos en la jerarquía).

Ahora bien, existe cierta controversia acerca del carácter fuerte de una aspirada frente a su correspondiente simple (*vid.* Szigetvári, 2008). Por ejemplo, en el marco del método verbotal para la corrección de la pronunciación en una lengua extranjera (Renard, 1971), se atribuye una menor tensión a las aspiradas. En efecto, mi práctica docente con estudiantes anglohablantes, con quienes aplico este método, me corrobora que, para eliminar la aspiración en la pronunciación de mis alumnos de español, debo conseguir que aumenten su tensión articulatoria. Para ello, se puede recurrir a ejercicios encaminados a un aumento global de la tensión corporal, como pedirles que hablen mientras hacen fuerza empujando una pared.

¿Cómo es posible, entonces, que el producto de un refuerzo sea un sonido aparentemente menos fuerte? La solución que propongo alude al concepto de base de articulación, en el sentido de los ajustes y hábitos articulatorios propios de una lengua que no atañen a la pronunciación de un tipo de segmento concreto, sino que se dan de manera generalizada en un enunciado⁸⁵ (para más detalles, *vid.* Gil Fernández, 2007, cap. 3; Puigvert Ocal, 2015). Una de las características especificadas por la base de articulación es, precisamente, el nivel de tensión general empleado en la articulación. Las lenguas difieren en el nivel de tensión «soportado» y así, por ejemplo, el inglés presenta una menor tensión articulatoria que el español⁸⁶.

Pues bien, dado el refuerzo sustantivo del gesto de abducción glotal en la posición inicial de pie, se predice un aporte de un mayor flujo de aire hacia las cavidades supraglóticas, con el consiguiente aumento de la presión oral (P_o). Si la base de articulación de una lengua establece un nivel de tensión que tolere esa presión intrabucal, los órganos implicados en el cierre oclusivo a nivel oral podrán mantener el contacto durante más tiempo. En ese caso, la distensión —cuando ocurra— puede coincidir con el receso del gesto glotal y, por ello, no se apreciará aspiración. En cambio,

⁸⁵ La base de articulación se puede modificar conscientemente bajo determinadas circunstancias, por ejemplo, para imitar a una persona en casos de parodia. Mi intuición me dice que, en estos casos, el ámbito en el que se inscribe la modificación puede ser un constituyente inferior al enunciado, como la frase entonativa.

⁸⁶ En el marco de la teoría de la optimidad, Kirchner (2001) propone una familia de restricciones, LAZY, que me recuerda a esta idea, aunque LAZY sirve para minimizar el esfuerzo articulatorio en términos de desplazamiento con respecto a una supuesta posición de reposo.

en una lengua como el inglés, con una base de articulación relativamente menos tensa, cuando el nivel de P_0 supere el umbral de tolerancia de presión establecido por esa base de articulación, los órganos articulatorios encargados del cierre oral se separarán para permitir que disminuya la presión. En ese momento, con los pliegues vocales aún separados, se percibirá una distensión aspirada.

Esta explicación permite conciliar la idea de que las aspiradas son sonidos menos tensos con el hecho de ser el producto de un refuerzo. La explicación se puede extender a las africadas. Al fin y al cabo, las aspiradas y las africadas son subtipos de oclusivas que tan solo difieren en su distensión (Martínez Celdrán y Fernández Planas, 2013). Frente a las oclusivas simples, con una distensión más rápida, las oclusivas africadas presentan una distensión más lenta, en la que se forma un estrechamiento que produce el ruido de fricción característico. A su vez, las oclusivas africadas y aspiradas solo se diferencian en que en las primeras el punto más estrecho, en el que se produce la fricción, es el mismo que el punto de articulación de la fase oclusiva, mientras que en las aspiradas la fricción es invariablemente glotal, ya que en este caso es ese el punto más estrecho del tracto vocal en ese momento.

La lingüística histórica ofrece datos de oclusivas que evolucionan a africadas en posiciones fuertes y, de ahí, se simplifican en fricativas. Por ejemplo, las siguientes evoluciones ocurrieron durante el cambio consonántico del alto germánico (Bybee, 2001, p. 80):

(35) $p > pf > f$

(36) $t > ts > s$

(37) $k > kx > x$

El paso de oclusiva a fricativa (en su estadio final) suele considerarse una lenición. Puede resultar llamativo el desencadenamiento de una lenición en una posición fuerte, como la inicial de palabra. Sin embargo, si el refuerzo lleva a un aumento de la P_0 por encima del nivel tolerado por la base de articulación, sí se entiende el cambio a africada. Una vez en este estadio, puede ocurrir que las claves perceptivas del momento fricativo se impongan y se recategorice el segmento como una simple fricativa.

No obstante, no todos los casos de africación merecen el mismo análisis. Para Lavoie (2001), «affricates [...] play different phonetic roles, representing either weakening or strengthening depending on the environment» (p. 6). En efecto, las africadas (normalmente de tipo sibilante) que proceden de un proceso de palatalización no se explican por refuerzo posicional, sino por asimilación contextual y, como tales, su aparición se ve favorecida en posiciones débiles, como el interior de palabra y las sílabas no acentuadas, ya que son estas posiciones donde se produce un mayor solapamiento entre gestos articulatorios sucesivos.

3.3.2. Gestos glotales: refuerzo por prominencia

En este apartado analizaremos brevemente los efectos de la prominencia prosódica sobre los gestos glotales, aunque, por lo que parece, la bibliografía no le ha prestado

toda la atención necesaria a este aspecto. Incluso, la revisión de Fougeron (1999b) no incluye, en el apartado dedicado al refuerzo por prominencia, ninguna mención a los gestos de la laringe (sí a otros, como el gesto velar, los gestos de la lengua o los de los labios). T. Cho (2003), en cambio, sí que comenta el comportamiento de la /t/ y la /d/ del neerlandés cuando aparecen afectadas por el acento léxico y por el acento de frase. En general —dice—, uno de los cambios más notorios es el aumento de la duración acústica. Por otro lado, la [t] ve reducido su VOT. A primera vista, esto puede sorprender, en comparación con otros estudios mencionados anteriormente, donde el aumento del VOT era, precisamente, una de las consecuencias inmediatas del refuerzo. Sin embargo, hay que tener en cuenta que el neerlandés es una lengua cuyas oposiciones laríngeas se basan en la presencia o ausencia de voz, no en la aspiración, como demuestra el hecho de que la [d] presente presonorización. Por tanto, la /t/ se caracteriza por la ausencia de especificación de rasgo glotal. El reajuste intergestual, característico de las sílabas prominentes (*vid.* § 3.2), explica la mayor duración de los gestos articulatorios de la consonante (y, por ello, de todo el segmento como tal). La prolongación de los gestos de oclusión oral puede reducir el desfase entre la distensión y el comienzo de la vibración laríngea propia de la vocal siguiente y disminuir el VOT. En cuanto a la [d], en condiciones de prominencia aumenta la duración de su fase presonorizada. Por un lado, esto tiene sentido a la luz de la prolongación del gesto de aducción glotal propio del miembro sonoro de la pareja. Por otro, en cambio, se supone que también aumenta la duración del gesto de oclusión oral, que resulta aerodinámicamente contradictorio con la continuación de la vibración fonatoria. Seguramente, dada la tendencia de los efectos de prominencia a poner de relieve las distinciones paradigmáticas entre los elementos del sistema, es posible que aquí estén entrando en juego alguna de las estrategias compensatorias que pueden prevenir el ensordecimiento pasivo (*vid.* § 4.1.4), para asegurar la prolongación de la sonorización, pero T. Cho (2003) no aporta más datos, así que no podemos estar seguros.

Ninguna de las referencias aquí manejadas hace mención explícita al efecto de la prominencia en una lengua de aspiración, pero la predicción parece seguirse de manera sencilla: el miembro aspirado aumentaría el VOT, mientras que el no aspirado quedaría sin cambios o incluso lo reduciría (recuérdese siempre que el VOT no solo depende del gesto glotal, sino también del oral y, si el primero permanece inalterado pero el segundo se alarga, el resultado final es una disminución del VOT). El caso no aspirado es, de hecho, equivalente al sordo en una lengua de sonorización (*i. e.* son los miembros no marcados, que carecen de especificación de gesto glotal, y donde los pliegues vocales se acomodan, pues, a una posición tan solo de ligera separación). En cuanto a la consonante aspirada, el retraso de la vocal con respecto a la consonante desembocará en un mayor VOT, lo que además consigue aumentar la diferencia paradigmática con respecto al miembro no aspirado.

Por su parte, Y. Lee (2011) estudia los efectos del acento de frase en el VOT de las oclusivas sordas del inglés en contexto *sC* y encuentra una interacción con el refuerzo posicional de inicio de frase entonativa. Aunque la posición no tiene un efecto principal (sí lo tiene la prominencia), el VOT se reduce tanto en la condición acentuada como en

la posición inicial no acentuada. La reducción del VOT se interpreta como una exacerbación de los efectos de la regla de no aspiración tras /s/:

Prosodic strengthening [...] is closely intertwined with the phonological system of a given language, not simply through enhancing the phonemic contrast between sounds in the language, but at the level of phonetic implementation of a phonological rule (p. 1193).

3.4. Gesto velar

3.4.1. Gesto velar: refuerzo por posición

La posición del velo del paladar y, en consecuencia, el grado de abertura velofaríngea (v-f) también están sujetos a variación por causa de factores prosódicos⁸⁷. Veamos, en primer lugar, las consecuencias del debilitamiento propio de final de sílaba, para ver luego el refuerzo progresivo en posición inicial de los distintos niveles.

Krakow (1993) utiliza el sistema Velotrace (Horiguchi y Bell-Berti, 1987) para monitorizar la posición vertical del velo en dos hablantes de inglés. Según esta autora, las consonantes nasales en la posición débil de coda se caracterizan por una posición del velo más baja que en ataque silábico. Además, dado un segmento [m], la coordinación entre el movimiento de descenso del velo y el del cierre labial también es sensible a la posición que ocupa la consonante en la estructura silábica. En posición inicial, el objetivo articulatorio del velo descendido se obtiene de manera simultánea a la consecución del cierre oral. En cambio, en coda, el final del movimiento del velo se sincroniza con el inicio del movimiento labial; es decir, que el descenso del paladar blando se anticipa en posición final de sílaba, y es por eso que en esta posición se produce una mayor coarticulación con la vocal precedente, que se nasaliza. Esta constatación empírica sustenta la idea de algunos modelos actuales que atribuyen a la posición de ataque un tipo de coordinación gestual en fase (simultánea), y a la coda en antifase (secuencial), *vid.* § 2.1.1.6.

En cuanto al comportamiento de las consonantes nasales en contexto intervocálico interior de palabra, Krakow (1993) detecta lo siguiente. Cuando la sílaba acentuada es la precedente, se dan las condiciones típicas que en inglés desencadenan el carácter ambisilábico de una consonante (*vid.* § 2.1.1.4). Sin embargo, las nasales del estudio de Krakow no parecen mantener características híbridas entre la coda y el ataque. Por el contrario, se asemejan por completo a las nasales de final de sílaba (con lo que eso implica, según se acaba de explicar), aunque bajo ciertas circunstancias pueden mantener las propiedades típicas de la posición de ataque (y, en ese caso, se comportan como tal a todos los efectos).

⁸⁷ Indudablemente, también existen factores segmentales (aparte de la diferencia obvia entre segmentos orales y nasales), relacionados sobre todo con el modo de articulación de las consonantes. No obstante, este tipo de causas de la variación quedan fuera del objetivo de esta tesis y, por ello, no serán analizadas con más detalle. Para un estudio más pormenorizado sobre la posición del velo en función de la identidad de cada segmento y también de su contexto, *vid.* Vaissière (1988) y, sobre todo, Bell-Berti (1993).

These data provide little support for the notion of ambisyllabicity as a combination of syllable-initial and syllable-final attributes, but the data do suggest that some word-medial consonants are likely to show free variation with respect to syllable affiliation. Considering the two that varied here, i.e., the nasal consonants in *helmet* and *Seymour*, the following observations can be made: A speaker may affiliate a medial consonant with the upcoming syllable instead of the preceding syllable which bears primary stress (1) when the stressed syllable already contains a post-vocalic consonant, as in the case of *helmet*, or (2) when the following syllable receives secondary stress, as in the case of *Seymour* (Krakow, 1993, p. 97).

Pasemos ahora a analizar el efecto del refuerzo en posición inicial de constituyente. A este respecto, Krakow llega a la conclusión de que la posición inicial de enunciado (la más fuerte) es la que presenta una mayor elevación del velo, mientras que, a medida que se acerca el final del enunciado, se detecta una posición más baja de este órgano, en una especie de declinación comparable a la declinación tonal en el ámbito de la entonación.

En efecto, una serie de estudios han relacionado las posiciones prosódicas altas con una menor amplitud (RMS) de la señal acústica en los segmentos nasales como la [n], lo que se interpreta como una posición más elevada del velo: *vid.* p. ej. Gordon (1996) para el estonio, o T. Cho y Keating (2001) para el coreano. En este último artículo, la menor energía se obtiene cuando la nasal está al inicio de una frase entonativa. En cambio, no hay diferencias estadísticamente significativas entre los niveles de la frase acentual y la palabra, pero estos a su vez sí se diferencian de la sílaba, en cuyo inicio se encuentran las nasales con mayor energía (por tanto, la posición velar más descendida).

Además de la energía, algunos autores han abordado este tipo de estudios mediante otro parámetro acústico: la mera duración del segmento. El propio Gordon, en su artículo sobre el estonio anteriormente citado, encuentra una duración de la [n] progresivamente mayor según se sitúe en inicio de sílaba, palabra, frase entonativa o enunciado, en dos de sus tres sujetos, pero el otro presenta la tendencia contraria. Por su parte, los dos sujetos francófonos de Keating *et al.* (2003), al igual que el último sujeto de Gordon (1996), presentan una menor duración acústica de la [n] en los constituyentes superiores (las unidades estudiadas en ese artículo son enunciado, frase entonativa, frase fonológica, palabra y sílaba). Tiene sentido pensar que la posición más elevada del velo y la menor energía nasal hayan inducido una segmentación acústica, a la hora de hacer el análisis, tal que el segmento nasal resultara más breve.

Los resultados correspondientes a estas medidas acústicas concuerdan con las mediciones del flujo nasal (U_n) realizadas por el propio Gordon, que evidencian un menor valor en posición inicial de las unidades prosódicas superiores. Fougeron (2001) encuentra la misma tendencia en francés con tres de sus cuatro informantes (el otro se comporta justo al contrario), aunque solo halla diferencias significativas entre el nivel de la frase entonativa y los inferiores a él, pero no entre estos (frase acentual, palabra, sílaba). Además de estudiar la [n], Fougeron también escoge una vocal nasal del francés,

la que la tradición transcribe como [ã]⁸⁸. También con la vocal obtiene menor U_n en frase entonativa, aunque en este caso el efecto solo es significativo en dos sujetos.

Al igual que pasaba con los estudios sobre los gestos glotales, las medidas más directas —más invasivas— las encontramos en las publicaciones más antiguas. Vaissière (1988) coloca a dos informantes anglófonos un sensor de posición sobre la parte móvil del paladar blando, introduciéndolo a través de la cavidad nasal, pegado a una fina tira de plástico flexible. Mediante esta técnica, constata que, en presencia de una frontera prosódica el velo se sitúa más alto, también en las consonantes orales. Además, se produce una menor coarticulación anticipatoria de la bajada del velo si hay un segmento nasal cercano en el contexto; es decir, que en una secuencia CVN, la consonante oral (C) se ve menos afectada por la nasal (N). En consecuencia, al haber una menor anticipación, la velocidad del movimiento de descenso velar es mayor, para alcanzar su objetivo articulatorio a tiempo. Esto es así tanto si la C es inicial de palabra como si la C es final de un constituyente superior y la secuencia CVN se establece a través de su frontera.

En definitiva, existen evidencias articulatorias, aerodinámicas y acústicas de que el refuerzo posicional del gesto velar en inicio de constituyente prosódico se manifiesta como un ascenso del velo —tanto en segmentos orales como nasales—, mientras que el debilitamiento propio de la posición final de sílaba (y, al parecer, también de enunciado) aflora como un velo más descendido. Llegados a este punto, cabe formularse varias preguntas. ¿Por qué el refuerzo de un segmento nasal debería hacerlo menos nasal (y el debilitamiento, más nasal)? ¿Por qué el efecto del refuerzo es el mismo para dos tipos de segmentos contrapuestos, a saber, orales y nasales? ¿A qué puede deberse el efecto de declinación, que se traduce en un debilitamiento en frontera final de enunciado (una posición típica de refuerzo)? En lo que sigue, intentaré argumentar que la particular fisiología del velo palatino y el carácter monovalente (no binario) de la especificación de la nasalidad son los responsables de estos comportamientos.

Piñeros (en prensa) reconoce que en la tradición lingüística hispánica normalmente se ha aludido a la oposición entre segmentos nasales y orales como aquella entre los dos valores, positivo y negativo, de un rasgo binario, [+ nasal] y [– nasal] (Alarcos Llorach, 1965; James W. Harris, 1969; Núñez Cedeño y Morales-Front, 1999; y otras referencias citadas por Piñeros). Sin embargo, a continuación argumenta que la oralidad es fonológicamente inactiva, puesto que no hay ningún proceso fonológico que deba generalizarse a la totalidad de los segmentos orales. Dicho de otro modo, ninguna regla precisa aludir a [– nasal] (por ejemplo, el carácter nasal se puede propagar por asimilación, pero no ocurre lo mismo con su contrario).

Steriade (1993) va un paso más allá y defiende la inactividad fonológica de la oralidad como característica universal. En efecto —dice—, existe una explicación más adecuada para los casos de aparente asimilación (preoralizaciones y postoralizaciones), así como, en general, para los contornos segmentales con una parte nasal y otra oral.

⁸⁸ Pero *vid.* Montagu (2004), que justifica la mayor adecuación del timbre [ɔ̃] para esta vocal, mientras que la tradicional [ɔ] resulta más cerrada, [õ], y la [ẽ] pasaría a [ã], a juzgar por la cualidad tímbrica de la porción vocálica inmediatamente precedente al inicio de la nasalidad, al menos en el francés parisino contemporáneo.

Basándose en una propuesta de Ladefoged (1971) y de Kingston (1985), Steriade afirma que todos los segmentos [+ interrupto] poseen, inherentemente, una estructura interna de tipo contorno, con dos momentos: $A_0 + A_{\max}$ (donde A = apertura), *i. e.* cierre + distensión, si bien es cierto que la distensión puede estar ausente bajo ciertas circunstancias⁸⁹. En su análisis, las nasales postoralizadas del guaraní, por ejemplo, se caracterizan fonológicamente como oclusivas sonoras orales, y la nasalidad es tan solo un rasgo coadyuvante de la sonoridad (por motivos aerodinámicos) durante la fase de cierre, mientras que la presencia de la distensión oral facilita la identificación de los rasgos subyacentes. Este enfoque permite explicar por qué, en semejantes condiciones contextuales y prosódicas, las consonantes [– interrupto] no experimentan la misma postoralización —por no poder crear contornos— a pesar de ir seguidas de vocal oral (algo que en el análisis asimilatorio carece de justificación).

Otros posibles usos de [– nasal], como el bloqueo de la armonía nasal por parte de segmentos orales, también se pueden descartar, ya que existen otras explicaciones posibles, como la incompatibilidad entre los rasgos sordo y nasal, sumada a la necesidad de colindancia para la propagación del rasgo.

Por último, la presencia del término marcado (las nasales) en un inventario fonológico necesariamente implica la existencia de orales (término no marcado). Para Steriade (1993), este tipo de universal solo es aplicable a rasgos privativos, pero no a equipolentes como, por ejemplo, $[\pm \text{ATR}]$ ⁹⁰, donde puede haber sistemas con uno solo de los valores (cualquiera de ellos) sin que se impliquen uno al otro.

Por todas estas razones, asumimos el carácter privativo del rasgo [nasal], que es el que puede aparecer en las representaciones fonológicas y al que pueden aludir las distintas reglas.

En términos fisiológicos, también parece ser que no hay una oposición entre una acción muscular y otra contraria, sino simplemente entre la activación o desactivación de una misma acción muscular. Para Bell-Berti (1993, pp. 65-67), el movimiento de subida del velo, responsable del cierre del paso v-f, depende de la acción de un músculo llamado *levator veli palatini* o, simplemente, *levator palatini*. Por otro lado, aunque reconoce que el debate está todavía abierto, puede que también se encuentren implicados algunos músculos que se encargan de estrechar la faringe, con lo que acercan al velo la pared trasera y los laterales de la faringe, y contribuyen así al cierre del paso del aire. En cuanto al movimiento de descenso del velo, hay cierta controversia acerca de si interviene la acción directa de un músculo, pero los estudios electromiográficos de Bell-Berti apuntan más bien a la simple relajación del *levator palatini*, y el propio efecto de la gravedad. Es decir, que no hay un movimiento activo de ningún músculo para abrir el paso v-f, ni siquiera en los casos en que alternan los

⁸⁹ De aquí se sigue la existencia de distintos tipos de distensiones en una oclusiva (simple, aspirada, africada). Además, otros tipos de contornos consonánticos coinciden justamente con los tipos de segmentos interruptos (nasales y laterales). Mielke (2005) plantea que las laterales se comportan fonológicamente como [+ interrupto] en algunas lenguas, y como [– interrupto] en otras. Teniendo en cuenta el análisis de $A_0 + A_{\max}$, predigo que las lenguas con contornos de lateralidad necesariamente implican un valor [+ interrupto] de las laterales.

⁹⁰ Siglas en inglés del rasgo *Advanced Tongue Root* (raíz lingual adelantada).

sonidos orales y nasales en la cadena lineal. Algunos estudios sí reconocen la intervención secundaria del *palatoglossus* —en realidad un músculo relacionado con los movimientos de la lengua, sobre todo en el caso de la articulación de las vocales posteriores—. Aunque este músculo se active con otra intención, su movimiento puede tirar del velo hacia abajo, una vez que el *levator palatini* deja de hacer esfuerzo (Dixit, 1991; Dixit, Bell-Berti y Harris, 1987). Es decir, si el *levator palatini* actúa, el *palatoglossus* no puede contrarrestarlo pero, si el primero deja de actuar, el otro puede acelerar el descenso natural del velo, que se produciría de todas maneras por un mecanismo pasivo.

Ahora bien, por un lado tiene sentido que un rasgo fonológico privativo se corresponda con una acción muscular del mismo carácter. Sin embargo, debe notarse que el rasgo fonológico marcado, [nasal], equivale a la desactivación muscular, desde el punto de vista fisiológico, y esto puede parecer paradójico. Chomsky y Halle (1976 [1968], p. 166) dicen que en posición de reposo el velo está descendido (para permitir la respiración nasal), mientras que justo en la preparación inmediatamente anterior a una locución se eleva, y es esa su posición por defecto para el habla. Esto tiene sentido desde el punto de vista aerodinámico y acústico: si el paso velofaríngeo estuviera abierto, se debilitaría la intensidad general de todos los sonidos, y además introduciría antirresonancias. El caso del velo palatino es, probablemente, uno de los mejores ejemplos de que la posición por defecto —o de retorno— para el habla no necesariamente coincide con la posición de reposo⁹¹. Si la activación del *levator palatini* se considera la opción por defecto para el habla, se entiende que lo que debe especificar la fonología son los momentos de desactivación de dicho músculo, que es lo que codificamos mediante el rasgo privativo [nasal].

Dicho esto, ahora podemos reconsiderar por qué el efecto que tiene la posición inicial de constituyente prosódico es el de aumentar la elevación del velo. En el § 3.1 se describió cómo la dinámica de tareas y, en concreto, el modelo de gestos prosódicos (gestos π) de Byrd y Saltzman (2003) predicen el comportamiento por efecto de una frontera prosódica. En el caso de un gesto seguido de otro gesto diferente pero referido a

⁹¹ Gil Fernández (2006) aporta numerosos ejemplos de ello. En su apartado § 2.7, afirma, a propósito de la base de articulación, que esta «sería aquella [disposición] que naturalmente adoptarían los articuladores inmediatamente antes de comenzar toda emisión en la lengua en cuestión y a la que volverían en cada pausa producida durante la misma». A continuación, añade en nota a pie de página: «Esta posición no debe confundirse en ningún caso con la posición de reposo absoluto mantenida en la respiración normal en la cual el velo del paladar está bajado, los labios juntos, la mandíbula inferior ligeramente bajada, y el ápice de la lengua apoyado en la región alveolar». Para apoyar esta afirmación, cita a Alonso (1974, pp. 263-264), quien testimonia lo siguiente:

En las inscripciones quimográficas, ya preparado el sujeto de la experiencia con la bocina colocada, pero sin haber iniciado la pronunciación, la plumilla marca sobre el cilindro una línea recta (posición de descanso). Pero unas 40 centésimas de segundo antes de iniciar la pronunciación, la plumilla da un salto de un par de milímetros, aunque la línea sigue siendo recta en su nueva altura hasta la aparición del primer fonema. La membrana del tamborcillo ha sufrido un minúsculo desplazamiento (no debido a salida alguna del soplo) que se mantiene invariable hasta la aparición de las articulaciones, y éstas se marcan sobre el papel ahumado partiendo de esta nueva altura y no de la de total descanso. ¿Qué significa ese desplazamiento? Hasta ese instante el sujeto tenía los órganos articuladores en reposo; en ese instante los pone en disposición de hablar.

la misma variable de tracto, la ralentización en la ejecución de los gestos de constricción en las proximidades de dicha frontera conlleva, además de un aumento en la duración gestual, también una mayor magnitud. Un ejemplo de esto sería el comportamiento laríngeo descrito en el § 3.3.1: una consonante especificada para la abducción glotal verá aumentado tal gesto (con todo lo que eso comporta, como el incremento del VOT) siempre y cuando vaya seguida de otro segmento especificado para la aducción, tal como una vocal. En cambio, ese mismo modelo de gestos π predice que, en inicio de constituyente, un gesto articulatorio alcanza una menor magnitud (y lo hace de manera más tardía) cuando no va seguido de otro gesto correspondiente a la misma variable de tracto. Este es necesariamente el caso de los gestos responsables de rasgos privativos como [nasal]. Es decir, la propia dinámica de tareas y la manera en que los gestos prosódicos afectan a los gestos articulatorios anticipa una menor magnitud de la nasalización en las proximidades de una frontera prosódica.

En definitiva, los estudios sobre el refuerzo del gesto velar han señalado la mayor elevación del velo en inicio de constituyente prosódico, lo que tiene por efecto una menor nasalización, pero no se ocupaban de explicar las causas de tal comportamiento. Por otro lado, el estudio de Byrd y Saltzman (2003) predice, mediante un modelado teórico, que la posición inicial induce una disminución en la magnitud de un gesto si este no va seguido de otro que afecte a su misma variable de tracto. Lo que he hecho en este apartado ha sido relacionar ambas informaciones para aportar una explicación plausible al —de otro modo— paradójico caso de las nasales.

Por otro lado, la causa de una mayor elevación del velo no solo en las consonantes nasales —algo que se acaba de justificar— sino también en las orales, como se ha descrito en la revisión bibliográfica previa, no se desprende de una manera tan clara a partir del modelo, teniendo en cuenta que no hay ninguna especificación gestual para la oralidad. Llegados a este punto, debemos retomar la idea de Chomsky y Halle (1976 [1968]) de que se produce una elevación del velo en el momento de la preparación de una locución, y la combinaremos con el fenómeno de la declinación del velo reseñada más arriba. A la luz de estos hechos, parece sensato proponer que, aun siendo la posición por defecto, el mantenimiento del velo en alto exige una tensión muscular que se va mermando poco a poco, y debe recibir un renovado impulso periódicamente. Por tanto, es plausible que el movimiento de elevación que se distingue en los momentos previos al habla se repita también en los inicios de constituyentes prosódicos (al menos, en los superiores). Eso sí, seguramente no recibe el tratamiento propio de un gesto articulatorio (no aparece en la partitura de gestos), sino que se trataría más bien de un ajuste articulatorio, en el sentido habitual de la base de articulación.

3.4.2. Gesto velar: refuerzo por prominencia

Pasemos ahora a revisar las variaciones en el movimiento del velo por efecto de la prominencia prosódica. Krakow (1993) señala que las consonantes orales en sílabas del tipo CVN adoptan una posición del velo más elevada en la sílaba que recibe el acento léxico. Sin embargo, se produce una interacción con el efecto de declinación que se acaba de explicar (§ 3.4.1), ya que dicha declinación se presenta de manera más acusada

si se comparan las sucesivas sílabas acentuadas de un enunciado que si se comparan las átonas entre sí. Por consiguiente, el efecto principal del acento se atenúa hacia el final del enunciado, donde se reduce la diferencia en la posición del velo entre sílabas acentuadas y no acentuadas.

Del mismo modo, el estudio de Vaissière (1988), también sobre el inglés, detecta que las consonantes orales presentan una posición del velo más elevada en las sílabas acentuadas que en las no acentuadas, mientras que el acento léxico sí se refleja en una posición más descendida del velo cuando se trata de consonantes nasales. En general, además, y al igual que pasaba por efecto de la posición, las sílabas acentuadas del tipo CVN se prestan a una mayor resistencia de la consonante oral a la coarticulación anticipatoria del gesto velar, lo que redundará en una mayor rapidez del movimiento de descenso del velo una vez que este se desencadena, máxime teniendo en cuenta que la posición que debe alcanzarse es más extrema (más baja). En cambio, tal y como señala Krakow (1993), en esas secuencias CVN, la vocal sí que se asimila más fácilmente al rasgo nasal de la consonante siguiente cuando aparece en sílaba acentuada, y lo atribuye no solo a la mayor magnitud de la abertura v-f de la consonante nasal, sino también a la mayor duración de la misma. T. Cho y Keating (2009) también señalan, para el inglés, la mayor duración nasal como una característica reforzada por la prominencia prosódica (aunque no por la posición inicial de constituyente).

Gordon (1996) mide un mayor flujo nasal, una mayor duración acústica y una mayor amplitud (RMS) cuando la [n] aparece en sílaba léxicamente acentuada en estonio. Sin embargo, si esa sílaba aparece, además, en posición inicial de algún constituyente prosódico superior a la propia sílaba, se anula el efecto de la prominencia y vence el efecto posicional, con el consiguiente descenso del U_n , acorde con las explicaciones que se han hecho antes al tratar dicho efecto.

En definitiva, tal y como se vio en el § 3.2, aunque a veces la variación segmental inducida por la prominencia puede ir en la misma dirección que aquella causada por la posición prosódica, lo normal es esperar efectos parciales o totalmente distintos, dada la diferencia en los mecanismos subyacentes. El caso de los movimientos del velo del paladar es una buena ilustración de esta diferencia. En el § 3.4.1 se ha argumentado que el efecto de la posición inicial de constituyente es la mayor elevación del velo, con independencia de que se trate de consonantes orales o nasales. En cambio, en este apartado hemos visto que la prominencia exagera la posición del velo propia de cada tipo de segmento, de modo que está más alto de lo normal en consonantes orales y más bajo de lo normal en consonantes nasales. De este modo, el carácter acentuado de una sílaba aumenta la distancia entre segmentos que se oponen paradigmáticamente. El aumento de la nasalidad de los segmentos nasales en las sílabas prominentes se sigue del característico ligero desfase entre gestos (*vid.* § 3.2), que así tienen más tiempo de desarrollo y pueden alcanzar una mayor magnitud en la dirección de su objetivo articulatorio.

3.5. Movimientos de la mandíbula

3.5.1. Movimientos de la mandíbula: refuerzo por posición

Descontando los gestos glotales y el gesto velar, el resto de los gestos articulatorios considerados por la fonología articulatoria (Browman y Goldstein, 1986, 1989, 1992) podrían denominarse, en conjunto, gestos orales, e incluyen los gestos labiales, los del cuerpo de la lengua y los del ápice de la lengua (estos dos últimos, a su vez, se pueden englobar bajo la etiqueta de gestos linguales). La característica común de todos los gestos orales es que, en su ejecución, intervienen los movimientos de más de un órgano articulatorio y, en concreto, siempre implican la participación de la mandíbula. Así, en los gestos labiales, se activan la mandíbula y los labios superior e inferior; en el gesto del cuerpo de la lengua, la mandíbula y dicho cuerpo; y en el gesto del ápice, la mandíbula, el cuerpo lingual y el propio ápice. El modelo de la fonología articulatoria no considera que haya un gesto de la mandíbula, propiamente, pero sí reconoce un papel fundamental de los movimientos de este órgano en la ejecución de los gestos orales.

La definición de los gestos articulatorios en términos de constricciones que hay que lograr en determinados puntos del tracto vocal conlleva que, en los gestos orales, la implicación de más de un órgano articulatorio confiere a dichos gestos varios grados de libertad. Es decir, que hay distintas maneras de combinar las acciones de los órganos para conseguir un mismo resultado. Como veremos en los próximos apartados, esto permite una variación en las estrategias empleadas entre distintos individuos y lenguas.

Por otro lado, los movimientos de los órganos participantes pueden llegar a ser contradictorios entre sí, lo que da lugar a efectos de compensación. Por ejemplo, lo habitual es pensar en una /a/ como una vocal que exige un notable grado de abertura de la boca. Sin embargo, es posible articular una [a] perfectamente aceptable, con su cualidad acústica esperada, sujetando un bolígrafo entre los dientes (postura que, lógicamente, no permite abrir demasiado la boca). En condiciones normales, la bajada de la mandíbula —o más bien la retracción con la que va aparejada (Vatikiotis-Bateson y Ostry, 1995)— provoca la constricción faríngea responsable de un alto valor de F1. En la situación descrita con el bolígrafo, la mandíbula no puede bajar, pero la raíz de la lengua sí puede bajar y retrasarse más de lo normal, para conseguir así el mismo efecto de constricción en la faringe⁹².

Pues bien, en este apartado y en los siguientes veremos que los efectos del refuerzo sobre los movimientos de la mandíbula a veces pueden ir en contra del objetivo articulatorio esperado para un determinado gesto en esas condiciones de refuerzo. Por ello es importante describir el comportamiento mandibular antes de pasar a la explicación de los demás gestos, en los que será necesario dar cuenta de las interacciones producidas por dicho órgano.

La mandíbula juega un papel importante en la alternancia de vocales y consonantes, con especificaciones de posición descendida para las primeras y elevada para las segundas. De este modo, hay un ciclo de apertura y cierre para cada sílaba (Nitttrouer,

⁹² Para más información sobre algunos efectos de compensación motivados por la posición de la mandíbula, *vid.* Stone y Vatikiotis-Bateson (1995).

1991; Nitttrouer, Munhall, Kelso, Tuller y Harris, 1988), lo que se ha relacionado con las alternancias de sonoridad en el seno de la sílaba (§ 2.1.1.3).

Como era de esperar, el gesto de elevación para las consonantes sufre algunas modificaciones por efecto de la posición prosódica. En inglés, el debilitamiento propio de final de sílaba hace que la posición mandibular sea más baja que en ataque (Macchi, 1988). En cambio, el refuerzo generalizado en frontera prosódica se manifiesta como un alargamiento temporal que afecta sobre todo al último gesto del constituyente (p. ej. en una secuencia VC] afecta al gesto de cierre). Eso es lo que comprobaron Beckman *et al.* (1992) al comparar, también en inglés, el final de frase entonativa con el final de palabra, con mayores efectos para el primero. Además, como producto de la ralentización inducida por la frontera, la velocidad de dicho cierre es menor. Por ser posición final de constituyente, no se observan efectos espaciales.

En cuanto a la magnitud, Vayra y Fowler (1992) proponen de modo provisional que esta experimenta una declinación en todos los gestos orales a lo largo de un enunciado, del mismo modo que son conocidas la declinación de f_0 (que correspondería no exactamente al gesto de constricción glotal, pero sí a los gestos tonales de la laringe) y la declinación del gesto velar (Krakow, 1993). En su experimento sobre diversas variedades dialectales del italiano, Vayra y Fowler informan de una declinación en el movimiento de descenso de la mandíbula, que se refleja en los valores de F1 de las vocales, especialmente de /a/⁹³ (también Stone, 1981, encuentra una posición más extrema —más elevada— para las consonantes de inicio de enunciado). En cualquier caso, se hace notar una diferencia con respecto a la declinación de f_0 , y es que la declinación mandibular afecta solo a las sílabas con acento primario, mientras que las sílabas no acentuadas presentan un patrón en forma de V (con el mayor debilitamiento en posición medial). Se da, por tanto, una interacción entre el efecto de la posición y el de la prominencia.

3.5.2. Movimientos de la mandíbula: refuerzo por prominencia

Stone y Hamlet (1982) descubrieron que las /d/ inglesas que sufrían aleteo (*flapping*), para lo que deben estar en inicio⁹⁴ de sílaba sin acento (ni siquiera secundario), no solo tenían menor duración sino que presentaban una menor magnitud del gesto de elevación de la mandíbula que las /d/ largas, propias de sílaba acentuada. Por otro lado, Stone (1981) detecta una mayor velocidad de los movimientos mandibulares en las sílabas con acento léxico, donde también se encuentran efectos espaciales, con una posición más elevada de lo normal para las consonantes y más descendida para las vocales. Además, hay una interacción con el efecto de declinación a lo largo del enunciado (*vid.* § 3.5): por ejemplo, una vocal no acentuada en sílaba inicial absoluta será más cerrada de lo habitual, porque la primera consonante es más cerrada (al ser inicial de enunciado) y la velocidad de apertura es menor por ser una sílaba no acentuada.

⁹³ También midieron F2 para estudiar la centralización en el eje horizontal de /i/ y /u/, que correspondería a movimientos de la lengua en lugar de la mandíbula, pero los resultados no fueron tan concluyentes.

⁹⁴ En realidad, parece ser su carácter ambisilábico lo que desencadena el cambio fonético (*vid.* § 2.1.1.4), pero en cualquier caso es cierta su asociación con el ataque de una sílaba no acentuada.

Sin embargo, el estudio de la prominencia sobre el movimiento mandibular se ha centrado en el acento de frase (también en inglés), al menos en los trabajos consultados (Beckman *et al.*, 1992; De Jong, 1995; Harrington, Fletcher y Beckman, 2000; Harrington, Fletcher y Roberts, 1995). Todos ellos apuntan a una mayor duración y velocidad del movimiento de apertura, que permite alcanzar una mayor magnitud durante la vocal (lo que aumenta la energía acústica radiada al exterior) y una mayor duración del movimiento de cierre, al tener que producir el retorno desde una posición más extrema. La modificación espacial alcanza posiciones más altas en el caso de las consonantes y más bajas en las vocales (De Jong, 1995).

Esto está en consonancia con la subida del valor de F1 vocálico que se ha señalado para los casos de sílabas prominentes (*vid.* p. ej. T. Cho y Keating, 2009) y encaja en los principios generales del refuerzo por prominencia (*vid.* § 3.2). Además, como ya se ha adelantado, la modificación del comportamiento de la mandíbula provoca alteraciones en el movimiento de otros articuladores para conseguir los objetivos de constricción propios de cada gesto oral, como se verá en los próximos apartados.

3.6. Gestos linguales

3.6.1. Gestos linguales: refuerzo por posición

Desde un punto de vista estricto del modelo de la fonología articulatoria existe una especificación independiente para el gesto del cuerpo de la lengua (TB, *tongue body*) y otra para el del ápice (TT, *tongue tip*). Sin embargo, en ambos tipos de gesto interviene efectivamente, como órgano, el cuerpo lingual, ya que no se puede hablar del movimiento del ápice sin referencia al mismo. A efectos de la exposición que sigue, trataremos todos los gestos de la lengua de manera conjunta en un mismo apartado.

En su revisión del estado de la cuestión de 1999 (b), Fougeron cita a Rousselot (1901) y a Straka (1963) como algunos de los estudios más antiguos que confirman el incremento del área de contacto linguopalatal en posición inicial de palabra y, en general, en ataque silábico frente a la coda. Además —continúa citando—, la presión ejercida por el ápice lingual sobre la zona dentoalveolar en consonantes como /n/, /t/, /d/, /l/ es mayor si ocupan el ataque silábico que la coda, y el pico de presión se alcanza más hacia el principio de la consonante en ataque, y más hacia el final en coda (según McGlone, Profitt y Christiansen, 1967; McGlone y Profitt, 1967).

Muchas de las investigaciones centradas en los gestos linguales se han basado en la medición del contacto linguopalatal mediante electropalatografía (EPG) en consonantes corales como la [n] o la [t], incluyendo la aspirada [t^h] y la fortis [t*] en los estudios de coreano. Así, se ha señalado un mayor porcentaje de electrodos activados, lo que se interpreta como un aumento en la magnitud gestual, y una mayor duración del periodo de contacto. Tal refuerzo espaciotemporal se ha constatado en hablantes de inglés (T. Cho y Keating, 2009; Fougeron y Keating, 1997); de chino minnán (taiwanés) (Keating *et al.*, 2003); de francés, aunque en este caso la [n] presenta más bien refuerzo espacial y la [t] refuerzo temporal, mientras que la nasal parece disminuir su duración (Keating *et al.*, 2003); y de coreano (Keating *et al.*, 2003). T. Cho y Keating (2001),

también sobre el coreano, encuentran efectos similares. Además notan que, en general, se da la siguiente progresión en cuanto al pico de magnitud de contacto: $t^* > t^h > t > n$. Las dos consonantes más fuertes de la serie son menos propensas a la variación por efecto de la posición prosódica, de modo que en niveles jerárquicos altos las cuatro tienden a igualarse. Sin embargo, no encuentran interacción entre el nivel prosódico y la identidad segmental, de modo que no se puede decir que el efecto prosódico afecte más a unas consonantes que a otras, sino que se trata tan solo de una tendencia.

Diversos estudios apuntan que, en posiciones de refuerzo, el incremento de la magnitud del gesto hace que aumente el área de contacto linguopalatal, que puede extenderse hacia posiciones más adelantadas, cubriendo filas del EPG que, en otras condiciones, no se activarían. Esto puede provocar incluso un cambio en la denominación del punto de articulación, tal que las alveolares se hacen dentoalveolares en coreano (Keating *et al.*, 2003). En neerlandés sucede algo similar, según deduce T. Cho (2003) basándose en el hecho de que el centro de gravedad de la explosión distensiva se sitúa más alto cuando la [t] o la [d] aparecen en frontera de frase entonativa, lo que refleja una reducción de la longitud de la cavidad de resonancia anterior, provocada por el adelanto del punto de articulación. En francés, la [k], cuya superficie de contacto queda registrada tan solo parcialmente por el EPG, adelanta su punto de articulación y activa un mayor porcentaje de electrodos, siempre en las filas traseras, claro está (Fougeron, 1999a).

Esta misma autora, en su estudio sobre las consonantes francesas, descubre un efecto interesante sobre la articulación de la lateral [l]. En los niveles prosódicos más bajos suele haber cierta asimetría, ya que la lengua suele dejar abierto solo uno de los canales laterales, mientras que se apoya en el otro costado. En cambio, en los niveles altos hay una mayor bilateralidad, con mayor contacto en el lado que solía estar abierto y menor contacto en el lado que antes estaba cerrado.

Por otro lado, T. Cho y Keating (2009) informan de una menor amplitud de la explosión distensiva de las oclusivas, y la atribuyen a la mayor lentitud en los movimientos de apertura. Esta variación en la velocidad responde al propio refuerzo temporal, aunque también destaca el hecho de que las consonantes con mayor superficie de contacto suelen presentar una distensión más lenta⁹⁵.

Keating *et al.* (1999) extienden el uso de la técnica de EPG al estudio de todas las obstruyentes con actividad lingual del inglés: [t], [d], [tʃ], [dʒ], [s], [z], [ʃ], [ʒ]⁹⁶. Todas estas consonantes presentan una mayor duración en inicio de palabra que en posición final. Sin embargo, en cuanto al refuerzo espacial, el comportamiento varía según el tipo de obstruyente. Las oclusivas simples también ven aumentada su magnitud en posición inicial de palabra; en cambio, las fricativas no muestran efectos espaciales, lo cual

⁹⁵ De hecho, Martínez Celdrán y Fernández Planas (2013) justifican así la tendencia de las oclusivas palatales a presentar una distensión africada, dada la mayor área de contacto en las consonantes de este punto de articulación.

⁹⁶ Las labiales y labiodentales quedan excluidas por no poder observarse su comportamiento mediante este procedimiento. De las velares, el artículo incluye también la [k] (no la [g]), pero reconoce las limitaciones de la EPG para estudiar articulaciones tan posteriores, a pesar de que la vocal que se usó en el contexto en ese caso fue la /i/, para propiciar el adelantamiento lingual por coarticulación.

puede explicarse por la escasa libertad posicional exigida por este tipo de sonidos para conseguir el grado de estrechamiento crítico (Recasens, 1999, también defiende que los segmentos sometidos a un mayor control articulatorio son menos propensos a la coarticulación, algo que podría explicar el menor grado de coarticulabilidad de las fricativas, especialmente de las sibilantes, aunque esto no quiere decir —claro está— que no puedan coarticularse en absoluto). En cuanto a las oclusivas africadas, su fase fricativa nos puede hacer pensar que también deberían ceder poco al refuerzo posicional. En efecto, la [dʒ] no presenta ningún aumento de magnitud. Por el contrario, inesperadamente, la [tʃ] muestra más contacto en posición final de palabra.

El comportamiento de las africadas quizá deba explicarse teniendo en cuenta la distinción de los dos momentos: la fase tensiva, de cierre, y la distensiva, de fricción. Degenshein (2004) documenta que la frontera de frase entonativa provoca un aumento en la duración de las africadas [tθ] y [dð] del lúo, una lengua nilo-sahariana de Kenia. Sin embargo —advierte—, existe una variación en las estrategias empleadas por cada individuo, ya que unos refuerzan la duración de la fase oclusiva y otros, de la fricativa. Aunque Degenshein habla de duración (no de magnitud), si relacionamos esta idea con el caso anterior, de la [tʃ] inglesa, tiene sentido pensar que un alargamiento temporal de la fase de cierre se pueda haber interpretado como un incremento de la magnitud del contacto sobre la duración global del segmento africado.

En cambio, cuando la fricción y la oclusión forman parte de dos segmentos diferentes, como sucede en las secuencias *s* + *C* de inicio de palabra en inglés, la adyacencia a una frontera de frase entonativa provoca un aumento en la duración y en la magnitud de ambos segmentos (Y. Lee, 2011). El mayor estrechamiento de la constricción durante la fricativa se refleja en una pérdida de energía (RMS). Por otro lado, aunque también se supone un refuerzo del gesto de abducción glotal en esa posición, el incremento de la duración de los gestos linguales correspondientes a *dos* segmentos hace que, en global, se reduzca el VOT de la oclusiva, algo que Lee interpreta como una exacerbación de la regla de no aspiración de oclusiva sorda tras /s/.

Recientemente, diversos trabajos basados en la técnica de la articulografía electromagnética (EMA) han precisado que, entre la fase tensiva y la distensiva de un determinado segmento, se ve más afectada por el refuerzo posicional aquella fase estrictamente adyacente a la frontera prosódica: *vid.* estudios sobre el tamil (Byrd *et al.*, 2000) y sobre el inglés (Byrd, Krivokapić y Lee, 2006; Byrd *et al.*, 2005). Es decir, que el refuerzo no afecta de manera homogénea a todo el desarrollo espaciotemporal de los gestos correspondientes a un segmento. En una secuencia C₁][C₂, la primera consonante experimenta más refuerzo en su fase distensiva y la segunda, en su fase tensiva. La fase distensiva de C₂ a veces también puede presentar cierto nivel de refuerzo, aunque sus efectos son menores y no se presentan en todos los hablantes (Byrd *et al.*, 2006).

El refuerzo temporal, en estos estudios, no solo se refleja en el incremento de la duración total del segmento, o de la fase (tensiva o distensiva) que resulta colindante con la frontera. También supone un menor grado de solapamiento entre los gestos de los segmentos contiguos y, en general, una mayor variabilidad en la coordinación intergestual, como muestra el aumento de la desviación estándar del intervalo entre los

picos de magnitud de C_1 y C_2 (Byrd *et al.*, 2000). Ambos picos se encuentran más atrasados de lo normal cuando se hallan bajo el efecto del refuerzo posicional (Byrd *et al.*, 2006), pero el de C_2 se retrasa más todavía, provocando el aumento del intervalo y el menor solapamiento. El retraso del pico de magnitud tiene sentido a la luz de la deceleración relativa en la ejecución de los gestos próximos a la frontera, que también deja una huella en la forma de un mayor intervalo de aceleración, *i. e.* se tarda más en alcanzar el pico de velocidad (Byrd *et al.*, 2000, 2005).

3.6.2. Gestos linguales: refuerzo por prominencia

Como se justificó en el § 3.2, en ocasiones pueden darse interacciones entre los efectos de la prominencia y los de la posición prosódica. A veces pueden respaldarse mutuamente, como ocurre con el porcentaje de electrodos (EPG) activados en el momento inmediatamente previo a la distensión de [t] y [n] en inglés (equivalente a una ventana de 15 ms): T. Cho y Keating (2009) encuentran un refuerzo solo cuando el acento primario coincide con el inicio de enunciado. Otras veces, en cambio, se halla una especie de efecto techo, de modo que ambos refuerzos no pueden acumularse, como cuando Y. Lee (2011) detecta una mayor duración en las consonantes del ataque tanto por posición como por prominencia, pero tal que el efecto posicional es proporcionalmente mayor cuando se trata de una sílaba átona.

En el § 3.2 también se adelantó que en sílaba acentuada las vocales se ejecutan con unos movimientos linguales más extremos, que las sitúan en posiciones más periféricas del espacio articulatorio (De Jong, 1995). A la inversa, en las sílabas no acentuadas, en las que el gesto de cierre se produce antes, puede surgir un conflicto entre objetivos articulatorios. Dicho de otro modo, puede que a la lengua no le dé tiempo de alcanzar su objetivo correspondiente a la abertura vocálica, y eso puede ser motivo de una reducción vocálica (centralización del timbre).

De Jong (1995) argumenta que el efecto de la prominencia sobre los distintos gestos consiste en exagerar los objetivos articulatorios y que esta hiperarticulación aumenta la distancia paradigmática entre los distintos elementos que se oponen en el sistema fonológico de la lengua. Ahora bien, en los gestos articulatorios de la lengua interviene también la mandíbula. Los movimientos de este órgano experimentan sus propios efectos de refuerzo (*vid.* § 3.5.2), y en ocasiones parece haber una contradicción entre ambos movimientos (mandibular y lingual).

En una vocal baja como la /a/, el descenso mandibular contribuye a esa diferenciación paradigmática, y la lengua también baja más para conseguir un F1 más alto. En cambio, en las vocales altas (/i/, /u/), la mandíbula baja pero la lengua no. Además, el requisito de extremar las características diferenciales pone en juego a la lengua en un plano ortogonal al movimiento mandibular. Es decir, que son los movimientos horizontales de la lengua los que pueden hacer que una vocal sea más anterior o más posterior de lo habitual, y eso modifica los valores de los otros formantes (p. ej. F2 y F3). En cualquier caso, la participación conjunta de varios órganos articulatorios en un mismo gesto de constricción implica necesariamente la posibilidad de que existan efectos de compensación, lo que se constata a la vista de las diferentes

estrategias empleadas por cada individuo (De Jong, 1995; Harrington *et al.*, 2000). El mayor ascenso y adelantamiento de la lengua para las vocales anteriores también ha sido probado mediante electromiografía (EMG) —*vid.* Stone (1997, pp. 31-32)—, que permite detectar un aumento en el nivel de actividad muscular del *genioglossus* (Tuller, Harris y Kelso, 1982).

Este efecto de refuerzo sobre la acción lingual también se encuentra, por ejemplo, para las alveolares, con una posición más adelantada, incluso una vez sustraído el efecto de adelantamiento provocado por el ascenso de la mandíbula (De Jong, 1995). Dicha modificación se traduce en un centro de gravedad con una frecuencia más elevada, por ejemplo en la [t] del inglés (T. Cho y Keating, 2009) o en la [s] en grupos *s + C* (Y. Lee, 2011) —en este caso, siempre y cuando la sílaba tónica sea además inicial de frase—, aunque estos resultados contrastan con el de T. Cho (2003) para el neerlandés, donde el centro de gravedad solo subía por efecto de la posición, no de la prominencia. Además, Cho y Keating miden mediante EPG una mayor duración de la oclusión linguopalatal en estos casos.

En cambio, los segmentos fricativos en sílaba prominente (con foco) aumentan de duración y su potencia de ruido, pero no cambian su magnitud ni el punto de articulación, calculado en función del pico espectral (Silbert y De Jong, 2008). Recuérdesse que las fricativas son menos propensas a la modificación de la magnitud, dada la exigencia de obtener un estrechamiento crítico que cumpla las condiciones aerodinámicas necesarias para generar la fricción.

El acento enfático, como cualquier prominencia, también se traduce en una menor coarticulación CV en el caso de las consonantes linguales, por la necesidad de extremar las posiciones de la lengua entre ambos tipos de segmento, como consecuencia del requisito de diferenciación paradigmática ligado a las sílabas prominentes (Lindblom, Agwuele, Sussman y Cortes, 2007). En cambio, el grado de coarticulación puede ser mayor cuando entra en juego una consonante labial.

Esta coarticulación se refiere al solapamiento entre gestos consecutivos. Por otro lado, dada la intervención de más de un articulador en los gestos orales, también cabe estudiar el solapamiento entre los movimientos de los distintos órganos participantes de un mismo gesto (p. ej. la mandíbula, el cuerpo de la lengua, y el ápice). Nittrouer (1991) descubre que, cuando se reduce la duración de un ciclo mandibular (de V a V), se adelanta la fase relativa del movimiento del ápice lingual. Dicho de otro modo, para ciclos mandibulares más cortos, hay una mayor proporción de la duración ocupada por movimientos simultáneos del otro articulador. Este acortamiento del ciclo mandibular puede ocurrir por diversos motivos, como la propia identidad vocálica o el contexto segmental, pero también, claramente, cuando se trata de una sílaba no acentuada.

3.7. Gestos labiales

3.7.1. Gestos labiales: refuerzo por posición

Al igual que sucede con los gestos linguales, los movimientos de los labios ejercen efectos de compensación con respecto al movimiento de la mandíbula, siempre que sea

necesario para la consecución de la constricción labial especificada para un determinado tipo de segmento (De Jong, 1995; Macchi, 1988). También, como es habitual, se ha observado que distintos hablantes pueden emplear distintas estrategias (Hinton, 1996).

En posición de coda silábica, se produce un aumento de la duración del gesto de cierre labial de una oclusiva como la [p] del inglés, o del gesto de estrechamiento crítico labiodental de la [f], mientras que la magnitud disminuye con respecto al ataque por efecto del debilitamiento propio de esta posición (Byrd *et al.*, 2005).

En el resto de fronteras prosódicas, se produce el refuerzo esperado, tanto en duración como en magnitud —que se puede reflejar en una mayor presión interlabial—, y de manera incremental según se asciende en la jerarquía de constituyentes, como muestra la comparación entre posición límite y medial de enunciado en inglés (Byrd *et al.*, 2005; Hinton, 1996), o entre los niveles de la palabra prosódica y la frase fonológica —menor y mayor— en tamil (Byrd *et al.*, 2000). Además, este último estudio, mediante la técnica de EMA, señala una mayor separación entre picos de magnitud y, por tanto, menor solapamiento gestual y menor coarticulación entre dos segmentos nasales dispuestos a través de la frontera —*m*][*n*, o *n*][*m*—, así como un mayor retraso en los picos de velocidad.

3.7.2. Gestos labiales: refuerzo por prominencia

Una serie de estudios de electromiografía (EMG) han detectado que, en las consonantes labiales, la prominencia de acento léxico se refleja en una mayor actividad muscular en el *orbicularis oris*, responsable del movimiento de los labios (Tuller *et al.*, 1982; *vid.* también —*apud* Fougeron, 1999b— K. S. Harris, Gay, Sholes y Liberman, 1968; Slis, 1971). Además, el artículo de Tuller *et al.* (1982) no encuentra un aumento de actividad similar en el pterigoideo medial, encargado del ascenso de la mandíbula, lo que podemos interpretar como un movimiento neto aún mayor por parte de los labios.

Una idea similar, pero para el acento de frase, aparece en De Jong (1995), quien documenta un mayor movimiento labial, tanto vertical como horizontal (de protrusión), incluso descontando el movimiento de la mandíbula, que sí sube en este caso.

Beckman y Edwards (1994) estudian, en inglés, los ciclos del movimiento labial atendiendo a distintos niveles de prominencia. Por un lado, comparan la sílaba no acentuada (con vocal reducida) con la sílaba con acento secundario o primario (con vocal plena); por otro, comparan la sílaba plena con y sin acento de frase. En general, encuentran un refuerzo acumulativo en la duración, magnitud y velocidad del movimiento de apertura labial. Las diferencias son más notorias y coherentes entre sílaba plena y reducida que entre tónica y átona. Para estas autoras, el acento secundario se realiza como una mayor tirantez (en inglés, *stiffness*) de los articuladores. El argumento es que, aunque en ambos niveles de prominencia aumenta la velocidad, aquí lo hace en mayor proporción, teniendo en cuenta el desplazamiento logrado. En cambio, el acento de frase se implementaría como un desfase del movimiento de cierre con respecto al de apertura. No obstante, debemos advertir aquí que es posible que esta estrategia no sea igual en otras lenguas donde no hay un cambio de cualidad vocálica tan acusado como en inglés, entre la sílaba no acentuada y la acentuada.

4. Fonología de la fortición y la lenición en español

Los efectos del refuerzo por posición y por prominencia, si bien introducen modificaciones perceptibles, suelen limitarse a cuestiones de detalle fonético subsegmental («for example, a stop remains a stop in all positions even if the degree of stricture increases», Fougeron, 1999b, p. 40). Es decir, no alteran la identidad del segmento —que transcribiríamos igual—, y solo se pueden investigar mediante técnicas instrumentales como las empleadas por los estudios que hemos revisado en el capítulo 3. Sin embargo, en ocasiones se producen cambios en la identidad segmental que pueden tomar carta de naturaleza e incorporarse a la fonología de la lengua como alternancias alofónicas a nivel sincrónico, o incluso consolidarse como cambios fonémicos en la forma subyacente, desde una perspectiva diacrónica. Lógicamente, estos cambios son particulares de cada sistema lingüístico.

En una lengua cabe distinguir (al menos) dos grandes tipos de variaciones alofónicas:

(38) Tipos de alófonos en función de sus causas

a) Alófonos posicionales (motivados por la posición prosódica que ocupan)

i. Caracterizados por un debilitamiento

En muchas variedades del español, el fonema /s/ (entre otros) se debucaliza (*i. e.* pierde la constricción oral) y se realiza como [h] en posición de coda silábica (p. ej. *asma* ['ahma]).

ii. Caracterizados por un refuerzo

En inglés, los fonemas /p/, /t/, /k/ se pronuncian con distensión aspirada, [p^h], [t^h], [k^h], cuando aparecen en posición inicial de pie acentual (p. ej. *canonical* [k^hə'nonɪkəl] ‘canónico’; *apologize* [ə'p^hɒlədʒaɪz] ‘pedir disculpas’).

b) Alófonos contextuales (motivados por el contexto segmental)

i. Caracterizados por una asimilación

En español, el fonema /n/ se asimila al punto de articulación de la consonante siguiente (p. ej. *ángel* ['aŋxɛl]; *infame* [ɪm'fame]). Las posiciones débiles son más proclives a la asimilación porque los gestos articulatorios se solapan más debido a su menor magnitud y duración pero, en última instancia, es el contexto el que condiciona la variante que resulta.

Dada la influencia de la estructura prosódica sobre la selección de los alófonos posicionales, estos se convierten a menudo en indicios de dicha estructura, y pueden servir, por ejemplo, como apoyo a los mecanismos de acceso al léxico (como se ha visto en el § 2.3.4). Además, otro sentido en el que los procesos de refuerzo posicional influyen sobre la fonología de una lengua es en el de que las posibilidades fonotácticas de la periferia de las palabras suelen ser distintas a las de las posiciones mediales (*vid.* § 2.1.3.2). En definitiva, como dice Keating (2006), «there is a tendency for prosodic

domains, such as words, in a given language to have a phonological and/or phonetic shape conditioned by the prosodic structure of the language» (pp. 6-7).

Normalmente, los cambios que se fonologizan están desencadenados por los niveles prosódicos léxicos, *i. e.* los que pueden tener una extensión igual o menor a la palabra morfosintáctica, a saber, σ , Σ , ω . Este es, desde luego, el caso de los cambios diacrónicos, que deben quedar reflejados en la nueva forma subyacente de la palabra. También suele ser habitual que las variaciones alofónicas prosódicamente condicionadas dependan de esos niveles; sin embargo, pueden existir alternancias motivadas por los niveles superiores (*vid.* p. ej. § 4.3.1).

Este capítulo estudia la influencia de la estructura prosódica sobre la fonología de los cambios segmentales que se producen en las palabras del español (aunque se comparará con procesos de otras lenguas cuando sea conveniente para enriquecer la explicación). Se incluyen tanto casos de alternancias alofónicas sincrónicas como ejemplos de cambios históricos en las representaciones léxicas. Como adelanté en la Introducción, usaré aquí los términos *lenición* y *fortición* para referirme a los efectos del debilitamiento o el refuerzo (respectivamente) que van más allá del mero detalle fonético y tienen un reflejo en un cambio segmental. Esto está en consonancia con el empleo de dichos términos en los estudios de corte más tradicional, que se centran, precisamente, en este tipo de cambios segmentales, muchas veces desde una perspectiva histórica, mientras que en los trabajos más modernos, experimentales, que atañen a pequeñas variaciones subsegmentales, se suelen usar, de un modo más genérico, las palabras *debilitamiento* y *refuerzo*⁹⁷.

Como se verá, las obras que tratan de la lenición y la fortición en español no le conceden ningún papel al nivel prosódico del Σ , lo que seguramente se puede relacionar con la escasa importancia del acento secundario en esta lengua. En cuanto a la prominencia de acento primario, Lavoie (2001) afirma que no suele tener tampoco una gran repercusión sobre los cambios segmentales que se producen en el seno de la palabra. Desde luego, puede tener sentido que el acento no modifique los segmentos en las formas subyacentes, ya que, según lo visto en el § 3.2, más bien parece que el efecto de la prominencia es afianzar los segmentos ya presentes en la forma léxica mediante una hiperarticulación local y un incremento de la distancia paradigmática. Por su parte, Ségéral y Scheer (2008) opinan que el acento puede introducir pequeños matices en los niveles relativos de fuerza, pero añaden que no les consta que exista ningún idioma en el que el acento constituya un criterio principal en la determinación de la fuerza de la posición sobre la que recae:

Stress and positional factors seem to entertain an implicational relationship: the former may or may not come to bear in addition of the latter, but systems where lenition is only defined by stress without the position of the segment playing any role do not appear to exist (p. 132).

⁹⁷ No obstante, como se dijo en el § 1, los conceptos asociados a estas palabras varían mucho de un modelo a otro (*vid.* nota 3 en la p. 29).

Por este motivo, la revisión que sigue se centrará fundamentalmente en las distintas posiciones silábicas y las posiciones dentro de la palabra. En general, la posición de coda es una posición débil, aunque puede no resultar tan débil si coincide con el final de la palabra, como se justificó en el § 3.1. La posición de ataque silábico es más fuerte que la de coda, aunque aún es una posición relativamente débil, por tratarse de un nivel prosódico tan bajo en la jerarquía. Por este motivo, es una posición que presenta comportamientos dispares. Su debilidad relativa la deja más expuesta a la influencia del contexto segmental. En efecto, las consonantes en posición de ataque pueden verse afectadas en un contexto intervocálico (o, de un modo más general, intersonántico), y se debe interpretar que esas modificaciones han sido desencadenadas por el carácter débil de la posición. En estos casos habrá que ver cómo interactúan la posición prosódica y el contexto segmental, y qué efectos específicos aporta cada uno. En otras ocasiones, en cambio, las consonantes de ataque experimentan procesos de fortición. La mayor incidencia de los cambios fortitivos se produce en los ataques que coinciden con el inicio de palabra.

Para unificar las condiciones en que suele producirse la fortición, Ségéral y Scheer (2008) se refieren a esas posiciones como la imagen especular de las codas silábicas (en inglés, *coda mirror*). En efecto, son fuertes la posición inicial de palabra y, según ellos, la que llaman posición de poscoda, es decir, un ataque interior de palabra, pero que no se sitúa en contexto intervocálico sino que va precedido por otra consonante (la coda de la sílaba anterior). En nuestra notación, eso corresponde a las condiciones (39):*b*.

(39) Posiciones de *coda* y de *coda mirror* según Ségéral y Scheer (2008)

a) Coda

coda final	___] _ω
coda interior	___] _σ [C

b) Coda mirror

inicial de palabra	[_ω ___
poscoda	C] [_σ ___

Sin embargo, conviene insistir en que, al separar las posiciones de ataque silábico (interior de palabra) en dos condiciones diferentes —la llamada posición intervocálica y la poscoda—, se están mezclando dos criterios: el de la posición prosódica y el del contexto segmental, y es el primero el que realmente está en la base de los procesos de refuerzo o debilitamiento.

A continuación se verán los distintos fenómenos de lenición y fortición asociados a sus posiciones prosódicas, y se intentará dar una explicación a algunos casos conflictivos que han dado lugar a una disparidad de interpretaciones y clasificaciones a lo largo de la historia.

4.1. Coda silábica

La posición de coda silábica es la posición débil por excelencia (*vid.* § 3.1). En ella se producen típicamente fenómenos de lenición, tales como la debucalización (que

Tabla 4.1. Porcentaje de conservación de [s] (vs. aspiración [h]) en coda real y en coda resilabeada. Datos de Terrell (1980).

	s.C	s.V
Buenos Aires	11 %	88 %
Cuba	2 %	18 %
Puerto Rico	4 %	22 %
Panamá	5 %	20 %

puede desembocar en elisión), el rotacismo, y el ensordecimiento (pasivo). Esporádicamente, se pueden encontrar aparentes casos de fortición, como veremos, pero son bastante aislados, y no exentos de dudas.

Por otro lado, la coda final de palabra puede ser relativamente menos débil que la coda interior, por el efecto de refuerzo en la frontera de un constituyente mayor. Además, bajo las circunstancias apropiadas, una coda final puede quedar asociada a un ataque silábico gracias a los procesos de resilabeo. En el § 2.1 se explicó que, incluso en contextos de resilabeo, se siguen encontrando pruebas de la afiliación de esa consonante a una posición de coda, pues experimenta las alternancias propias de esta posición (p. ej. la debucalización, como en la secuencia *los amigos* [lo.ha.'mi.ɣoh]). Aun así, es cierto que, dado el carácter de variación libre que entraña esta alternancia alofónica (*i. e.* la debucalización no es obligatoria, incluso en coda sin resilabear), sí se predice una menor frecuencia o probabilidad de aparición de la aspiración en casos de resilabeo. En efecto, Terrell (1980) aporta datos que confirman tal predicción (*vid.* Tabla 4.1)⁹⁸.

Sirva esto de momento como ejemplo del distinto comportamiento de las codas finales de palabra, asunto al que volveremos más tarde cuando veamos los distintos fenómenos que afectan a la posición de coda.

La coda como posición propicia a la neutralización

La coda es una posición especialmente proclive a la neutralización de oposiciones fonológicas. Indudablemente, los debilitamientos de los gestos articulatorios que en ella se pueden producir son una de las posibles causas de tales neutralizaciones, ya que las acciones requeridas sobre el tracto vocal pueden no llegar a diferenciarse lo suficiente de otras opciones. Además, la reducción (y el solapamiento) gestual hace que las consonantes en coda a menudo sufran cambios por asimilación a la consonante que encabeza la sílaba siguiente. Por otro lado, el hecho de ir seguidas por otra consonante en lugar de una vocal constituye una circunstancia que dista de ser óptima para la correcta percepción de todas las claves acústicas que permiten reconocer su identidad segmental (Wright, 2004). No en vano, la mayor modulación de la señal es la que se

⁹⁸ En la tabla se aprecia que, para cada contexto, Buenos Aires presenta un mayor porcentaje de conservación que las otras variedades analizadas. Esto puede entenderse teniendo en cuenta que en Argentina el proceso de debucalización (y, sobre todo, la posible elisión posterior) está mucho más estigmatizado que en el Caribe. Sin embargo, dentro de la propia variedad bonaerense se confirma que la retención de [s] es mucho más frecuente en los contextos que permiten el resilabeo que en aquellos que no lo permiten.

produce en el contraste entre una consonante y una vocal, y esa modulación consecutiva de la onda portadora es crucial para la perceptibilidad (*vid.* § 2.1.1.3).

En español, por ejemplo, algunas neutralizaciones en coda se pueden atribuir al efecto lenitivo (pérdida de la tensión articulatoria) condicionado por dicha posición prosódica. En este grupo se encontrarían las siguientes neutralizaciones:

(40) Neutralizaciones en español por efecto de la lenición en coda

/p/ ~ /β/ ⁹⁹	cápsula	[¹ kaβsula]	ábside	[¹ aβsiðe]
/t/ ~ /ð/	atmósfera	[að ¹ mosfera]	admirable	[aðmi ¹ raβle]
/k/ ~ /ɣ/	acción	[aɣ ¹ θi ¹ on]	signo	[¹ siɣno]
/r/ ~ /r̄/	carta	[¹ kaɾ̄ta]		

En un modelo fonológico basado en rasgos distintivos, estas parejas de fonemas se oponen por el rasgo [$\pm$ tenso]. En los tres primeros casos, además, suele darse una concomitancia del carácter sordo con el valor [+tenso] y del carácter sonoro con [–tenso], aunque el hecho de que ocurra así en español no significa que esa relación sea universal (Martínez Celdrán y Fernández Planas, 2013, p. 74)¹⁰⁰.

Dada esta concomitancia, se ha debatido mucho cuál de los rasgos —tensión o sonoridad— es el relevante desde el punto de vista fonológico en español. El autor que más ha defendido la primacía de la tensión ha sido Martínez Celdrán (1985, 1991, 1993). *Vid.* Gómez Asencio (2000 [1994], p. 162 y ss.) para un resumen de los razonamientos propuestos a favor de la tensión; y también Martínez Celdrán y Fernández Planas (2013), cap. 2, § 3, para un estado de la cuestión actualizado. A continuación se repasan algunos de los argumentos que se han esgrimido en defensa de la primacía de la tensión.

(41) Primacía de la tensión sobre la sonoridad en las obstruyentes del español

- a) Martínez Celdrán (1985) presenta unos experimentos perceptivos en los que los hablantes reconocían la diferencia entre «sordas» y «sonoras», a pesar de haber eliminado por completo la barra de sonoridad. No obstante, debe advertirse que este no es uno de sus artículos más concluyentes, pues el estudio está basado en la percepción de palabras y, aunque algunas de ellas forman pares mínimos entre sí, esto no ocurre siempre, de modo que el conocimiento léxico puede ser suficiente para identificar de manera correcta los estímulos.
- b) En cambio, la metodología utilizada por el mismo autor en sus estudios de 1991 y de 1993 permite alcanzar resultados más fiables. En el primero de esos trabajos emplea secuencias del tipo VCV, carentes de significado. En el segundo, se basa en estímulos sintéticos situados a lo largo de un continuo

⁹⁹ Para una justificación de la elección de la aproximante como forma más representativa del fonema, *vid.* § 4.3.1. Por otro lado, en las transcripciones fonéticas se han elegido las realizaciones más características, aunque existe más variedad.

¹⁰⁰ En efecto, hay idiomas en que esto no es así, p. ej. en algunas lenguas caucásicas, como el tabasarán (Kodzasov y Muravjeva, 1982), el archí (Kodzasov, 1977) o el agul (Kodzasov, 1990), *apud* Ladefoged y Maddieson (1996, p. 96).

fonético, y cuyas posibilidades de categorización constituyen un par mínimo (*cava*, *capa*). En ambos estudios se manipula la duración de la fase tensiva (correspondiente al silencio acústico) y se observa que los estímulos de mayor duración se perciben como sordos, mientras que los más breves se perciben como sonoros. Se concluye que la reducción temporal se interpreta como pérdida de tensión, lo que justifica esa manera de categorizar los estímulos. Otros autores, p. ej. García Santos (2002), han realizado estudios semejantes, con resultados equiparables.

- c) Se suele decir que «en el habla cuchicheada, donde desaparece o, al menos, queda neutralizada la vibración de las cuerdas vocales» (Gómez Asencio, 2000 [1994], p. 163), se sigue distinguiendo entre sordas y sonoras. Eso debe ser a causa del factor tensión¹⁰¹.
- d) En el habla real, existen realizaciones sonorizadas de las supuestas sordas, y realizaciones ensordecidas de las supuestas sonoras, y ello no impide su correcta identificación.
- e) Cuando un español escucha una pronunciación sonora pero tensa, como pueda ser la del alemán, o la del chino, tiende a interpretar una sorda, porque a lo que está dando primacía es a la tensión.

Como se ha dicho, en efecto, las tensas se caracterizan, entre otras cosas, por una mayor duración. Además, el incremento de la tensión muscular y de la fuerza espiratoria suele tener como consecuencia una mayor acumulación de presión intraoral (P_o), lo que se puede manifestar en una mayor intensidad de la explosión distensiva. Otra repercusión aerodinámica de la tensión es que, en las obstruyentes, a mayor duración, resulta más difícil mantener la vibración laríngea. Esto es así debido al progresivo aumento de la P_o , que hace más difícil conseguir la presión transglotal (P_t) necesaria para mantener el proceso vibratorio (Titze, 2000). Este condicionante hace que se termine asociando sonidos largos con sonidos sordos y sonidos breves con sonidos sonoros. Por este motivo, la percepción de un sonido de corta duración desata una cadena de inferencias por la cual este sonido se puede tomar como relajado y, por tanto, también como sonoro, aunque no presente el componente de baja frecuencia asociado a la vibración de los pliegues vocales (que se manifestaría gráficamente como barra de sonoridad en un espectrograma). En cualquier caso, Martínez Celdrán (1991, p. 139) advierte que la duración no debe tomarse como parámetro de la sonoridad, sino de la tensión, y tampoco estos son sinónimos, pues la tensión es la causa y la duración, la consecuencia.

En definitiva, los argumentos anteriores muestran que las parejas /p/ ~ /β/, /t/ ~ /ð/, /k/ ~ /γ/ en español se oponen por la tensión (siendo la sonoridad un rasgo

¹⁰¹ Según Malécot (1970, p. 1589), en el lenguaje susurrado, aunque todos los sonidos suenan sordos, hay una diferente configuración laríngea entre los que deberían ser sordos y los que deberían ser sonoros. Para los primeros, los pliegues vocales se separan; para los últimos, se juntan. En este sentido —argumenta Malécot—, puede seguir siendo el gesto glotal el responsable de la diferenciación. Sin embargo, debe tenerse en cuenta que no basta con que haya una diferencia articulatoria conocida, sino que esta tiene que reflejarse en alguna clave acústica perceptible.

concomitante). Esto explica que en una posición que favorece el debilitamiento o relajación, como es la coda silábica, las manifestaciones de ambos fonemas converjan en realizaciones laxas, con la consiguiente neutralización de las oposiciones.

En el caso de la pareja /r/ ~ /r/, también la distinción recae sobre el rasgo [$\pm$ tenso] (Alarcos Llorach, 1965). Nótese que, a diferencia de las anteriores parejas, en esta son fonológicamente sonoros los dos miembros de la oposición¹⁰². Por este motivo propongo que no se puede considerar completamente redundante el rasgo [$\pm$ sonoro], y debe incluirse junto con [$\pm$ tenso] en el sistema fonológico del español. Seguramente, la necesidad de ambos y su casi total pero no absoluta predictibilidad mutua haya sido abono de las disputas académicas acerca de la primacía de uno u otro rasgo en la caracterización del inventario de fonemas de esta lengua.

Ambos fonemas, /r/ y /r/, se oponen en español únicamente en contexto intervocálico. En posición inicial de palabra, posición fuerte por excelencia, las restricciones fonotácticas del español solo permiten la vibrante (como se ha explicado en el § 2.1.3.2, el alargamiento propio de esta posición puede ser contradictorio con la brevedad de la percusiva). También hay vibrante en ataque interior tras consonante heterosilábica. El segundo miembro de los ataques complejos o ramificados es, en cambio, la percusiva¹⁰³ (vid. § 2.1.1.6 para una justificación de las restricciones sobre los ataques complejos). En coda, la neutralización de la tensión también impide la oposición. La gran variedad de manifestaciones fonéticas que alternan en esta posición (vid. nota 103) hace difícil la elección del elemento subyacente. Por un lado, a la vista de formas flexivas de plural como [a'mores] (y no [a'mores]), hay que deducir que el segmento subyacente en *amor* es una /r/. Sin embargo, una de las posibles realizaciones de las róticas en coda, aunque enfática, es la vibrante. Por ejemplo, en *harto* ['ar̥to] o en *arca* ['arka], el movimiento vibratorio de la lengua durante la rótica es involuntario¹⁰⁴. No es el aleteo controlado propio de una percusiva, sino que simplemente es una consecuencia aerodinámica de la obstrucción interpuesta al paso del aire: la P_o acumulada vence la resistencia de la lengua y, al pasar rápidamente la corriente de aire, se produce una succión por el *efecto Bernoulli*. Es decir, se trata del mismo proceso que el responsable de cualquier otra vibrante en posición de ataque silábico. Ahora bien, resulta sorprendente encontrar una producción próxima a la pronunciación prototípica del miembro tenso de la oposición precisamente en la coda, que es la posición débil por excelencia. En efecto, se trata de

¹⁰² No obstante esta afirmación, debemos reconocer que en el habla real existen frecuentemente realizaciones ensordecidas de la vibrante /r/, especialmente en algunas variedades americanas. Esto se justifica fácilmente por las complejas condiciones aerodinámicas requeridas para su producción: la obstrucción a la salida del aire durante los periodos de cierre de la vibrante, por breves que estos sean, puede provocar un aumento de la P_o que dificulte el mantenimiento de la fonación. En la percusiva /r/ también puede llegar a ocurrir lo mismo. Aunque solo se produzca un contacto y este sea, por definición, ultrabreve, el cambio repentino de las condiciones aerodinámicas provoca descensos micromelódicos (Buenafuentes de la Mata, Madrigal Montoya, y Garrido Almiñana, 2000; Di Cristo, 1982), cuando no el ensordecimiento completo.

¹⁰³ Nótese que siempre me refiero, aquí, a la representación fonológica subyacente, aun admitiendo que las manifestaciones fonéticas en distintas posiciones y contextos son mucho más variadas. Para un estudio pormenorizado de estas posibilidades, vid. Blecua Falgueras (2001).

¹⁰⁴ «Trills [are] caused by aerodynamic forces, as opposed to taps and flaps, which involve active muscular movements of the tongue» (Solé, 2002, p. 656).

una posición débil por naturaleza, pero también es una posición propicia al alargamiento (los gestos articulatorios deben sucederse de manera secuencial y relativamente espaciados entre sí para asegurar su correcta percepción). Así, aun asumiendo una forma subyacente percusiva, /r/, para las róticas en coda, la mayor duración del contacto apicoalveolar provocado por el alargamiento propio de esta posición puede tener como consecuencia la acumulación de P₀ por detrás del punto de contacto, tal que el aire vence la resistencia y se pone en marcha el mecanismo de vibración. En otras palabras, las realizaciones vibrantes de coda no son el producto de un refuerzo (que difícilmente podría explicarse en esta posición) sino que son la consecuencia aerodinámica del alargamiento sí atribuible a la coda. En cualquier caso, como posición débil que es, también se documentan otras variantes alofónicas en la coda y, de hecho, suelen ser muy frecuentes las aproximantes, [ɹ].

Este debilitamiento puede justificar la neutralización de /r/ y /l/ de algunas variedades del español, pues la [l] es también una aproximante (lateral). Lo único que las diferencia es la presencia, en la [l], de las antirresonancias propias del modo de articulación lateral. Esta similitud puede dificultar la percepción de la identidad del segmento en cuestión. De hecho, un argumento a favor del origen perceptivo de la confusión es que la neutralización se resuelve a veces a favor de una y a veces de la otra realización. En Castilla-La Mancha y en Murcia, por ejemplo, es más frecuente una realización rótica en interior de palabra, pero la lateral en posición final; en Extremadura y en Andalucía triunfa más la opción rótica, en general (Alvar, 1996). En los países caribeños se ha optado preferentemente por la lateral, aunque también se encuentran opciones debucalizadas, con resultado [h], o incluso con elisión, lo que prueba el origen lenitivo del fenómeno. La neutralización también se da en determinadas zonas de Colombia, Venezuela, Ecuador, Argentina y Uruguay, sobre todo en las hablas populares (Real Academia Española, 2011, p. 257 y ss.). En efecto, debe tenerse en cuenta que esta neutralización suele estar bastante estigmatizada en todas las regiones afectadas, de modo que no es generalizable a la población culta o de posición socioeconómica alta.

Ahora bien, aunque todas las neutralizaciones vistas hasta ahora son, efectivamente, fruto de una lenición, también hay otras características de la coda que la hacen una posición poco favorecedora al mantenimiento de las oposiciones fonológicas. Al fin y al cabo, lo habitual es que la consonante de coda vaya seguida por otra consonante (la del ataque de la sílaba siguiente). En español, prácticamente solo es posible una coda seguida de vocal si se trata de una coda final de palabra y coincide que la palabra siguiente empiece por vocal, como se ha explicado al hablar del resilabeo¹⁰⁵. La secuencia de dos consonantes puede desencadenar cambios por asimilación y estos pueden tener como consecuencia la neutralización de alguna oposición.

Un caso muy habitual en las lenguas del mundo, y presente en español, es la asimilación del punto de articulación de una nasal a la consonante siguiente. Por

¹⁰⁵ Existe la posibilidad adicional de que una consonante en coda vaya seguida de vocal, dentro de una misma palabra, cuando existe una gran conciencia de que existe una frontera morfológica entre ambos segmentos, p. ej. en *subacuático* o en *subinspector*, donde la asociación de la /β/ con una u otra sílaba puede variar según los hablantes.

ejemplo, *envío* [em'bio], *infiel* [im'f̥iel], *pinza* ['piɲθa], *entero* [en't̥ero], *pensar* [pen'saɾ], *ancho* ['an't̥ʃo], *ancla* ['aɲkla]. Como consecuencia de este condicionamiento contextual, se pierde la posibilidad de mantener la oposición entre /m/ y /n/ (el fonema /ɲ/ está directamente excluido de la coda, según las restricciones fonotácticas del español, por ser palatal). En este caso, pues, se trata de una neutralización por asimilación, no por lenición.

Otras neutralizaciones se producen por otros motivos. Por ejemplo, en la neutralización que afecta a los rasgos de la acción glotal, es cierto que la pérdida de las oposiciones laríngeas puede deberse al ensordecimiento pasivo propio de esta posición, que sí es una lenición (*vid.* § 4.1.4). Pero además se puede justificar por las malas condiciones para la percepción de las claves acústicas que sustentan dichas oposiciones, como el VOT, cuyas condiciones óptimas de percepción implican que la consonante vaya seguida de una vocal, no de otra consonante.

En definitiva, aunque el efecto lenitivo de la coda silábica contribuye decisivamente a un buen número de neutralizaciones, en ocasiones la pérdida de una oposición fonológica en esta posición se debe a otros motivos, o incluso a una mezcla de factores que conviene diferenciar.

A continuación se analizarán los fenómenos de lenición más característicos de la posición de coda.

4.1.1. Debucalización (I)

La debucalización consiste en una pérdida de los rasgos supralaríngeos, de modo que el resultado es un elemento glotal, bien fricativo [h], bien oclusivo [ʔ].

En español, este proceso está muy generalizado a diversas variedades, aunque la consideración sociolingüística puede variar de una a otra. Las zonas donde se da la aspiración incluyen Extremadura, Castilla la Mancha, Murcia, Andalucía y Canarias (Alvar, 1996), así como numerosas regiones de América (islas caribeñas; costa oriental de México; países de Centroamérica; zonas costeras de Venezuela, Colombia, Ecuador y Perú; Chile; este de Bolivia; Paraguay; Uruguay; y Argentina excepto el noroeste; *vid.* Zamora Munné y Guitart, 1982, pp. 182-183).

La solución en todos estos casos es la fricativa [h] (la llamada aspiración). El caso más conocido es la aspiración de /s/ (p. ej. *autobús* [au̯to'βu^h]). El efecto de debilitamiento de la coda lleva a una hipoarticulación, manifestada como reducción de la magnitud de los gestos articulatorios orales. La lengua debería crear, en la zona alveolar, un estrechamiento crítico que diera lugar a las condiciones aerodinámicas necesarias para la generación de un flujo turbulento (fuente de fricción), *vid.* Catford (1977, p. 36 y ss.). Sin embargo, al no alcanzar el suficiente grado de tensión que asegure una posición precisa, la fricción no se produce en dicha zona alveolar sino que cambia de afiliación a la glotis, dado que las propiedades acústicas del tubo (*i. e.* del tracto vocal) determinan que la fricción solo puede darse en el punto más estrecho del mismo en cada momento¹⁰⁶.

¹⁰⁶ Solé (2010) comprueba que, efectivamente, el debilitamiento propio de la posición de coda hace que, durante la pronunciación de las fricativas, la P_o crezca con menor velocidad y alcance un valor máximo

Este tipo de debucalización afecta fundamentalmente a las obstruyentes, así como a la /r/, aunque Mondéjar (1979) afirma que «toda consonante del español, que vaya en posición implosiva, puede aspirarse en andaluz» (p. 391). Algunos ejemplos de debucalización son los siguientes: *chef* [ˈtʃe^h], *arroz* [aˈro^h], *aragonés* [araˈɣoˈne^h], *reloj* [reˈlo^h], *adaptar* [aða^hˈtaɾ], *atlántico* [a^hˈlantiko], *anécdota* [aˈne^hðota], *pierna* [ˈpie^hna].

Méndez Dosuna (1987) argumenta que la aspiración no empezó en posición final absoluta (ante pausa), sino que primero se debió de dar en las secuencias de contacto de sílabas, tanto en interior de palabra como por *sandhi* externo. Este contacto se refiere a grupos heterosilábicos de dos consonantes (C₁.C₂). Desde una postura que acepta la validez de las escalas de sonoridad —o su equivalente de polaridad inversa, las escalas de fuerza de las consonantes, EFC— (vid. § 2.1.1.3), Méndez Dosuna afirma que, en las lenguas que admiten este tipo de grupos, se prefiere que C₂ sea más fuerte que C₁ según la EFC. Esta preferencia se formaliza como una Ley de Contacto de Sílabas (LCS). Si se da el caso contrario, se puede terminar produciendo un cambio fonético que corrija la situación, que en este caso comporta la sustitución de C₁ por [h], ya que la [h] se encuentra al mismo nivel que las semiconsonantes en la EFC¹⁰⁷. En ese trabajo, defiende una extensión progresiva a lo largo de los siguientes contextos:

(42) Orden de difusión de la aspiración de /s/ según el contexto segmental siguiente (Méndez Dosuna, 1987)

- s + semiconsonante
- s + líquida
- s + nasal
- s + fricativa sonora
- s + fricativa sorda / oclusiva sonora

Una vez puesto en marcha, el proceso se puede propagar a otros contextos, aunque no contravengan la LCS. Así, se atestigua —en orden de difusión— la aspiración de /s/ ante oclusiva sorda, ante pausa, y ante vocal (a través de frontera de palabra).

Según Méndez Dosuna, dentro de las oclusivas también hay un orden de difusión condicionado por el grado de homorganicidad de la /s/ y la oclusiva siguiente. Cuando esta es una [p], dado que su articulación no implica a la lengua sino a los labios, se anticipa la producción de dicha [p], y es más fácil que se debilite (debucalice) la /s/. La lenición ocurre en menor medida cuando lo que sigue es una [k], ya que su articulación implica al dorso de la lengua, pero no compromete la zona coronal. En cambio, seguida de [t] es el contexto en el que menos probabilidades hay de que se produzca la

más reducido. En consecuencia, es necesario que pase más tiempo hasta que se produzca un ruido de fricción audible. Dicho ruido es, en todo caso, menos intenso que el de las fricativas en otras posiciones, e incluso puede que no llegue a producirse tal fricción en la cavidad oral.

¹⁰⁷ Toda su justificación del fenómeno se basa exclusivamente en dicho contacto, mientras que no concede toda la importancia necesaria al hecho de que C₁ se sitúe en una posición de coda silábica. Personalmente, creo que el desencadenante hay que buscarlo en la posición prosódica, aunque no cuestiono que el contexto pueda suponer un factor coadyuvante que matice los efectos de la posición.

debucalización, puesto que la preparación y transición al sonido siguiente favorece que se mantenga el estrechamiento adecuado durante la pronunciación de la [s].

La debucalización de una consonante en coda puede desencadenar algunos fenómenos de asimilación posteriores. El elemento debucalizado solo conserva los rasgos laríngeos y, para subsanar la carencia de otros rasgos, estos se pueden propagar desde la consonante siguiente. Por ejemplo, seguido de consonante oclusiva, puede dar una especie de resultado geminado (*los palos* [lo^{hp}ˈpalo^(h)], *los toros* [lo^{ht}ˈtoːro^(h)], *las casas* [la^{hki}ˈkasa^(h)])¹⁰⁸. Cuando lo que sigue es una espirante, el caso es algo más complejo, ya que comporta una propagación mutua de rasgos y, al extenderse el valor [–sonoro] de la aspirada [h] a la consonante siguiente, caracterizada como [–interrumpido], el resultado ha de ser necesariamente una fricativa, que es el único tipo de sonido que combina esos dos valores de rasgos, [–interrumpido, –sonoro]. El punto de articulación es el mismo que el de la espirante original: *los barcos* [lo^{fi}ˈfaɪko^(h)], *las dunas* [la^{θi}ˈθuna^(h)], *los gatos* [lo^{xi}ˈxaɾo^(h)].

En otras lenguas, el resultado de la debucalización puede ser a veces la oclusiva glotal [ʔ]. Como apunta Blevins (2004, pp. 120-121), en el chino hablado en Fuzhou se produce la debucalización de las oclusivas orales /p/, /t/, /k/ en final de palabra, que se pronuncian [ʔ], lo que conlleva la neutralización del punto de articulación de dichas oclusivas.

El mismo tipo de debucalización también se puede dar en inglés. A propósito de ello, Blevins (2004) se refiere al orden de difusión del cambio a los distintos puntos de articulación en los siguientes términos:

The fact that closure durations may differ for stops at different points of articulation means that debuccalization itself may occur first with the stops of shortest duration, and only later with those whose closure durations are longer. Coronal closure durations can be shorter than those for labial and velar stops, and tongue-tip movements show higher velocities than tongue-dorsum or lip movements (p. 121).

A continuación, veremos un fenómeno de lenición que tradicionalmente se ha descrito como un cambio de un punto de articulación a otro, pero que recibe un mejor tratamiento como instancia de debucalización. Se trata de la velarización de nasal en posición prepausal, como ocurre por ejemplo en algunas zonas de Andalucía; en las Antillas; en Centroamérica; en la costa de Venezuela, Colombia, Ecuador y Perú; y también en Paraguay y en Chile. Nótese que este aparente cambio de punto de articulación no afecta a las consonantes orales, sino solo a las nasales. Si se pierde, por debilitamiento, la articulación coronal que se esperaría en esta posición, la impresión de velarización puede proceder del mero hecho de que el velo se encuentra en posición descendida, más cercano al dorso de la lengua. No se trata, por tanto, de un movimiento activo del dorso para acercarse a la zona de articulación velar, sino más bien al contrario, de una aproximación del velo al dorso, justificada por la propia nasalidad del segmento.

¹⁰⁸ Para subsiguientes evoluciones relacionadas con la presencia de una postaspiración, *vid.* Ruch y Harrington (2014).

Blevins (2004, p. 115) señala que, por este motivo, la [ŋ] es un resultado habitual de la neutralización entre nasales, junto con una vocal (o paravocal) velar no redondeada nasalizada [ũ] cuando entre el velo y el dorso no llega a producirse un contacto pleno, sino tan solo una aproximación.

Ohala y Ohala (1993, pp. 234-235) señalan algunos motivos acústicos y perceptivos por los que una consonante nasal velar puede confundirse con una vocal o paravocal nasalizada. Cuando el punto de articulación está muy retrasado, la cavidad oral, que actúa como cámara de resonancia secundaria, es muy pequeña, y las antirresonancias que se generan son de frecuencias muy altas. Como quiera que las altas frecuencias son, en todos los sonidos sonoros, menos intensas que las bajas, es muy probable que los oyentes desprecien la antirresonancia y la impresión auditiva sea más bien la de una vocal nasalizada, pero no la de una consonante nasal.

Finalmente, hay otro fenómeno que tradicionalmente también se ha descrito como un cambio en el punto de articulación, pero que parece que se explica mejor como debucalización. Se trata de la evolución de la fricativa prepalatal del español medieval (/ʃ/, una vez que se produjo la neutralización con /ʒ/, *vid.* § 4.3.5) a la fricativa velar (sorda) /x/ de los Siglos de Oro, que permanece hasta nuestros días. García Santos (2002, p. 129 y ss.) opina que un desplazamiento tan grande en el punto de articulación resulta poco natural. Por otro lado, le llama la atención que, precisamente, el cambio se produzca con más facilidad en los contextos de vocal anterior que posterior. Por eso prefiere buscar una explicación alternativa al cambio de punto articulatorio, sobre todo teniendo en cuenta que el proceso es relativamente común y está volviendo a producirse en el español porteño. Ante estas consideraciones, este autor defiende que el primer paso del cambio consiste en una debucalización de la [ʃ] en [h], y el contexto de vocal anterior podría facilitar la asimilación de la consonante a la abertura de la vocal. Solo posteriormente se produciría el paso de [h] a [x]. Según él, además,

esto llevaría a una revisión de la concepción tradicional de los dialectos castellanos: los meridionales serían los conservadores, mientras que el norte habría innovado al pasar de [h] a [x] (*op. cit.*, p. 133, nota 115).

4.1.2. Elisión (I)

Con frecuencia, el resultado de una debucalización termina siendo elidido, lo que supone el último paso de la ruta lenitiva (recuérdese la definición de Hyman, 1975, *vid.* nota 2 en la p. 28). La falta de rasgos supralaríngeos del segmento en cuestión contribuye a su vulnerabilidad y este puede llegar a desaparecer por completo, a menos que adopte los correspondientes rasgos de la consonante siguiente por medio de alguno de los procesos de asimilación señalados en el apartado anterior. Por este motivo, las consonantes en posición final absoluta son las más propensas a la elisión (Alarcos Llorach, 1958; Poplack, 1980)¹⁰⁹.

¹⁰⁹ López Morales (1984) también aduce argumentos de tipo morfosintáctico. En el caso de la pérdida de /s/ en coda en español, si esa /s/ es el exponente del morfema de plural en la flexión nominal, o bien del de segunda persona del singular en la flexión verbal, entonces es más probable que su presencia resulte

Según la propuesta de Méndez Dosuna (1987), presentada en el § 4.1.1, la posición final no debería favorecer la debucalización, porque en ella no se produce ninguna situación de contacto silábico ni, por tanto, ninguna secuencia de sonoridad que haya que corregir. Si bien esto es cierto, la idea que se pretende defender en este apartado es que, si, a pesar de todo, llega a producirse la debucalización, es muy probable que no se mantenga la aspiración, sino que el proceso de lenición se complete hasta el borrado del segmento. Es decir, en posición final es frecuente la retención de /s/ y también su elisión; lo menos habitual es que el proceso se detenga en el estadio intermedio [h].

4.1.3. Rotacismo

En español hay un fenómeno, asociado con hablantes rurales e incultos, por el que una /s/ en coda puede convertirse en una rótica, p. ej. [ɾ], especialmente cuando va seguida de otra consonante coronal, p. ej. *las dos* [laɾ ˈðo(h)], *los zapatos* [loɾ θaˈpaɾo(h)]. Este fenómeno se explica como una reducción sustantiva (i. e. de la magnitud de los gestos) de la /s/, sin llegar a una debucalización: no se mantiene únicamente la presencia del nodo laríngeo, sino que también hay especificaciones supralaríngeas (estas son más fáciles de conservar gracias a la homorganicidad con la consonante siguiente).

4.1.4. Ensordecimiento pasivo (I)

La producción de los sonidos fonológicamente sordos comporta la abducción de los pliegues vocales, de modo que el aire puede pasar libremente a través de la glotis, sin provocar la vibración de los pliegues. Por el contrario, la vibración laríngea característica de los sonidos sonoros se consigue mediante la aducción aritenoides. Al quedar juntos los pliegues vocales, se crea un diferencial de presión entre la presión subglotal (P_s) y la presión supraglotal (P_o)¹¹⁰. La compresión del volumen del sistema subglotal en la espiración hace que aumente la presión por debajo de la laringe, hasta que finalmente el aire logra vencer la resistencia opuesta por los pliegues vocales. Estos, entonces, se separan y dejan pasar la corriente de aire rápidamente a través de la glotis. Por el *efecto Bernoulli*, la velocidad de dicho flujo provoca un descenso en la presión intraglotal y los pliegues, dada su elasticidad y su especial dinámica, vuelven a juntarse,

redundante si previamente han aparecido otras palabras con esa marca, en una cadena de concordancia. Eso hace que las /s/ de final de frase sean más prescindibles y, por tanto, más propensas a la elisión.

En cambio, Labov (1996 [1994]) recela de este tipo de explicación funcionalista, pues esta «supone que los oyentes procesan la información en cada palabra de una frase en el orden en que la reciben. Éste es un supuesto dudoso, y hay muchas pruebas psicolingüísticas contra él. Además, [... si] aceptamos este planteamiento, entonces tenemos que aceptar la idea de que la lengua inglesa es masivamente contrafuncional, pues en inglés la marca de plural aparece normalmente sobre el último elemento de la frase nominal» (p. 851). Además, los datos empíricos que maneja Labov demuestran que la realidad es mucho más compleja, pues a menudo se produce un efecto de perseverancia; es decir, si una secuencia de palabras que concuerdan comienza con una marca morfológica explícita, esta tiende a mantenerse en los vocablos subsiguientes, mientras que si la marca no aparece desde el principio, es posible que esta no se manifieste en las palabras posteriores.

¹¹⁰ Aunque, técnicamente, la presión supraglotal se puede descomponer en la presión de la cavidad faríngea (P_f), la de la cavidad oral (P_o), y la de la cavidad nasal (P_n), aquí se utilizará P_o de una manera laxa para hacer referencia a la presión supraglotal, en su conjunto, siempre que no sea necesaria una mayor especificación.

con lo que se completa un ciclo. La repetición de este proceso constituye, precisamente, la fuente cuasiperiódica propia de los sonidos sonoros¹¹¹.

En ocasiones, sin embargo, a pesar de que la configuración glótica es propicia a la vibración, esta no llega a producirse, porque la P_s no supera a la P_o , de modo que las condiciones aerodinámicas de la vibración no se cumplen. Una P_s insuficiente puede venir ocasionada por el menor remanente de aire pulmonar hacia el final de un grupo espiratorio. Por extensión, esta característica de un tipo de posición final se puede transferir al final de palabra y, ulteriormente, al final de sílaba¹¹². Además, el debilitamiento característico de las codas suele manifestarse como una menor tensión articuladora, y también puede verse disminuida la tensión espiratoria.

A la inversa, el motivo de este ensordecimiento puede ser una P_o elevada, que dificulte el mantenimiento del diferencial de presión necesario a través de la glotis. Tal acumulación de aire en la cavidad oral ocurre cuando hay un cierre o estrechamiento acusado que impide su salida normal al exterior, *i. e.* cuando el sonido es una obstruyente. También las róticas, especialmente la vibrante [r], se prestan a este tipo de ensordecimiento, debido al aumento de la P_o con cada ciclo de obturación. Blevins (2004, pp. 105-106) sostiene que el alargamiento (aumento de la duración) típico de la posición final de constituyente (*vid.* § 3.1) contribuye al incremento de la P_o y, por tanto, favorece el ensordecimiento.

Algunas lenguas ponen en marcha algunas estrategias para asegurar el mantenimiento de la vibración glotal, las cuales incluyen

contracting the expiratory muscles; decreasing average area of the glottis and/or tension of the vocal folds; decreasing the level of activity in muscles which underlie the walls of the supraglottal cavity; actively enlarging the volume of that cavity; or creating a narrow opening between the posterior pharyngeal wall and soft palate [...]. These manoeuvres, occurring singly or in combinations, will have their greatest effect on duration of closure voicing when they occur during the closure interval itself, in concert with the rise in pressure above the glottis which naturally accompanies vocal tract occlusion (Westbury y Keating, 1986, pp. 151-152).

Sin embargo, estos autores continúan diciendo que cualquiera de esas estrategias comporta un mayor esfuerzo articulador que el mínimo requerido para producir una secuencia VCV y que, por tanto, si los hablantes tienden a buscar la reducción del coste articulador, será más fácil que se incurra en un ensordecimiento pasivo.

Por ejemplo, en español es frecuente que se produzca el ensordecimiento de las aproximantes espirantes en posición de coda. Recuérdese, además, que la neutralización de la tensión en dicha posición hace que /p/, /t̪/, /k/ en final de sílaba se equiparen a las

¹¹¹ El fenómeno es, en realidad, algo más complejo, ya que no solo basta con iniciar la vibración, sino que hay que mantenerla. Para una discusión de los modelos propuestos al respecto, *vid.* Titze (2000).

¹¹² A propósito del modelo aerodinámico de Westbury y Keating (1986), estos autores puntualizan lo siguiente: «The expectations from [P_s] modelling do not account for positional effects other than those associated with position in utterance. Other effects might be considered language-specific generalizations of the utterance patterns» (p. 161).

espirantes. Cualquiera que sea el origen (tanto si vienen del fonema oclusivo, tenso, sordo, como si vienen del espirante, laxo, sonoro), los alófonos en posición de coda pueden presentarse ensordecidos, como [β̥], [ð̥], [ɣ̥].

También algunas pronunciaciones muy extendidas por el centro peninsular, como *admirar* [aθmi'raɪ] en lugar de [aðmi'raɪ], *ignorar* [ixno'raɪ] en lugar de [iɣno'raɪ], pueden interpretarse como causadas por el tipo de ensordecimiento que estamos describiendo, con el añadido de que el segmento [– interrupto, – sonoro] se reinterpreta como fricativa, debido a la estructura del sistema fonológico de esta lengua.

Este tipo de ensordecimiento que nos ocupa se denomina ensordecimiento pasivo, puesto que, en origen, no es producto de la realización voluntaria de un gesto articulatorio (que debería ser de apertura glotal). A este respecto, señalan Ladefoged y Maddieson (1996):

there is a conflict between an acoustic and an articulatory definition of voicelessness. For some linguists, voicelessness invariably implies an open glottis, whereas for others it means the absence of vibration, whether produced by active laryngeal control or not. [...] it is important to bear this distinction between active and passive devoicing in mind (p. 49, énfasis añadido).

Para Locke (1983, pp. 178-179), el ensordecimiento pasivo es tanto más probable si la vocal precedente es una vocal alta. Puesto que el volumen de la cavidad bucal es ya reducido durante la pronunciación de la vocal, la P_0 aumenta antes, una vez que comienza la articulación de la consonante. De hecho, en las lenguas con vocales ensordecidas, estas son predominante o exclusivamente vocales altas¹¹³. Si el peligro de ensordecimiento existe para la propia vocal, más aún para la consonante siguiente, puesto que la obstrucción articulatoria eleva todavía más la presión.

También el punto de articulación de la consonante influye en la mayor o menor probabilidad de que se produzca el ensordecimiento pasivo. Si el punto de articulación es muy retrasado, como por ejemplo el de las velares, la cavidad que queda entre la glotis y la obstrucción oral es muy pequeña. Además, no solo el volumen de la cavidad resultante es menor, sino que también la superficie de las paredes de dicha cavidad es más reducida, de modo que la posible contribución de la cesión (en inglés, *compliance*) por parte de dichas paredes ante la presión ejercida sobre ellas también se ve limitada. Entonces la P_0 aumenta mucho más rápidamente y es más fácil que se dé el ensordecimiento. En cambio, esto es mucho menos frecuente con puntos de articulación adelantados, que permiten un mayor volumen de la cavidad oral (Blevins, 2004; Laver, 1994; Locke, 1983; Westbury y Keating, 1986). Tipológicamente, se comprueba que /b/ es más frecuente que su respectiva sorda /p/, mientras que /k/ es más frecuente que /g/. Si una lengua tiene un sistema fonológico defectivo en cuanto a las oclusivas, lo más probable es que falte la /p/ o la /g/ (Locke, 1983, p. 146)¹¹⁴. Según Maddieson (2006),

¹¹³ Vid. nota 122 en la p. 149.

¹¹⁴ Este autor señala, además, que tanto los niños como los afásicos tienden a sustituir [g] por [k] en la mayoría de las ocasiones. En el caso de los niños, el pequeño tamaño de su boca puede explicar que este tipo de condicionantes aerodinámicos sean todavía más fuertes.

en la muestra de 451 lenguas llamada UPSID¹¹⁵ hay 40 inventarios que no tienen /g/, cuando se esperaría que la tuvieran, porque tienen otras velares y otras oclusivas sonoras; 38 no tienen /p/, cuando también se esperaría que la tuvieran. En cambio, hay 15 que no tienen /d/, 9 que no tienen /b/, 3 que no tienen /k/ y tan solo 1 que no tiene /t/. Mielke (2009, pp. 704-706) añade lo siguiente:

the recurrent lack of /g/ can be attributed to a difficulty in maintaining voicing when the space between the glottis and the stop closure is small, because intraoral pressure increases rapidly, but must be lower than subglottal pressure in order for voicing to be maintained (Ohala 1983, Keating 1984). The absence of /p/ may also have a phonetic explanation, but Maddieson (2001) shows that there is much more of an areal phenomenon than the lack of /g/.

Atendiendo a los datos y explicaciones anteriores, el orden de difusión del ensordecimiento pasivo es tal que el fenómeno afecta primero a los puntos de articulación más retrasados y luego a lo más adelantados, *i. e.* antes a las velares, después a las dentoalveolares y por último a las labiales. Por ejemplo, Blevins (2004, p. 106) recoge la noticia de ensordecimientos pasivos en tonkahua y en frisón, donde tan solo las velares se han visto afectadas. Si se compara con el orden de difusión en cuanto al punto de articulación de otro fenómeno descrito anteriormente —la debucalización (*vid.* § 4.1.1)—, se observará que no hay una coincidencia exacta. En aquel caso las coronales ocupaban un extremo de la secuencia (el más propenso al cambio), mientras que aquí se encuentran en un punto intermedio. Esta divergencia encuentra una explicación natural en las diferentes causas de ambos fenómenos. Si en la debucalización eran la menor área de la obstrucción y la mayor velocidad articulatorias las que justificaban una mayor tendencia a la debucalización, en cambio en el ensordecimiento pasivo la explicación es de tipo aerodinámico, de modo que un punto de articulación más adelantado implica mayor volumen de la cavidad y menor probabilidad de ensordecimiento.

Como se ha dicho antes, para compensar la tendencia al ensordecimiento pasivo, los hablantes disponen de una serie de estrategias. Por ejemplo, en algunas lenguas, para conseguir que una sonora sea sonora todo el tiempo, se baja la laringe, con el consiguiente descenso de la P_0 (Ladefoged y Maddieson, 1996, p. 51)¹¹⁶. Otra posibilidad es dejar muy relajados los músculos de la boca, para que, si el aire que se va acumulando en la cavidad oral hace mucha presión, se pueda aumentar el volumen de la cavidad, gracias a la elasticidad de los músculos. Y también sirve adelantar la raíz de la lengua o bajar un poco la mandíbula inferior. Sin embargo, para una misma configuración glotal, no hay ninguna lengua que distinga fonológicamente entre realizar gestos complementarios que aseguren la vibración real de los pliegues vocales o no

¹¹⁵ UCLA Phonological Segment Inventory Database.

¹¹⁶ A veces puede ocurrir que la P_0 descienda incluso por debajo de la presión atmosférica, que es la que hay fuera de la boca, y, al liberar la oclusión, se produzca un ingreso de aire del exterior. Este es el mecanismo empleado para pronunciar consonantes implosivas. Ladefoged y Maddieson (1996, *ibid.*) señalan que hay una especie de continuo entre las oclusivas sonoras y las implosivas, pero aclaran inmediatamente que no se trata en absoluto de la misma cosa.

realizarlos (Ladefoged y Maddieson, 1996, p. 52). En cualquier caso, aun existiendo estos posibles mecanismos de compensación, el ensordecimiento pasivo es un fenómeno frecuente en las lenguas en posición de coda y, especialmente, en final de palabra.

Tradicionalmente se ha considerado que los sonidos sordos son más tensos (más fuertes) que sus correspondientes sonoros. En esta consideración intervienen tanto descripciones estáticas de dichos sonidos como, por supuesto, los estudios sobre la sonorización en ataque intervocálico, una posición que favorece la lenición (*vid.* § 4.1.5). Por eso Lavoie (2001, p. 7) considera que este tipo de ensordecimiento en posición final le acarrea problemas teóricos, al tener que aceptar que tanto la sonorización como el ensordecimiento pueden ser fruto de un debilitamiento consonántico. Prefiere, entonces, atribuirlo a una mera neutralización en la coda. Sin embargo, decir que se trata de una neutralización no soluciona el problema, pues habría que explicar por qué en esa posición la dirección de la neutralización favorece sistemáticamente a la sorda. Las neutralizaciones se producen en las posiciones en que es habitual un cambio de motivación fonética, mientras que las oposiciones se mantienen allí donde no se dan esos cambios (Blevins, 2004, p. 97). La fuerza de las distintas posiciones prosódicas puede condicionar la estructura fonológica de la lengua, al determinar las neutralizaciones que se producirán más probablemente.

Otros autores, como por ejemplo Iverson y Salmons (2007), defienden que un ensordecimiento es siempre un tipo de fortición, aunque se dé en posición de coda. Sin embargo, hay sobrados argumentos para considerar que el ensordecimiento pasivo es un caso de lenición (John Harris, 2009); en esto resulta crucial la diferencia que hemos trazado entre el ensordecimiento activo y el pasivo. Desde luego, parece mucho más adecuado catalogar un cambio por las características del proceso y por sus causas que no simplemente por su resultado. La posición de coda, como se ha argumentado hasta ahora, es una posición lenitiva por excelencia por sus propias características prosódicas. Además, teniendo en cuenta el contexto segmental que sigue habitualmente a una posición de coda, se puede entender que las claves perceptivas de las oposiciones basadas en la voz —entre ellas el VOT— se debiliten cuando lo que sigue es otra consonante. A esto se suma el hecho de que el ensordecimiento a menudo se produce a la par que otros procesos que son claramente leniciones (piénsese, por ejemplo, en la pérdida de tensión de /p/, /t/, /k/ en coda, que se neutralizan con /β/, /ð/, /ɣ/, y que pueden presentar alófonos laxos ensordecidos [β̥], [ð̥], [ɣ̥]).

John Harris (2009) reflexiona sobre la adecuación del término «ensordecimiento» en lenguas como el alemán, donde la oposición no se da entre sonora y sorda sino entre no aspirada y aspirada. Entre las causas de la pérdida de aspiración se encuentra la poca estabilidad de las claves perceptivas (VOT) en contexto preconsonántico, así como el debilitamiento articulatorio típico de la posición de coda, que lleva consigo una reducción sustantiva del gesto de apertura glotal. Continúa argumentando que lo que parece que ocurre en las codas es la pérdida de las oposiciones basadas en rasgos laríngeos, ya sean estos [$\pm$ sonoro] o [$\pm$ glotis ensanchada]¹¹⁷, lo que aboca a un

¹¹⁷ Para una primera justificación de este rasgo, *vid.* Halle y Stevens (1971).

resultado común (la sorda no aspirada). Esta pérdida de todos los rasgos de un nodo¹¹⁸, en este caso el laríngeo, recuerda a la pérdida de otro nodo, el supralaríngeo, en otro fenómeno que hemos visto anteriormente: la debucalización.

4.1.5. Vocalización

En algunas variedades del español de América, las consonantes en posición de coda se pronuncian de manera tan relajada que llegan a convertirse en una paravocal posnuclear, p. ej. *absoluto* [aʊso'luʔo], *cuerpo* ['kʰeɪpo]. Recasens (2002) sostiene que es posible que el cambio no se deba tanto a un debilitamiento articulatorio como a una similitud espectral y una recategorización perceptiva de la transición consonántica. Pienso que ambas explicaciones son perfectamente compatibles.

En el paso del latín al castellano también se produjeron vocalizaciones de consonantes en coda. Por ejemplo, *factu(m)* > *faito*, o *nocte(m)* > *noite*, con las subsiguientes evoluciones provocadas por la yod —ya por asimilación—¹¹⁹. Por su parte, la lateral /l/, cuyos gestos de aproximación dorsal y contacto coronal se producen secuencialmente cuando el segmento ocupa la posición de coda silábica (*vid.* § 2.1.1.6), experimentó un debilitamiento del gesto coronal y una reinterpretación de su otro gesto como un gesto vocálico. El resultado concreto de dicha vocalización depende del contexto. Tras la vocal /a/, la lateral vocaliza en [u], p. ej. *alteru(m)* > *aʊtru* > *otro*, mientras que tras /u/ dio lugar a una [i], p. ej. *multu(m)* > *mui̯to* > *mucho* (Recasens, 2002).

Finalmente, existen casos de vocalización del primer elemento de un ataque complejo, p. ej. *hablar* [aʊ'laɾ], *madre* ['maɪre], como ocurre, por ejemplo, en Chile (*op. cit.*). Aunque, técnicamente, no corresponderían a este apartado sobre las codas, lo cierto es que se ha producido una reinterpretación de la estructura silábica (quizá incluso antes de la vocalización pero, indudablemente, después de ella), pues la nueva (para)vocal pasa a la coda de la sílaba anterior. En estos casos parece que cobra más peso la idea de que el proceso no consiste tanto en un debilitamiento articulatorio (pues el segmento ocupaba originalmente una posición de ataque) como en una recategorización de las transiciones espectrales. Ahora bien, este tipo de vocalizaciones de ataques solo se produce cuando se trata de ataques complejos, y de este modo la segunda sílaba no queda desprovista de ataque, porque sigue habiendo una consonante en dicha posición.

¹¹⁸ El término «nodo» se utiliza aquí en el sentido habitual de los trabajos sobre geometría de rasgos (Clements, 1985).

¹¹⁹ En este tipo de grupo /kt/ seguramente se produjo, en un primer momento, una fricativización de la oclusiva. De hecho, parece que el debilitamiento afectó especialmente a las oclusivas que eran velares pues son en las que se acumula más P_o, debido al reducido volumen situado por detrás del punto de cierre. Si se pronuncia con una tensión insuficiente, es probable que no se resista el empuje del aire y se termine liberando la oclusión. Por tanto, la evolución sería [kt] > [xt]. Posteriormente, por asimilación a la consonante siguiente, es probable que la fricativa adelantara su punto de articulación: [çt]. Por último, se produciría la vocalización. Para más detalles sobre esta evolución, *vid.* p. ej. Recasens (2002).

4.2. Ataque silábico en contexto intersonántico

Aunque en la tradición pesa la denominación de *posición intervocálica*, se debe matizar que ni la palabra «posición» ni la palabra «intervocálica» son las más acertadas.

En primer lugar, muchos de los típicos fenómenos que ocurren entre vocales también se dan cuando alguno de los segmentos contiguos no es una vocal sino una consonante sonante. Por ejemplo, la sonorización *patrem* > *padre* es un buen testigo de ello, como reconoce Martínez Celdrán (1993, p. 622), siguiendo a Martinet (1974 [1955]). Ségéral y Scheer (2008) añaden: «it appears that for any given position, obstruents engaged in *muta cum liquida* behave exactly as their simplex peers» (p. 146). Es decir, que los ataques ramificados (al menos los de sonoridad creciente) siguen el mismo patrón que cualquier ataque silábico simple (a excepción, claro está, de los casos de vocalización que se han visto en el § 4.1.5). No obstante —admiten Ségéral y Scheer—, hace falta corroborar si esta afirmación es de alcance universal para todas las lenguas.

En segundo lugar, y como se ha explicado anteriormente (*vid.* § 4), la tradicional denominación de posición intervocálica trasluce un criterio mixto. Por un lado, implica una referencia a la posición prosódica de ataque silábico. Aunque esta mención no es explícita, es la única posibilidad para una consonante flanqueada por vocales dentro de una misma palabra, dado el Principio de Maximización del Ataque (Selkirk, 1982)¹²⁰. Por otro lado, le otorga un papel al contexto segmental adyacente (sendas vocales a cada lado).

En definitiva, frente a la consagrada etiqueta de *posición intervocálica*, conviene incidir en la idea de que lo que es una verdadera posición es el ataque silábico, mientras que la mención a las vocales —o, mejor aún, a las sonantes— contiguas debería designarse como contexto. Ahora bien, hay que reconocer que «posición de ataque en contexto intersonántico» es una etiqueta un poco larga, pero «ataque intersonántico» debería ser suficiente para separar ambas especificaciones —posicional y contextual—. Por otro lado, «ataque intervocálico» puede estar justificado en los casos en que un fenómeno realmente se restrinja a ese contexto que, en cualquier caso, sigue siendo el más prototípico de los intersonánticos.

La posición de ataque tiene motivos para ser considerada una posición fuerte, al ser inicial de constituyente prosódico. De hecho, es una posición más fuerte que la coda, por ejemplo, y en ocasiones puede albergar procesos de fortición. Sin embargo, por tratarse de un inicio de constituyente menor (la sílaba), no es tan fuerte como la posición (de ataque) inicial de palabra. Comparativamente, pues, suele caracterizarse por una menor duración de los gestos articulatorios. Tal reducción temporal, obviamente, puede ocasionar una reducción sustantiva de dichos gestos, ya que los órganos articulatorios no disponen de mucho tiempo para alcanzar su meta, y pueden incurrir en una hipoarticulación. Ambas reducciones suponen, claro está, un debilitamiento, que puede traducirse en un cambio segmental por lenición.

¹²⁰ Incluso las consonantes ambisilábicas ocuparían una posición de ataque, simultáneamente a la de coda de la sílaba anterior.

Aunque la lenición como tal se justifica por la posición prosódica de que se trata, existe en este caso un factor coadyuvante en el contexto segmental. Teniendo en cuenta que las vocales se caracterizan por gestos articulatorios de apertura oral, es comprensible que un gesto consonántico en dirección opuesta (el cierre) no puede alcanzar, ni mucho menos mantener, una gran magnitud si no dispone de tiempo suficiente para realizarse antes de que los órganos articulatorios tengan que volver a apartarse para pronunciar la siguiente vocal.

El efecto fundamental de la lenición en este caso es, por tanto, la reducción temporal y sustantiva de los gestos articulatorios orales, como veremos, por ejemplo, en la espirantización. Sin embargo, los fenómenos lenitivos típicos del ataque intersonántico se suelen caracterizar, además, por la sonorización de la consonante afectada. Al igual que sucede con los gestos orales, el gesto de apertura glotal que caracteriza a una consonante sorda puede no llegar a alcanzar su objetivo dada su menor duración (lo que vuelve a suponer una instancia de lenición).

Ahora bien, la inherente independencia de los distintos gestos articulatorios puede hacer que la realización del gesto glotal se presente desalineada temporalmente con respecto a los gestos orales. Como consecuencia, parte de la sonorización característica de las sonantes circundantes se puede solapar con el segmento sordo y, dada la brevedad de este, cualquier solapamiento de este tipo puede suponer un porcentaje importante de su duración total. Por otro lado, la menor duración y magnitud de los gestos orales determina un menor aumento de la P_0 durante la pronunciación de la sorda. Esta circunstancia aerodinámica puede ser causa de una sonorización pasiva, como se verá enseguida.

Estos factores coadyuvantes se fundamentan en el hecho de que sean sonantes los segmentos contiguos a la consonante afectada y, aquí sí, entramos en el terreno de la asimilación, aunque debe quedar claro que la propagación de rasgos (en este caso, la sonoridad) es posible tan solo porque en primer lugar se ha producido una reducción de los gestos consonánticos (*i. e.* una lenición).

Algunos autores (p. ej. Szigetvári, 2008, § 4) subrayan el problema que supone considerar que la sonorización intervocálica es una asimilación, entre otras cosas porque sería difícil justificar la necesidad de dos desencadenantes, uno a cada lado. Tienen razón en esto; pero quizá la complementariedad de la lenición y la asimilación, en la línea de los argumentos descritos arriba, pueda explicar lo atípico de este fenómeno. De hecho, el mismo autor reconoce: «assimilation is a separate phenomenon, which may be triggered off by a previous instance of lenition» (p. 106).

4.2.1. Sonorización pasiva

La sonorización de las sordas en posición intervocálica es el fenómeno de lenición en el que primero se centraron todos los trabajos clásicos, asociado a otro tipo de cambio —la espirantización— al que suele acompañar. No obstante, nótese que, tal y como se acaba de explicar, a pesar de haber sido considerado el ejemplo de lenición por

antonomasia, en realidad constituye un fenómeno híbrido, desencadenado por una lenición inicial pero acompañado por una posterior asimilación¹²¹ y ¹²².

Al igual que se ha hablado de un ensordecimiento pasivo, el tipo de sonorización que se produce en contexto intersonántico se ha denominado sonorización pasiva.

Passive voicing means that without the devoicing effect of glottal spreading, stops can be voiced during most or all of closure if they occur in the context of sonorant sounds, even if there are no active voicing gestures (vocal fold slacking, tongue-root advancement, etc.) on the part of the speaker (Jessen y Ringen, 2002, p. 190).

La sonorización pasiva puede afectar a las obstruyentes que, sin estar especificadas activamente como [+ sonoro], tampoco lo están como [+ glotis ensanchada]. Es decir, actúa sobre el elemento no marcado de la oposición de rasgos laríngeos, que será el elemento relativamente *fortis* en las lenguas de sonorización y el elemento relativamente *lenis* en las lenguas de aspiración (Gósy y Ringen, 2009). El hecho de que la sonorización pasiva pueda afectar por igual a un elemento *fortis* o *lenis* puede alertar de que el fenómeno no se puede justificar exclusivamente con referencia a la fuerza, sino que, efectivamente, interviene otro proceso (una asimilación). No obstante, se debe admitir que desde el punto de vista fonético esa *fortis* de un sistema y la *lenis* de otro son equiparables (son sordas no aspiradas), y que el cambio es hacia más *lenis* aún.

En lo que sí ayuda la posible lenición vinculada al ataque silábico (que es un inicio de constituyente relativamente menor) es en la reducción de la duración del segmento, lo que facilita, desde el punto de vista aerodinámico, la sonorización pasiva. Westbury y Keating (1986) señalan que «a stop sandwiched between two identical vowels should naturally be largely voiced if its closure is relatively short» (p. 152) y continúan diciendo que «under certain conditions [...] it is more difficult to terminate than to allow (or maintain) voicing during a stop» (p. 158).

Es sabido que la presión transglotal necesaria para *mantener* la vibración de los pliegues vocales es menor que la requerida para *iniciar* dicha vibración: en torno a 2000 dyn/cm² y 4000 dyn/cm² respectivamente, según Baer (1975), *apud* Westbury y Keating (1986). Por tanto, cuando el proceso vibratorio ya está en marcha (por la sonante precedente), este puede continuar durante una obstruyente siguiente si la duración de la misma es lo suficientemente breve como para que la constricción oral no llegue a provocar tal aumento de P_o que se pierda el umbral de P_t requerido. Todo esto —bien entendido— siempre que los pliegues vocales permanezcan también lo bastante

¹²¹ Esta idea de la sonorización como asimilación ya la han defendido diversos autores desde hace tiempo. En el ámbito concreto de la lingüística histórica del español, *vid.* p. ej. Bustos Tovar (1960).

¹²² Cabe advertir que la asimilación entre consonantes y vocales no siempre toma la dirección de la sonorización en todas las lenguas. Por ejemplo, las vocales altas del japonés, /i/ y /u/ —*vid.* Nogita, Yamane, y Bird (2013) para una justificación de esta transcripción frente a la clásica /ui/—, se ensordecen entre consonantes sordas o entre sorda y final de palabra, p. ej. [arʌkimaʃiʈa] ‘caminar+PASADO’, [ʈsʌkaimasʌ] ‘utilizar+PRESENTE’. El mayor incremento de la P_o, motivado por el menor volumen de la cavidad oral, puede explicar que la alternancia únicamente afecte a las vocales altas (mediante una especie de ensordecimiento pasivo); pero indudablemente es un caso de asimilación, puesto que solo se produce en presencia de otros segmentos sordos.

juntos, como se presentarían en el estado fonatorio de la vocal, y esto es posible, igualmente, si la brevedad de la consonante impide desarrollar un gesto de apertura glotal de la magnitud requerida para conseguir una sorda activa. Además, el nivel de P_0 se mantendrá más bajo si los músculos de las paredes del tracto vocal permanecen relajados y elásticos, permitiendo que las paredes cedan ante el aumento de presión, compensándolo. En tales condiciones, la vibración de los pliegues puede continuar de manera espontánea hasta unos 100 ms después de que se produzca una oclusión; en cambio, si las paredes están tensas, la fonación apenas se mantendrá unos 10 ms (Westbury y Keating, 1986).

Además, teniendo en cuenta que la P_s crece o decrece abruptamente al inicio y al final de un enunciado, respectivamente, se espera que la P_t necesaria para la sonorización pasiva se dé más bien en posición media de enunciado, que es justo donde esperamos la aparición de un ataque silábico si este debe ir flanqueado por sonantes (*ibid.*).

El carácter aerodinámico de la causa que subyace a la sonorización pasiva justifica que su orden de difusión, en cuanto al punto de articulación, sea primero a las labiales, luego a las coronales y luego a las dorsales; es decir, justo el orden contrario al del ensordecimiento pasivo (*vid.* § 4.1.4 para más detalles). No obstante, Bustos Tovar (1960) y Recasens (2002) coinciden en señalar algunos casos paradójicos en los que son las dorsales las primeras afectadas por la sonorización. Cabe pensar que, si la base de articulación de una lengua es poco tensa, resulte más difícil resistir la mayor P_0 acumulada tras el punto de articulación velar y sean estas las primeras en reducir la magnitud del gesto de obstrucción, de modo que disminuya la P_0 . Como consecuencia de esto, aumentaría la P_t y sería más probable la sonorización pasiva. Nótese, en cualquier caso, que esta sonorización necesariamente debe ir acompañada de una (previa) espirantización.

Por otro lado, Recasens (2002) apunta que también hay una explicación aerodinámica plausible al hecho de que las oclusivas del latín no sonorizaran tras paravocal posterior /u/ —p. ej. *paucu(m)* > *poco*—, pero sí lo hicieran las fricativas —p. ej. *causa(m)* > /'kauza/, con /z/ sonora en castellano medieval—. El punto de articulación posterior de la /u/ contribuye a un nivel relativamente alto de P_0 en el momento de inicio de la consonante siguiente. Si esta se articula con un cierre oral completo, la subida de P_0 impide la sonorización pasiva; por el contrario, si hay un escape de aire, como en la fricativa, la sonorización es posible.

Por último, se predice que la sonorización pasiva será más difícil en la sílaba acentuada de la palabra, ya que esta sílaba se caracteriza por una mayor duración (*vid.* § 3.2).

4.2.2. Degeminación

Uno de los primeros cambios que se asociaron con la lenición, ya desde la formulación de las leyes neogramáticas del cambio fónico, es la simplificación de las consonantes geminadas en contexto intervocálico, como en los siguientes casos:

(43) Ejemplos de degeminación en el paso del latín al castellano

cappa(m)	> capa
gutta(m)	> gota
vacca(m)	> vaca
affirmare	> afirmar
ingressu(m)	> ingreso

Las geminadas se analizan como una configuración tal que un mismo material segmental —o melódico, como se dice a menudo— aparece asociado a una posición de ataque silábico y, a la vez, a la coda de la sílaba precedente, a la que se propaga íntegramente (*vid.* p. ej. Perlmutter, 1995). Teniendo esto en cuenta, podemos afirmar que encajan en la descripción de consonantes en ataque intervocálico. Por un lado, es cierto que la consonante ocupa la posición de ataque (la coda, en realidad, solo adquiere su interpretación segmental debido a la propagación de los rasgos, procedentes de dicho ataque). Por otro lado, en el nivel (*tier*) segmental solo hay una posición (no dos), y está efectivamente flanqueada por vocales a ambos lados. Si hubiera dos posiciones, cada consonante lindaría con otra consonante en uno de sus costados, y no podría decirse con total propiedad que se situara en un contexto intervocálico, pero no es ese el caso: no hay aquí dos consonantes, sino una única consonante, que es larga por ir asociada a dos posiciones del esqueleto.

Pues bien, dada la tendencia lenitiva a reducir la duración de las consonantes en esta posición, se puede entender que la degeminación sea una consecuencia directa de tal lenición, y ella provoca, a la larga, un cambio en la representación, con la desaparición de la posición de coda precedente.

4.2.3. Espirantización

La espirantización consiste en la pérdida del contacto total durante el cierre de una oclusiva, y su conversión en una aproximante de tipo espirante, como [β], [ð], [ɣ]¹²³. Se trata, por tanto, de una reducción sustantiva de los gestos articulatorios, aunque esta viene normalmente motivada por una reducción temporal, verdadera causante de la hipoarticulación. Ambos tipos de reducción se encuadran bien en los procesos de tipo lenitivo que caracterizan a la posición de ataque silábico en interior de palabra. Además, la mayor abertura de las vocales (o sonantes) del contexto puede justificar que, si el tiempo de ejecución disponible para la consonante meta se reduce, no pueda llegar a alcanzarse el objetivo articulatorio del cierre completo.

Los trabajos de los neogramáticos y de buena parte de la tradición historicista identificaron la lenición de una manera prácticamente exclusiva con los procesos de espirantización, junto con la degeminación y la sonorización de las sordas intervocálicas. Se trata, pues, de una concepción de la lenición bastante restringida. Pero no solo lo es por los fenómenos que incluye, sino también por considerar que la lenición era privativa del contexto intervocálico y por asociarla fundamentalmente a una familia de lenguas: la

¹²³ Para una justificación de la denominación de este subtipo de aproximante, *vid.* Martínez Celdrán y Fernández Planas (2013, cap. 2, § 2).

romance¹²⁴. Además, los neogramáticos consideraban que esos tres tipos de cambios (a saber, espirantización, degeminación y sonorización) eran procesos totalmente inconexos entre sí¹²⁵ y suponían que, por ejemplo, en el paso del latín al español, la degeminación aplicó a las geminadas, la sonorización a las sordas simples y la espirantización a las oclusivas que ya eran sonoras. Sin embargo, lo cierto es que las oclusivas sordas del latín han terminado dando aproximantes (espirantes) sonoras en español; es decir, han sufrido dos tipos de cambios¹²⁶. Por ejemplo, la espirantización ha terminado acompañando a la sonorización en un caso como el siguiente:

(44) totu(m) > todo [ˈtoðo]

A pesar de esta evidencia, muchos estudios tradicionales de diacronía se limitan a hablar de espirantización en los casos en los que el étimo latino ya contenía una (oclusiva) sonora:

(45) habere > haber [aˈβeɾ]

Es cierto que este sería un ejemplo «puro», digamos, de espirantización, al no ir acompañado de sonorización. Pero, como se acaba de exponer, no es la única posibilidad, sino que estos fenómenos se encuentran muy relacionados entre sí, por tratarse de instancias de lenición propiciadas en la misma posición y los mismos contextos.

García Santos (2002), aplicando métodos experimentales al estudio de la fonología histórica, ha reproducido en el laboratorio los tres tipos de cambios que la tradición consideraba lenición. En sus investigaciones, comprueba que, recortando la parte adecuada de un sonido para reducir su duración mediante un programa de tratamiento computacional de audio, podemos pasar sucesivamente de una geminada (sorda) a una oclusiva (simple) sorda, después a una oclusiva sonora y por último a una espirante

¹²⁴ Ni siquiera todas las lenguas romances sufrieron lenición, sino que esta se circunscribe a la Romania occidental. En oriente se conservaron las sordas: rumano *ripă, spata, spica*; dálmata *raipa, spuota, spaica*; dialectos del sur de Italia *catena, fuoco, amica* (Lapesa, 1981, p. 85). Esto hizo que se postulara la hipótesis de que la lenición se debía al influjo del sustrato celta, como sigue mencionando el propio Lapesa (*op. cit.*). Lloyd (1993), por el contrario, presenta tanto argumentos a favor como en contra de la hipótesis celtista, y dice que lo más sensato es admitir una posible influencia de tal sustrato en la parte occidental de la Romania, pero solo como elemento coadyuvante, combinado con otro tipo de explicación, de corte estructuralista.

¹²⁵ La irrupción del estructuralismo, con su idea de la lengua como sistema, trajo consigo un notable cambio de enfoque. Los cambios, que antes requerían una ley fonética independiente para cada uno, se explicaron de manera unificada como inercias generales de la lengua, motivadas por su estructura y por la tendencia a mantener el rendimiento funcional, de modo que dichos cambios se producen en cadena (de empuje o de tracción), *vid.* Martinet (1974 [1955]). El propio Martinet está de acuerdo con Thurneysen en que los tres tipos de cambios, englobados bajo la etiqueta general de lenición, se caracterizan por ser «una mutación de consonantes que normalmente tuvo su origen en una reducción de la energía empleada para articularlas» (*op. cit.*, p. 365).

¹²⁶ Lloyd (1993) puntualiza que el resultado sonoro de las sordas intervocálicas se equipararía en el sistema a las sonoras originales, y es por esto que también aquellas terminarían espirantizando, pero solo por paralelismo con estas. Consúltense esa obra para un análisis detallado de todas las implicaciones estructurales de estos cambios.

(sonora). Todos estos cambios históricos se explican de manera muy simple como progresivas reducciones en la duración de los gestos¹²⁷. En definitiva, estos experimentos reproducen, mediante la manipulación de sonidos en el laboratorio, las condiciones articulatorias que debieron de darse en la base de estos cambios, y comprueba que el resultado induce el mismo efecto perceptivo que el sonido que, en efecto, terminó resultando en cada uno de estos cambios históricos.

Para los sonidos velares y bilabiales, las relaciones entre la duración y lo percibido son como sigue. Las geminadas tienen su punto óptimo en torno a los 210 ms, aunque se distinguen hasta por encima de los 160-150 ms. Entre 140 y 70 ms se identifica la oclusiva sorda, cuyo punto óptimo se sitúa en torno a los 100-90 ms. «Por debajo de 60-50 ms se identifica el sonido sonoro. A medida que se elimina duración se percibe más nítidamente, situándose el punto óptimo entre 20-30 ms, aunque en nuestras pruebas se sigue identificando [...] hasta prácticamente 0 ms, frente a las mediciones de Martínez Celadrán» (García Santos, 2002, p. 89). La serie dental, por su parte, difiere bastante de las otras dos. A los 210 ms hay problemas ya para la identificación de la geminada; la sorda simple se distingue bastante clara en los 150 ms; y la sonora se percibe a partir de los 95-90 ms.

Como se ve, esto favorece el debilitamiento de las dentales por encima de las otras dos series. Sin embargo, se recordará que las causas aerodinámicas de la sonorización pasiva favorecen que ese otro cambio se propague en primer lugar a las labiales (§ 4.2.1). No hay contradicción en esto, pues el orden de difusión puede ser diferente según el fenómeno del que se trate en cada caso. Es cierto que la sonorización afecta primero a las labiales, como consecuencia de su naturaleza aerodinámica, por ser esos sonidos los que presentan un mayor volumen de la cavidad supraglotal. En cambio, cuando el fenómeno en cuestión es una degeminación o una espirantización, las coronales son las más débiles, por presentar el órgano activo una menor superficie de contacto y ser así más proclive a la reducción de la misma en el momento en que se produce una disminución de la duración.

4.2.4. Elisión (II)

El progresivo debilitamiento de las sonoras intervocálicas puede dar, como último paso en la cadena lenitiva, la pérdida total o elisión del segmento (recuérdese la importancia de la elisión como final de la cadena evolutiva, en la definición de lenición de Hyman, 1975, *vid.* nota 2 en la p. 28). En el paso del latín al español, la elisión afectó a instancias de [ð] y [ɣ] en contacto con alguna vocal anterior o con una yod y, en menor medida, a casos de [β], preferentemente —aunque no solo— en contacto con vocal posterior (labializada), como en los siguientes ejemplos:

(46) crudele(m)	> cruel
comedere	> comer
fide(m)	> fe

¹²⁷ Esta es la misma idea que defendía ya Martínez Celadrán (1993) y Lavoie (2001) reitera que el debilitamiento viene de la mano de una reducción temporal.

cadere	> caer
litigare	> lidiar
regale(m)	> real
fumigare	> humear
digitu(m)	> dedo
fastidiu(m)	> hastío
pulegiu(m)	> poleo
rivu(m)	> río
gingiva(m)	> encía

Las condiciones para que se produzca el cambio apelan, en este caso, no solo a la posición prosódica, sino también al contexto, pues la pérdida de la espirante se ve facilitada por la presencia de una vocal homorgánica. En cualquier caso, la difusión léxica del fenómeno presenta idiosincrasias, como discute, por ejemplo, Lloyd (1993).

Otro factor que contribuye a la pérdida es la pertenencia de la consonante a una sílaba no acentuada, especialmente si es la postónica (*vid.* p. ej. Pensado Ruiz, 1984), aunque esta misma autora puntualiza que en el caso de las esdrújulas la pérdida es menos frecuente, ya que interactúa con el fenómeno de la síncope vocálica.

Posteriormente, el fenómeno de la elisión de espirante intervocálica ha afectado de nuevo al español, con mayor o menor impacto según las zonas donde se habla esta lengua. El segmento que más ha sufrido dicha elisión es la [ð], que, por ser coronal, pone en juego una menor masa de su órgano articulatorio y eso favorece su pérdida. Este cambio se documenta en el español de España, en palabras como *cada* o *nada*, y sobre todo en los participios¹²⁸ (aun siendo esta [ð] procedente de sorda latina). Pensado Ruiz afirma que «la pérdida actual de la *-d-* en el participio es un fenómeno sin relación con la pérdida antigua» (*op. cit.*, p. 26). La pérdida contemporánea de la consonante también se manifiesta en el español de América, en concreto en las Antillas, y en la costa caribeña de Panamá y Colombia; en la costa pacífica de toda Sudamérica (pero no en las zonas interiores de los países andinos); y en Paraguay. Lógicamente, la difusión léxica y la valoración sociolingüística del fenómeno puede cambiar de un lugar a otro. Además, en las zonas de América mencionadas, el proceso no se ha limitado a la elisión de [ð], sino que también afecta a la [β] y a la [ɣ], sobre todo si el contexto vocálico lo favorece (p. ej. *agua* ['awa]).

4.2.5. Debucalización (II)

En algunas variedades del español se produce la llamada aspiración de /x/, pronunciada como [h]. Esto ocurre, por ejemplo, en parte de Castilla-La Mancha, y también en Extremadura, Murcia, o Andalucía; y, dentro de las variedades americanas, en Centroamérica, las Antillas, Venezuela, Colombia, y la costa de Ecuador y de Perú. Se trata de un caso de lenición, más frecuente en sílaba no acentuada y cuando la /x/ está flanqueada por dos vocales del mismo timbre, lo cual se puede interpretar como

¹²⁸ La elisión es más probable cuando el segmento es morfológicamente predecible (*vid.* nota 109 en la p. 140).

que el debilitamiento basado en las condiciones prosódicas lleva a una más fácil asimilación al grado de abertura de las vocales y provoca la debucalización.

4.3. Otras instancias de ataque silábico

En este apartado se describirá el comportamiento de los ataques silábicos que no se encuentran en contexto intersonántico, o bien de aquellos ataques que, aun siendo intersonánticos, no presentan fenómenos en los que tal contexto juegue un papel determinante ni coadyuvante.

Típicamente, esto incluye los ataques en posición inicial de palabra, una posición fuerte, por ser inicial de constituyente prosódico y por ser este un constituyente relativamente mayor (en comparación con la sílaba, por ejemplo). Además, un ataque inicial de palabra puede llegar a ser también inicial de otros niveles prosódicos superiores, por ejemplo, en el caso de que la palabra en cuestión encabece un enunciado.

Debe entenderse que, en una secuencia de palabras, un ataque en posición inicial puede perfectamente hallarse en contexto intersonántico si la palabra precedente acaba en vocal, por ejemplo. La fonología de las distintas lenguas puede variar en el tratamiento de los fenómenos posléxicos que afectan a los segmentos en situación de juntura de palabras. En cualquier caso, aunque exista la posibilidad de que un ataque inicial aparezca precedido por una vocal, esa circunstancia convive necesariamente con la contraria; es decir, también es perfectamente posible que aparezca precedido por consonante, o incluso por ningún segmento, sino por una pausa, en el caso de que el inicio de palabra sea también un inicio de enunciado.

Otro caso de ataque silábico no intersonántico es lo que Ségéral y Scheer (2008) llaman la posición de poscoda, también una posición fuerte según ellos (*vid.* § 4). Sin embargo, el próximo apartado ilustra un fenómeno del español en el que se ve que no todas las consonantes que van detrás de una coda se comportan del mismo modo, por lo que no siempre tiene sentido proponer la «poscoda» como una clase natural de posiciones en cuanto a los fenómenos que desencadena (*vid.* especialmente la nota 130).

4.3.1. Oclusivización de aproximantes espirantes

En la mayor parte de variedades del español hay una alternancia entre los alófonos aproximantes espirantes —[β], [ð], [j], [ɣ]—, que aparecen en la mayoría de posiciones y contextos, y los alófonos oclusivos —[b], [d], [t̪], [g]—, que aparecen tras pausa, tras consonante nasal, o también tras lateral en el caso de [d] y [t̪]¹²⁹. Así, por ejemplo, tenemos la espirante [β] en *haber, abrir, hablar* (*i. e.* encabezando sílaba precedida por vocal, tanto en ataque simple como complejo); *club, obsceno* (en coda —simple o compleja—, precedida por vocal); *alba, árbol, esbozo* (precedida por buena parte de las consonantes del español¹³⁰, incluyendo la lateral, en el caso de [β], pero no así en el de

¹²⁹ Hay ciertas variedades en las que también se produce una alternancia similar, pero con algunas particularidades, como se verá enseguida.

¹³⁰ Esto parece desmentir, al menos para este fenómeno del español, la idea defendida por Ségéral y Scheer (2008) de que las consonantes situadas tras una coda son siempre más fuertes. Como se ve, en este caso, solo se produce el refuerzo (con carácter alofónico) detrás de algunas consonantes. No obstante, se

la dentoalveolar: p. ej. *toldo* [ˈtoɫˤdo]). En cambio, precedida por nasal (*ámbito*), tenemos la oclusiva [b]. En inicio de palabra, aparece la espirante también en la mayoría de los casos: *una balsa*, *unas botas*, *el barato*. Sin embargo, la realización es oclusiva cuando la palabra precedente acaba en nasal (*un balón*). En el caso de la dentoalveolar y de la palatal, también cuando precede lateral (*el décimo*, *el yate*). Finalmente, si el inicio de palabra coincide con inicio absoluto de enunciado, también se pronuncia la oclusiva.

Como se ha visto, el contexto tras lateral solo afecta a dos de las cuatro obstruyentes en cuestión, pero no a las otras dos. En aras de una mayor generalización, este contexto se pueden aunar con el de tras nasal (que afecta a todas) bajo la siguiente descripción: tras interrupta homorgánica. En efecto, las nasales se asimilan al punto de articulación de la consonante siguiente y, por tanto, serán homorgánicas con respecto a cualquiera de las cuatro obstruyentes que nos ocupan. En cambio, la lateral /l/ solo es homorgánica con respecto a las dos obstruyentes con especificación coronal, esto es, [d] y [t̪]¹³¹, pero no la labial [b] o la dorsal [g]. Por otro lado, esta generalización implica aceptar que, desde el punto de vista fonológico, en español la lateral se comporta como interrupta. Esta visión es avalada por James W. Harris (1969, p. 39 y ss.). Mielke (2005) defiende que las laterales de distintas lenguas se comportan de manera diferente en cuanto al valor continuo/interrupto.

Las descripciones tradicionales (Alarcos Llorach, 1965; Navarro Tomás, 1996 [1918]; Quilis, 1993) no consideran algunos otros contextos marginales que pueden surgir en interior de palabra o a través de frontera léxica, fundamentalmente en presencia de préstamos y extranjerismos.

Por su parte, al centrarse únicamente en la nasal y en la lateral, algunos autores, como James W. Harris (1969) restringen aún más la especificación del contexto, y hablan de *sonantes interruptas homorgánicas*, en lugar de *consonantes interruptas homorgánicas*. Sin embargo, esta segunda descripción es más completa y exacta.

En el léxico patrimonial del español no hay ninguna palabra en la que alguna de estas obstruyentes laxas vaya precedida por una obstruyente interrupta homorgánica (que, necesariamente, habría de ser una oclusiva tensa /p/, /t̪/, /k/), excepto algunas derivadas

han documentado pequeñas variaciones subsegmentales en el grado de abertura de las obstruyentes laxas en contexto posconsonántico, y con mayor o menor efecto según cuál sea la consonante precedente (*vid.* Carrasco, Hualde, y Simonet, 2012, y las referencias allí incluidas). Estos datos apuntan a que en el proceso de variación de estos segmentos hay ciertas contribuciones de carácter coarticulatorio.

¹³¹ Hay autores, como Keating (1988), que defienden que las palatales son segmentos complejos, con una doble especificación: como coronal y como dorsal. En cambio, Recasens y Romero (1997), a partir de un estudio articulográfico, sostienen que esta circunstancia puede variar de una lengua a otra. En ruso, por ejemplo, hay cierto desfase entre los momentos en que se alcanza la máxima magnitud de constricción en las zonas correspondientes a la articulación coronal y a la articulación dorsal. Por el contrario, en catalán, el lapso de tiempo transcurrido entre ambos picos de magnitud es mínimo, lo que se interpreta como que en realidad no hay dos movimientos, sino uno solo, de la zona coronal, que produce una constricción palatoalveolar. No obstante, hay que tener en cuenta que es perfectamente posible que dos lenguas presenten sendas especificaciones de dos gestos y que solo se diferencien en su coordinación intergestual (en una de ellas, los gestos podrían llegar a ser casi simultáneos). En cualquier caso, como se verá, los datos presentados en este apartado muestran un patrón fonológico de alternancias que parece confirmar que las palatales en español están especificadas simultáneamente como coroneales y como dorsales, pues están sujetas a procesos de alternancia desencadenados por motivos coarticulatorios tanto cuando la consonante adyacente es dentoalveolar como cuando es velar.

con el prefijo *post-*, como *postdata* o *postdorsal*. Es cierto que la /t̪/ de dicho prefijo no suele pronunciarse, y en ese caso tenemos una pronunciación aproximante de la obstruyente laxa ([poʃˈð̪a̠a], [poʃˈð̪oɾˈsal]). Pero en caso de llegar a pronunciarse la /t̪/ final del prefijo, entonces la laxa se articularía como oclusiva ([poʃˈt̪a̠a], [poʃˈt̪oɾˈsal]).

Algo parecido ocurre cuando se encuentra una oclusiva tensa del final de una palabra (de origen extranjero, habitualmente) con una obstruyente laxa del inicio de la palabra siguiente. Es posible que la consonante final, extraña a la fonotáctica del español, no se pronuncie. Pero, si llega a pronunciarse, entonces la obstruyente laxa se realizará como oclusiva. Los siguientes ejemplos recogen algunas oraciones en las que esto puede suceder.

- (47) Ahora el ketchp viene en un envase nuevo.
- (48) Había un bufet de ensaladas.
- (49) El día del debut llegó más rápido de lo esperado.
- (50) Ese brick lleva abierto más de un mes; tíralo.
- (51) Es un cómic gallego.

No debe achacarse la oclusivización en estos casos a un mayor énfasis o cuidado en la dicción con motivo del esmero en pronunciar el segmento previo, extraño a la fonotáctica patrimonial, pues el segundo elemento sigue pronunciándose como espirante cuando la consonante previa no es homorgánica.

- (52) Creo que hoy me pondré el anorak verde.

Lo mismo puede decirse de grupos no homorgánicos de interior de palabra.

- (53) El fútbol es muy popular en España.
- (54) Te voy a contar una anécdota.

Por tanto, la generalización adecuada es que las realizaciones oclusivas se dan tras pausa y tras consonante interrumpida homorgánica, o simplemente tras interrumpida homorgánica, puesto que el carácter interrumpido presupone que es una consonante. En cualquier caso, no cabe restringir las interrumpidas homorgánicas al subconjunto de las sonantes, como hace James W. Harris (1969).

Esta descripción refleja el patrón de alternancias más extendido en la mayoría de las variedades del español. No obstante, hay zonas en las que tiende a encontrarse la pronunciación oclusiva detrás de cualquier consonante. Quilis (1993, p. 221) documenta este tipo de alófono tras /s/, /l/, /r/, /ɲ/ y /ʎ/ —o algún subconjunto de ellos, según las zonas— en diversas regiones de México, Centroamérica, las Antillas, Colombia, Ecuador, Bolivia y Argentina (*vid. op. cit.* para más detalles). Carrasco, Hualde y Simonet (2012) lo confirman para Costa Rica, y lo comparan con los datos bibliográficos que manejan, que sitúan el mismo fenómeno en las zonas de la sierra de

Colombia, así como en el judeoespañol de Estambul. Estos autores sugieren que este hecho puede apuntar a una distribución similar en el español de los siglos XV y XVI. Consideran que el fenómeno comenzaría como una lenición de /b/, /d/, /g/ en contexto intervocálico y solo después empezaría a extenderse al contexto posconsonántico, con vacilaciones. Las variedades americanas en cuestión y la judeoespañola terminarían por rechazar la espirantización posconsonántica.

En cualquier caso, el estudio del español costarricense firmado por Carrasco *et al.* refleja que la alternancia no sigue una distribución tajante, sino que está sujeta a variación subsegmental (*i. e.* de detalle fonético) por distintas causas. En general, se encuentra una mayor constricción en inicio de palabra y en sílaba acentuada (en cambio, la posición inicial de palabra no tuvo efecto significativo en un grupo de control de madrileños). También se encuentra un efecto del contexto precedente, de modo que la /s/ suele favorecer una mayor constricción, seguida de las líquidas, las paravocales y las vocales. Además, se observa una interacción entre el contexto y el punto de articulación de la obstruyente laxa en cuestión, lo que corrobora la influencia de factores no solo prosódicos sino también coarticulatorios.

Volviendo a la distribución alofónica más generalizada (la que presenta oclusivas solo tras pausa o tras consonante interrumpida homorgánica), existe un problema de análisis que tiene que ver con la dirección del proceso.

Los autores de la tradición filológica española (*vid.* p. ej. Alarcos Llorach, 1965; Quilis, 1993) reconocen que los fonemas correspondientes no se pueden describir ni como oclusivos ni como aproximantes (ellos dicen fricativos), precisamente por incluir alófonos de ambos tipos; pero no dudan en optar por la transcripción oclusiva para la representación de /b/, /d/, /g/¹³². En cambio, anotan la palatal como /y/ (/j/ en la versión actual del AFI), en lugar de utilizar el símbolo /ɣ/ (/ʝ/), de modo que se pierde la generalización del comportamiento afín de los cuatro fonemas. Posiblemente no quisieron asemejarlos del todo por el hecho de que la palatal no presenta una oclusión simple sino una doble articulación —similar, en ese sentido, a una africada, aunque su segundo elemento suele ser una aproximante más que una fricativa, *vid.* Martínez Celdrán y Fernández Planas (2013)—. Sin embargo, desde una óptica más moderna, las africadas son un tipo más de oclusivas, solo que su distensión dura más tiempo, generalmente debido a la mayor superficie de contacto articulatorio (algo especialmente acusado en las consonantes palatales); y por ello se forma, durante la distensión, un estrecho canal en el que se produce el ruido de fricción (Martínez Celdrán y Fernández Planas, 2013)¹³³.

Quizás por paralelismo con los procesos diacrónicos, el análisis sincrónico ha asumido de manera habitual que el punto de partida era el oclusivo y de ahí se llegaba al espirante en los casos indicados. Es posible que este hecho también pueda atribuirse a la idea prescriptivista de que las hiperarticulaciones son realizaciones más próximas a la

¹³² Según el alfabeto de Navarro Tomás y la *Revista de Filología Española*, *RFE*, la velar sonora se transcribe con una /g/ de imprenta. En cambio, el AFI destina para esta cualidad una /g/ manuscrita.

¹³³ Desde el punto de vista fonológico, la tradición generativista propuso el rasgo [± distensión tardía] para dar cuenta de la diferencia entre las oclusivas simples y las africadas.

norma, mientras que las hipoarticulaciones se alejan de ella. Según esta visión, la forma subyacente debe ser, por tanto, la más fuerte, y sobre ella se aplica un debilitamiento, en lugar de partir de la forma débil y reforzarla. Además, hay quien defiende el carácter lenitivo del proceso apoyándose en el hecho de que existen variaciones en el detalle fonético de las realizaciones que reflejan un menor grado de constricción consonántica en las tareas más espontáneas (como muestran los datos de Carrasco *et al.*, 2012, para el grupo de Madrid).

James W. Harris (1969) fue el primer fonólogo generativista en formalizar el proceso mediante una Regla de Espirantización; es decir, toma la pronunciación oclusiva como punto de partida, como la más representativa del fonema, y aplica un debilitamiento en todos los contextos en los que aparecen las espirantes, a saber, (a) tras obstruyente o, en su defecto, si es una sonante, (b) tras (sonante) continua o, en su defecto, si es una sonante interrupta, (c) tras (sonante interrupta) que no comparta el mismo valor del rasgo [coronal] con respecto al segmento objeto del cambio. Esto queda formalizado en la regla (55).

(55) Regla de Espirantización (James W. Harris, 1969, p. 40)

$$\begin{bmatrix} + \text{obstr} \\ - \text{tense} \end{bmatrix} \rightarrow \begin{bmatrix} + \text{cont} \\ - \text{strid} \end{bmatrix} / \left\{ \begin{bmatrix} + \text{obstr} \\ + \text{cont} \end{bmatrix} \right\} \quad \frac{\quad}{<[\alpha \text{ cor}]>}$$

Otros autores después de Harris, y desde diversos modelos fonológicos, han mantenido la idea de tomar la oclusiva como punto de partida de un proceso que, por tanto, resulta clasificado como una instancia de lenición (*vid.* p. ej. Goldsmith, 2000 [1981], para un enfoque autosegmental; Parrell, 2014, desde un enfoque de fonología articulatoria y dinámica de tareas).

Parrell, en su tesis doctoral, asume que, aunque una menor duración lleva consigo una reducción en la magnitud gestual (pues no da tiempo a alcanzar el objetivo), en cambio una mayor duración no implica un refuerzo espacial que sobrepase el objetivo articulatorio del gesto. En la dinámica de tareas, los movimientos de los órganos implicados en los gestos articulatorios se modelan mediante un sistema de muelles que tienden a una posición de reposo. Para cada gesto, el reposo se alcanza cuando se consigue el objetivo articulatorio implicado por el propio gesto. Al activarse el gesto siguiente, se establece un nuevo objetivo y eso desencadena nuevos movimientos para lograr el reposo. Este modelo supone la existencia de una amortiguación crítica en el sistema de muelles, de modo que los articuladores se aproximan de manera asintótica al objetivo. De lo contrario, rebasarían el punto pretendido y comenzarían a oscilar a su alrededor con una amplitud decreciente en el tiempo. Pues bien, es esta aproximación asintótica la que justifica la afirmación de Parrell de que, por más hiperarticulación que se produzca, no se puede rebasar el objetivo inicialmente establecido por el gesto.

Por otro lado, se sabe que las oclusivas normalmente tienen un objetivo articulatorio que se sitúa más allá del punto de mero contacto. Es decir, si no fuera por la no interpenetrabilidad de la materia, los órganos activos alcanzarían un desplazamiento

mayor al que obtienen. En el caso de los labios, por ejemplo, esto se nota en que el labio superior y el inferior no simplemente entran en contacto, sino que se produce una compresión del tejido muscular, como refleja la siguiente cita de Löfqvist y Gracco (1997):

The lips were moving at high velocities at the instant of oral closure. As a consequence, the lower lip was continuing its upward movement after the closure had occurred. During this time, the upper lip often showed an upward movement that appeared to result from a mechanical interaction between the two lips. [...] This suggests that the virtual target for the lips in making the stop is a region of negative lip aperture—that is, to reach the virtual target the lips would move beyond each other (p. 891).

Teniendo esto en cuenta, Parrell (2014) realiza un modelado computacional que le permite reproducir los datos empíricos encontrados en cuanto a la alternancia entre realizaciones oclusivas y espirantes de las obstruyentes laxas, frente a la producción siempre oclusiva de las tensas (sordas). Para ello toma como ejemplo las bilabiales. En el caso de la sorda /p/, propone un objetivo articulatorio de una abertura de labios (LA, por sus siglas en inglés) negativa, es decir, que el objetivo se sitúa más allá del punto de mero contacto labial. En concreto, postula una LA de -2 mm; aunque reconoce que la cantidad puede ser arbitraria, es la que, tras un tanteo, le permite modelar resultados invariablemente oclusivos (para la /p/) incluso en condiciones de reducción temporal e hipoarticulación. Por ejemplo, una duración de 200 ms permite alcanzar el objetivo espacial de -2 mm; en cambio, aunque una duración de 80 ms lleva a una reducción de la magnitud obtenida ($LA = -0.8$ mm), se sigue produciendo un contacto, es decir, el sonido es oclusivo. Por su parte, para modelar la variación entre [b] y [β], postula un objetivo de $LA = -0.5$ mm. En estas circunstancias, solo en condiciones de gran duración (200 ms) se asegura la aparición de la oclusiva [b], con la consecución efectiva del objetivo propuesto. En cambio, con duraciones menores (p. ej. 80 ms), se obtiene una LA de 0.6 mm. La magnitud de la hipoarticulación es similar a la que se produce con la sorda (al articulador le falta en torno a 1.2 mm para alcanzar su objetivo en ambos casos), con la diferencia de que en el caso de la sorda eso la sigue situando en un rango de valores negativos (oclusión), mientras que la sonora se espirantiza.

Dados los resultados de este modelado, Parrell concluye que la forma subyacente de la obstruyente laxa debe ser una oclusiva (el objetivo es una LA negativa). Por otro lado —admite—, /p/ y /b/ también difieren en otros parámetros, como la especificación del gesto glotal, para conseguir la diferencia entre sordo y sonoro, o la duración intrínseca de los gestos articulatorios, menor en la /b/ que en la /p/ (recuérdese que en la fonología articulatoria los gestos tienen una especificación inherente de duración).

Ahora bien, podemos señalar algunos problemas en el enfoque de este estudio. Por un lado, al negar la posibilidad de que un gesto rebase su objetivo articulatorio, está negando la existencia de procesos de fortición que sabemos que existen (o al menos los deja sin explicar). Por ejemplo, en el § 4.3.4 veremos el proceso evolutivo que convirtió la vocal latina /i/ en una consonante /ʒ/ en castellano medieval. Evidentemente, no tiene sentido pensar que el objetivo articulatorio implícito en la /i/ ya era el de obtener una

constricción consonántica. Por el contrario, tenemos que dar cabida a la posibilidad de que un gesto se vea reforzado más allá de su propio objetivo siempre que se den las condiciones propicias, que pueden ser del tipo de las que se han detallado en el § 3.1.

Además, si una /b/ es intrínsecamente oclusiva pero también notablemente breve, la propia especificación fonológica del segmento impide que se logren ambos objetivos de manera simultánea. Es decir, se trata de características antagónicas, ya que, para obtener la magnitud pretendida se debe aumentar la duración (tanto que este es el resultado menos frecuente); en cambio, si la /b/ se mantiene en su rango de duración más habitual, nunca logra el contacto oclusivo.

Por último, al proponer la existencia de dos objetivos LA negativos pero de distinta magnitud (−2 mm para /p/ y −0.5 mm para /b/), se está implicando la existencia de dos categorías de oclusivas: oclusivas de primera y de segunda clase, si se me permite la expresión. Esta idea no encuentra apoyo en ninguna descripción tipológica de los diferentes modos de articulación.

Estas razones me llevan, en principio, a descartar la suposición de una oclusiva subyacente en el caso de la obstruyente laxa. Pero volvamos por un momento al formalismo SPE y a la propuesta de James W. Harris (1969) que hemos visto en (55). Para Lozano (2000 [1978], p. 325), no queda justificado que «las oclusivas sean las básicas [...]. Se podría asimismo suponer que /β, ð, ɣ/ son las básicas y explicar la alternancia por medio de una regla de Formación de Oclusiva tal como la siguiente»:

(56) Regla de Formación de Oclusiva

$$\left[\begin{array}{l} + \text{ obstr} \\ + \text{ sonoro} \end{array} \right] \rightarrow [- \text{ cont}] / \left\{ \begin{array}{l} \# \text{ —————} \\ \left[\begin{array}{l} - \text{ obstr} \\ - \text{ cont} \\ \alpha \text{ p. art.} \end{array} \right] \text{ —————} \\ \left[\alpha \text{ p. art.} \right] \end{array} \right\}$$

Esta regla sigue sin poder generalizarse a todas las consonantes interrumpidas homorgánicas, pero es mejor que la propuesta de Harris, pues tiene más sentido formular la regla en esta dirección. En efecto, eso permite generalizar el cambio a un contexto con una única descripción, descontando el caso de ir precedido por pausa, al que volveremos enseguida. Este tipo de descripción es más simple que la de Harris, ya que la Regla de Espirantización de este se formula sobre múltiples contextos (todos aquellos en los que no se produce la oclusivización), que no presentan una característica común que los unifique (no forman una clase natural). En definitiva, los contextos de la Regla de Espirantización responden a los contextos de la condición alternativa (*elsewhere condition*, en inglés, *vid.* Kiparsky, 1973).

Por su parte, Gómez Asencio (2000 [1994]) advierte sobre la inconveniencia de «identificar el fonema con una de sus variantes» (p. 162), especialmente si estas son «las realizaciones menos frecuentes, menos *normales*, de los respectivos fonemas» (*ibid.*). En efecto, los procesos cognitivos conducentes a la creación de prototipos —en definitiva, las categorías abstractas o fonemas— tienen especialmente en cuenta la

frecuencia de los distintos elementos que se categorizan («the most frequently occurring variants [...] constitute the prototype for the category», Bybee, 2001, p. 42)¹³⁴.

Por estas consideraciones, propongo la utilización de los símbolos de las aproximantes espirantes, en lugar de las oclusivas, para la representación fonológica de las obstruyentes laxas (sonoras): /β/, /ð/, /j/, /ɣ/. Es decir, que utilizaré estos símbolos no solo para la transcripción de los alófonos aproximantes (entre corchetes) sino también para la de los fonemas (entre barras).

A continuación se exponen otros datos, procedentes de la adquisición del español como L1, que avalan esta postura. Macken y Barton (1980) encontraron que, en el habla infantil, las obstruyentes laxas se pronunciaban como aproximantes en muchos casos, incluso cuando aparecían en posición inicial de enunciado y cuando el tempo de habla era lento, y a pesar de la tendencia general de los niños a la oclusivización. Estas obstruyentes fueron pronunciadas como aproximantes en un 45.7% de los casos por los niños de 2 años, pero solo en un 31.4% de los casos por los niños de 4 años. Esta progresión se puede interpretar como que lo que aprenden los niños es una regla de formación de oclusiva, y no una regla de espirantización.

En definitiva, consideraremos que la forma subyacente es una aproximante espirante, que se realiza como oclusiva en determinadas circunstancias, y no al revés. La dirección del cambio es relevante, porque además tiene más sentido postular que se produce un refuerzo en posición inicial de enunciado —la posición prosódica más fuerte de todas las posibles— en lugar de suponer un debilitamiento en posición inicial de todos los demás constituyentes prosódicos, por debajo del nivel U.

Ahora bien, antes se ha dicho que la descripción de uno de los contextos de cambio como «tras consonante interrupta homorgánica» permitía unificar los casos tras nasal (para las obstruyentes de todos los puntos de articulación) con los casos tras lateral (solo para [d] y [t̪]), y los casos marginales discutidos más arriba, precedidos de oclusiva tensa. Sin embargo, el contexto tras pausa sigue obligando a formular una disyunción en la aplicación de la regla —*vid.* p. ej. la regla (56)—. Quiero proponer que estos cambios, que siempre se han tratado de forma conjunta, en realidad responden a dos fenómenos totalmente distintos y, como tales, deberían formularse por separado, aunque den lugar al mismo resultado.

El contexto tras pausa corresponde al inicio de enunciado fonológico; es decir, más que de un contexto se trata propiamente de una posición prosódica. Es aquí donde realmente podemos hablar de fortición. En cambio, los contextos tras consonante interrupta homorgánica se caracterizan en referencia al segmento precedente y, por tanto, los cambios que ocurren en esos casos deben analizarse como asimilación de la obstruyente al carácter interrupto de dicho segmento previo.

Por tanto, las realizaciones oclusivas de los fonemas obstruyentes laxos se pueden deber a la aplicación de una de las dos reglas siguientes.

¹³⁴ Para una visión diferente, *vid.* Boersma (2006); Johnson, Flemming, y Wright (1993).

(57) Regla de fortición de obstruyentes laxas

$$\begin{bmatrix} + \text{obstruyente} \\ - \text{tenso} \end{bmatrix} \rightarrow [+ \text{interrupto}] / [_{\text{U}} _]$$

(58) Regla de asimilación de obstruyentes laxas

$$\begin{bmatrix} + \text{obstruyente} \\ - \text{tenso} \end{bmatrix} \rightarrow [+ \text{interrupto}] / \left[\begin{array}{c} + \text{interrupto} \\ \alpha \text{ labial} \\ \{ \beta \text{ coronal} \} \\ \gamma \text{ dorsal} \end{array} \right] \quad \overline{\begin{bmatrix} \alpha \text{ labial} \\ \beta \text{ coronal} \\ \gamma \text{ dorsal} \end{bmatrix}}$$

Como puede comprobarse, el resultado es el mismo en ambas reglas: la oclusivización de la espirante. Sin embargo, el desencadenante es posicional en un caso y contextual en otro.

En ellas se ha seguido un formalismo que recuerda al de SPE por mantener el paralelismo con las dos propuestas que se han analizado (la de James W. Harris, 1969, y la de Lozano, 2000 [1798]), y de ese modo subrayar las semejanzas y las diferencias entre dichas propuestas y la alternativa que aquí se defiende. Sin embargo, eso no significa que necesariamente esa sea la mejor formalización posible. De hecho, por ejemplo, la expresión de un proceso de asimilación se captaría de una forma más natural (menos arbitraria) en un modelo autosegmental.

Se ha adoptado el rasgo $[\pm \text{interrupto}]$ en lugar de $[\pm \text{continuo}]$ y, por tanto, un cambio de polaridad en el valor del rasgo, puesto que el valor no marcado de los rasgos es el que típicamente aflora en casos de neutralización (Gussenhoven y Jacobs, 2011, § 5.5) y, efectivamente, las obstruyentes tensas interrumpidas (oclusivas sordas) y las obstruyentes laxas (aproximantes espirantes) se neutralizan en coda silábica, y el resultado más habitual es $[- \text{tenso}, - \text{interrupto}]$, *vid.* § 4.1.

En la regla (57), la notación $[_{\text{U}} _]$ expresa inicio de enunciado fonológico. En la regla (58), la disyunción marcada por la llave establece la prioridad de los rasgos en la determinación de la homorganicidad. En primer lugar, es necesario que coincida el valor del rasgo [labial] (presente/ausente, si aceptamos el carácter monovalente de los rasgos de punto de articulación). Una vez discriminados los segmentos labiales de los no labiales, entre estos últimos es necesario que además coincida al menos uno de los dos rasgos, [coronal] o [dorsal], para que se cumpla el grado de homorganicidad requerido por esta regla. Recuérdese que los segmentos dentoalveolares son coroneles, y los velares dorsales, pero que los palatales son a la vez coroneles y dorsales (*vid.* nota 131 en la p. 156) y en ese caso basta que coincida uno de los dos rasgos con la interrumpida precedente para que se aplique la regla. No es necesaria, por tanto, una homorganicidad absoluta en el caso de las palatales, sino que es suficiente con una homorganicidad simple.

Por último, hay otro contexto más en el que puede darse la oclusivización de espirantes, y que no queda recogido en las reglas anteriores. Se trata de secuencias en las que la obstruyente laxa aparece precedida por otra obstruyente laxa (homorgánica). Por ejemplo, tenemos una doble /β/ en palabras como *obvio*, *subvención*, o *subversión*. También pueden aparecer secuencias similares a través de frontera de palabra:

(59) Es una **web** boliviana.

(60) Madrid **donó** más sangre que otras comunidades.

(61) Quieren mejorar el **airbag** gastando poco.

Estos ejemplos presentan algunas características interesantes. Por un lado, el primer segmento de la secuencia no es [+ interrupto]. De hecho, las obstruyentes laxas no están especificadas como [+ interrupto] ni como [– interrupto], precisamente porque presentan un alófono con un valor y un alófono con el otro. Por tanto, no pueden desencadenar la regla (58).

Sin embargo, la pronunciación de estas palabras y frases está sujeta a variación estilística. Aunque lo más habitual es que la secuencia se pronuncie como una aproximante, en una pronunciación enfática también sería aceptable la oclusiva. Mientras que un oyente nativo percibiría como rara (fuera de contexto) una pronunciación oclusiva en una palabra como, por ejemplo, *Cuba* ['kuba], en cambio no le sonaría raro *obvio* ['o(b)b̥io]: solo le sonaría inusualmente enfático. Es decir, que en las alternancias analizadas hasta ahora hay una distribución complementaria entre los alófonos oclusivo y espirante, pero aquí hay una distribución equivalente, porque simplemente depende de una variación estilística.

Por otro lado, el caso precedido por obstruyente laxa también difiere de los casos recogidos en la regla (58) en que exige una homorganicidad absoluta: nótese que la oclusivización no se produce en palabras como *adyacente* [að̥ja'tɛnt̪e], *coadyuvante* [koað̥ju'βant̪e], a pesar de que la /ð/ y la /j/ comparten el rasgo [coronal].

Por estas razones, parece que la posible aparición estilística (enfática) de la oclusiva en casos como el de *obvio* debe responder a otra regla diferente, que seguramente debe hacer mención al hecho de que se trata de una secuencia de dos segmentos (obstruyentes laxos) totalmente idénticos¹³⁵.

4.3.2. Oclusivización de aproximantes laterales

En diferentes zonas de las penínsulas ibérica e itálica, la /l/ inicial de palabra del latín experimentó un refuerzo, primero en términos de duración, [l:], y más tarde reinterpretado en términos de magnitud, de modo que en algunos casos dio la palatal [ʎ]

¹³⁵ Probablemente no podamos hablar de verdaderas geminadas, porque la secuencia se da a través de frontera de morfema. En latín la palabra *obvio* estaba formada con un prefijo *ob-*, si bien es cierto que esa consciencia se ha perdido en español. En cambio, la composicionalidad es transparente en *subvención* y *subversión* (si no desde el punto de vista semántico, sí al menos desde el formal, puesto que es reconocible el prefijo *sub-*). Quizá la propia consciencia morfológica puede ser incluso un factor que propicie la pronunciación enfática, precisamente para recalcar ese origen derivado. Por otro lado, los casos en juntura de palabra de los ejemplos (59)–(61) se dan, obviamente, a través de frontera de morfema.

(*vid.* § 4.3.3 para una mayor discusión de los posibles condicionantes de este cambio), pero en otros dio lugar a oclusivas retroflejas, [d̠], [t̠], o incluso [t̠ʂ]¹³⁶ (Lapesa, 1981, pp. 96-97). En la zona de dominio astur-leonés encontramos, por ejemplo, pronunciaciones como [ˈduna] ‘luna’, [ˈt̠soβu] ‘lobo’, [ˈt̠siŋɣa] ‘lengua’ (*ibid.*).

García Arias (2003, p. 201 y ss.) confirma la existencia de variantes [d̠], [d̠ʂ] y [t̠ʂ] en la zona sur-occidental de Asturias, que corresponden tanto a /l/ inicial, al igual que en los ejemplos anteriormente mencionados, como a /l:/ geminada del latín, y así por ejemplo palabras como *cebolla* o *muelle* también muestran producciones retroflejas de este tipo. Además, basándose en las grafías «r» y «ri», con las que se transcriben en unos villancicos de los siglos XVII y XVIII, argumenta que la pronunciación en esa época aún debía de ser sonora, y sitúa el proceso de ensordecimiento que daría lugar a la [t̠ʂ] entre los siglos XVIII y XIX.

Por otro lado, el resultado [t̠ʂ] —denominado con frecuencia *che vaqueira*— se puede terminar convirtiendo en [t̠ʃ], como reflejan palabras del tipo *chorimingar* ‘lloriquear’ o *chargato* ‘lagarto’ (*op. cit.*, p. 217).

4.3.3. Palatalización de sonantes interrumpidas coronales /l/ y /n/

A grandes rasgos, la evolución de las obstruyentes latinas equiparó los resultados en posición inicial de palabra y en inicio de sílaba interior de palabra en contexto intersonántico. Por ejemplo, dio lugar a oclusivas sordas en ataque intersonántico, procedentes de la degeminación (§ 4.2.2), y respetó las oclusivas sordas originales del latín en inicio de palabra. Del mismo modo, surgieron sonoras (primero oclusivas y luego espirantes) en contexto intersonántico, provenientes de la sonorización (y espirantización) de las oclusivas sordas simples del latín (§ 4.2.1 y 4.2.3), y se mantuvieron las sonoras originales en posición inicial de palabra.

Sin embargo, el mantenimiento de tal paralelismo entre ambas posiciones en el caso de las sonantes requería algunas diferencias de tratamiento. La simplificación de las geminadas /l:/ y /n:/ del latín en sus respectivas simples /l/ y /n/ en posición de ataque intersonántico supone una confluencia con la /l/ y /n/ originales del latín en esa misma posición. Las simples originales no tienen la opción de modificarse para mantener la distinción fonológica: no pueden recurrir a la sonorización (puesto que ya son sonoras) ni a la espirantización (porque no son obstruyentes).

Cada uno de los primitivos romances peninsulares resolvió este problema de una manera diferente. Por ejemplo, en el dominio gallego-portugués se procedió a la elisión de las simples intervocálicas: *vontade* ‘voluntad’, *lúa* ‘luna’. Eso permitió mantener la diferencia con respecto a las antiguas geminadas, ahora simplificadas (lat. *capillu(m)* > gall. *cabelo* ‘cabello’, lat. *annu(m)* > gall. *ano* ‘año’), con la consiguiente equiparación de su resultado a las simples iniciales de palabra.

El resto de los romances, incluido el castellano, conservaron generalmente la simple intervocálica sin modificaciones. Para mantener la oposición, la solución más extendida fue la palatalización de las geminadas. Las geminadas se caracterizaban por una mayor

¹³⁶ Recuérdesse que una africada es un tipo de oclusiva con distensión alargada, como defienden, por ejemplo, Martínez Celdrán y Fernández Planas (2013). *Vid.* nota 133 en la p. 158.

duración del gesto lingual de contacto en la zona dentoalveolar y, en algún momento, la necesidad de preservar la oposición pudo propiciar que se reinterpretara la mayor duración como una mayor magnitud gestual, esto es, un mayor grado de contacto entre los órganos articulatorios. La extensión de la zona de contacto daría lugar a una articulación alveolopalatal [ʎ] o [ɲ] (lat. *capillu(m)* > cast. *cabello*, lat. *annu(m)* > cast. *año*). Nótese que tal reinterpretación fue solo posible en el caso de las coronales, ya que un refuerzo sustantivo del gesto lingual podía ensanchar la zona de contacto articulatorio y, al fin y al cabo, las palatales tienen una especificación coronal (a la vez que dorsal). En cambio, la geminada bilabial /m:/ se simplificó y se unificó con su correspondiente simple /m/ (p. ej. lat. *communicare* > cast. *comunicar*).

Como se ha dicho, las soluciones palatalizadas /ʎ/ y /ɲ/ se extendieron por toda la península, excepto en el dominio gallego-portugués, y también a excepción del habla de Bielsa (Huesca), donde se mantuvo la pronunciación de una doble consonante [l.l] [n.n], como en *bellota*, *ninno*, *penna* (Lapesa, 1981, § 120₃).

No está claro si se puede considerar esta palatalización un caso de fortición, puesto que aparece en una posición que, precisamente, suele desencadenar leniciones. Tampoco parece que se trate de una asimilación, dado que no hay ningún elemento palatal en el contexto. Es evidente que el intento de preservación de la oposición fonológica entre las simples y las geminadas proporciona una explicación, de corte estructuralista-funcionalista, a la mera existencia de un cambio. Ahora bien, el porqué de la dirección concreta adoptada por ese cambio debe achacarse más bien a las peculiaridades de las geminadas (mayor duración y la posibilidad de reinterpretar esta como mayor magnitud). También pudo afectar, como factor coadyuvante, el desarrollo independiente de pronunciaciones palatales en esa posición, provenientes de otros orígenes (*l* + *yod*, *n* + *yod*), estos sí explicables por asimilación.

Algunos romances también adoptaron pronunciaciones palatales en inicio de palabra, correspondientes a las sonantes simples /l/ y /n/ del latín (p. ej. astur-leonés *llobu* ‘lobo’, *ñariz* ‘nariz’¹³⁷), o a veces solo en la serie lateral (catalán *llop* ‘lobo’). La /l/ simple llegó a palatalizar incluso, aunque con menor intensidad, en interior de palabra en astur-leonés: *allargar*, *baillar* (Lapesa, 1981, p. 488).

También se encuentran algunas palabras en castellano, hacia el siglo XIII, como *llenguaje* (*op. cit.*, p. 178); en aragonés, entre los siglos XI y XIV, como *lliçença*; y se encuentran algunos restos en el andaluz de Ronda, en el siglo XVIII: *llucero*, *llengua*, *llance*, o en posición interior: *cullebra*, *rellatar* (*op. cit.*, p. 512).

Su aparición en topónimos (Lloreda, San Llorente, Los Llamosos, etc.) demuestra que fue un fenómeno de mucho arraigo, puesto que la onomástica es siempre un tipo de léxico más conservador de cara a los cambios. Esta difusión es reconocida por Lapesa, quien, sin embargo, señala que se trataba de un fenómeno «considerado vulgar fuera del dominio catalán, por lo que en el resto de España permaneció fuera del uso escrito y fue relegado al claramente dialectal» (*op. cit.*, p. 179).

¹³⁷ En cualquier caso, la palatalización de /n/ inicial es menos sistemática que la de /l/ en las hablas leonesas.

En los casos en posición inicial de vocablo, no se puede achacar el aumento de magnitud a una mayor duración propia de la geminada, puesto que no hay tal geminada. Sin embargo, el refuerzo temporal sí es característico de la posición prosódica que ocupa (inicial de palabra). La solución adoptada debe explicarse, en cualquier caso, como un intento de mantener la similitud entre los resultados de las geminadas (intervocálicas) y las simples iniciales de palabra, en paralelo con el tratamiento dado a las obstruyentes. Además, pudo influir, como factor coadyuvante, la aparición de palatales iniciales de palabra por motivos independientes, como la evolución de los grupos /kl/, /pl/ y /fl/ (el primero de ellos, explicable por asimilación).

En definitiva, estos cambios por palatalización de las sonantes interrumpidas coronales /l/ y /n/ (o de sus correspondientes geminadas) no son tan nítidamente adscribibles a un tipo de cambio —fortición o asimilación— como otros que hemos visto hasta ahora, y puede que no sean genuinamente ni lo uno ni lo otro. La duración de las geminadas intervocálicas o el refuerzo temporal de las simples iniciales de palabra puede hacer pensar en una fortición, idea también avalada por los resultados oclusivos retroflejos (§ 4.3.2). Pero todos estos cambios forman parte de un intrincado juego de relaciones estructurales entre posiciones y resultados de distintos elementos del sistema, que deben de ser los verdaderos motores del cambio. Por otro lado, además, se relacionan con otros resultados palatales, de origen claramente asimilatorio, de modo que las mencionadas presiones estructurales pudieron propiciar una confluencia de soluciones para cambios de muy distintos orígenes.

4.3.4. Fricativización de la espirante palatal

En latín clásico solo se consideraban diptongo las combinaciones vocálicas *ae*, *oe*, *au*, y también la combinación *eu*, procedente de préstamos del griego. Todas las demás secuencias formaban hiato. Sin embargo, en el paso del latín al castellano, se desarrolló una fuerte tendencia antihiática, que favoreció el desarrollo de diptongos siempre que interviniera una vocal alta que no fuera núcleo de sílaba acentuada. Así, por ejemplo, la palabra latina *iuvene(m)* ‘joven’ pasó de pronunciarse con un hiato [i.ˈu] a presentar un diptongo [ˈju].

El elemento paravocálico [j] ocuparía entonces una posición prenuclear en una cima compleja. Sin embargo, la tendencia universal a preferir sílabas con ataque desencadenaría un reanálisis de la posición estructural ocupada por dicho elemento palatal. Hay que tener en cuenta que para poder ocupar una posición de ataque es necesaria una especificación [+consonántico], de modo que podemos suponer una sustitución del fonema vocálico /i/ ¹³⁸ en posición paranuclear por el fonema

¹³⁸ Podríamos hablar de /j/, aunque, a diferencia del modelo *SPE*, a día de hoy es mayoritaria la visión de que el carácter no silábico (no nuclear) de un elemento no se expresa como un rasgo [–silábico] inherente al fonema, sino como producto de la asociación entre el segmento y una determinada posición en la estructura silábica. No obstante, es cierto, a menudo se sigue usando la especificación [–silábico] o una notación del tipo /j/ (entre barras) para indicar de una manera más abreviada o directa que el conocimiento fonológico acerca de ese segmento incluye sus propiedades inherentes, así como su adscripción a una determinada posición estructural en el seno de la sílaba.

consonántico /j/ en posición de ataque, ya que esa es la consonante más parecida posible a la vocal original¹³⁹.

Sea como sea, todos los cambios mencionados son una cuestión de estructura silábica. Ahora bien, llegó un momento en que ese elemento consonántico palatal se reforzó y dio lugar a una fricativa sibilante prepalatal sonora: /ʒ/, p. ej. cast. medieval ['ʒoβen]. Este cambio supone un refuerzo sustantivo de los gestos articulatorios de dicha consonante, y se debe a la propia posición prosódica que ocupa: la posición inicial de palabra, de donde se extendería a otros ataques silábicos en interior de palabra (p. ej. [maʒeʃ'tað] 'majestad', del latín *maiestatem*).

Este resultado fricativo /ʒ/ a partir de un origen aproximante (o incluso vocálico, si nos remontamos aún más atrás) es un caso claro de refuerzo si tenemos en cuenta que el único desencadenante posible del cambio es la propia posición prosódica que ocupa el segmento. No obstante, el castellano medieval también presentaba /ʒ/ en otros casos que, por el contrario, deben ser analizados como asimilación (p. ej. los causados por el contacto de una *yod* con determinadas consonantes).

De manera independiente (poligenética) con respecto a este fonema medieval, el español moderno presenta, en algunas variedades dialectales, un segmento fricativo prepalatal sonoro [ʒ] como realización del resultado del yeísmo.

El yeísmo es el nombre que recibe el fenómeno de neutralización de la /ʎ/ (aproximante lateral palatal sonora) y la /j/ (aproximante espirante palatal sonora). Dada la pérdida de distintividad en todas las posiciones y contextos posibles, más que de neutralización, cabe hablar ya de desfonologización, puesto que se produce la pérdida del fonema /ʎ/ a favor de /j/. Entre las causas del fenómeno se encuentra el escaso rendimiento funcional de la oposición, pues hay muy pocos pares mínimos reales. Además, existen factores articulatorios que favorecen la pronunciación central frente a la lateral en casos de menor tensión articulatoria. En efecto, dada la desigual fuerza de la musculatura de los laterales y el centro de la lengua, resulta difícil mantener dichos laterales suspendidos en el aire, sin contacto con los bordes del paladar, a menos que se ejerza una gran tensión articulatoria. En el § 3.6.1 se vio que, incluso con la /l/, es habitual apoyar uno de los laterales, sobre todo en las posiciones débiles, y suele ser precisamente el lateral sobre el que disponemos de un menor control (el lateral izquierdo en las personas diestras, y viceversa). Podría pensarse, a modo de analogía, que también es habitual que nos apoyemos con una mano u otra parte del cuerpo en algún objeto cercano cuando sentimos una falta de equilibrio. Pues bien, esto explica que la consonante central /j/ sea preferida sobre la lateral /ʎ/. Por último, parece que hay un grado notable de similitud acústica, que puede estar en la base de una dificultad para percibir adecuadamente la diferencia, que a su vez justificaría la reorganización de las categorías fonológicas por parte de los oyentes. Para una revisión actualizada de todos estos factores, *vid.* Rost Bagudanch (2014) y la bibliografía allí citada.

El yeísmo es un fenómeno en progresiva extensión que, a día de hoy, afecta casi a la totalidad de los territorios hispanohablantes. Son excepción algunos núcleos del tercio norte de la península ibérica, especialmente zonas rurales de Castilla y León y Asturias.

¹³⁹ Lo mismo ocurre a día de hoy en español (*vid.* § 2.1.1.2).

También se mantiene la distinción en zonas de contacto con el catalán, ya que en esta lengua aún se conserva la diferencia fonológica. En América también son yeístas casi todos los hablantes, excepto en las tierras altas de los países andinos (Colombia, Ecuador, Perú); y en Bolivia y Paraguay, por influencia del sustrato guaraní, que favorece una pronunciación reforzada de la palatal central, y permite así aumentar el margen de seguridad entre ambos fonemas. *Vid.* Rost Bagudanch (2011) y Quilis (1993, p. 316 y ss.) para una descripción detallada de la distribución geográfica del fenómeno.

Un subconjunto de las zonas en las que se da el yeísmo presenta, en lugar de una solución /j/, una evolución de la espirante a un segmento fricativo /ʒ/ ¹⁴⁰. Tal pronunciación es característica de la variedad porteña, por ejemplo, y se extiende por Uruguay y la mayor parte de Argentina (excepto el centro-oeste y el noroeste del país, así como las zonas limítrofes con Paraguay). En España, también se encuentran pronunciaciones fricativas en algunas zonas de Andalucía, Extremadura y Castilla-La Mancha.

Como argumentan Martínez Celdrán y Fernández Planas (2013), la pronunciación [ʒ] queda bien caracterizada con la denominación de fricativa, y no es necesario describirla como *rehilada*, término empleado por Navarro Tomás (1996 [1918]), quien hubo de recurrir a esta estrategia para diferenciarla de la pronunciación que hoy sabemos que es aproximante ([j]), pero que él consideraba también fricativa ¹⁴¹. Del mismo modo, el fenómeno descrito es, simplemente, una fricativización, aunque a menudo se sigue hablando de rehilamiento por tradición. Desde una perspectiva actual, parece preferible hablar de yeísmo fricativo frente a yeísmo espirante.

4.3.5. Ensordecimiento pasivo (II)

En las variantes americanas con resultado fricativo [ʒ] del yeísmo (§ 4.3.4), este segmento tiende a ensordecerse y dar como resultado [ʒ̥] ¹⁴². Este ensordecimiento es especialmente habitual en posición inicial de enunciado, no tanto en posición inicial de palabra, y menos en posición inicial de sílaba (interior de palabra). Por ejemplo, afecta con mucha frecuencia al pronombre personal de 1 Sg, que además aparece a menudo en inicio de enunciado, de modo que no es extraño oír a un argentino o a un uruguayo referirse a sí mismo como [ʒ̥o].

Al igual que el ensordecimiento típico de las codas (*vid.* § 4.1.4), este ensordecimiento que ahora nos ocupa también tiene un origen pasivo, aunque en este caso se produce en inicios de constituyente, en lugar de en finales. Como se explicó en el § 1, es posible llegar a un mismo resultado mediante procesos o circunstancias diferentes.

En el modelo aerodinámico de Westbury y Keating (1986) se aprecia lo siguiente:

¹⁴⁰ Para los hablantes que nunca pronuncian [j], proponer un fonema /j/ resulta poco ajustado a su realidad psicológica. Por eso se propone aquí que el símbolo de la fricativa (/ʒ/) se use no solo para la transcripción fonética sino también para la fonológica, pues este es su verdadero prototipo.

¹⁴¹ En ese mismo sentido, *vid.* también Barbón Rodríguez (1975, 1978), *apud* Quilis (1993).

¹⁴² Se ha optado por transcribir [ʒ̥], con el diacrítico, en lugar de [ʃ], para poner de relieve su relación con la forma subyacente, que es la sonora /ʒ/ (*vid.* § 4.3.4 y, en particular, la nota 140). La variante ensordecida, al menos a día de hoy, aún no se ha generalizado, sino que sigue existiendo variación sincrónica. Quizá algún día llegue a tomar carta de naturaleza y entonces sí deberá transcribirse como /ʃ/.

air pressure below the glottis remains generally high and stable over the middle segments of an utterance, but not over beginning (or ending) segments. Rather, in those respective environments, subglottal pressure rises above and falls toward the air pressure exerted by the atmosphere [...] in a generally linear fashion. Since the incidence of voicing depends in large part upon the difference between subglottal and supraglottal pressures, stop voicing should be more likely utterance-medially than initially or finally, simply because that pressure difference in the former environment tends to be somewhat greater than in the latter environments (p. 153).

Más tarde (p. 159) añaden que, incluso en las lenguas en las que se mantiene la oposición de sonoridad en posición inicial de palabra, el habla de los niños que están adquiriendo su L1 manifiesta una clara preferencia por el ensordecimiento en inicio de enunciado.

En el § 4.1.4 se vio que el ensordecimiento pasivo en coda viene propiciado, en un primer momento, por el descenso de la P_s (y, consiguientemente, de la P_t) en posición final de enunciado, de donde se puede extender a la posición final de otros constituyentes prosódicos de menor nivel. Pues bien, del mismo modo, el ensordecimiento pasivo en posición de ataque se produce, en primera instancia, cuando coincide con un inicio de enunciado, y de ahí se extiende a los inicios de otros constituyentes menores. Esta extensión debe entenderse de dos maneras. Desde el punto de vista diacrónico, el cambio se producirá antes en las posiciones más propicias y, más tarde en la evolución de la lengua, puede llegar a afectar a otras posiciones. Desde el punto de vista sincrónico, dada una situación en que el fenómeno puede afectar a varias posiciones, se da una relación implicativa tal que, si se puede producir el ensordecimiento en la frontera de un determinado tipo de constituyente, necesariamente se podrá producir en la de los constituyentes mayores, y lo hará en estos con más frecuencia que en los menores. En cualquier caso, hay que destacar que, frente a la motivación natural del ensordecimiento en las fronteras de enunciado, el ensordecimiento en otros constituyentes es una característica más restringida, específica de las lenguas que lo presentan (*vid.* nota 112 en la p. 142).

Dadas las exigentes condiciones aerodinámicas que permiten la vibración de los pliegues vocales, la sonorización es, ya de entrada, antagónica con la producción de obstruyentes¹⁴³. De hecho, los datos tipológicos muestran que las obstruyentes sordas son menos marcadas que las sonoras (Ladefoged y Maddieson, 1996; Mielke, 2009). Las obstruyentes se caracterizan por la producción de un obstáculo a la salida del aire, suficiente para que aumente la P_0 . Por otro lado, la vibración laríngea requiere una P_t

¹⁴³ El mismo antagonismo, por motivos aerodinámicos, se produce en las vibrantes. En diversas variedades americanas (p. ej. en Centroamérica; islas del Caribe; zonas interiores de Colombia y Ecuador; Bolivia, Paraguay y Chile, *vid.* Zamora Munné y Guitart, 1982, pp. 182-183) es habitual encontrar realizaciones ensordecidas de la rótica, p. ej. [ɾ̥], que pueden llegar a asibilarse ([ɾ̥ʰ]). Como explica Solé (2002), el aumento de la P_0 durante las fases de cierre de una vibrante puede impedir que se alcance la P_t necesaria para la vibración laríngea, y eso provoca el ensordecimiento parcial del segmento, durante dichas fases de cierre. Además, las vibrantes sordas se caracterizan por un mayor flujo de aire translingual, y a consecuencia de ello se puede generar un ruido de fricción en la zona de constricción.

positiva (es decir, que la P_o no debe superar a la P_s). Teniendo en cuenta que los gestos articulatorios presentan una mayor duración en los inicios de constituyente prosódico (por refuerzo temporal), se entiende que estas posiciones fuertes favorezcan una mayor acumulación de P_o como consecuencia de la prolongada obstrucción, y eso contribuye al ensordecimiento pasivo.

Este mismo ensordecimiento afectó a las sibilantes sonoras del español medieval, que acabaron neutralizándose con sus correspondientes sordas, y desfonologizándose en su paso al español clásico. No obstante, *vid.* Lloyd (1993, p. 426 y ss.) para una posible influencia del adstrato vasco durante la formación del primitivo romance castellano, que podría haber dado inicio al cambio desde mucho antes de lo que se había pensado, aunque no llegara a consolidarse hasta el siglo XVI.

5. Un corpus para el estudio del refuerzo subsegmental en español (metodología)

En el capítulo anterior se han descrito los procesos de fortición (segmental) más destacados del español. A partir de este capítulo nos centraremos en el estudio del refuerzo consonántico desde una perspectiva subsegmental o de detalle fonético, de una manera comparable a los trabajos sobre otras lenguas que se revisaron en el capítulo 3. En aquel momento se apuntó que, aunque existen tendencias similares entre las distintas lenguas investigadas, también es cierto que se observan particularidades de cada lengua en el modo o en el grado en que se modifica el comportamiento espaciotemporal de los gestos articulatorios como consecuencia del efecto de la posición o la prominencia prosódicas. Tiene sentido, por tanto, preguntarse cómo se manifiesta el refuerzo articulatorio en español.

El estudio que presento aquí se circunscribe a los fenómenos de refuerzo que afectan a una serie de consonantes del español en el seno de unidades léxicas. No se incluye ningún compuesto sintagmático ni ningún adverbio en *-mente* (que serían unidades léxicas correspondientes a dos palabras fonológicas, *vid.* § 2.1.3.1). Por tanto, para todas las palabras estudiadas, los límites de la palabra morfosintáctica coinciden con los límites de una misma y única palabra fonológica (ω).

Resulta interesante investigar el refuerzo provocado por los niveles prosódicos léxicos (*i. e.* iguales o inferiores a ω) por distintos motivos. Por un lado, las marcas fonéticas del refuerzo en estos niveles son una ayuda importante al reconocimiento de los distintos constituyentes prosódicos inferiores, que, a su vez, juegan un papel decisivo en el acceso al léxico (§ 2.3.3). Por otro lado, dado el carácter acumulativo del refuerzo y que hay algunas marcas que, en algunas lenguas, no afloran de manera significativa hasta niveles superiores, es interesante conocer qué tipos de fenómenos se manifiestan ya desde los primeros niveles de la jerarquía prosódica en español. Además, tal y como se ha argumentado en el § 4, el refuerzo que acompaña a los niveles prosódicos léxicos es susceptible, llegado el caso, de una hipotética consolidación como alternancia o cambio segmental, que afectaría a la estructura fonológica de la lengua.

El objetivo, por tanto, es describir las posibles modificaciones en la velocidad, duración y magnitud de los distintos gestos articulatorios según las consonantes aparezcan en posición inicial de sílaba —interior de palabra— (σ) o inicial de palabra (ω), y según estén incluidas en una sílaba acentuada ($^1\sigma$) o no acentuada (σ)¹⁴⁴. Se ha excluido del estudio la unidad pie (Σ) por diversos motivos. En el § 2.2.1 se vio que hay autores como B. Hayes (1995) que defienden que el pie (en concreto, el troqueo moraico) debe de jugar un papel importante en la asignación del acento primario en español, ya que la posición de este es sensible al peso silábico. Sin embargo, también se

¹⁴⁴ En la formalización gráfica, he optado por un punto suscrito al símbolo de la sílaba para explicitar el carácter no acentuado. A la hora de escoger esta representación, me he inspirado en las parrillas métricas, donde se suele alternar entre un asterisco (*) para las sílabas prominentes y un punto (·) para las no prominentes.

argumentó que existe un número suficiente de idiosincrasias como para defender que el acento primario en español (o, mejor dicho, su posición) no se puede predecir por regla, sino que es una característica ligada al léxico. En cambio, la localización de los posibles acentos secundarios sí se sigue automáticamente de la ubicación del primario, pero en este caso la asignación no parece sensible al peso, por lo que resulta más razonable proponer un algoritmo basado en el troqueo silábico. Esta disparidad y el hecho de que el acento primario sea una propiedad de cada palabra, que se debe aprender de memoria, constituyen un primer motivo para descartar el estudio de los posibles efectos del pie sobre el refuerzo consonántico. Por otro lado, además, el pie no juega ningún papel en el acceso al léxico en español (*vid.* § 2.3.3). Por último, como se vio en el capítulo 3, el pie no se ha contemplado en las investigaciones sobre el refuerzo subsegmental ni siquiera en lenguas como el inglés, a pesar de que la importancia fonológica (y psicolingüística) del pie en esta lengua es indiscutible.

Este capítulo describe la metodología de la confección y del análisis fonético del corpus de habla en el que se basa este estudio, que incluye datos articulográficos, aerodinámicos y acústicos, como se detallará en los próximos apartados. No obstante, antes que nada, es ineludible presentar las hipótesis de este trabajo, en la medida en que estas condicionan dicha metodología.

5.1. Hipótesis

Las hipótesis que aquí se plantean se basan en los descubrimientos acerca de los efectos del refuerzo en otras lenguas (*vid.* capítulo 3). Cada una de ellas se puede relacionar con un objetivo parcial de esta tesis, que busca verificar (o falsar) si esos mismos efectos se constatan en español. La Tabla 5.1 recoge de manera esquemática todas las hipótesis propuestas, para facilitar una visión de conjunto de las mismas. Las hipótesis van numeradas en función del tipo de efecto esperado (p. ej. sobre la velocidad, la duración, etc.); por otro lado, la letra A designa las hipótesis sobre el refuerzo por posición, mientras que la B se reserva para las hipótesis sobre el refuerzo por prominencia.

En primer lugar, esperamos una ralentización de los gestos articulatorios en las proximidades de una frontera prosódica. Suponemos que esa ralentización se produce en mayor medida cuando el gesto es estrictamente adyacente a la frontera; así, por ejemplo, en posición inicial de constituyente, el gesto de constricción de la consonante puede verse más afectado que el subsiguiente gesto de distensión. Como consecuencia, esperamos un pico de velocidad menor en la constricción de una consonante en inicio de palabra que en inicio de sílaba (H_{1A}) y, posiblemente, también una menor velocidad en la distensión, aunque, en este caso, quizá se observen menos efectos sobre los distintos órganos articulatorios (H_{2A}).

En cambio, el refuerzo por prominencia se caracteriza por un aumento en la velocidad del gesto de distensión consonántica que da paso a la vocal de la sílaba acentuada. Por tanto, esperamos un mayor pico de velocidad en la distensión en sílaba

Tabla 5.1. Hipótesis sobre el refuerzo consonántico como consecuencia de la posición prosódica y de la prominencia.

ω designa a las consonantes en inicio de palabra, σ en inicio de sílaba; ${}^1\sigma$ en sílaba acentuada y σ en sílaba no acentuada. %t se refiere al momento de aparición del pico de magnitud, calculado como proporción sobre la duración total del segmento. U_n = flujo nasal, P_n = presión nasal.

REFUERZO POR POSICIÓN			REFUERZO POR PROMINENCIA		
Pico de velocidad de la constricción					
H _{1A}		$\omega < \sigma$			
Pico de velocidad de la distensión					
H _{2A}		$\omega < \sigma$	${}^1\sigma > \sigma$		H _{2B}
Duración del gesto de constricción					
H _{3A}		$\omega > \sigma$			
Duración del gesto de distensión					
H _{4A}		$\omega > \sigma$	${}^1\sigma > \sigma$		H _{4B}
Solapamiento temporal entre los movimientos articulatorios de la constricción					
H _{5A}		$\omega < \sigma$			
Duración acústica del segmento consonántico					
H _{6A}	(a) orales	$\omega > \sigma$	${}^1\sigma > \sigma$	(a) orales	H _{6B}
	(b) nasales	$\omega < \sigma$	${}^1\sigma > \sigma$	(b) nasales	
VOT					
H _{7A}		$\omega = \sigma$	${}^1\sigma = \sigma$		H _{7B}
Altura de la mandíbula					
H _{8A}		$\omega > \sigma$	${}^1\sigma > \sigma$		H _{8B}
Altura del ápice					
H _{9A}		$\omega > \sigma$	${}^1\sigma > \sigma$		H _{9B}
Abertura labial					
H _{10A}		$\omega < \sigma$	${}^1\sigma < \sigma$		H _{10B}
Flujo y presión nasal					
H _{11A}		$\omega < \sigma$	${}^1\sigma > \sigma$		H _{11B}
% t pico de magnitud, y pico U _n / P _n					
H _{12A}		$\omega > \sigma$	${}^1\sigma > \sigma$		H _{12B}
Intensidad relativa					
H _{13A}	(a) fricativas	$\omega > \sigma$	${}^1\sigma > \sigma$	(a) fricativas	H _{13B}
	(b) espirantes	$\omega < \sigma$	${}^1\sigma < \sigma$	(b) espirantes	
	(c) nasales	$\omega < \sigma$	${}^1\sigma > \sigma$	(c) nasales	

acentuada que en sílaba no acentuada (H_{2B}). No se atribuye ningún efecto de la prominencia sobre el gesto de la constricción (*i. e.* no hay una supuesta H_{1B}).

Por otro lado, la extensión temporal de los gestos articulatorios de constricción se caracterizará por una mayor duración como efecto del refuerzo posicional (H_{3A}): la causa directa de este incremento temporal es la deceleración en la ejecución de los gestos que se acaba de comentar. También los gestos de distensión de la consonante podrían aumentar de duración si, efectivamente, se reduce su velocidad, y esta expectativa se recoge en la H_{4A}.

Por lo que se refiere al efecto de la prominencia sobre la duración gestual, no se espera ninguna alteración durante la fase de constricción. En cuanto a la distensión, en un primero momento podría pensarse que, si en el caso anterior una ralentización conlleva un aumento temporal, en este caso el aumento de velocidad previsto por la H_{2B} debería traer consigo una disminución de la duración de ese gesto distensivo. No obstante, como se explicó en el § 3.2, las sílabas acentuadas se caracterizan por un desfase en el gesto de cierre que subsigue a la vocal. Esto provoca un aumento del tiempo disponible para la ejecución del gesto de formación de dicha vocal, o lo que es lo mismo, del gesto de distensión de la consonante que la precede. Por consiguiente, la H_{4B} formula la expectativa de que la distensión consonántica dure más en sílaba acentuada que en sílaba no acentuada.

Las posiciones contiguas a fronteras prosódicas, debido a la ralentización de los gestos y el aumento temporal de los mismos, llevan aparejado un menor grado de solapamiento entre los distintos movimientos articulatorios. Esto se ha documentado no solo como una reducción de la coarticulación entre segmentos contiguos, sino también entre gestos de un mismo segmento (*vid.* la descripción detallada del fenómeno en el capítulo 3). Aquí se comprobará si también se ve afectada la organización intragestual en el caso de los gestos orales, en los que participan conjuntamente los movimientos de varios órganos articulatorios. Así, la H_{5A} predice que en posición inicial de palabra (con respecto a la inicial de sílaba) se observará un menor grado de solapamiento temporal entre el movimiento de la mandíbula y el del ápice (en el caso del gesto apical) y entre el movimiento de la mandíbula y el de los labios (en el caso del gesto labial).

Una consecuencia lógica del refuerzo temporal de los gestos (tanto de constricción como de distensión) es el aumento concomitante de la duración acústica de los segmentos consonánticos correspondientes (H_{6A} y H_{6B}). Como excepción, encontramos el caso de las nasales en posición inicial de constituyente. En el § 3.4.1 se argumentó por qué las características particulares del gesto velar hacen que el refuerzo posicional se manifieste como una menor magnitud del movimiento de descenso del velo. Este comportamiento puede provocar que, a la hora de delimitar la extensión acústica del segmento nasal, se obtenga una menor duración del mismo, como recoge la predicción (b) de la H_{6A} en la Tabla 5.1.

También en relación con la duración relativa de los gestos hemos visto que el aumento temporal del gesto de abducción glotal provoca un aumento del VOT en las lenguas que basan sus oposiciones laríngeas en el rasgo [$\pm$ glotis ensanchada]. Sin embargo, el español no es una lengua de aspiración sino de sonorización. Por supuesto, esto no quiere decir que el VOT en español siempre sea exactamente igual a cero en una consonante sorda; existen pequeños desajustes en la coordinación intergestual que pueden provocar un VOT ligeramente positivo pero no por ello llega a percibirse la distensión como aspirada. Ahora bien, el grado de abertura glotal propio de las sordas del español no exige un ensanchamiento acusado de la glotis (Martínez Celdrán y Fernández Planas, 2013, p. 67 y ss.). En condiciones de refuerzo, por tanto, no esperamos una exacerbación de la abducción, porque esta no estaría especificada como tal en la partitura de gestos. En definitiva, no se predice una diferencia significativa en

el VOT de las consonantes en posición fuerte frente a otras posiciones relativamente débiles (H_{7A} y H_{7B}).

En cuanto a la magnitud de los gestos, como se acaba de recordar, no todos los gestos se comportan igual en la condición de frontera prosódica. En el caso general, esperamos que el aumento de duración pueda llevar consigo un aumento de magnitud, al permitir que el gesto se desarrolle durante más tiempo, como se explicó en el § 3.1. Por ejemplo, en posición inicial de palabra se predice una mayor altura de la mandíbula (H_{8A}); una mayor magnitud de los movimientos de la lengua en las consonantes que implican tales movimientos, p. ej. una mayor altura del ápice (H_{9A}); y un mayor cierre labial o, lo que es lo mismo, una menor abertura (H_{10A}). En cambio, el gesto del velo se comporta de una manera singular, como se justificó en el § 3.4.1, y puede acusar un menor movimiento de descenso en posición inicial de palabra que en inicio de sílaba. Por ello en inicial de palabra esperamos un menor nivel de flujo y de presión del aire nasal (H_{11A}).

Las variaciones de la magnitud por efecto de la prominencia son más uniformes, y esperamos una mayor altura mandibular (H_{8B}) y apical (H_{9B}), una menor abertura labial (H_{10B}), y un mayor flujo y presión nasales (H_{11B}).

El pico de magnitud del gesto (o el pico de flujo y de presión nasal) se alcanzará de manera más tardía tanto en inicio de constituyente mayor como en posición prominente, si se calcula como proporción sobre el total de la duración del segmento, lo que abreviaremos como %t (donde la t significa tiempo) aunque técnicamente se ha calculado como proporción, no como porcentaje (H_{12A} y H_{12B}).

Por último, se puede calcular la proporción de la energía acústica radiada por la consonante con respecto a la de la vocal siguiente. Para las oclusivas (sordas), lógicamente, la intensidad durante la fase de cierre debería ser nula, descontando el posible ruido de fondo, cuyo efecto queda controlado, en todo caso, si lo que se mide es la intensidad relativa con respecto a la vocal. No se espera ningún cambio al respecto en función de la posición o la prominencia, de modo que, si hubiera una modificación real en la intensidad relativa, el cambio estaría motivado únicamente por un aumento de la intensidad en la vocal (algo que sí se espera en sílaba acentuada). Si, en vez de centrarnos en la fase de cierre, nos fijamos en la explosión distensiva, su intensidad podría aumentar en condiciones de refuerzo, debido a la mayor acumulación de P_o durante una fase de oclusión más larga. No obstante, hay que tener en cuenta que habrá muchas consonantes oclusivas con un VOT cercano a cero, en las que será muy difícil delimitar una fase de explosión para medir su intensidad. Esto aboca a una distribución bimodal, con dos subgrupos de oclusivas (con y sin barra de explosión reconocible en el espectrograma), y a un consiguiente desequilibrio en el análisis estadístico. Por este motivo, se ha decidido prescindir del análisis de la intensidad de las oclusivas en este estudio, aunque el tema merece una investigación aparte.

Veamos qué pasa con la intensidad en otros modos de articulación. En el caso de las fricativas, es posible que sean más resistentes al aumento de magnitud (por la necesidad de mantener un nivel de estrechamiento muy concreto) pero, si llega a haber una modificación espacial, se puede traducir en un aumento de la intensidad del ruido de

fricción, al tener que pasar el aire por un canal más estrecho (supuesto *a*). Las espirantes, en cambio, presentarán una menor intensidad relativa debido a la menor abertura oral, lo que dificulta la radiación de la energía acústica (supuesto *b*). Por último, las nasales se caracterizarán por una menor intensidad relativa en posición inicial de palabra, debido a la menor magnitud del gesto de descenso velar ($H_{13A:c}$), pero la intensidad será mayor en sílaba acentuada ($H_{13B:c}$).

5.2. Sujetos

Todos los datos analizados en este estudio se grabaron en el Laboratorio de Fonética de la Universidad de California en Los Ángeles (UCLA) durante la primavera de 2010. La captación de participantes se realizó mediante anuncios colgados en distintos lugares del campus de la universidad y en las residencias de los alumnos de grado. De este modo se reclutó a 6 sujetos mexicanos, hablantes nativos de español, utilizando un método de muestreo consecutivo, es decir, que podía participar todo aquel que se presentara voluntario siempre que cumpliera los requisitos de elegibilidad que se describen más abajo.

Los sujetos potenciales contactaron conmigo mediante correo electrónico por iniciativa propia después de haber visto los anuncios. Se eligió este procedimiento, recogido en el protocolo aprobado por el comité ético institucional (IRB, por sus siglas en inglés) con el número G09-10-100-01 C, para respetar la privacidad de los sujetos lo más posible. De este modo, yo no tenía acceso a los datos de contacto de ningún candidato hasta que ellos daban el primer paso de escribirme. Por otro lado, yo solo les escribí a sus respectivos correos electrónicos en los días inmediatamente previos a su participación, para determinar su elegibilidad, mandarles un documento explicativo del experimento, y acordar fecha y hora para la sesión de grabación.

Después de mandar su mensaje, cada sujeto recibió otro correo con las instrucciones para determinar su elegibilidad. Puesto que algunas preguntas implicaban información personal, no se pidió a los potenciales participantes que las respondieran una a una, sino simplemente contestar con una de entre estas tres opciones: «Soy elegible», «No soy elegible», o «Ya no me interesa». Por supuesto, también podían simplemente no contestar.

A continuación se detalla la lista de criterios de elegibilidad. Para ser elegible, un sujeto debía cumplir todos los requisitos de la lista.

(62) Criterios de elegibilidad

- 1) Ser mayor de edad.
- 2) Ser hablante nativo de español.

Si el voluntario había nacido en EE. UU. o había vivido allí desde pequeño, *no* era elegible.

- 3) Contestar NO a *todas* las preguntas siguientes:
 - a.* ¿Tienes algún problema de habla o de audición?
 - b.* ¿Estás embarazada?

- c. ¿Llevas un marcapasos, un desfibrilador automático implantable, un clip para aneurisma cerebral, un implante coclear, o cualquier otro tipo de implante o prótesis de metal?
- d. ¿Tienes algún empaste en los incisivos?
- e. ¿Llevas barba?
- f. ¿Necesitas gafas para leer una fuente Helvética, de 58 puntos, a dos metros de distancia? (Las lentes de contacto están permitidas.)

El requisito (62):2 refleja el intento de descartar a sujetos bilingües, para evitar interferencias lingüísticas.

Por otro lado, los voluntarios con algún tipo de problema de habla o de audición eran excluidos por motivos obvios.

Los campos electromagnéticos se han relacionado, aunque sin llegar a conclusiones firmes, con abortos espontáneos en mujeres embarazadas. Esta posible relación es rechazada por muchos investigadores (*vid.* p. ej. Kavet y Tell, 1991) y desde luego no parece estar motivada para un tiempo de exposición tan breve como el del articulógrafo. Por otro lado, algo mucho más cotidiano como la exposición a las ondas de un ordenador no está normalmente contraindicado para las mujeres embarazadas en ningún tipo de regulación. Sin embargo, parece que la práctica habitual es excluir, por prudencia, a las mujeres embarazadas como sujetos en experimentos con articulografía electromagnética (EMA, por sus siglas en inglés), *vid.* Perkell *et al.* (1992).

Del mismo modo, aunque los resultados de Joglar, Nguyen, Garst y Katz (2009) sugieren que la tecnología EMA se puede usar con seguridad con pacientes que lleven marcapasos, los propios autores del estudio recomiendan tener precaución en estos casos, ya que no todos los modelos de marcapasos ni de articulógrafo han sido probados. Por este motivo, es preferible evitar sujetos con tales dispositivos, o con cualquier otro tipo de implante o prótesis de metal, como aparece recogido en el punto (62):3:c.

Los sujetos con empastes en los incisivos fueron excluidos porque el adhesivo empleado para pegar los sensores de EMA a las encías podía dañarlos.

También se excluyó del estudio a los hombres con barba, dada la dificultad de pegar los sensores correspondientes a los labios, y también puesto que los cables de los sensores se recogían y fijaban en la mejilla con cinta adhesiva para que entorpecieran lo menos posible y para evitar que se engancharan y, al tirar, se despegaran los sensores.

Por último, los sujetos que necesitaban usar gafas para leer las frases mostradas en el ordenador no resultaban aptos para el estudio, ya que al llevar puesta la máscara nasal no era posible usar gafas. En cambio, las lentes de contacto no suponían ningún problema.

Después del proceso de selección, se envió un documento a los sujetos con información detallada acerca de los procedimientos incluidos en el experimento. Si seguían queriendo participar, se fijaba una cita en un momento conveniente para ambas partes para realizar el experimento.

Finalmente participaron en el estudio 6 sujetos mexicanos (3 hombres y 3 mujeres), como se muestra en la Tabla 5.2.

Tabla 5.2. Descripción de los sujetos participantes en el estudio.

Nombre	Sexo	Edad
Adriana Ramírez	Mujer	19
Rebecca Barrera	Mujer	19
Lourdes Solís	Mujer	23
Pablo Quintero	Hombre	18
David Torres	Hombre	23
Eduardo L. Reyes	Hombre	19

Cada sujeto recibió 25 dólares estadounidenses por hora de participación. Dados los estándares locales, se consideró que esta cantidad era apropiada para la tarea requerida, teniendo en cuenta los diferentes dispositivos y aparatos que se pedía usar. Por otro lado, era una cantidad suficientemente atractiva como para reclutar el número esperado de sujetos, pero sin ser coercitiva para las personas más desfavorecidas económicamente.

Se informó a todos los participantes de que el dinero provenía de una beca concedida por el Ministerio de Educación y Ciencia de España para financiar mi investigación.

5.3. Lista de palabras

Los sujetos tenían que leer en voz alta una serie de palabras españolas en una frase portadora, con dos repeticiones cada una. Las palabras estaban escogidas para incluir las siguientes ocho consonantes: [p], [f], [β], [m], [t], [s], [ð], [n], suficientes para poner a prueba todas las hipótesis planteadas en el § 5.1. Todas ellas implican gestos orales en los que participa el movimiento de la mandíbula. Además, hay cuatro que se caracterizan por un gesto labial y cuatro, por un gesto apical. En cuanto al modo de articulación, hay dos oclusivas, dos fricativas, dos espirantes, y dos nasales (en cada subgrupo, una labial y una coronal). De cara al análisis estadístico, además, es beneficioso el diseño factorial (*i. e.* hay el mismo número de consonantes con gesto labial o apical, una por cada modo de articulación; y hay el mismo número con cada modo, una labial y otra apical).

En parte, esta última consideración está en la base de la exclusión de otras posibles consonantes. Por ejemplo, si se hubieran considerado consonantes palatales o velares, habría sido imposible mantener el equilibrio en la combinación de los distintos factores (p. ej. no existe una fricativa palatal [ç] en español, ni es posible encontrar una nasal velar [ŋ] en posición inicial de sílaba). Además, en el caso de las velares, especialmente, habría que pegar sensores de EMA en una posición demasiado trasera del dorso de la lengua, con el consiguiente riesgo de producir náuseas a los sujetos. Por otro lado, otros posibles modos de articulación —como las laterales¹⁴⁵, vibrantes y percusivas— están restringidos a las consonantes corales. A eso se suma el hecho de que el modelo del articulógrafo utilizado (el AG100) solo permite grabar los movimientos de puntos localizados en el eje medio sagital, de modo que se puede captar el contacto apical de

¹⁴⁵ Todos los participantes eran yeístas, de modo que no podrían pronunciar una lateral palatal [ʎ].

una [l], por ejemplo, pero no la posición de los laterales de la lengua, dato relevante para la articulación de este segmento. Por último, la inclusión de la percusiva [r] también habría introducido un importante desequilibrio en el diseño estadístico, ya que no se puede comparar el efecto de la posición inicial de sílaba con la inicial de palabra, pues en español la /r/ está restringida al interior de unidades léxicas. En este caso, solo se podría haber estudiado el efecto de la prominencia ($^1\sigma$ vs. σ), pero quizá la inclusión de estos casos en el modelado estadístico habría dado más peso relativo al efecto del acento (frente al efecto de la frontera de ω), inclinando el equilibrio de manera artificial.

Cada una de las ocho consonantes estudiadas aparece el mismo número de veces (dos) en cada una de las posibles condiciones resultantes de la combinación del factor posicional y acentual, *i. e.* inicial de palabra en sílaba acentuada, inicial de palabra en sílaba no acentuada, inicial de sílaba en sílaba acentuada, inicial de sílaba en sílaba no acentuada^{146 y 147}.

Con todas estas combinaciones factoriales en mente, finalmente resultó una lista de 47 palabras (*vid.* Tabla 5.3). En algunas de ellas se analizaron las dos consonantes de los respectivos ataques silábicos (inicial e interior), mientras que en otras solo se analizó una consonante, para dar lugar a un número equilibrado de casos en cada condición factorial, como se ha explicado más arriba.

Las palabras se presentaron en un orden aleatorio, diferente para cada sujeto, y estos debían pronunciarlas insertas en una frase portadora: «Yo repito ____ como pícaro» [$^1j\text{ɔ}$ re ^{1} pi ɔ ____ komo ^{1} pikaro]¹⁴⁸.

¹⁴⁶ En ningún caso se analizaron consonantes en coda, ya que el objetivo de este experimento es describir los efectos del refuerzo, no del debilitamiento.

¹⁴⁷ Como se ha dicho, cada consonante aparece dos veces en cada una de estas cuatro condiciones. Una de las veces está seguida de la vocal [a] y otra, de la vocal [i]. La inclusión de este factor en el diseño responde a la intención de estudiar, más adelante, los efectos coarticulatorios derivados del contexto vocálico y su posible interacción con los fenómenos de refuerzo. Sin embargo, tal y como se viene defendiendo en esta tesis, el desencadenante primero de los efectos de refuerzo debe buscarse en la posición de la estructura prosódica en la que se inserta el segmento. Por este motivo, el factor contextual no será analizado aquí.

¹⁴⁸ La mera presencia de un marco proporciona un contexto más natural para la pronunciación. Sin embargo, la elección de esta frase en concreto requiere alguna explicación adicional. Esta se justifica por el propósito de utilizar este mismo corpus más adelante para estudiar la influencia de los distintos gestos articulatorios sobre algunos parámetros aerodinámicos, como la P_s y la P_o . La P_s se puede calcular de un modo indirecto equiparándola al pico de P_o durante la producción de una oclusiva sorda como la [p]. En efecto, al cerrarse la vía oral y estar abierta la glotis, las cavidades subglotal y supraglotal constituyen un único volumen y sus presiones prácticamente se equiparan (Hertegard, Gauffin, y Lindestad, 1995). En este caso, la P_s puede inferirse a partir de los valores de P_o correspondientes a la [p] de *repito* y *pícaro*. Además, el carácter tónico de la sílaba y la coarticulación con [i] producen un mayor incremento de P_o , haciendo que así resulte más fácil que alcance el nivel de la presión subglotal, ya que esta es una premisa clave para este cálculo. Por otro lado, estas sílabas se sitúan en la tercera posición de la frase (por el inicio y por el final, respectivamente), de modo que se evita cualquier interferencia que pudiera suponer la proximidad a la frontera de enunciado.

En líneas generales, se asume que la P_s está sujeta a una declinación más o menos constante (Ohala, 1990b). Por tanto, si se obtienen las medidas de P_s correspondientes a dos puntos de tiempo, se pueden inferir los valores de P_s de cualquier punto intermedio por interpolación, que en general resulta adecuada, excepto durante la distensión de las oclusivas sordas, momento en que la estimación suele superar al valor real. En cualquier caso, es necesario tener en cuenta un buen número de precauciones metodológicas para que la interpolación sea fiable (Hertegard *et al.*, 1995; Koike y Hirano, 1973; Netsell, 1969; Rothenberg, 1982; Solomon, Ramanathan, y Makashay, 2007).

Tabla 5.3. Lista de palabras. Las consonantes analizadas en cada una de ellas aparecen subrayadas.

Palabra	Transcripción	Palabra	Transcripción
<u>b</u> asar	[βa'saɾ]	<u>n</u> ada	[ˈnaða]
<u>b</u> a <u>t</u> a	[ˈβaɾa]	<u>n</u> adar	[na'ðaɾ]
<u>b</u> ic <u>i</u>	[ˈβisi]	na <u>v</u> al	[na'βal]
<u>c</u> ima	[ˈsima]	<u>N</u> ina	[ˈnina]
<u>c</u> iv <u>i</u> l	[si'βil]	<u>n</u> ivel	[ni'βel]
<u>d</u> a <u>b</u> a	[ˈðaβa]	<u>p</u> ana	[ˈpana]
<u>d</u> atar	[ða'taɾ]	<u>p</u> arar	[pa'raɾ]
de <u>l</u> <u>f</u> in	[ðel'fin]	pa <u>t</u> in	[pa'tin]
<u>d</u> ir <u>a</u>	[ði'ra]	pe <u>d</u> ir	[pe'ðiɾ]
<u>d</u> iv <u>a</u>	[ˈðiβa]	<u>P</u> ili	[ˈpili]
<u>f</u> am <u>a</u>	[ˈfama]	<u>p</u> ip <u>i</u>	[pi'pi]
<u>f</u> atal	[fa'tal]	Ra <u>f</u> i	[ˈrafi]
<u>f</u> ina	[ˈfina]	<u>s</u> anar	[sa'naɾ]
<u>f</u> inal	[fi'nal]	<u>S</u> ara	[ˈsara]
ga <u>f</u> as	[ˈɣafas]	se <u>n</u> il	[se'nil]
hip <u>p</u> y	[ˈxipi]	so <u>f</u> á	[so'fa]
<u>J</u> av <u>i</u>	[ˈxaβi]	ta <u>m</u> iz	[ta'mis]
ma <u>m</u> á	[ma'ma]	<u>t</u> apa	[ˈtapa]
ma <u>m</u> i	[ˈmami]	<u>t</u> apar	[ta'paɾ]
<u>m</u> a <u>s</u> a	[ˈmasa]	<u>t</u> ilín	[ti'lin]
<u>m</u> a <u>t</u> ar	[ma'taɾ]	<u>t</u> isis	[ˈtisis]
mi <u>d</u> i	[ˈmiði]	ti <u>t</u> i	[ti'ti]
<u>m</u> in <u>i</u>	[ˈmini]	<u>v</u> iril	[βi'ril]
<u>m</u> is <u>i</u> l	[mi'sil]		

Ahora bien, dada la manera de constituir frases fonológicas explicada en el § 2.1.4.1, la frase portadora, junto con la palabra clave, admite dos estructuras prosódicas diferentes, según se produzca o no la reestructuración opcional. Los ejemplos (63) y (64) representan los aspectos relevantes de ambas posibilidades.

(63)[Yo repito]_φ [[____]_ω]_φ [como pícaro]_φ

(64)[Yo repito [____]_ω]_φ [como pícaro]_φ

Como se puede ver, en el primer caso la consonante inicial de la palabra morfosintáctica coincide con un inicio de φ, mientras que en el segundo, tras la reestructuración, la ω correspondiente a la palabra meta ha quedado absorbida en la φ precedente y la consonante inicial ahora es verdaderamente inicial de ω. Dado el contexto metalingüístico en el que aparece la palabra meta, como complemento de un

verbo que aquí se puede interpretar como verbo de lengua (*repetir*), algunos hablantes pueden sentir la necesidad de enfatizar dicha palabra meta, lo que hace que esta constituya su propia ϕ . Dado que el objetivo de este estudio no incluye los niveles prosódicos supraléxicos, durante las sesiones de grabación se pidió a los sujetos que no hicieran ningún tipo de pausa ni dieran énfasis a la palabra meta, con el fin de asegurar que la consonante inicial estaba alineada con una frontera ω pero no ϕ . Si, en algún caso, algún sujeto no seguía esta indicación, se le pedía que volviera a decir ese enunciado de manera adecuada y se grababa de nuevo¹⁴⁹.

5.4. Desarrollo de una sesión experimental

Cuando un participante venía a grabar, recibía una explicación detallada de los distintos procedimientos experimentales, y se le pedía que confirmara que había entendido el documento informativo enviado por correo electrónico. El sujeto podía hacer preguntas o pedir alguna demostración previa y, si quería participar, firmaba la hoja de consentimiento informado y se quedaba con una copia. Acto seguido, se le colocaban los diferentes instrumentos de medición, como se detalla en los siguientes subapartados, y se procedía a la obtención simultánea de datos de distinto tipo: articulográficos, aerodinámicos y acústicos.

5.4.1. Obtención de datos articulográficos

La articulografía es un método basado en las propiedades de los campos electromagnéticos que sirve para detectar el movimiento de los órganos articulatorios (p. ej. la mandíbula, la lengua o los labios) durante la producción de habla.

En el caso del modelo utilizado, el AG100 de Carstens Medizinelektronik, el sujeto se pone un casco con tres transmisores equidistantes entre sí, que emiten un campo electromagnético alterno de forma radial, cada uno de ellos con una frecuencia diferente. Estos campos inducen un voltaje en unas bobinas diminutas alojadas en una serie de sensores que se pegan en distintos puntos de la lengua, los labios, etc. Según el voltaje medido en cada caso, se puede calcular la distancia de cada sensor a cada uno de los transmisores y, con ello, determinar sus coordenadas cartesianas en el plano definido

¹⁴⁹ Por otro lado, puesto que la palabra meta aparece en la posición final de la ϕ , es esperable que reciba el acento de frase correspondiente a la prominencia de dicho nivel. En este sentido, hay que reconocer que en los resultados no será posible deslindar los efectos atribuibles exclusivamente al acento léxico de los debidos al acento de frase o a una mezcla de ambos (esto debería ser objeto de una futura investigación). No obstante, dado el carácter acumulativo de los fenómenos de refuerzo, en todo caso no esperamos una modificación de la tendencia propia del acento de palabra, sino, como mucho, una magnificación de sus efectos característicos sobre la dinámica gestual. Los efectos exclusivos del acento de frase, como la aparición de objetivos tonales, no se estudian aquí y no interfieren con nuestras variables de interés.

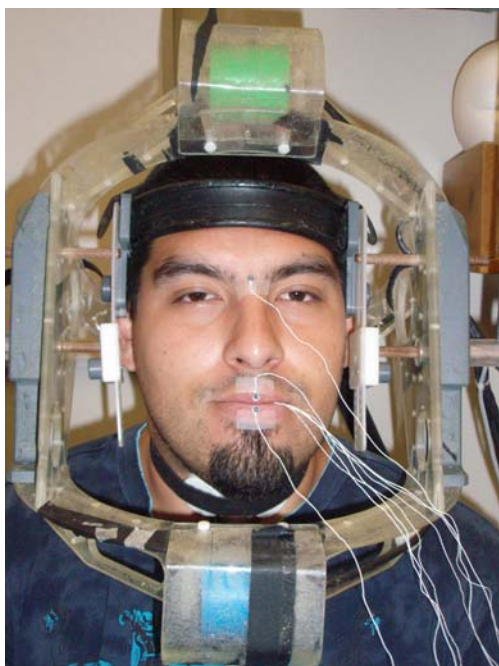


Figura 5.1. Colocación de los sensores de EMA



Figura 5.2. Colocación de la máscara y los tubos de mediciones aerodinámicas

por dichos transmisores, que, en este modelo, necesariamente coincide con el plano medio sagital^{150 y 151}.

Los procedimientos empleados tanto en la preparación y calibración del equipo como durante las sesiones de adquisición de datos siguieron el protocolo recomendado por el Laboratorio de Fonética de la UCLA¹⁵².

A cada sujeto se le colocaron dos sensores de referencia —uno en la encía superior y otro en el puente de la nariz— y se le pidió que se introdujera momentáneamente en su boca y mordiera una tarjeta plastificada con dos sensores pegados a su línea central, separados unos 4 cm entre sí, para medir el plano de su mordida. Toda esta información se utilizó para la posterior corrección y rotación de los datos, según se indica en el § 5.5.2.

A continuación, se le colocaron los sensores restantes, en los puntos de interés para este experimento: uno en la encía inferior —para medir el movimiento de la mandíbula—, otro en el ápice de la lengua (aproximadamente a 0.5 cm de la punta), y otros dos en el labio superior y en el labio inferior, respectivamente. La Figura 5.1 muestra a un sujeto con los sensores de EMA ya colocados.

¹⁵⁰ Para más detalles sobre el principio de medición que subyace a este sistema, *vid.* Perkell *et al.* (1992), y también el manual de uso casero del Laboratorio de Fonética de la UCLA, publicado en la siguiente página web: <http://www.linguistics.ucla.edu/faciliti/facilities/physiology/Emamual.html>.

¹⁵¹ La restricción al plano medio sagital, como la impuesta por el modelo AG100 que se ha usado en este experimento, se especifica en algunos trabajos por medio de las siglas EMMA, con doble M (por *electromagnetic midsagittal articulography*). En esta tesis, sin embargo, se usa la denominación más amplia de EMA como nombre general de la técnica empleada.

¹⁵² *Vid.* nota 150 y también la siguiente página web: <http://www.linguistics.ucla.edu/faciliti/facilities/physiology/ema.html>

5.4.2. Obtención de datos aerodinámicos

Como se ha visto en el § 3.4, uno de los métodos indirectos para detectar los movimientos del velo del paladar y el grado de abertura velofaríngea consiste en la medición de datos aerodinámicos como el flujo de aire nasal (U_n). Para obtener esta información, al sujeto se le coloca alrededor de la nariz una máscara, de la que sale un tubo que se conecta a un transductor. Este transductor detecta el descenso de presión causado por la resistencia que oponen al flujo del aire tanto la máscara como el tubo y, a partir de ese dato, se puede calcular el valor de U_n . El transductor, a su vez, se conecta con una interfaz que digitaliza la señal y la manda al ordenador, donde se graba¹⁵³.

Aunque en las investigaciones sobre aerodinámica del habla es habitual medir la presión oral (P_o) mediante la conexión de un transductor a un tubo introducido en la boca, en cambio no es tan frecuente la obtención de datos de presión nasal (P_n). Sin embargo, esta se puede medir fácilmente mediante un tubo introducido en un orificio nasal, sin necesidad de la intermediación de ninguna máscara. En este experimento, decidí grabar la información correspondiente a la P_n en paralelo a la del U_n , para comparar los resultados de ambos procedimientos.

El tubo empleado para medir la P_n necesita una preparación especial para evitar que se salga del orificio nasal, y también para impedir fugas de aire entre el tubo y la pared interna de la nariz. Para conseguir estos dos objetivos, se intentó sujetar el tubo entre dos bolitas de gomaespuma (a los efectos, tapones para los oídos). Este fue el procedimiento seguido en el caso de David y Eduardo. Sin embargo, la sospecha de que, aun así, había alguna fuga de aire, me llevó a modificar el procedimiento con los demás sujetos. Con ellos usé un solo tapón de espuma, pero más ancho, y perforé un agujero en el centro, a través del cual introduje el tubo hasta que sobresaliera por el otro extremo. La Figura 5.3 compara la densidad de los valores de los picos de P_n medidos durante los segmentos nasales ([m] y [n]) para cada sujeto. Como se puede observar, los datos de David y Eduardo se concentran alrededor de 0 cm de H_2O , lo que podría ser una confirmación de la fuga de aire producida. No obstante, hay que tener en cuenta que también los datos de Lourdes muestran una distribución similar, aunque no tan acusada. Además, el grado de P_n captada por el dispositivo en cada caso puede variar por otros motivos ajenos al diseño experimental, como la propia anatomía de las fosas nasales de cada sujeto, o dependiendo de si ese día estaban más o menos congestionados. En cualquier caso, me parece que la solución del tapón único y perforado responde mejor a las condiciones experimentales que se pretende obtener.

Ahora bien, en condiciones normales, la máscara destinada a medir U_n se interpondría en el recorrido del tubo de P_n hasta la nariz. Para sortear este problema, se practicó un pequeño orificio en la máscara a través del cual se introdujo el tubo. Luego se añadió un poco de masilla adhesiva para evitar fugas de aire por el minúsculo hueco que quedaba entre la pared externa del tubo y los bordes de dicho orificio. La Figura 5.2 muestra a una de las participantes con la máscara de U_n y el tubo de P_n ya colocados.

¹⁵³ Para más información sobre este procedimiento, *vid.* p. ej. Baken (1996, p. 277 y ss.) y Rothenberg (1977).

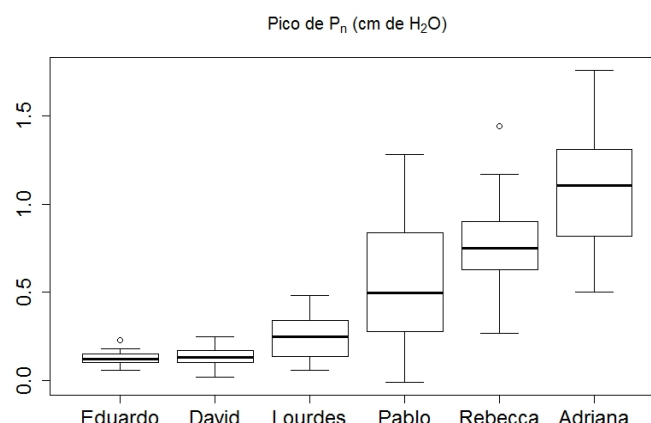


Figura 5.3. Distribución de los valores del pico de P_n durante los segmentos nasales

Finalmente, ambos dispositivos se conectaron a una caja de transductores modelo EIPF-4 de Scicon R&D Inc. y se grabaron los datos a través del sistema de adquisición multicanal PCquirer 516, también de la misma empresa, conectado a un ordenador con el programa PCquirerX. Las señales se grabaron a una frecuencia de muestreo de 11 000 Hz y se les aplicó, mediante el mismo programa, un filtro pasabajo con el punto de corte en 80 Hz para eliminar los ruidos innecesarios que resultarían de las oscilaciones de alta frecuencia (dado el punto de corte, se descartan las variaciones que se produzcan en un periodo menor a $T = 0.0125$ s).

Por otro lado, para conocer la equivalencia entre los voltajes transmitidos por los transductores y las unidades de medida normalmente utilizadas para los aspectos aerodinámicos del habla (presión en cm de H_2O , flujo en ml/s), se realizó la calibración con los instrumentos de Scicon destinados a tal efecto, siguiendo los procedimientos habituales de calibración con un manómetro de agua (tubo en forma de U) y con una espita que insufla aire en un tubo y empuja hacia arriba una bolita encerrada en el mismo (*vid.* p. ej. Baken, 1996, caps. 7 y 8).

5.4.3. Obtención de datos acústicos

La interfaz PCquirer se conectó a un pequeño micrófono de clip, de condensador y omnidireccional, modelo Sony ECM44B, protegido con una pantalla de espuma. Según las indicaciones del fabricante, la respuesta frecuencial es de 40 Hz a 15 kHz, la sensibilidad es de $-53 \text{ dB} \pm 3 \text{ dB}$, el rango dinámico de 90 dB o más, y la relación señal/ruido de 62 dB o más. Dicho micrófono se sujetaba por medio del clip al casco de EMA, señalando a la boca del sujeto, y capturaba el audio sobre el que se han realizado los análisis acústicos.

Dado que el casco de EMA también disponía de su propio micrófono incorporado, las señales de ambos audios sirvieron para alinear temporalmente los datos articulográficos con el resto de datos —acústicos y aerodinámicos—, grabados con PCquirer (*vid.* último epígrafe del § 5.5.2).

5.5. Procesamiento preliminar de los datos

Cada grabación de datos realizada por medio del programa PCquirerX produce como resultado un único archivo de extensión *.pmf* en el que se encuentran empaquetados los distintos canales de información obtenidos. Para extraer dichos canales por separado, convirtiéndolos en formato *.wav*, se utilizó EGGWorks, un programa desarrollado por Henry Tehrani que se puede descargar gratuitamente de internet¹⁵⁴. Así, de cada *.pmf* se extrajeron las correspondientes pistas de audio, U_n y P_n .

Por otro lado, antes de poder analizar los distintos canales de información correspondientes a cada enunciado, fueron necesarios algunos pasos intermedios, que se describen a continuación.

5.5.1. Segmentación y etiquetado del audio

El proceso de segmentación y etiquetado se realizó semiautomáticamente con la ayuda de la versión española de EasyAlign¹⁵⁵ (Goldman, 2011; Goldman y Schwab, 2014), un paquete de *scripts* diseñado como extensión del programa Praat (Boersma y Weenink, 2011). Posteriormente, se hizo una revisión manual de todos los archivos *TextGrid* generados y se afinó la alineación con el audio cuando fue necesario.

Por otro lado, en los *TextGrids* resultantes solo se retuvieron las líneas de transcripción correspondientes a los segmentos y las palabras, y se añadieron marcas a las consonantes concretas que se iba a analizar en cada palabra. Además, se insertaron sendos niveles de transcripción para codificar el sujeto, la unidad prosódica (palabra o sílaba) en cuya posición inicial se encontraba la consonante en cuestión, y el carácter acentuado o no acentuado de la sílaba en la que aparecía. Un último nivel recogía las marcas temporales, insertadas tras un análisis manual, para el cálculo del VOT en las oclusivas sordas.

5.5.2. Ajustes en los archivos de EMA

Los datos originales grabados por el dispositivo EMA presentan unos formatos de archivo solo reconocidos por el programa del propio articulógrafo. Para hacerlos legibles por Praat, primero se leyeron en Matlab como puras matrices de datos y se extrajeron las distintas señales, correspondientes a las coordenadas x o y de cada uno de los canales de EMA. A partir de esas matrices, también con Matlab, se realizó una serie de transformaciones, que se detallan a continuación.

Filtrado

A todas las señales de EMA se les aplicó un filtro pasabajo de respuesta finita (FIR), de orden $N = 100$, con una ventana de Kaiser ($\beta = 5$). La frecuencia de corte seleccionada fue de 10 Hz ($T = 0.100$ s) para los puntos de referencia (la encía superior y el puente de la nariz), cuya posición a lo largo de la grabación se supone bastante

¹⁵⁴ www.linguistics.ucla.edu/faciliti/sales/software.htm (acceso: 2010-01-15).

¹⁵⁵ <http://latlcui.unige.ch/phonetique/easyalign.php> (versión del 2011-04-18).

estable. En cambio, para el resto de los puntos, el corte se aplicó a 40 Hz ($T = 0.025$ s), para evitar perder movimientos lingüísticamente relevantes.

Dado que el efecto de la ventana produce un retraso de $N/2$ muestras en los vectores de datos, dicho retraso debe ser compensado. Por ello se eliminaron las primeras 50 muestras, y se añadieron otras tantas (con valor cero) al final, para mantener la duración intacta.

Por último, hay que tener en cuenta que la utilización de un filtro de orden 100 produce cierta distorsión en las primeras 99 muestras, de modo que, incluso después de eliminar las 50 primeras como se acaba de explicar, sigue habiendo otras 49 que no resultan muy fiables. Dada la frecuencia de muestreo inicial de 500 Hz, una pérdida de 50 muestras equivale a los primeros 0.100 s de cada pista grabada, lo que no afecta a los segmentos de interés para nuestro análisis, que nunca se encuentran tan al inicio de la grabación.

Reducción de la frecuencia de muestreo

Este paso se realizó después de haber aplicado el filtro para evitar la distorsión de muestreo (en inglés, *aliasing*). La frecuencia original de 500 Hz se redujo a 250 Hz tomando consecutivamente, sin solapamientos, parejas de valores del vector inicial y transfiriendo al educto un único valor: el de la media aritmética. Si el vector original presentaba un número impar de muestras, se perdía la última.

Estabilización de los puntos de referencia

Como se recordará, dos de los canales de EMA se emplearon para grabar otros tantos puntos de referencia, fijos en la estructura ósea. La primera referencia, que llamaremos R1, corresponde a la encía superior; y la segunda referencia, R2, al puente de la nariz. A su vez, aludiremos a las coordenadas x e y de dichos puntos como R1x, R1y, R2x, R2y. El resto de puntos serán designados colectivamente como P a efectos de la siguiente explicación. Por último, expresiones del tipo R2', P', etc. aludirán a la posición a la que deben ser desplazados dichos puntos, mediante una serie de transformaciones matemáticas, con el objeto de facilitar la interpretación de los mismos.

Durante una grabación, la cabeza del sujeto puede moverse ligeramente dentro del casco que lleva amarrado, haciendo que R1 y R2 se muevan dentro del sistema de coordenadas provisto por el casco. Sin embargo, todos los demás puntos (lengua, mandíbula, labios) se pueden referir de una manera relativa a estas referencias, para neutralizar dichos movimientos de la cabeza y quedarnos exclusivamente con la información del movimiento de los articuladores. Para ello se realiza una serie de operaciones sobre las matrices de datos, de modo que se modifica la interpretación geométrica de los mismos.

Dada una matriz, llamaremos fotograma al conjunto de los datos de los distintos canales que corresponden a un mismo instante de la grabación. Si se plasman gráficamente esos datos, efectivamente se puede considerar que constituyen un fotograma de toda una secuencia dinámica de movimientos: son como una imagen instantánea de las distintas posiciones de los sensores en un punto de tiempo

determinado. Por ejemplo, la Figura 5.4 en la p. 190 es una representación idealizada y simplificada de un fotograma, donde se representan R1, R2 y tres puntos más.

Pues bien, se trata de que las posiciones de R1 y R2 permanezcan inmóviles a lo largo de todos los fotogramas, estabilizando así los puntos que se consideran referencia para la interpretación de los otros puntos, P. Para ello, de entre toda la secuencia, se escoge arbitrariamente un fotograma de referencia (FR), y se realizan las transformaciones oportunas sobre los demás fotogramas para asimilarlos a ese estado de cosas. El procedimiento implica trasladar y rotar geométricamente el conjunto de puntos de modo que las posiciones de los puntos de referencia en esos otros fotogramas (R1 y R2) terminen coincidiendo exactamente con las posiciones tomadas como objetivo (R1' y R2'), que son las del fotograma FR.

Para ello, en cada fotograma se designa R1 como el nuevo centro de coordenadas (0,0), sobre el que después se hace pivotar el resto de los puntos para rotarlos hasta la posición adecuada. En términos matemáticos, la traslación de R1 al centro consiste en restar a todas las coordenadas x de un fotograma el valor de R1x de ese mismo fotograma, y a todas las coordenadas y se les resta el valor de R1y. Así, al restarse sus valores a sí mismo, R1 se iguala a (0,0): $R1x - R1x = 0$, $R1y - R1y = 0$. Esos valores se restan también a los demás puntos P para que se desplacen acordemente y mantengan su posición relativa con respecto a R1. Por supuesto, esta misma operación también se realiza sobre el propio fotograma FR, para que R1 (que aquí coincide ya con R1') también se interprete como centro de coordenadas en ese caso¹⁵⁶.

Por ejemplo, en la Figura 5.4, se observa que R1 se proyecta sobre el punto correspondiente a +3 unidades del eje x y +4 unidades del eje y . Por ello, en la Figura 5.5 se resta 3 a las coordenadas x de todos los puntos, y en la Figura 5.6 se resta 4 a las coordenadas y de todos los puntos, de modo que dichos puntos se desplazan al unísono, manteniendo las posiciones relativas entre sí, pero consiguiendo como resultado la traslación del punto R1 al centro de coordenadas.

Una vez establecido el nuevo sistema de coordenadas, se calculan por separado dos ángulos, para luego hallar su diferencia. Por un lado se averigua, para cada fotograma, el ángulo θ que forma sobre la horizontal el segmento que une R2 con R1. El ángulo se computa como el arcotangente del resultado de dividir R2y entre R2x¹⁵⁷ (*vid.* Figura 5.7).

(65) Cálculo del ángulo θ

$$\theta = \arctan \left(\frac{R2y - R1y}{R2x - R1x} \right)$$

¹⁵⁶ Puesto que, después de esta transformación, en todos los fotogramas $R1 = (0,0)$, no se seguirá haciendo la distinción entre R1 y R1', por simplificar.

¹⁵⁷ En realidad hay que calcular la longitud de los dos catetos del triángulo rectángulo cuya hipotenusa es el segmento que une R2 con R1. El hecho de que R1 valga (0,0) facilita los cálculos de las longitudes de los catetos: $R2y - R1y = R2y$; $R2x - R1x = R2x$.

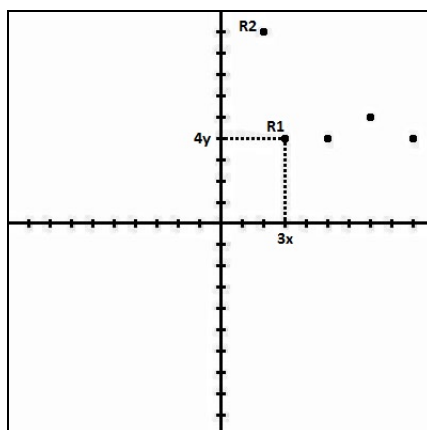


Figura 5.4. Ejemplo simplificado de un fotograma.

R1 (encía superior) y R2 (puente de la nariz) constituyen las referencias fijas para los demás puntos, correspondientes a estructuras articulatorias móviles, como por ejemplo, la lengua. Se calculan las proyecciones de R1 y R2 sobre los ejes x e y para proceder a la traslación de todos los puntos de modo que R1 se sitúe sobre el centro de coordenadas (0,0).

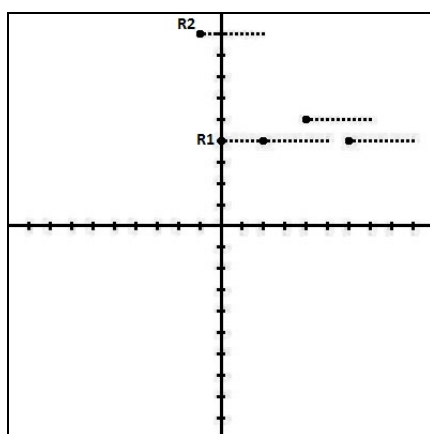


Figura 5.5. Traslación de las coordenadas x mediante sustracción del valor inicial R1x

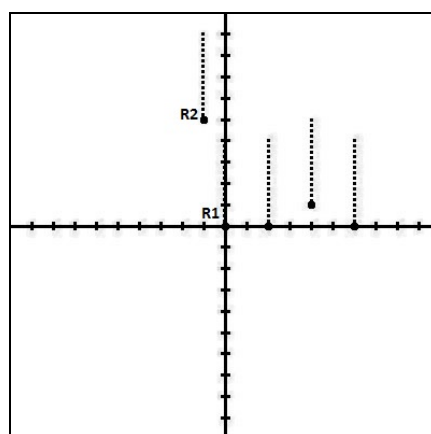


Figura 5.6. Traslación de las coordenadas y mediante sustracción del valor inicial R1y: ahora R1 coincide con el centro (0,0)

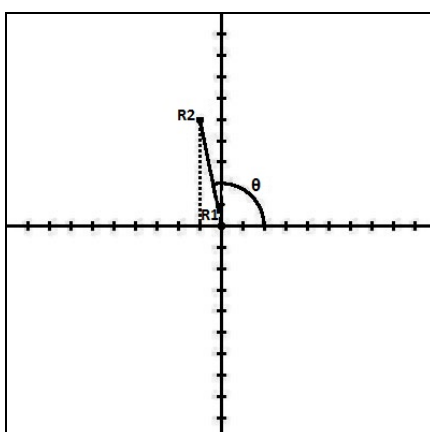


Figura 5.7. El ángulo θ que forma el segmento R2-R1 sobre la horizontal se calcula como el $\arctan (R2y/R2x)$

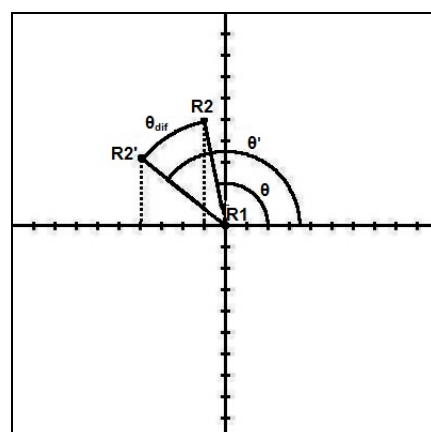


Figura 5.8. El ángulo $\theta_{dif} = \theta' - \theta$ es el ángulo que debe aplicarse en la rotación para hacer coincidir R2 (el puente de la nariz) con R2' (posición de convergencia tomada del fotograma de referencia, FR)

En un sistema de coordenadas, se reconoce la existencia de cuatro cuadrantes en función de los valores positivos o negativos de x e y (*vid.* Figura 5.9). Así, el cuadrante I presenta valores positivos para ambos ejes y, desde él, en sentido antihorario, se numeran el resto de los cuadrantes (II, III, IV). Pues bien, dado que sabemos que R2 (el puente de la nariz) siempre estará por encima de R1 (encía superior), esto apunta a valores positivos de y ; por tanto, a los cuadrantes I o II. En un ciclo de 360° hay dos ángulos que comparten el mismo valor de tangente, opuestos 180° entre sí. Por tanto, si R2x es negativo, el cálculo del ángulo θ_1 dará un resultado negativo, pero no debe interpretarse como si R2 estuviera en el cuadrante IV, sino que simplemente hay que sumarle 180° para recolocarlo en el cuadrante II.

Además del valor de θ , es preciso conocer el ángulo θ' al que queremos llegar, es decir, el que forma sobre la horizontal el segmento R2'-R1 en el fotograma de referencia, FR. El cálculo se realiza, igualmente, mediante la función arcotangente. Después de conocer θ' y θ , se puede calcular la diferencia entre los dos ángulos ($\theta_{\text{dif}} = \theta' - \theta$), *vid.* Figura 5.8. Todo lo que hay que hacer a continuación es efectuar una rotación geométrica del conjunto del fotograma, pivotando sobre R1, de modo que R2 termine coincidiendo con la posición objetivo R2', y que el resto de los puntos P se muevan también acordeamente hasta su localización definitiva, P' (*vid.* Figura 5.10).

Por la manera de calcular θ_{dif} , siempre sabremos la dirección de la rotación, porque simplemente tenemos que partir del θ original del fotograma y sumarle θ_{dif} para llegar al deseado θ' ($\theta + \theta_{\text{dif}} = \theta + \theta' - \theta = \theta'$), y que así todos los fotogramas presenten el mismo ángulo θ' (en otras palabras, que R2' coincida siempre en la misma posición). Las siguientes fórmulas matemáticas sirven para calcular las coordenadas x e y de los puntos

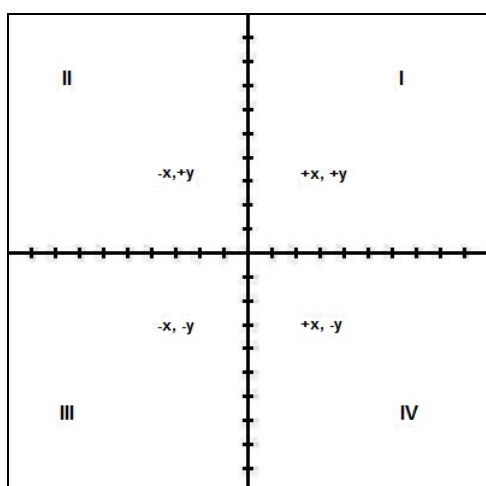


Figura 5.9. Cuadrantes (I-IV) de un sistema de coordenadas en función de los valores positivo o negativo de x e y .

Empezando en el cuadrante I (+x, +y), el resto de cuadrantes se distribuyen en sentido antihorario. Cada cuadrante abarca 90° sexagesimales o $\pi/2$ radianes.

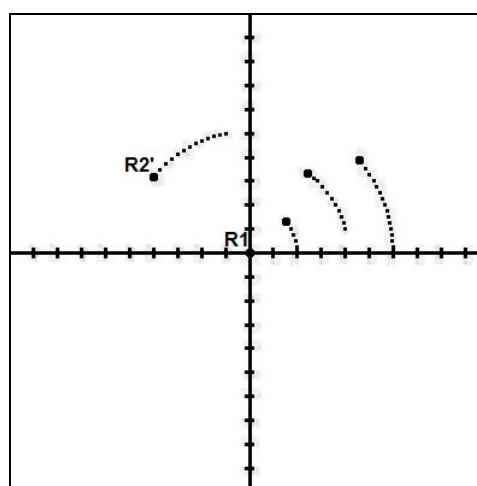


Figura 5.10. Rotación de todos los puntos sobre el centro de coordenadas (R1) en función del ángulo θ_{dif} calculado previamente.

Como resultado, R2 se reubica en la posición que tenía como objetivo, R2', que es la posición tomada como referencia del fotograma FR.

P' a partir de sus coordenadas originales P y el seno y el coseno del ángulo de rotación θ_{dif} . Si se quisiera restar θ_{dif} , se usaría $\sin(-\theta_{dif})$ y $\cos(-\theta_{dif})$, respectivamente.

$$(66) P'_x = P_x \cdot \cos(\theta_{dif}) - P_y \cdot \sin(\theta_{dif})$$

$$(67) P'_y = P_x \cdot \sin(\theta_{dif}) + P_y \cdot \cos(\theta_{dif})$$

Rotación del plano de la mordida para hacerlo horizontal

En el apartado anterior se ha explicado cómo se transforman los datos para fijar los puntos de referencia en una misma posición a lo largo de toda una sesión de grabación para el mismo sujeto. Es decir, es un procedimiento que ayuda a la interpretación intrasujeto. Ahora se verá cómo, junto a la fijación de R1 en el origen de coordenadas, la rotación del plano de la mordida para hacerlo horizontal contribuye a la comparabilidad entre sujetos, aunque hay que tener en cuenta, por supuesto, que siempre hay diferencias anatómicas insalvables, en el tamaño y forma exacta de la boca, o en la distancia que hay entre R1 y R2.

A partir del fotograma FR se puede calcular el ángulo α que forma el plano de la mordida con respecto a la horizontal en el sistema de coordenadas resultante de todas las traslaciones y rotaciones del apartado anterior. Para ello, se recurre de nuevo a la función arcotangente. Cuando se grabó el plano de la mordida, como se recordará, se tomaron los valores correspondientes a dos sensores pegados sobre la línea media de la tarjeta que había que morder. Llamemos M1 al sensor situado más atrás en la boca del sujeto y M2 al sensor más adelantado. En la representación gráfica del sistema de coordenadas, el sensor M1 queda más a la derecha. Por ello, el cálculo pertinente es el expresado en la siguiente fórmula:

(68) Cálculo del ángulo α

$$\alpha = \arctan \left(\frac{M1_y - M2_y}{M1_x - M2_x} \right)$$

Para saber el signo del ángulo con respecto a la horizontal y, por tanto, conocer la dirección en que se debe rotar, se debe ver el signo resultante de la diferencia $M2_y - M1_y$. Una vez obtenidos estos dos datos, se deben rotar todos los fotogramas de todas las pistas usando α en lugar de θ_{dif} en las fórmulas de rotación (66) y (67) del apartado anterior¹⁵⁸.

Conversión de las unidades de coordenadas en milímetros

El sistema de medición articulométrica expresa los resultados medidos en centésimas de milímetro. Dicho de otro modo, la precisión de medición del aparato es tal que el

¹⁵⁸ Nótese que en este caso no hace falta medir el ángulo α' para calcular la diferencia, puesto que el objetivo al que se quiere llegar es conocido de antemano y es un ángulo de 0° (la horizontal).

error intrínseco no llega más que a ± 0.01 mm (por supuesto, puede haber errores de medición de mayor magnitud por otras causas, pero no por la precisión del sistema).

El hecho de que el aparato mida en centésimas de milímetro es, pues, una característica bienvenida. Sin embargo, una vez registradas esas mediciones, y para facilitar la interpretación, se decidió convertir las medidas a una unidad con la que estamos más acostumbrados a trabajar: el milímetro.

Alineación temporal de los archivos de EMA con los datos acústicos y aerodinámicos

Como se adelantó en el § 5.4.3, el audio captado por el propio casco de EMA se usó para alinear temporalmente los archivos articulográficos con el resto de pistas de datos (audio, U_n , P_n)¹⁵⁹. Con este fin, se calculó el lapso temporal entre las dos versiones del audio tomando como referencia el instante en el que empezaba la primera barra de explosión de cada pista (Yo repito...), y ese lapso se restó o sumó —según correspondiera— al inicio de los correspondientes archivos de EMA.

5.6. Extracción de los datos de interés

Una vez realizadas las adaptaciones preliminares pertinentes, descritas en el anterior apartado, todas las pistas de datos fueron analizadas en Praat y se extrajeron los valores necesarios para poner a prueba las distintas hipótesis. Aunque algunas mediciones son muy directas, en otros casos conviene aclarar el método de cálculo seguido para así proporcionar una definición operativa de las distintas variables.

En el modelo de la dinámica de tareas, el movimiento de los articuladores normalmente sigue un patrón cuasi-sinusoidal a lo largo del tiempo. Esto se representa de forma idealizada en la Figura 5.11 de la p. 194, donde la línea negra continua puede representar la evolución temporal de la posición de un articulador, o del objetivo de constricción de un determinado gesto. Por ejemplo, puede pensarse que esa línea representa la variable *abertura labial*. En el punto 1, la abertura es máxima: puede corresponder al momento de consecución de un objetivo articulatorio propio de una vocal. En cambio, en el punto 3, la abertura es mínima: corresponde a una consonante, p. ej. [p]. A continuación, la abertura vuelve a aumentar de valor en la siguiente vocal (punto 5), y así sucesivamente.

Podemos calcular la derivada de esta señal para medir la velocidad del movimiento articulatorio (en la figura, la derivada corresponde a la línea de círculos blancos). Dada una señal digital, una manera habitual de calcular la derivada (*i. e.* el crecimiento o decrecimiento de la señal original) es restar a cada valor el de la muestra inmediatamente anterior, según la siguiente fórmula:

(69) Cálculo diferencial retrospectivo de primer orden

$$y[n] = x[n] - x[n - 1]$$

¹⁵⁹ Este ajuste es necesario debido al pequeño desfase producido entre ambos sistemas de grabación por la imposibilidad de pulsar el botón de grabar exactamente al mismo tiempo.

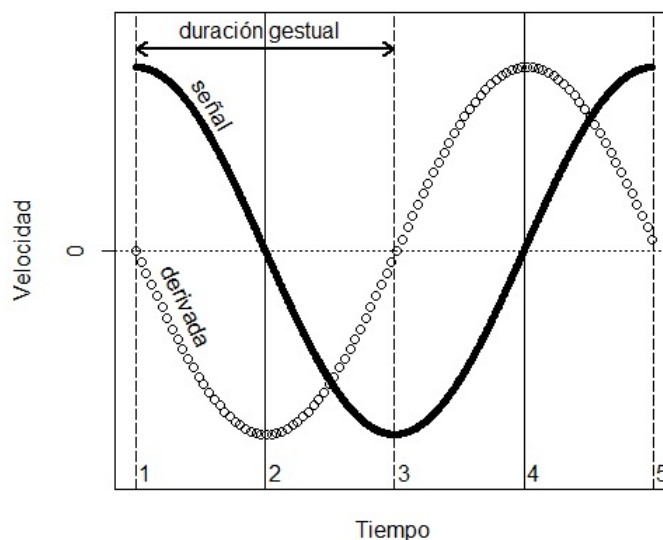


Figura 5.11. Representación idealizada de una señal articulográfica y su derivada.
Vid. la explicación en el texto para más detalles.

Con este método, a la primera muestra del vector original no se le puede restar la anterior (porque no hay anterior), y esto devuelve un vector de valores con una muestra menos en la señal derivada, lo que provoca un pequeño desajuste temporal con respecto a la señal primaria. Una posible solución es introducir una primera muestra con valor cero. Sin embargo, hay otra solución que parece preferible, porque además hace cálculos más estables del grado de variación de los datos a lo largo del tiempo (Harrington, 2010, p. 147 y ss.). La diferenciación centrada de tres puntos, *vid.* fórmula (70), se computa como la mitad de la diferencia entre el valor siguiente y el valor anterior al punto que se está calculando en cada momento. En este caso, se pierden dos valores (uno al principio, ya que no hay valor previo, y otro al final, ya que no hay valor subsiguiente), pero la señal derivada queda alineada temporalmente con la original. En esta tesis se optó por este método para calcular la velocidad de los distintos movimientos articulatorios.

(70) Cálculo diferencial centrado con una ventana de tres puntos

$$y[n] = \frac{1}{2} (x[n + 1] - x[n - 1])$$

Sigamos asumiendo que en nuestro ejemplo de la Figura 5.11 la señal representa la abertura labial. Los puntos 1, 3 y 5 representan los momentos correspondientes a los picos de magnitud del gesto. Dado que en esos puntos, momentáneamente, la señal ni crece ni decrece, el valor de la derivada es igual a cero. En los intervalos en que la señal decrece (p. ej. entre 1 y 3), la derivada presenta valores negativos; cuando la señal crece (entre 3 y 5), la derivada presenta valores positivos. Por otro lado, 2 y 4 se corresponden con los momentos de pico de velocidad del movimiento (con un mínimo y un máximo en la derivada, respectivamente): esos son los instantes de mayor decrecimiento o crecimiento en la señal original. En concreto, el punto 2 señala el momento en que se

alcanza la mayor velocidad en la constricción, mientras que el punto 4 refleja el momento de mayor velocidad en la distensión. Téngase en cuenta que una cosa es el momento en que se produce el pico de velocidad (su proyección sobre el eje x) y otra distinta es el propio valor del pico, que es el valor de la derivada en ese momento (o su proyección sobre el eje y)¹⁶⁰.

En el § 4.3.1 se explicó que la dinámica de tareas se modela como un sistema de muelles que tienden a una posición de reposo y que, al activarse el gesto siguiente, se establece un nuevo objetivo de constricción que desencadena una serie de movimientos encaminados a obtener una nueva posición de reposo. Por este motivo, se puede considerar que la duración de un gesto articulatorio abarca desde el momento de reposo anterior hasta el nuevo momento de reposo. En una secuencia como p. ej. [apa], la duración del gesto labial encaminado a conseguir la constricción propia de la [p] se calcula como la diferencia temporal entre el momento en que se alcanza ese mínimo de abertura y el momento del máximo correspondiente a la vocal previa (*i. e.* la duración del intervalo entre los puntos 1 y 3). A partir del punto 3 y hasta el punto 5, tenemos un gesto encaminado a la nueva abertura propia de la vocal siguiente (en ese sentido, el gesto de la distensión consonántica coincide con el de la formación de la vocal).

Esto significa que la extensión del gesto de cierre labial de la [p] (su duración articulatoria) no coincide con el intervalo que medimos como duración acústica, sino que la duración acústica está desfasada (es más tardía) con respecto a la duración articulatoria y, en cualquier caso, no tiene por qué presentar el mismo valor en milisegundos que esta. La duración acústica equivale a la suma de dos momentos: el intervalo en el que el valor de la abertura labial es suficientemente pequeño como para producir la oclusión que se traduce en silencio acústico, y el pequeño intervalo después de ese, ya durante el movimiento de apertura, que se corresponde con el ruido de la distensión.

Puesto que el pico de magnitud del gesto se obtiene en algún momento incluido en la duración acústica del segmento, el etiquetado basado en la segmentación acústica se utilizó para localizar en cada consonante el pico de magnitud de cada una de las pistas de datos articulográficos relevantes para las hipótesis establecidas en el § 5.1. En concreto, la coordenada y del sensor situado en la encía inferior se tomó como medición de la altura de la mandíbula y la coordenada y del sensor situado en la punta de la lengua sirvió para medir la altura del ápice. La abertura labial se calculó como la distancia euclidiana entre los sensores del labio superior y el labio inferior.

¹⁶⁰ Dado el método de cálculo de la derivada, los valores de esta en un momento determinado representan la magnitud del cambio entre muestras consecutivas —con las precisiones de la fórmula (70)—. Por ejemplo, un pico de velocidad de constricción de 0.4 significa un cambio de 0.4 mm por muestra. Teniendo en cuenta la frecuencia de muestreo de 250 Hz (*i. e.* 250 muestras por segundo), una sencilla regla de tres permite calcular la velocidad en términos de espacio recorrido en unidad de tiempo, relación que, seguramente, nos resulta más intuitiva ($0.4 * 250 = 100$ mm/s). Obviamente, esto no equivale a afirmar que los respectivos órganos articulatorios realmente lleguen a desplazarse 10 cm en un segundo, en primer lugar porque el gesto puede durar mucho menos tiempo, y en segundo lugar porque esa es solo la velocidad alcanzada en un instante (el del pico de velocidad), pero la velocidad del movimiento en otras fases del gesto es menor.

Además del valor (en milímetros, según el sistema de coordenadas EMA) de esos puntos, se localizó el instante temporal en que aparecía el pico de magnitud, y se calculó en qué momento proporcional se alcanzaba dicho punto, sobre el total de la duración acústica del segmento.

Ahora bien, no todas las señales articulográficas presentan el aspecto idealizado que se ha ilustrado en la Figura 5.11. En ocasiones, existen microoscilaciones que provocan varios puntos de cambio de tendencia dentro de un mismo segmento. En estos casos se tomó como referencia el punto de inflexión más cercano al centro acústico de la consonante (el centro de su *duración* acústica, se entiende). Otras veces, en cambio, no hay ni un solo punto de cambio de tendencia en todo el segmento. Esto puede ocurrir por influencia del contexto, por ejemplo, cuando la vocal previa presenta una tendencia que se mantiene durante la consonante y continúa en la misma dirección durante la vocal siguiente (como podría pasar con el movimiento de la mandíbula en una secuencia como [aβi]). Ante esta circunstancia, parece que cabrían dos soluciones. La primera de ellas implica reconocer que, puesto que no hay un cambio de tendencia, el pico de magnitud dentro de un segmento coincide con uno de sus extremos acústicos (inicial o final, según corresponda). No obstante, esto puede provocar medidas también extremas (p. ej. de la duración de un gesto), frente al caso general con inflexiones gestuales en torno al centro acústico de la consonante. En consecuencia, algunas de las diferencias detectadas estadísticamente podrían deberse en realidad a efectos contextuales, algo que no coincide con los intereses de este estudio. Una solución alternativa que permite controlar los efectos del contexto consiste en considerar, en esos casos sin cambio de tendencia, que el objetivo articulatorio se alcanza exactamente en el centro acústico del segmento. Se trata de una opción más conservadora, ya que, si acaso, tiende a diluir los efectos diferenciales que podrían tener algunas de las variables estudiadas, al situar las medidas de esos segmentos en el centro, punto en torno al cual se sitúan también los casos en que sí hay momento de inflexión. Es decir, puede que esta solución impida encontrar algunos efectos realmente presentes, pero garantiza que los efectos que sí se detecten efectivamente están ahí y no se deben al influjo del contexto. Por este motivo, esta fue la opción escogida aquí.

De manera análoga a la información articulográfica, en el caso de los datos aerodinámicos (U_n y P_n) se calculó el valor del pico de magnitud, *i. e.* el máximo de flujo y de presión alcanzados durante las consonantes nasales, y también el punto proporcional sobre la duración acústica en el que se alcanzaba ese pico.

En cuanto al VOT, se midió, como es habitual desde Lisker y Abramson (1964), el tiempo transcurrido desde el momento de la distensión de las oclusivas ([p] y [t] en nuestro caso), localizado en el espectrograma mediante la barra de explosión, hasta el momento en que empieza la vibración laríngea, detectada visualmente, también en el espectrograma, mediante claves como la barra de sonoridad, la estriación típica correspondiente a los pulsos glotales, y la curva de detección de la f_0 . Además, también se midió la proporción que ese lapso suponía sobre el total de la duración acústica de la consonante.

Por último, la intensidad relativa se midió, siguiendo a Carrasco *et al.* (2012), como la proporción entre el mínimo de intensidad de la consonante y el máximo de intensidad de la vocal siguiente. Para ello se usaron los parámetros de configuración recomendados como opciones por defecto por el manual de Praat. En cuanto al parámetro de la f_0 mínima esperada, se usó un valor de 75 Hz para los hombres y de 100 Hz para las mujeres.

5.7. Análisis estadístico

Las distintas hipótesis fueron puestas a prueba mediante el cálculo de modelos de regresión lineal de efectos mixtos —en adelante, me referiré a ellos como modelos mixtos, como suele hacerse para abreviar— (*vid.* p. ej. Baayen, 2008, cap. 7; Gelman y Hill, 2007, especialmente los caps. 11-13). Los modelos mixtos presentan numerosas ventajas con respecto al cálculo clásico de la ANOVA con medidas repetidas (*vid.* p. ej. Harrington, 2010, p. 22), y son más resistentes a los desequilibrios que se pueden producir cuando el número de observaciones no es exactamente el mismo en cada una de las condiciones determinadas por la combinación de los distintos niveles de las variables independientes.

Datos excluidos del análisis

En efecto, en el corpus usado para este estudio hay algunos casos de datos perdidos, correspondientes a segmentos que, tras un análisis visual en Praat, fueron descartados por presentar algún tipo de anomalía en alguna de las pistas de datos. El total de observaciones esperadas era de 768 (6 sujetos * 8 consonantes * 2 unidades [ω/σ] * 2 niveles de prominencia [$^1\sigma/\sigma$] * 2 palabras con esas características¹⁶¹ * 2 repeticiones de cada una de las combinaciones anteriores). El número de observaciones que quedó después de eliminar los datos perdidos fue de 746 (un 97.1 % del total).

Por otro lado, en el análisis de los datos aerodinámicos sobre la nasalidad, se excluyeron las observaciones correspondientes a David y a Eduardo, que, como se advirtió en el § 5.4.2, presentaban valores anormalmente cercanos a cero, posiblemente debido a una fuga de aire durante la grabación.

Puntualmente, para dos de los modelos, se descartaron de manera adicional las observaciones que presentaban valores matemáticamente anómalos, seguramente debidos a algún problema durante la extracción automática de los mismos. En el análisis del pico de P_n se eliminó 1 caso (de un total de 123 observaciones de segmentos nasales) por presentar un valor negativo. En el modelado de la velocidad de la distensión de la mandíbula se eliminaron 2 observaciones (de 746) por tener un valor igual a cero.

5.7.1. Efectos fijos

En cada uno de los modelos mixtos que se realizaron se incluyeron las siguientes variables intrasujetos como efectos fijos: Unidad (con dos niveles: ω/σ), Acento (con dos niveles: $^1\sigma/\sigma$) y Consonante (la identidad del segmento, con un número de niveles

¹⁶¹ *Vid.* nota 147 en la p. 181.

variable según el modelo concreto¹⁶²). El hecho de incluir la Consonante en el modelo permite controlar las diferencias inherentes entre las distintas consonantes, de modo que si los efectos de la Unidad o del Acento resultan significativos es porque realmente esas condiciones prosódicas alteran el comportamiento de las consonantes en general, modificando a cada una de ellas a partir de su línea de base.

Además de los efectos principales referidos, se incluyeron todas las interacciones bidireccionales entre esas variables. Por ejemplo, la interacción Unidad x Acento puede ayudar a descubrir efectos techo que impidan la acumulación de efectos de refuerzo en la condición inicial de palabra cuando coincida con sílaba acentuada. Las interacciones Consonante x Unidad y Consonante x Acento son interesantes porque pueden poner de manifiesto que una determinada consonante se vea más o menos afectada que las demás por alguna de las condiciones prosódicas estudiadas (recuérdese por ejemplo lo dicho acerca de la mayor resistencia de las fricativas en cuanto a los cambios de magnitud).

Las variables Unidad y Acento se codificaron siguiendo el esquema que opone un grupo de control a un tratamiento (codificación conocida habitualmente como *dummy coding*). En este caso, se toma uno de los niveles de cada variable como nivel de referencia y se estudia la diferencia que el otro nivel presenta con respecto a aquel. El nivel de referencia de Unidad es σ y el de Acento es σ ; así, cuando en el modelo se consigna el efecto de cada una de estas variables, en realidad se está plasmando el efecto diferencial que supone, frente a esas referencias, estar en posición inicial de palabra o estar en sílaba acentuada, respectivamente.

Por su parte, la variable Consonante se codificó mediante el sistema llamado de desviación con respecto a la media, de manera que el efecto de cada nivel se refiere a cuánto se aparta la media de ese nivel (*i. e.* de esa consonante) con respecto a la media de todas las medias¹⁶³. Como es habitual en los modelos de regresión, en el modelo no se introducen todos los niveles de cada variable, sino que siempre se deja uno fuera para evitar problemas de colinealidad (una vez introducidos $k - 1$ niveles, el efecto del otro es predecible). En efecto, aunque el cálculo del modelo no devuelve ninguna estimación del efecto del nivel de referencia, dado que en la codificación de este tipo todos los coeficientes deben sumar cero, es muy fácil calcular el dato que falta. En la presentación de los resultados en el próximo capítulo, este valor se incluirá entre paréntesis para recordar esta circunstancia.

Los modelos se calcularon mediante el programa de análisis estadístico R, con la función `lmer` del paquete `lmerTest` (Kuznetsova, Brockhoff y Christensen, 2012).

¹⁶² Según la variable que se intenta modelar en cada caso, tiene sentido incluir unas consonantes u otras en el análisis. Por ejemplo, el estudio de la duración acústica se aplica a las ocho consonantes del experimento, mientras que las hipótesis referidas a la nasalidad solo se ponen a prueba con las dos consonantes nasales. Lo mismo puede decirse de las oclusivas, fricativas y espirantes. Por último, los movimientos de la mandíbula se estudian en todas las consonantes, pero los datos del ápice de la lengua solo se modelan para las cuatro consonantes corales, mientras que la abertura labial solo se analiza en las cuatro consonantes labiales.

¹⁶³ Téngase en cuenta que, aunque los resultados suelen coincidir, no es lo mismo calcular la media sobre todas las observaciones que calcular la media para cada nivel y luego hacer la media entre esas medias de nivel (especialmente, pueden diferir en los casos en que no hay un número equilibrado de observaciones en cada grupo).

Para cada modelo se obtuvo la estimación β del efecto de cada predictor (*i. e.* cada nivel de cada variable, de los que entran en el modelo, pues siempre queda uno fuera, como se acaba de explicar). También se calculó el error estándar (SE) del coeficiente estimado y su valor *t*. Dado el elevado número de observaciones introducidas en cada modelo (siempre superior a 100), la distribución del valor *t* se aproxima a la distribución normal, de modo que se puede considerar que hay un efecto significativo de la inclusión de un predictor en el modelo siempre que su valor *t* sea superior a 1.96, valor que normalmente se redondea a 2 por simplificar (Baayen, 2008, p. 270).

Asimismo, se calculó mediante la función `anova` la significatividad de la inclusión en el modelo de cada variable, en global (*i. e.* no nivel por nivel), significatividad alcanzada con valores de *p* (la probabilidad de cometer un error tipo I o de falsa alarma) por debajo de 0.05. A diferencia del paquete `lme4` (Bates, 2005; Bates, Maechler y Bolker, 2012), el paquete `lmerTest` permite realizar un test *F* sobre la inclusión de las variables en el modelo, calculando los grados de libertad mediante la aproximación de Satterthwaite (por eso, como se verá, no siempre figura un número entero de grados de libertad).

5.7.2. Efectos aleatorios

En cada modelo se incluyó el efecto aleatorio de la variable Sujeto sobre el término constante de dicho modelo (*i. e.* la ordenada en el origen). De esta manera, se controlan las posibles diferencias que cada sujeto presente, como línea de base, en la producción de las distintas consonantes, y permite aislar el verdadero efecto de la Unidad y el Acento, que son las variables de interés en este estudio.

En cada caso, se calcularon también modelos alternativos en los que no solo se incluye el efecto del Sujeto sobre la constante, sino también sobre la pendiente de Unidad, sobre la pendiente de Acento, o sobre ambas. Estos efectos reflejarían el distinto comportamiento de unos sujetos frente a otros en lo que se refiere al efecto de las variables prosódicas. Por ejemplo, puede darse el caso de que unos sujetos experimenten un refuerzo consonántico en posición inicial de palabra con respecto a un determinado parámetro, pero otros sujetos no.

Para cada variable modelada, se probaron, por tanto, cuatro modelos (todos ellos con efecto aleatorio del Sujeto sobre la constante, y además con posibles efectos sobre una, otra o ambas pendientes). La bondad del ajuste de esos modelos a los datos se comparó mediante diversos tests chi cuadrado, y se descartaron los efectos aleatorios que no incrementaban significativamente dicho grado de ajuste. Además, eliminar dichos efectos previene la aparición de problemas de convergencia computacional en el cálculo de los modelos, algo relativamente habitual cuando se intenta forzar un cambio de pendiente por el efecto aleatorio de una variable que sea categórica (y Sujeto lo es).

No se incluyó nunca el efecto del Sujeto sobre la interacción Unidad x Acento, que, en cualquier caso, sería muy difícil de interpretar. También se excluyó explícitamente la interacción entre el efecto aleatorio sobre la constante y el efecto aleatorio sobre la pendiente. Se asumió que los distintos sujetos podían tener distintas líneas de base en la pronunciación de cada consonante (efecto sobre la constante) y que también podían

comportarse de manera diferente en cuanto al efecto de las distintas condiciones prosódicas (efecto sobre las pendientes), pero no parece que tenga sentido pensar que una distinta línea de base de un sujeto condicione un distinto comportamiento en cuanto al refuerzo por causa prosódica.

Dicho esto, la formulación en R del modelo más amplio posible (*i. e.* con todos los efectos aleatorios) quedaría como en (71), que se lee algo así como «modelo lineal de efectos mixtos de la variable modelada en función de Unidad, Acento y Consonante, y sus interacciones hasta el grado de bidireccionales, más el efecto aleatorio del Sujeto sobre la constante, más el efecto aleatorio del Sujeto sobre la Unidad descartando explícitamente la interacción con la constante [mediante el 0], más el efecto aleatorio del Sujeto sobre el Acento, también sin interacción con la constante».

(71) `lmer (variable modelada ~ (Unidad + Acento + Consonante)^2 + (1|Sujeto) +
(0 + Unidad|Sujeto) + (0 + Acento|Sujeto))`

Por último, debe decirse que la variable Sujeto, al igual que Consonante, se codificó mediante el esquema de desviación con respecto a la media.

6. Resultados

Como se ha explicado en el § 5.7.2, la pertinencia de la inclusión de efectos aleatorios del Sujeto sobre las pendientes de Unidad y Acento se valoró comparando el grado de adecuación a los datos de dichos modelos con el del modelo básico (aquel que tan solo incluye el efecto aleatorio sobre el término constante). La Tabla 6.1 indica el resultado del test χ^2 con el que se realizó dicha comparación, y señala en qué casos se mantuvo el efecto sobre las pendientes (los números de los modelos corresponden a la secuencia de presentación en este capítulo y serán explicitados para cada caso en su debido momento). En los modelos sin efecto sobre las pendientes, la comparación que aparece en la tabla es entre el modelo básico y el modelo con las dos posibles pendientes aleatorias (por eso el número de grados de libertad es mayor).

Tabla 6.1. Resultados de la comparación χ^2 entre los modelos con efecto aleatorio del Sujeto sobre la pendiente (y la constante) y el modelo básico con efecto solo sobre la constante.

	df	χ^2	p		Pendientes afectadas por el efecto aleatorio del Sujeto
Modelo 1	6	52.79	p < 0.001	***	Unidad + Acento
Modelo 2	3	19.47	p < 0.001	***	Unidad
Modelo 3	6	19.39	p < 0.01	**	Unidad + Acento
Modelo 4	6	7.82	p = 0.251	ns	
Modelo 5	6	0.62	p = 0.996	ns	
Modelo 6	6	7.63	p = 0.267	ns	
Modelo 7	6	9.11	p = 0.167	ns	
Modelo 8	6	3.40	p = 0.757	ns	
Modelo 9	3	14.38	p < 0.01	**	Unidad
Modelo 10	3	11.56	p < 0.01	**	Unidad
Modelo 11	3	8.35	p < 0.05	*	Acento
Modelo 12	3	24.84	p < 0.001	***	Unidad
Modelo 13	3	104.26	p < 0.001	***	Unidad
Modelo 14	6	5.71	p = 0.457	ns	
Modelo 15	6	3.17	p = 0.788	ns	
Modelo 16	6	0.16	p = 1	ns	
Modelo 17	6	1.32	p = 0.970	ns	
Modelo 18	6	0.87	p = 0.990	ns	
Modelo 19	6	2.30	p = 0.890	ns	
Modelo 20	6	7.02	p = 0.319	ns	
Modelo 21	6	8.07	p = 0.233	ns	
Modelo 22	6	1.29	p = 0.972	ns	
Modelo 23	6	5.92	p = 0.433	ns	
Modelo 24	6	3.87	p = 0.694	ns	
Modelo 25	6	4.02	p = 0.674	ns	
Modelo 26	6	2.64	p = 0.853	ns	
Modelo 27	6	7.73	p = 0.259	ns	
Modelo 28	3	12.97	p < 0.01	**	Acento
Modelo 29	3	12.95	p < 0.01	**	Unidad
Modelo 30	6	1.08	p = 0.982	ns	
Modelo 31	3	8.63	p < 0.05	*	Acento
Modelo 32	6	2.76	p = 0.838	ns	

(df = grados de libertad; ns = no significativo)

6.1. H₁: pico de velocidad de la constricción

En este apartado se expondrán los resultados de los modelos conducentes a poner a prueba la hipótesis sobre el pico de velocidad de la constricción (H_{1A}), que predecía una menor velocidad en posición inicial de palabra que de sílaba. El modelo M₁ comprueba el efecto de la posición prosódica sobre la velocidad del movimiento de la mandíbula; M₂ hace lo propio con la velocidad del movimiento del ápice; y M₃ se refiere a la velocidad del movimiento de los labios.

La Tabla 6.2 recoge la media y la desviación estándar de las distintas observaciones sobre el pico de velocidad de la mandíbula y muestra, efectivamente, un promedio de velocidad menor en ω que en σ . Sin embargo, para dilucidar si la diferencia es significativa, incluso una vez descontado el posible efecto de los otros factores, es necesario interpretar los datos del modelo M₁ (Tabla 6.3).

La estimación β de la ordenada en el origen (el término constante del modelo) es el valor del pico de velocidad de la constricción cuando todas las variables adoptan su nivel de referencia. Teniendo en cuenta la codificación de las variables (*vid.* § 5.7), 0.20 es el valor del pico de velocidad promediado para todos los sujetos y todas las consonantes cuando estas se encuentran en posición inicial de sílaba (no acentuada).

Tabla 6.2. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de la mandíbula correspondiente a la fase de constricción (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.08 (0.06)	0.13 (0.09)	0.11 (0.08)	
No acento	0.08 (0.06)	0.20 (0.14)	0.14 (0.12)	
TOTAL	0.08 (0.06)	0.17 (0.12)	0.12 (0.10)	n = 746 (n = n.º de observaciones)

Tabla 6.3. M₁: pico de velocidad del movimiento de la mandíbula (fase de constricción).

Las β (no estandarizadas) de cada modelo se expresan en la unidad de medida propia de la variable modelada (en este caso, mm/muestra).

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.20	0.03	7.56			
Unidad (Sujeto)	-0.12	0.02	-6.06	(1, 5.1) = 20.35	p < 0.01	**
Acento (Sujeto)	-0.07	0.01	-5.69	(1, 5.1) = 10.26	p < 0.05	*
Consonante				(7, 707.7) = 11.74	p < 0.001	***
β	0.04	0.01	3.51			
∂	-0.08	0.01	-6.45			
f	0.05	0.01	4.12			
m	0.08	0.01	6.45			
n	-0.06	0.01	-5.60			
p	-0.01	0.01	-0.55			
s	-0.02	0.01	-1.62			
t	(0.00)					
Unidad x Acento				(1, 707.7) = 38.90	p < 0.001	***
Consonante x Unidad				(7, 707.7) = 8.17	p < 0.001	***
Consonante x Acento				(7, 707.7) = 7.82	p < 0.001	***

Al lado de cada coeficiente β se consigna su error estándar (SE) y la t , que calcula, en unidades de desviación estándar, cuán alejada está la β de cero. Una t con un valor absoluto superior a 2 se puede interpretar como que la estimación del coeficiente es significativamente distinta de cero y, por tanto, dicho coeficiente realiza una aportación significativa al modelo.

La estimación del efecto de la Unidad (con un coeficiente β de -0.12) comporta que, si se mantienen las demás condiciones iguales pero en vez de considerar el nivel de referencia (la posición inicial de sílaba) tomamos la inicial de palabra, entonces se predice un valor del pico de velocidad que decrece con respecto al término constante ($0.20 - 0.12 = 0.08$). Una ANOVA, con los grados de libertad¹⁶⁴ y el valor calculado de F indicados en la tabla, confirma que la probabilidad de equivocarse si afirmamos que la Unidad tiene un efecto significativo es menor que 0.01 (*i. e.* menor que el 1%). Por tanto, podemos rechazar la hipótesis nula de que la Unidad no tiene efecto, y aceptar la hipótesis alternativa de que sí lo tiene. El efecto, además, actúa en la dirección esperada; es decir, el pico de velocidad de la constricción es significativamente menor en posición inicial de palabra que en posición inicial de sílaba (si el resto de las condiciones se mantienen constantes). Debe señalarse que el efecto de Unidad es significativo incluso teniendo en cuenta que los distintos sujetos se comportan de manera distinta con respecto a esta variable (*vid.* nota 164).

La H_1 no incluía ninguna predicción sobre el efecto de la prominencia sobre la velocidad de la constricción. Sin embargo, el modelo refleja un efecto significativo de la sílaba acentuada, en la misma dirección que el de la posición inicial de palabra: la deceleración. La magnitud del efecto del Acento (-0.07), aun siendo significativa, es menor que la de la Unidad (-0.12).

El test F sobre la variable Consonante muestra que hay una diferencia intrínseca de la velocidad de constricción entre los distintos tipos de segmento consonántico estudiados. Si se observan con detenimiento los coeficientes β estimados para cada nivel de esta variable, se comprueba que, en general, las labiales son más rápidas (tienen valores positivos de β , que se suman a la velocidad de partida, 0.20) que las corales; la $[p]$ tiene un coeficiente negativo, pero tan solo ligeramente y, en cualquier caso, tiene un valor mayor que el de las corales. Tiene sentido pensar que, en las labiales, un movimiento más rápido de la mandíbula contribuya a la formación del gesto de constricción labial, dado el gran nivel de acoplamiento entre los movimientos de los labios y los de la mandíbula. Como se advirtió en el § 5.7, en el modelo tan solo se incluyen todos los niveles menos uno, para cada variable. Por eso no hay un valor de SE ni de t para la $[t]$, aunque su coeficiente β se puede calcular, ya que la suma de todos los coeficientes debe ser igual a cero. Puesto que todos expresan la diferencia entre la media de cada nivel y la gran media, si se toman en su conjunto, esa diferencia es cero.

¹⁶⁴ El M_1 es uno de los modelos que no solo incluyen el efecto aleatorio del Sujeto sobre la constante, sino también sobre la pendiente de Unidad y la de Acento, como se recuerda con la mención, entre paréntesis, «([Sujeto])» al lado de cada uno de esos factores. En cualquier caso, debe quedar claro que la información consignada en la tabla corresponde al efecto principal de la Unidad o el Acento, una vez controlada la variación entre sujetos. El número de grados de libertad calculado para estos factores es más reducido que el de los demás, también como consecuencia de haber tenido en cuenta el efecto aleatorio del Sujeto.

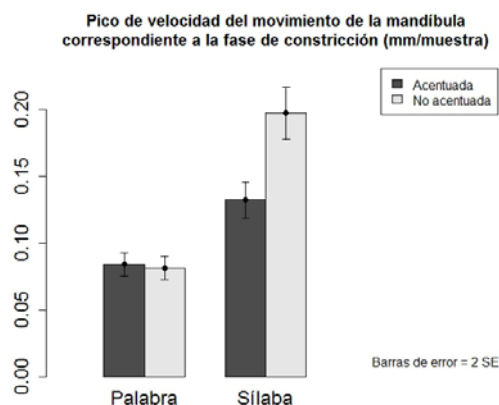


Figura 6.1. M_1 : Interacción Unidad x Acento en el pico de velocidad del movimiento de la mandíbula (fase de constricción).

En cuanto a las interacciones, se han omitido los coeficientes, que habría que desglosar para cada una de las posibles combinaciones entre los niveles de las variables implicadas, ya que eso alargaría considerablemente las tablas. Se aporta, en cambio, el test F sobre la inclusión en el modelo de la interacción, tomada en su conjunto. Cuando alguna de las interacciones resulte de especial interés, se presentarán sus datos de forma gráfica, que suele ser mucho más fácil de interpretar.

En general, distinguiremos tres tipos de interacciones, que llamaré «patrón cruzado», «efecto techo» y «diferencia de magnitud». En un patrón cruzado puede ocurrir, por ejemplo, que el efecto del Acento en posición inicial de sílaba vaya en una dirección pero en posición inicial de palabra vaya justo en dirección contraria. Otra manifestación típica del patrón cruzado es que uno de los factores prosódicos (bien la Unidad o bien el Acento) afecte a unas consonantes en una dirección y a otras consonantes en dirección contraria. En una interacción de efecto techo podemos encontrar, por ejemplo, que tanto la Unidad como el Acento modifican el comportamiento de un determinado parámetro, ambos en la misma dirección, pero que sus efectos no se acumulan. De este modo parece que uno de los factores prosódicos pierde su efectividad cuando el otro está activo. También podemos asimilar a este tipo de interacción de efecto techo aquellos casos en que la variable de estudio cambia bajo un determinado factor prosódico en algunas consonantes pero no en otras. El tercer tipo de interacción es aquel en que un efecto se manifiesta siempre como un cambio en la misma dirección, pero lo hace con diferente magnitud en unas condiciones que en otras (p. ej. afecta más a unas consonantes que a otras).

Por ejemplo, la Figura 6.1 muestra la interacción Unidad x Acento en el M_1 y es un ejemplo de efecto techo. En ella se aprecia claramente que el movimiento de ascenso de la mandíbula es mucho más lento cuando la consonante está en posición inicial de palabra que en interior de palabra. Por otro lado, la velocidad también tiende a ser menor en sílaba acentuada que en sílaba no acentuada. Sin embargo, se observa un límite en cuánto puede aminorarse la velocidad, ya que, en el caso de la posición inicial de palabra, el carácter acentuado de la sílaba no acumula su efecto sobre el factor

posicional. Gracias a esta figura también se hace evidente que la magnitud del efecto del Acento es menor que la del efecto de la Unidad. Las barras de error reflejan 2 unidades de error estándar; así, si las barras correspondientes a dos columnas no se solapan, se puede interpretar como una diferencia significativa entre las medias de los grupos a un nivel $\alpha = 0.05$.

Por su parte, la Figura 6.2 ilustra el diferente efecto del Acento sobre unas consonantes u otras. En general, todas las consonantes (excepto la [t]) experimentan un descenso en el pico de velocidad en sílaba acentuada, aunque las más afectadas son las labiales, especialmente la [β] y la [f]. Se trata, fundamentalmente, de una interacción de diferencia de magnitud, aunque hay un patrón cruzado en la [t].

Podemos analizar los efectos aleatorios del Sujeto sobre las pendientes de Unidad o Acento como interacciones Sujeto x Unidad o Sujeto x Acento, respectivamente. La Figura 6.3 es un reflejo del distinto comportamiento de cada sujeto en cuanto al efecto de la Unidad. Aunque todos los informantes presentan la misma tendencia (a la deceleración en posición inicial de palabra), el efecto es mucho más acusado en algunos de ellos, mientras que se presenta más tenuemente en otros como, por ejemplo, en el caso de Lourdes. En cualquier caso, la diferencia entre la posición ω y la posición σ es significativa. Esta figura ilustra un caso «puro» de diferencia de magnitud.

El M_1 se completa con interacciones de diferencia de magnitud Consonante x Unidad (la [ð] se ve menos afectada por la ralentización en inicio de palabra) y Sujeto x Acento (en dos sujetos la condición 'σ ejerce una ralentización más acusada, en otros dos algo más intermedia, y en los dos últimos, prácticamente insignificante).

Pasemos ahora a la velocidad del movimiento del ápice de la lengua (modelo M_2). En la Tabla 6.4 se resume la estadística descriptiva correspondiente a dicho pico de velocidad (en las consonantes corales). La Tabla 6.5 muestra el valor de los coeficientes y la significatividad de los distintos predictores introducidos en el modelo, y deja ver las diferencias intrínsecas de velocidad entre consonantes: [ð] y [s] (*i. e.* las consonantes en las que no debe llegar a haber un contacto) presentan una velocidad menor a la media, mientras que [n] y [t] desarrollan una velocidad mayor a la media. En cambio, no se encuentra ningún efecto de la posición prosódica ni de la prominencia. Puede que la posición no llegue a ser significativa precisamente porque existe un patrón cruzado en el comportamiento de los distintos sujetos (en todos tiende a disminuir la velocidad en ω , pero para Lourdes aumenta significativamente en esa posición).

Por su parte, la Tabla 6.6 y la Tabla 6.7 se refieren a la velocidad del movimiento de constricción labial (M_3). Al igual que ocurría en el M_2 , existe un efecto principal de la identidad de la Consonante. Las consonantes en las que debe llegar a haber un contacto completo entre los órganos articulatorios (en este caso, los labios) —[m] y [p]— presentan una mayor velocidad que el promedio de todas las medias grupales, correspondientes a [β], [f], [m] y [p]. No obstante, esa diferencia con respecto a la gran media no es significativa en el caso de la [m] (el valor *t* asociado a su coeficiente es menor que 1.96). En cambio, las consonantes en las que no llega a haber un contacto articulatorio completo —[β] y [f]— alcanzan un menor pico de velocidad en el movimiento de constricción labial.

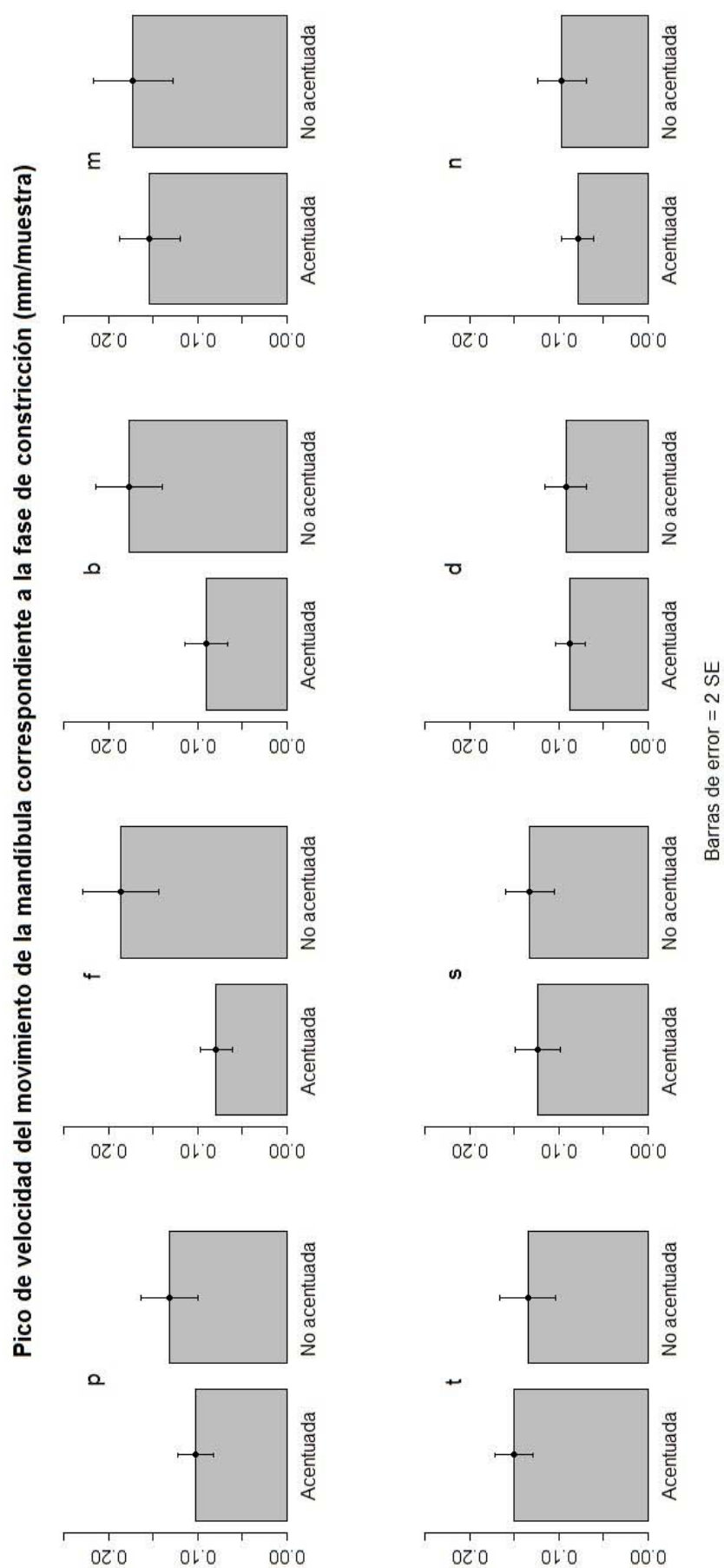


Figura 6.2. M_I: Interacción Consonante x Acento en el pico de velocidad del movimiento de la mandíbula (fase de constricción).

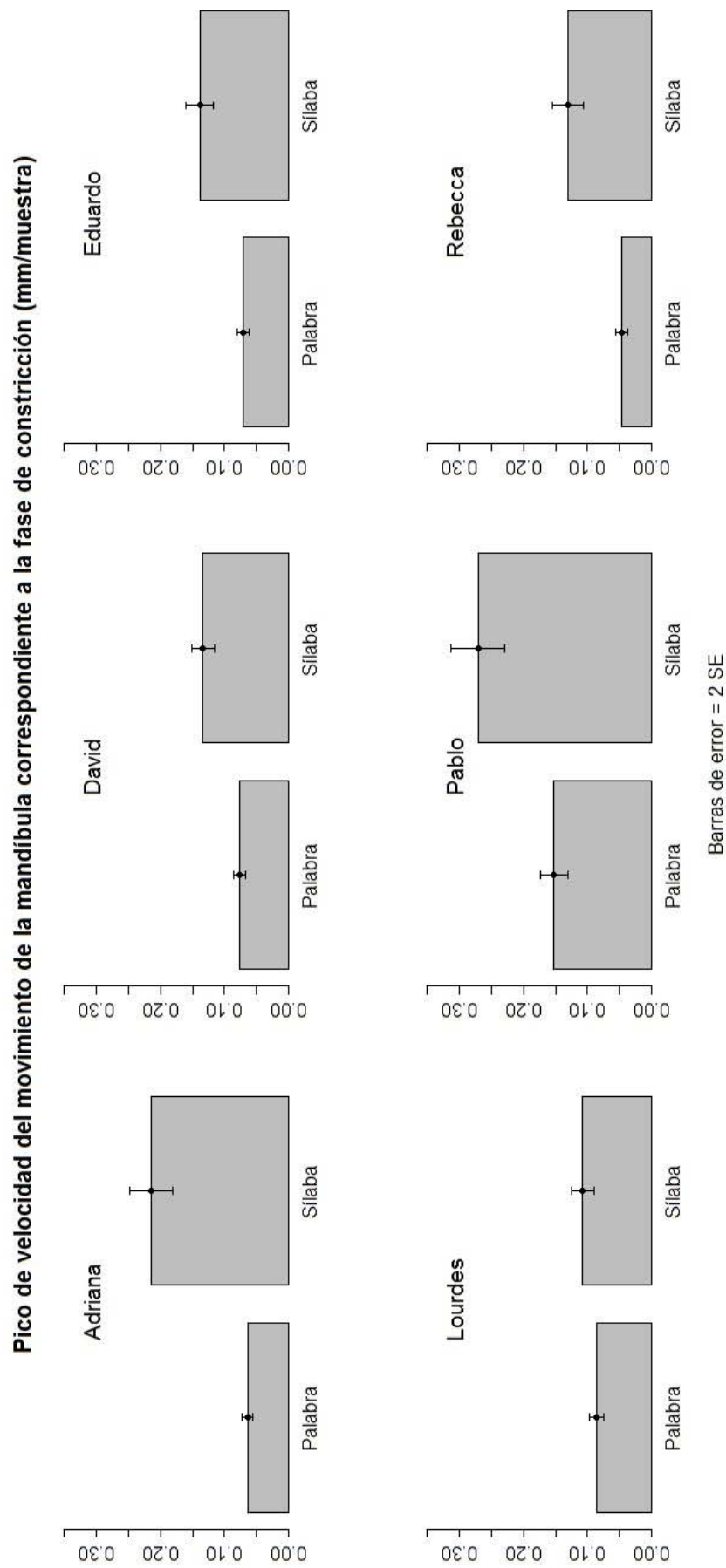


Figura 6.3. M_1 : Efecto aleatorio del Sujeto sobre la pendiente de Unidad en el pico de velocidad de la constricción correspondiente al movimiento de la mandíbula.

Tabla 6.4. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento del ápice correspondiente a la fase de constricción (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.39 (0.27)	0.42 (0.29)	0.41 (0.28)	
No acento	0.37 (0.25)	0.43 (0.34)	0.40 (0.30)	
TOTAL	0.38 (0.26)	0.43 (0.32)	0.40 (0.29)	n = 380

Tabla 6.5. M₂: Pico de velocidad del movimiento del ápice (fase de constricción).

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.43	0.05	8.01			
Unidad (Sujeto)	-0.06	0.07	-0.90	(1, 5) = 0.54	p = 0.495	ns
Acento	-0.01	0.03	-0.04	(1, 356.9) = 0.29	p = 0.592	ns
Consonante				(3, 356.9) = 49.57	p < 0.001	***
ø	-0.16	0.03	-4.74			
n	0.09	0.03	2.68			
s	-0.10	0.03	-2.75			
t	(0.17)					
Unidad x Acento				(1, 356.9) = 0.35	p = 0.553	ns
Consonante x Unidad				(3, 356.9) = 1.29	p = 0.278	ns
Consonante x Acento				(3, 356.9) = 2.58	p = 0.053	ns

Tabla 6.6. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de los labios correspondiente a la fase de constricción (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.37 (0.21)	0.57 (0.31)	0.47 (0.28)	
No acento	0.38 (0.22)	0.75 (0.28)	0.57 (0.31)	
TOTAL	0.38 (0.22)	0.66 (0.31)	0.52 (0.30)	n = 366

Tabla 6.7. M₃: Pico de velocidad del movimiento de los labios (fase de constricción).

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.75	0.08	9.61			
Unidad (Sujeto)	-0.37	0.05	-8.02	(1, 5.2) = 47.64	p < 0.001	***
Acento (Sujeto)	-0.18	0.04	-4.50	(1, 5.5) = 8.06	p < 0.05	*
Consonante				(3, 343.1) = 42.23	p < 0.001	***
β	-0.07	0.03	-2.32			
f	-0.10	0.03	-3.62			
m	0.05	0.03	1.77			
p	(0.12)					
Unidad x Acento				(1, 343.1) = 16.93	p < 0.001	***
Consonante x Unidad				(3, 343.1) = 1.46	p = 0.225	ns
Consonante x Acento				(3, 342.1) = 5.05	p < 0.01	**

Descontando el efecto de las diferencias entre consonantes, en el M₃ se constata que existe una diferencia significativa entre ω y σ , y entre ${}^l\sigma$ y σ . El refuerzo posicional lleva a una menor velocidad durante el gesto de formación de la consonante, tal y como predecía la H_{1A}. El refuerzo por prominencia afecta a la dinámica gestual en el mismo sentido.

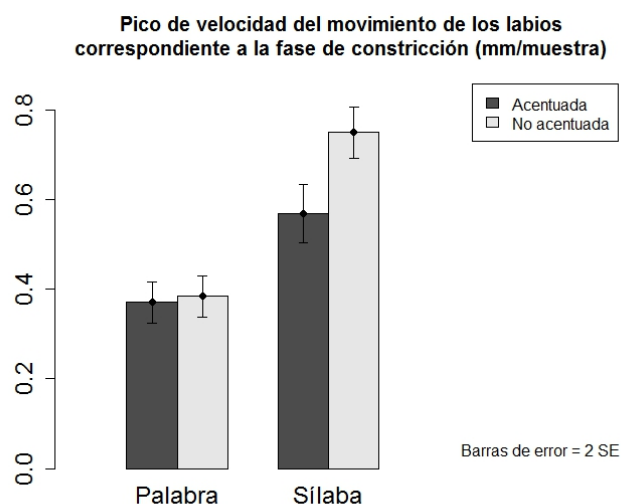


Figura 6.4. M₃: Interacción Unidad x Acento en el pico de velocidad del movimiento de los labios (fase de constricción).

Además, hay una interacción significativa entre la Consonante y el Acento (la [f] es la más afectada por la ralentización en 'σ), y otra entre la Unidad y el Acento. Esta última apunta a un efecto techo en la misma dirección que la que ya se observó para la velocidad del movimiento de la mandíbula un poco más arriba. En general, las consonantes labiales se producen con un menor pico de velocidad en el gesto de constricción cuando aparecen en posición inicial de palabra. En esta posición prosódica, no existe ninguna diferencia significativa en la velocidad asociada al carácter acentuado o no acentuado de la sílaba. En cambio, en interior de palabra, las consonantes se articulan con una velocidad significativamente menor cuando pertenecen a una sílaba prominente (Figura 6.4).

El modelo incluye, por último, el efecto aleatorio del Sujeto sobre ambas pendientes (Unidad y Acento), dada la existencia de diferencias de magnitud del efecto de dichos factores prosódicos entre unos sujetos y otros, aunque todos ellos apuntan en la misma dirección.

6.2. H₂: pico de velocidad de la distensión

La H_{2A} predecía una menor velocidad en el gesto de distensión de las consonantes situadas en posición inicial de palabra (en comparación con aquellas consonantes que tan solo son iniciales de sílaba en la jerarquía prosódica). En cambio, se espera que el refuerzo por prominencia actúe en sentido contrario, con una mayor velocidad en la distensión de las consonantes que pertenecen a sílabas acentuadas (H_{2B}). Es hora de poner a prueba estas hipótesis, tanto para el movimiento de la mandíbula (M₄), como el del ápice (M₅) y el de los labios (M₆).

Los datos correspondientes al modelo M₄ se presentan en la Tabla 6.8 y la Tabla 6.9. El efecto de Consonante es significativo. Si se observan los coeficientes estimados para cada uno de los niveles de esta variable, se puede ver que la [s] y la [t] son los sonidos

en los que la mandíbula alcanza una mayor velocidad durante la fase distensiva, mientras que [ð] y [n] son los más lentos. Todas las labiales presentan valores más cercanos a la gran media, y no difieren significativamente de esta, como refleja el cálculo de sus correspondientes *t* de Student.

Por otro lado, se encuentra un efecto significativo de la Unidad, tal que la velocidad distensiva de la mandíbula decrece en posición inicial de palabra, como se esperaba (H_{2A}). El Acento, también conforme a las expectativas (H_{2B}), se revela como un factor significativo que incrementa la velocidad del movimiento de distensión. Ninguna de las interacciones puestas a prueba alcanza el nivel requerido de significatividad.

En cuanto al movimiento del ápice (M_5), la Tabla 6.11 revela la existencia de diferencias intrínsecas a los distintos tipos de consonante, que resultan significativas. Los coeficientes β calculados para cada una de esas consonantes muestran que el gesto de la distensión de las consonantes que no implican contacto articulatorio total se desarrolla a una menor velocidad, aunque solo difieren significativamente de la gran media la [ð] y la [n]. Esto recuerda a la misma tendencia observada en la fase de constricción.

No se detecta, sin embargo, que la posición prosódica o la prominencia modifiquen significativamente la línea de base de cada consonante. En cambio, sí resulta significativa la interacción Unidad x Acento. Los datos presentados en la Tabla 6.10 permiten apreciar un patrón cruzado: la condición de sílaba acentuada incrementa la

Tabla 6.8. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad del movimiento de la mandíbula correspondiente a la fase de distensión (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.14 (0.15)	0.17 (0.14)	0.16 (0.15)	
No acento	0.10 (0.12)	0.12 (0.10)	0.11 (0.11)	
TOTAL	0.12 (0.13)	0.15 (0.13)	0.13 (0.13)	n = 744

Tabla 6.9. M_4 : Pico de velocidad del movimiento de la mandíbula (fase de distensión).

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.12	0.02	5.62			
Unidad	-0.02	0.01	-2.17	(1, 713.4) = 13.16	p < 0.001	***
Acento	0.05	0.01	4.32	(1, 713.4) = 30.04	p < 0.001	***
Consonante				(7, 713.4) = 21.61	p < 0.001	***
β	-0.03	0.02	-1.72			
ð	-0.04	0.02	-2.41			
f	-0.03	0.02	-1.84			
m	0.01	0.02	0.64			
n	-0.05	0.02	-2.88			
p	0.02	0.02	0.96			
s	0.07	0.02	3.57			
t	(0.05)					
Unidad x Acento				(1, 713.4) = 0.34	p = 0.560	ns
Consonante x Unidad				(7, 713.4) = 1.77	p = 0.090	ns
Consonante x Acento				(7, 713.4) = 0.99	p = 0.436	ns

Tabla 6.10. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad de la distensión correspondiente al gesto del ápice (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.40 (0.28)	0.33 (0.23)	0.37 (0.26)	
No acento	0.29 (0.24)	0.37 (0.28)	0.33 (0.27)	
TOTAL	0.35 (0.27)	0.35 (0.26)	0.35 (0.26)	n = 380

Tabla 6.11. M₅: Pico de velocidad de la distensión correspondiente al gesto del ápice.

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.37	0.05	7.50			
Unidad	-0.08	0.03	-2.50	(1, 361.9) = 0.07	p = 0.798	ns
Acento	-0.04	0.03	-1.13	(1, 361.9) = 2.81	p = 0.095	ns
Consonante				(3, 361.9) = 30.31	p < 0.001	***
ð	-0.16	0.03	-4.76			
n	0.12	0.03	3.85			
s	-0.06	0.03	-1.83			
t	(0.10)					
Unidad x Acento				(1, 361.9) = 10.64	p < 0.01	**
Consonante x Unidad				(3, 361.9) = 0.48	p = 0.695	ns
Consonante x Acento				(3, 361.9) = 0.73	p = 0.538	ns

Tabla 6.12. Media (y desviación estándar), por condiciones prosódicas, del pico de velocidad de la distensión correspondiente al gesto labial (mm/muestra).

	Palabra	Sílaba	TOTAL	
Acento	0.56 (0.26)	0.59 (0.25)	0.58 (0.26)	
No acento	0.47 (0.21)	0.50 (0.21)	0.48 (0.21)	
TOTAL	0.51 (0.24)	0.54 (0.24)	0.53 (0.24)	n = 366

Tabla 6.13. M₆: Pico de velocidad de la distensión correspondiente al gesto labial.

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.50	0.05	10.03			
Unidad	-0.03	0.03	-1.29	(1, 347.9) = 3.34	p = 0.069	ns
Acento	0.09	0.03	3.61	(1, 347.9) = 25.24	p < 0.001	***
Consonante				(3, 347.9) = 24.14	p < 0.001	***
β	-0.07	0.03	-2.44			
f	-0.07	0.03	-2.31			
m	0.07	0.03	2.40			
p	(0.07)					
Unidad x Acento				(1, 347.9) = 0.00	p = 0.977	ns
Consonante x Unidad				(3, 347.9) = 0.54	p = 0.656	ns
Consonante x Acento				(3, 347.9) = 0.58	p = 0.626	ns

velocidad de distensión de las consonantes situadas en inicio de ω , pero la disminuye cuando la consonante es inicial de σ .

Finalmente, el M₆ se refiere a la velocidad distensiva del gesto labial (Tabla 6.12 y Tabla 6.13). Al igual que ocurría en el gesto de constricción, existen diferencias intrínsecas entre las consonantes cuyo objetivo es un contacto total ([p] y [m]), que se

desarrollan a mayor velocidad) y aquellas que tan solo apuntan a un estrechamiento del canal de salida del aire ([β] y [f], más lentas).

El efecto principal de la Unidad no es significativo, aunque apunta en la dirección esperada (una aminoración de la velocidad). En cambio, sí es significativo el efecto principal del Acento, también en la dirección esperada según la hipótesis H_{2B}, pues la pertenencia de la consonante a una sílaba acentuada provoca un aumento en la velocidad de separación de los labios en el gesto que lleva a la vocal siguiente. Teniendo en cuenta que la β del término constante del modelo —la ordenada en el origen— es 0.50 y que el coeficiente estimado para el Acento es 0.09, el aumento de velocidad provocado por el carácter acentuado de la sílaba supone un incremento del 18% con respecto a la media de todos los sujetos y todas las consonantes cuando la consonante aparece en condiciones similares pero en sílaba no acentuada.

Ninguna de las interacciones resulta significativa.

6.3. H₃: duración del gesto de constricción

La H₃ predice una mayor duración de los gestos correspondientes a la fase de constricción de las consonantes por efecto del refuerzo posicional. Debemos comprobar esta hipótesis con respecto al movimiento de la mandíbula (M₇), el del ápice (M₈) y el de los labios (M₉).

Tabla 6.14. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de la mandíbula correspondiente a la fase de constricción.

	Palabra	Sílaba	TOTAL	
Acento	0.094 (0.041)	0.099 (0.035)	0.097 (0.038)	
No acento	0.087 (0.047)	0.112 (0.038)	0.100 (0.045)	
TOTAL	0.090 (0.044)	0.106 (0.037)	0.098 (0.042)	n = 746

Tabla 6.15. M₇: Duración del movimiento de la mandíbula (fase de constricción).

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.112	0.010	10.97			
Unidad	-0.025	0.003	-8.86	(1, 713.5) = 54.03	p < 0.001	***
Acento	-0.013	0.003	-4.86	(1, 713.5) = 3.05	p = 0.081	ns
Consonante				(7, 713.5) = 38.22	p < 0.001	***
β	-0.010	0.005	-2.04			
ø	-0.025	0.005	-5.21			
f	0.027	0.005	5.68			
m	-0.004	0.005	-0.81			
n	-0.028	0.004	-6.24			
p	0.005	0.005	0.99			
s	0.028	0.005	5.79			
t	(0.007)					
Unidad x Acento				(1, 713.5) = 25.84	p < 0.001	***
Consonante x Unidad				(7, 713.5) = 2.25	p < 0.05	*
Consonante x Acento				(7, 713.5) = 2.03	p < 0.05	*

Tabla 6.16. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento del ápice de la lengua correspondiente a la fase de constricción.

	Palabra	Sílaba	TOTAL	
Acento	0.107 (0.040)	0.105 (0.038)	0.106 (0.039)	
No acento	0.092 (0.040)	0.117 (0.039)	0.105 (0.041)	
TOTAL	0.100 (0.041)	0.111 (0.039)	0.105 (0.040)	n = 380

Tabla 6.17. M₈: Duración del movimiento del ápice de la lengua (fase de constricción).

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.117	0.011	10.94		
Unidad	-0.025	0.004	-6.32 (1, 360.7) = 16.89	p < 0.001	***
Acento	-0.012	0.004	-3.22 (1, 360.7) = 0.06	p = 0.802	ns
Consonante			(3, 360.7) = 38.23	p < 0.001	***
ø	-0.021	0.004	-5.07		
n	-0.012	0.004	-2.99		
s	0.026	0.004	6.09		
t	(0.007)				
Unidad x Acento			(1, 360.7) = 22.96	p < 0.001	***
Consonante x Unidad			(3, 360.7) = 0.73	p = 0.533	ns
Consonante x Acento			(3, 360.7) = 0.30	p = 0.824	ns

En el modelo referido a la mandíbula (Tabla 6.14 y Tabla 6.15)¹⁶⁵, el efecto principal de la Unidad ejerce una modificación significativa de la duración con respecto al término constante (nivel de referencia), pero en dirección contraria a la esperada, *i. e.* la duración articulatoria del movimiento de ascenso mandibular se reduce en la condición inicial de palabra. El efecto principal del Acento no resulta significativo, pero sí el de la Consonante. Además, hay interacciones Unidad x Acento (la tabla de promedios permite constatar un patrón cruzado), Consonante x Unidad (con una reducción temporal en ω que tan solo es significativa para la [f] y la [β]) y Consonante x Acento (también por diferencia de magnitud, a favor de la [β]).

Los datos correspondientes a la duración del movimiento del ápice de la lengua (M₈) están recogidos en la Tabla 6.16 y la Tabla 6.17. En esta última se observa que la duración gestual, en contra de lo que predice la H_{3A}, es significativamente menor si la consonante se sitúa en posición inicial de palabra. El efecto principal del Acento no resulta significativo (sí lo es el de la Consonante). Tan solo es significativa la interacción Unidad x Acento (con un patrón cruzado).

Para concluir las comprobaciones acerca de la H₃ (duración del gesto de constricción), falta estudiar el caso de la constricción labial (Tabla 6.18 y Tabla 6.19). La identidad de la Consonante influye notablemente en la duración gestual. La fricativa [f] es la que presenta una mayor duración articulatoria, seguida de la [p], la [β] y la [m], en ese orden. Las fricativas y las oclusivas tienen, efectivamente, un modo de articulación que se suele caracterizar por una mayor duración (también acústica).

¹⁶⁵ Si bien, como norma general, todos los resultados que se presentan aquí aparecen redondeados a dos posiciones decimales, los datos de duración muestran tres posiciones, para facilitar la lectura de los milisegundos, unidad habitual en los estudios de fonética.

Tabla 6.18. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de los labios correspondiente a la fase de constricción.

	Palabra	Sílaba	TOTAL	
Acento	0.108 (0.037)	0.110 (0.029)	0.109 (0.033)	
No acento	0.095 (0.040)	0.115 (0.031)	0.105 (0.037)	
TOTAL	0.101 (0.039)	0.113 (0.030)	0.107 (0.035)	n = 366

Tabla 6.19. M₉: Duración del movimiento de los labios (fase de constricción).

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.115	0.009	13.53			
Unidad (Sujeto)	-0.020	0.006	-2.98	(1, 5) = 2.95	p = 0.147	ns
Acento	-0.005	0.003	-1.65	(1, 342.5) = 1.80	p = 0.180	ns
Consonante				(3, 342.5) = 18.55	p < 0.001	***
β	-0.009	0.004	-2.42			
f	0.021	0.004	5.75			
m	-0.012	0.004	-3.33			
p	(0.000)					
Unidad x Acento				(1, 342.5) = 13.40	p < 0.001	***
Consonante x Unidad				(3, 342.5) = 5.30	p < 0.01	**
Consonante x Acento				(3, 342.5) = 1.09	p = 0.355	ns

Además, se trata de segmentos sordos, que también tienden a ser más largos que los sonoros (existen condicionantes aerodinámicos que dificultan el mantenimiento de la vibración laríngea durante mucho tiempo cuando se está produciendo una constricción en la cavidad oral).

Por su parte, ninguno de los factores prosódicos estudiados resulta significativo. Sí lo es la interacción Unidad x Acento. La Figura 6.5 refleja que la condición de referencia (la cuarta columna: inicial de sílaba, sílaba no acentuada) es la que presenta una mayor duración. Con respecto a ella, la condición $^1\sigma$ (tercera columna) y la condición ω (segunda columna) suponen cierto descenso, cuyas respectivas magnitudes recogen los coeficientes β del modelo. Sin embargo, cuando se da la circunstancia de ser a la vez ω y $^1\sigma$, el cambio no es igual a la suma de las dos β , sino que el efecto queda modificado, precisamente, por el coeficiente de la interacción (que no aparece en la Tabla 6.19). Así, la primera columna de la figura refleja el valor promedio que aparece en la primera casilla de la Tabla 6.18. En definitiva, la figura deja ver que las duraciones de las consonantes en posición acentuada y no acentuada no difieren significativamente si se mantiene la condición del constituyente prosódico, pero el hecho de que la tendencia del efecto del Acento sobre ω y σ apunte en direcciones contrarias es suficiente para que la interacción entre ambos factores sea significativa.

También merece una mención especial la interacción Consonante x Unidad. A juzgar por la Figura 6.6, parece que la posición inicial de palabra tiende a disminuir la duración del movimiento de constricción labial, pero lo hace en mayor medida en la $[\beta]$ y la $[f]$, aquellas en las que no llega a producirse un contacto completo.

Por último, los distintos sujetos se comportan de manera diferente ante el efecto de la posición ω (para dos de ellos implica una mayor duración; para el resto, una reducción).

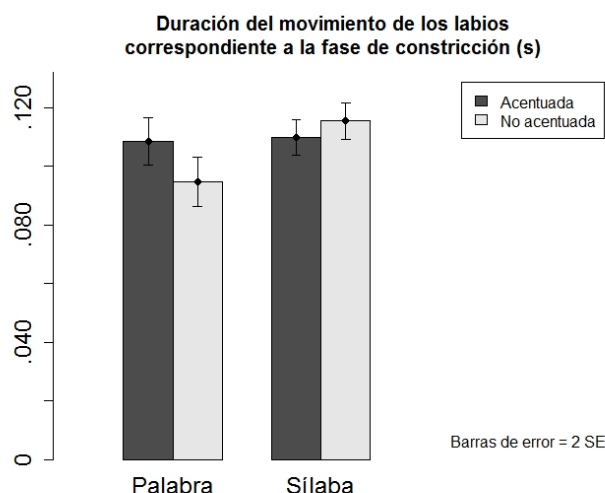


Figura 6.5. M₉: Interacción Unidad x Acento en la duración del movimiento labial de constricción.

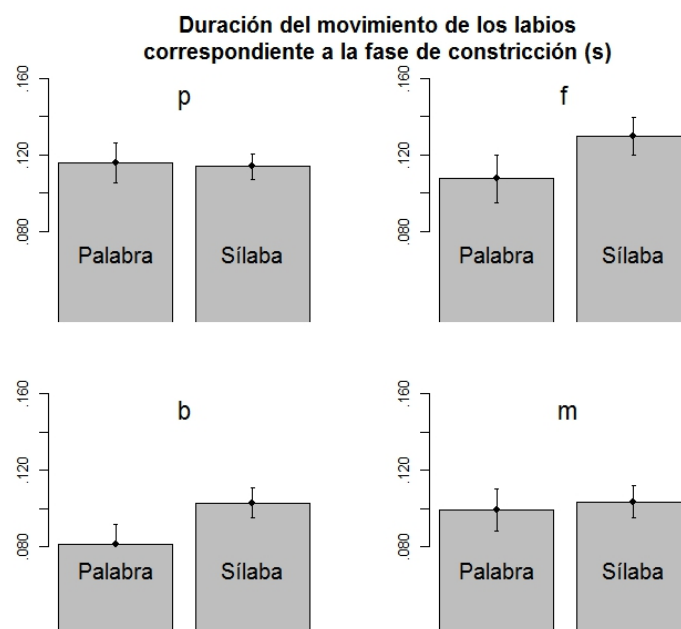


Figura 6.6. M₉: Interacción Consonante x Unidad en la duración del movimiento labial (fase de constricción).

6.4. H₄: duración del gesto de distensión

La H_{4A} prevé un aumento de la duración del gesto de distensión en posición inicial de palabra como consecuencia de la ralentización en la ejecución motora. La H_{4B} vincula el carácter acentuado de la sílaba a un incremento temporal del gesto de distensión consonántica (gesto de formación de la vocal), derivado del retraso en el inicio del gesto de constricción de la consonante posterior a la vocal.

El M₁₀ pone a prueba estas hipótesis en el caso del movimiento de la mandíbula (*vid.* Tabla 6.20 y Tabla 6.21). Los resultados muestran que hay una diferencia significativa entre unas consonantes y otras.

Tabla 6.20. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de la mandíbula correspondiente a la fase de distensión.

	Palabra	Sílaba	TOTAL	
Acento	0.123 (0.040)	0.121 (0.048)	0.122 (0.044)	
No acento	0.097 (0.039)	0.100 (0.047)	0.098 (0.044)	
TOTAL	0.110 (0.042)	0.110 (0.049)	0.110 (0.045)	n = 746

Tabla 6.21. M₁₀: Duración del movimiento de la mandíbula (fase de distensión).

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.100	0.015	6.71			
Unidad (Sujeto)	-0.003	0.004	-0.81	(1, 4.9) =	0.11	p = 0.753 ns
Acento	0.021	0.003	7.39	(1, 709.5) =	131.38	p < 0.001 ***
Consonante				(7, 709.5) =	30.94	p < 0.001 ***
β	-0.016	0.005	-3.49			
ϕ	-0.017	0.005	-3.61			
f	-0.024	0.005	-5.13			
m	-0.016	0.005	-3.36			
n	0.003	0.004	0.78			
p	0.016	0.005	3.41			
s	0.021	0.005	4.53			
t	(0.033)					
Unidad x Acento				(1, 709.5) =	1.16	p = 0.282 ns
Consonante x Unidad				(7, 709.5) =	1.30	p = 0.248 ns
Consonante x Acento				(7, 709.5) =	4.22	p < 0.001 ***

Por otro lado, los distintos informantes se comportan contradictoriamente con respecto al efecto de la Unidad, en un patrón cruzado que, seguramente, es la causa de que la posición prosódica no llegue a alcanzar nivel de significatividad en el modelo. En cambio, sí es significativo el efecto principal del Acento, que provoca un aumento de duración en torno al 20% con respecto al nivel de referencia. Este aumento es, por lo general, más acusado en las consonantes labiales que en las corales, lo que se refleja en una interacción Consonante x Acento por diferencia de magnitud (*vid.* Figura 6.7).

La Tabla 6.22 y la Tabla 6.23 corresponden a los resultados del M₁₁ sobre el movimiento del ápice de la lengua. La menor duración del movimiento distensivo corresponde a la [ø] y la mayor a la [t], con diferencias significativas con respecto a la gran media, como para hacer que el efecto global de la Consonante también contribuya sensiblemente al modelo.

No hay un efecto significativo de la Unidad, pero sí del Acento, que provoca un aumento de la duración distensiva (en todos los sujetos menos en uno, para quien disminuye levemente). Por otro lado, hay una interacción entre la Unidad y el Acento (por diferencia de magnitud). La Figura 6.8 muestra que, si bien la prominencia tiende a aumentar la duración del gesto de distensión, la diferencia entre 'σ y σ solo es significativa en posición inicial de palabra.

Por último, pasemos al M₁₂ (Tabla 6.24 y Tabla 6.25), que modela la duración de la distensión correspondiente a los labios (a partir de unas diferencias preexistentes atribuibles a la identidad de la consonante: dura más en la [p] y en la [f]).

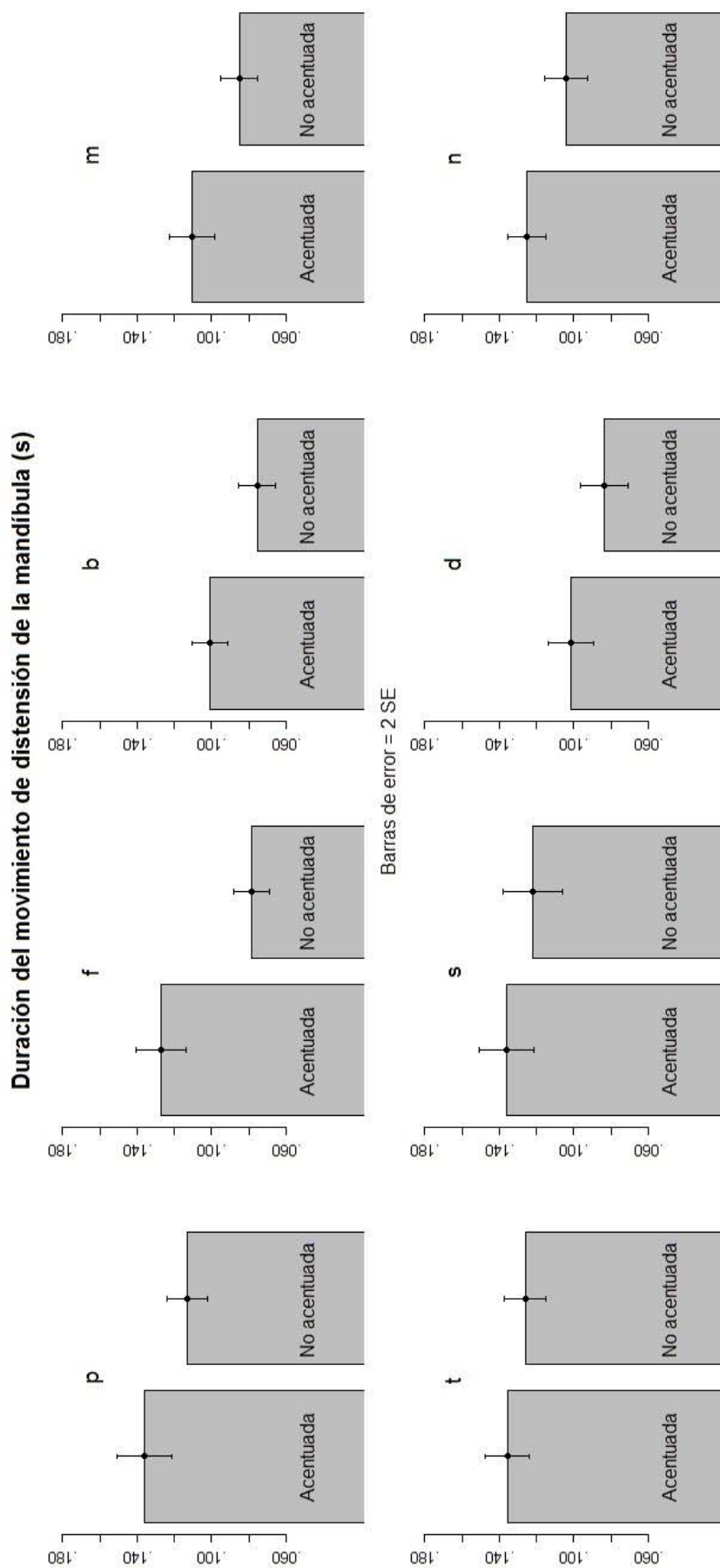


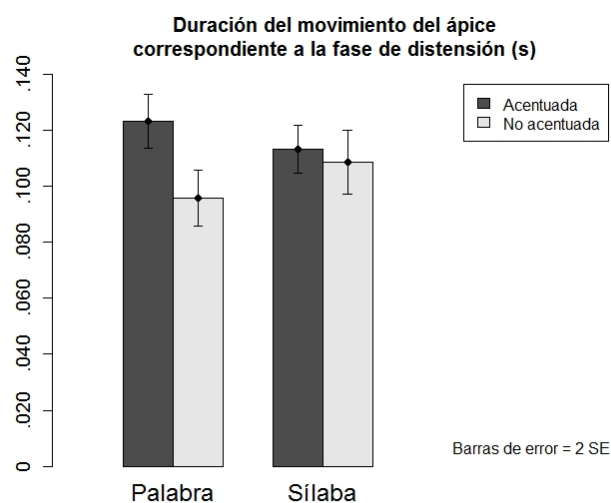
Figura 6.7. M₁₀: Interacción Consonante x Acento en la duración del movimiento de distensión de la mandíbula.

Tabla 6.22. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento del ápice de la lengua correspondiente a la fase de distensión.

	Palabra	Sílaba	TOTAL	
Acento	0.123 (0.047)	0.113 (0.041)	0.118 (0.044)	
No acento	0.096 (0.049)	0.109 (0.056)	0.102 (0.053)	
TOTAL	0.110 (0.050)	0.111 (0.050)	0.110 (0.050)	n = 380

Tabla 6.23. M₁₁: Duración del movimiento del ápice de la lengua (fase de distensión).

	β	SE	t	(df)	F	p (test F)	
Ordenada en el origen	0.109	0.017	6.29				
Unidad	-0.013	0.005	-2.77	(1, 360.9)	0.22	p = 0.641	ns
Acento (Sujeto)	0.004	0.006	0.63	(1, 5.8)	8.07	p < 0.05	*
Consonante				(3, 360.9)	21.87	p < 0.001	***
ø	-0.024	0.005	-4.82				
n	0.005	0.005	1.02				
s	0.004	0.005	0.76				
t	(0.015)						
Unidad x Acento				(1, 360.9)	11.76	p < 0.001	***
Consonante x Unidad				(3, 360.9)	0.21	p = 0.887	ns
Consonante x Acento				(3, 360.9)	0.64	p = 0.587	ns

**Figura 6.8. M₁₁: Interacción Unidad x Acento en la duración del movimiento del ápice (fase de distensión).**

El modelo incluye el efecto aleatorio del Sujeto sobre la pendiente de Unidad, y es que para dos de los informantes la duración se reduce significativamente en ω , mientras que para el resto no se modifica, o incluso aumenta ligeramente. Además, la interacción Consonante x Unidad muestra que esa reducción solo afecta a la [f] y, en menor medida, a la [p]. En cualquier caso, el efecto principal de la Unidad no realiza una contribución global al modelo que supere los niveles de significatividad requeridos.

Sí es significativo el efecto del Acento, que se relaciona con un aumento de la duración distensiva. Este aumento se produce tanto en inicial de palabra como en inicial de sílaba, aunque provoca mayores diferencias entre 'σ y σ precisamente en inicial de sílaba (lo que se refleja en una interacción Unidad x Acento por diferencia de magnitud).

En definitiva, la duración de la distensión no se ve afectada por la posición prosódica pero sí por la prominencia. En sílaba acentuada, se produce un refuerzo temporal del gesto de distensión de la consonante (gesto de formación de la vocal). Dada la desalineación entre los gestos articulatorios y la duración acústica de los segmentos (*vid.* § 5.6), el alargamiento temporal de este gesto puede manifestarse como un aumento en la duración acústica de la consonante, de la vocal, o de ambas. En el § 6.6 se aborda la duración consonántica, pues su modelado está previsto por una de las hipótesis (H_6). Sin embargo, aunque este es un estudio sobre las consonantes, parece pertinente aquí añadir un modelo (M_{13}) que estudie si, como parece previsible, el Acento provoca un aumento significativo en la duración acústica de las vocales cuando estas se encuentran en sílaba prominente, *vid.* Tabla 6.26 y Tabla 6.27.

Hay una diferencia significativa en la duración vocálica dependiendo de la identidad de la consonante que la precede (p. ej. las vocales son más breves después de las fricativas [f] y [s], que son segmentos consonánticos intrínsecamente largos; y son más largas después de las espirantes y de las nasales, aunque solo difieren de la gran media cuando siguen a [m]).

El efecto de la Unidad no es significativo, aunque apunta en la dirección de una reducción temporal de la vocal si aparece en la primera sílaba de la palabra. Esta tendencia es especialmente notable en dos de los seis sujetos.

Tabla 6.24. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del movimiento de los labios correspondiente a la fase de distensión.

	Palabra	Sílaba	TOTAL	
Acento	0.117 (0.037)	0.134 (0.046)	0.126 (0.043)	
No acento	0.096 (0.030)	0.103 (0.046)	0.099 (0.039)	
TOTAL	0.106 (0.035)	0.118 (0.048)	0.112 (0.043)	n = 366

Tabla 6.25. M_{12} : Duración del movimiento de los labios (fase de distensión).

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.103	0.016	6.37			
Unidad (Sujeto)	-0.009	0.007	-1.04	(1, 5) =	4.07	p = 0.100 ns
Acento	0.031	0.003	9.81	(1, 342.6) =	129.43	p < 0.001 ***
Consonante				(3, 342.6) =	35.63	p < 0.001 ***
β	-0.019	0.004	-5.50			
f	0.008	0.003	2.39			
m	-0.010	0.003	-2.82			
p	(0.021)					
Unidad x Acento				(1, 342.6) =	5.63	p < 0.05 *
Consonante x Unidad				(3, 342.6) =	5.33	p < 0.01 **
Consonante x Acento				(3, 342.6) =	1.77	p = 0.153 ns

Tabla 6.26. Media (y desviación estándar), por condiciones prosódicas, de la duración (en segundos) del segmento vocálico siguiente a la consonante.

	Palabra	Sílaba	TOTAL	
Acento	0.111 (0.034)	0.127 (0.049)	0.119 (0.043)	
No acento	0.087 (0.030)	0.100 (0.045)	0.093 (0.039)	
TOTAL	0.099 (0.034)	0.113 (0.049)	0.106 (0.043)	n = 746

Tabla 6.27. M₁₃: Duración acústica del segmento vocálico siguiente.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.100	0.018	5.56			
Unidad (Sujeto)	-0.013	0.008	-1.50	(1, 5) = 3.01	p = 0.144	ns
Acento	0.027	0.002	12.18	(1, 708) = 257.76	p < 0.001	***
Consonante				(7, 708) = 9.58	p < 0.001	***
β	0.006	0.004	1.64			
ϕ	0.006	0.004	1.73			
f	-0.010	0.004	-2.64			
m	0.008	0.004	2.25			
n	0.006	0.003	1.77			
p	0.002	0.004	0.49			
s	-0.011	0.004	-3.11			
t	(-0.007)					
Unidad x Acento				(1, 708) = 1.11	p = 0.292	ns
Consonante x Unidad				(7, 708) = 3.96	p < 0.001	***
Consonante x Acento				(7, 708) = 3.51	p < 0.01	**

Por otro lado, tal y como se esperaba, el Acento provoca un aumento de la duración acústica de la vocal del 27%. Puesto que no existe interacción Unidad x Acento, este incremento se atestigua en la misma medida tanto en posición inicial como en interior de palabra.

Las interacciones que sí resultan significativas (por diferencia de magnitud) son las que se producen entre la Consonante y la Unidad, y entre la Consonante y el Acento (la duración acústica de la vocal apenas aumenta en sílaba acentuada si la consonante precedente es una $[\beta]$).

6.5. H₅: solapamiento temporal entre los movimientos articulatorios de la constricción

Este apartado presenta los resultados correspondientes a los modelos M₁₄ y M₁₅, que analizan el grado de solapamiento temporal entre los movimientos de los distintos órganos articulatorios que se coordinan en un mismo gesto de constricción.

En primer lugar, se describe la duración del solapamiento entre el movimiento de la mandíbula y el movimiento del ápice en el gesto apical de las consonantes corales (Tabla 6.28 y Tabla 6.29). Una vez controladas las diferencias intrínsecas entre las distintas identidades de la Consonante, se comprueba que la duración del solapamiento se reduce en posición inicial de palabra con respecto a la inicial de sílaba (en consonancia con la predicción de la H_{5A}).

Tabla 6.28. Media (y desviación estándar), por condiciones prosódicas, de la duración (s) del solapamiento temporal entre los movimientos de la mandíbula y del ápice (fase de constricción).

	Palabra	Sílaba	TOTAL	
Acento	0.087 (0.037)	0.086 (0.034)	0.086 (0.036)	
No acento	0.073 (0.036)	0.095 (0.036)	0.084 (0.037)	
TOTAL	0.080 (0.037)	0.091 (0.035)	0.085 (0.036)	n = 380

Tabla 6.29. M₁₄: Solapamiento temporal entre los movimientos de la mandíbula y del ápice (fase de constricción).

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.095	0.009	11.00		
Unidad	-0.022	0.004	-6.47 (1, 360.5) = 20.02	p < 0.001	***
Acento	-0.009	0.004	-2.90 (1, 360.5) = 0.30	p = 0.584	ns
Consonante			(3, 360.5) = 67.51	p < 0.001	***
ø	-0.018	0.004	-4.71		
n	-0.021	0.004	-5.77		
s	0.028	0.004	7.51		
t	(0.011)				
Unidad x Acento			(1, 360.5) = 21.47	p < 0.001	***
Consonante x Unidad			(3, 360.5) = 0.96	p = 0.413	ns
Consonante x Acento			(3, 360.5) = 1.31	p = 0.271	ns

Tabla 6.30. Media (y desviación estándar), por condiciones prosódicas, de la duración (s) del solapamiento temporal entre los movimientos de la mandíbula y de los labios (fase de constricción).

	Palabra	Sílaba	TOTAL	
Acento	0.084 (0.035)	0.090 (0.030)	0.087 (0.032)	
No acento	0.073 (0.035)	0.103 (0.029)	0.088 (0.035)	
TOTAL	0.078 (0.035)	0.096 (0.030)	0.087 (0.034)	n = 366

Tabla 6.31. M₁₅: Solapamiento temporal entre los movimientos de la mandíbula y de los labios (fase de constricción).

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.103	0.009	11.24		
Unidad	-0.030	0.003	-8.60 (1, 346.5) = 53.51	p < 0.001	***
Acento	-0.013	0.003	-3.70 (1, 346.5) = 0.22	p = 0.642	ns
Consonante			(3, 346.5) = 21.36	p < 0.001	***
β	-0.013	0.004	-3.54		
f	0.019	0.004	5.51		
m	-0.006	0.004	-1.75		
p	(0.000)				
Unidad x Acento			(1, 346.5) = 22.25	p < 0.001	***
Consonante x Unidad			(3, 346.5) = 4.44	p < 0.01	**
Consonante x Acento			(3, 346.5) = 2.60	p = 0.052	ns

No hay efecto significativo del Acento, pero sí una interacción Unidad x Acento, con un patrón cruzado (*vid.* Tabla 6.28).

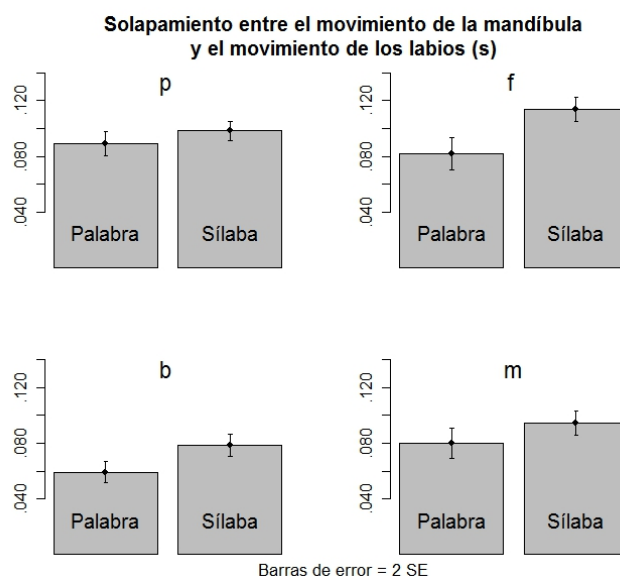


Figura 6.9. M_{15} : Interacción Consonante x Unidad en la duración del solapamiento entre el movimiento de la mandíbula y el movimiento de los labios (fase de constricción).

En cuanto al solapamiento entre el movimiento de la mandíbula y el de los labios (en las consonantes labiales), la Tabla 6.30 y la Tabla 6.31 aportan los datos relevantes del M_{15} . De las distintas variables introducidas en el modelo, resultan significativas las siguientes. La posición ω provoca una reducción en el grado de solapamiento intragesual (como predice la H_{5A}). También hay diferencias intrínsecas entre consonantes (con un menor solapamiento en el caso de la aspirante y de la nasal, y un mayor solapamiento en la fricativa y la oclusiva, exactamente igual que ocurría en el modelo anterior, sobre la coordinación entre la mandíbula y el ápice). Por último, hay interacciones entre la Unidad y el Acento (patrón cruzado, *vid.* Tabla 6.30), y entre la Consonante y la Unidad (por diferencia de magnitud, *vid.* Figura 6.9).

6.6. H_6 : duración acústica del segmento consonántico

La H_6 predice un aumento de la duración acústica por efecto de ω y también de ${}^1\sigma$, excepto en el caso de las consonantes nasales en inicio de palabra, que deberían ver reducida su duración. Dada esta disyunción, analizaremos por separado los efectos del refuerzo sobre las consonantes orales (M_{16}) y sobre las nasales (M_{17}).

Los datos de las consonantes orales se recogen en la Tabla 6.32 y la Tabla 6.33. La duración acústica media es diferente entre unas consonantes y otras (las sordas duran más que las sonoras), como indica el efecto principal de Consonante. La prominencia determinada por el Acento también tiene un efecto significativo, en la dirección esperada: las consonantes en ${}^1\sigma$ duran más que aquellas en σ . En cambio, no se encuentra efecto significativo de la Unidad, ni de ninguna de las interacciones.

La Tabla 6.34 y la Tabla 6.35 se refieren a la duración acústica de las nasales (M_{17}). Aunque las duraciones medias de [m] y de [n] no difieren significativamente de la gran media —el promedio de aquellas dos— (el valor absoluto de t es menor que 1.96), sí

Tabla 6.32. Media (y desviación estándar), por condiciones prosódicas, de la duración acústica (en segundos) de las consonantes orales.

	Palabra	Sílaba	TOTAL	
Acento	0.116 (0.045)	0.114 (0.042)	0.115 (0.043)	
No acento	0.102 (0.043)	0.106 (0.041)	0.104 (0.042)	
TOTAL	0.109 (0.044)	0.110 (0.042)	0.109 (0.043)	n = 556

Tabla 6.33. M₁₆: Duración acústica de las consonantes orales.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.106	0.009	12.25			
Unidad	-0.004	0.003	-1.58	(1, 530.6) = 0.54	p = 0.461	ns
Acento	0.008	0.003	2.82	(1, 530.6) = 29.79	p < 0.001	***
Consonante				(5, 530.6) = 211.46	p < 0.001	***
β	-0.045	0.004	-12.57			
ϕ	-0.041	0.004	-11.53			
f	0.010	0.004	2.78			
p	0.017	0.004	4.78			
s	0.028	0.004	7.72			
t	(0.031)					
Unidad x Acento				(1, 530.6) = 2.25	p = 0.134	ns
Consonante x Unidad				(5, 530.6) = 1.01	p = 0.413	ns
Consonante x Acento				(5, 530.6) = 1.98	p = 0.080	ns

Tabla 6.34. Media (y desviación estándar), por condiciones prosódicas, de la duración acústica (en segundos) de las consonantes nasales.

	Palabra	Sílaba	TOTAL	
Acento	0.104 (0.026)	0.099 (0.025)	0.102 (0.025)	
No acento	0.083 (0.030)	0.086 (0.024)	0.084 (0.027)	
TOTAL	0.094 (0.030)	0.092 (0.025)	0.093 (0.028)	n = 190

Tabla 6.35. M₁₇: Duración acústica de las consonantes nasales.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.086	0.009	9.37			
Unidad	-0.003	0.003	-1.38	(1, 176.9) = 0.09	p = 0.760	ns
Acento	0.013	0.003	3.99	(1, 176.9) = 60.43	p < 0.001	***
Consonante				(1, 176.9) = 6.23	p < 0.05	*
m	-0.003	0.002	-1.42			
n	(0.003)					
Unidad x Acento				(1, 176.9) = 4.87	p < 0.05	*
Consonante x Unidad				(1, 176.9) = 12.75	p < 0.001	***
Consonante x Acento				(1, 176.9) = 1.73	p = 0.190	ns

resulta significativo el efecto principal de la variable Consonante, tomada en su conjunto (la [n] dura más que la [m]).

La influencia de la posición inicial de palabra apunta en la dirección esperada (una reducción temporal), pero no llega a ser significativa. En cambio, el Acento sí produce

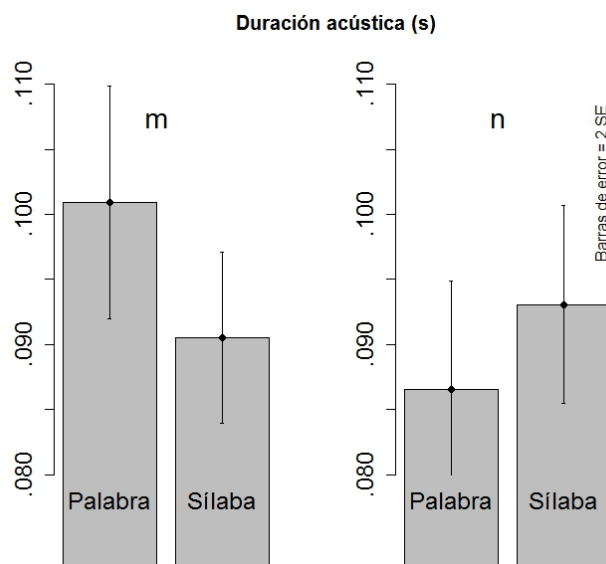


Figura 6.10. M_{17} : Interacción Consonante x Unidad en la duración acústica de las consonantes nasales.

de manera significativa un aumento de la duración acústica, en consonancia con las predicciones de H_6 .

También resulta significativa la interacción entre Unidad y Acento (la magnitud del refuerzo temporal ligado a la 'σ es algo mayor si además la consonante se sitúa en posición inicial de palabra). Además, hay una interacción entre Consonante y Unidad: tal como se puede apreciar en la Figura 6.10, la [n] dura menos en posición inicial de palabra que en interior. Aunque no llega a haber una diferencia significativa, la dirección es la apuntada por la hipótesis propuesta en $H_{6A,b}$ («el refuerzo por posición en las nasales se traduce en una menor duración acústica»). En el caso de la [m], en cambio, aumenta la duración acústica en posición inicial de palabra. Dada la disparidad en las direcciones del cambio, se entiende que el efecto principal de la Unidad no llegue a ser significativo.

6.7. H_7 : VOT

Veamos ahora el comportamiento del VOT (M_{18}) y de la proporción de tiempo que ocupa la distensión de las oclusivas sobre el total de su duración acústica (M_{19}).

La Tabla 6.36 muestra que el VOT promedio es prácticamente idéntico para las distintas condiciones prosódicas. La Tabla 6.37 corrobora, conforme a lo previsto en la H_7 , que no existe ningún efecto significativo de la Unidad ni del Acento. Tampoco se aprecian diferencias de VOT entre la [p] y la [t]. No hay interacciones significativas.

Lo mismo puede decirse del cálculo proporcional del lapso de tiempo que va desde la distensión hasta el inicio de la vibración laríngea sobre el total de la duración acústica de la oclusiva (Tabla 6.38 y Tabla 6.39). Ninguno de los predictores incluidos en el modelo condiciona significativamente dicha proporción. Como se ha dicho, este resultado confirma la hipótesis planteada al respecto.

Tabla 6.36. Media (y desviación estándar), por condiciones prosódicas, del VOT (en segundos) de las oclusivas sordas.

	Palabra	Sílaba	TOTAL	
Acento	0.006 (0.009)	0.008 (0.009)	0.007 (0.009)	
No acento	0.007 (0.009)	0.007 (0.010)	0.007 (0.010)	
TOTAL	0.007 (0.009)	0.007 (0.010)	0.007 (0.009)	n = 187

Tabla 6.37. M₁₈: VOT de las oclusivas sordas.

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.007	0.003	2.68		
Unidad	-0.001	0.002	-0.36 (1, 172.7) = 0.61	p = 0.435	ns
Acento	0.000	0.002	0.08 (1, 172.7) = 0.03	p = 0.871	ns
Consonante			(1, 172.7) = 0.35	p = 0.555	ns
p	0.000	0.001	0.30		
t	(0.000)				
Unidad x Acento			(1, 172.7) = 0.08	p = 0.784	ns
Consonante x Unidad			(1, 172.7) = 0.15	p = 0.699	ns
Consonante x Acento			(1, 172.7) = 0.21	p = 0.646	ns

Tabla 6.38. Media (y desviación estándar), por condiciones prosódicas, de la proporción de tiempo que ocupa la distensión de las oclusivas sobre el total de su duración acústica.

	Palabra	Sílaba	TOTAL	
Acento	0.04 (0.06)	0.05 (0.06)	0.05 (0.06)	
No acento	0.05 (0.07)	0.05 (0.07)	0.05 (0.07)	
TOTAL	0.05 (0.07)	0.05 (0.07)	0.05 (0.07)	n = 187

Tabla 6.39. M₁₉: Proporción de tiempo que ocupa la distensión de las oclusivas sobre el total de su duración acústica.

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.05	0.02	3.06		
Unidad	0.00	0.01	-0.08 (1, 174.7) = 0.31	p = 0.580	ns
Acento	0.00	0.01	-0.11 (1, 174.7) = 0.35	p = 0.553	ns
Consonante			(1, 174.7) = 2.01	p = 0.158	ns
p	0.01	0.01	0.84		
t	(-0.01)				
Unidad x Acento			(1, 174.7) = 0.19	p = 0.659	ns
Consonante x Unidad			(1, 174.7) = 0.00	p = 0.996	ns
Consonante x Acento			(1, 174.7) = 0.00	p = 0.961	ns

6.8. H₈: altura de la mandíbula

Después de haber estudiado la velocidad de los gestos (§ 6.1-6.2) y los parámetros relacionados con diversos aspectos de la duración en el ámbito articulatorio y acústico (§ 6.3-6.7), este apartado y varios de los siguientes se centran en el análisis de la magnitud alcanzada por los gestos articulatorios. Aquí se presentan los resultados del M₂₀, destinado a comprobar la hipótesis de que la mandíbula alcanza una posición más

elevada cuando la consonante se sitúa en una posición de refuerzo, tanto si se trata del inicio de palabra como si el refuerzo lo desencadena la prominencia acentual de la sílaba.

La Tabla 6.40 recoge los valores, promediados para cada condición prosódica, del pico de magnitud de la mandíbula (en mm). Teniendo en cuenta que la referencia del sistema de coordenadas se sitúa en la encía superior, y que todos los valores de la altura de la mandíbula son negativos, se entiende que la posición vertical de la mandíbula es tanto más alta cuanto más cercano a cero sea un dato.

Los parámetros del modelo y los resultados de sus respectivas pruebas de significatividad aparecen en la Tabla 6.41. El efecto principal de la Unidad condiciona una posición más elevada de la mandíbula, tal y como se esperaba (H_{8A}).

En cambio, no se encuentra una influencia significativa del Acento, frente a la predicción de H_{8B} , aunque la dirección de la tendencia (hacia la mayor elevación) es la adecuada.

La introducción de la variable Consonante en el modelo se revela significativa, puesto que las consonantes labiales alcanzan, en general, una menor altura mandibular que las coronales (aunque la [ð] también se caracteriza por una posición más baja que la media intergrupar). La mayor altura la alcanzan la [t] y, sobre todo, la [s].

Tabla 6.40. Media (y desviación estándar), por condiciones prosódicas, de la altura (en mm) de la mandíbula.

	Palabra	Sílaba	TOTAL	
Acento	-22.37 (4.67)	-22.78 (4.59)	-22.58 (4.63)	
No acento	-22.24 (4.80)	-22.86 (4.74)	-22.55 (4.77)	
TOTAL	-22.30 (4.73)	-22.82 (4.66)	-22.56 (4.70)	n = 746

Tabla 6.41. M_{20} : Altura de la mandíbula.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	-22.86	1.87	-12.25			
Unidad	0.62	0.18	2.68	(1, 716) = 14.88	p < 0.001	***
Acento	0.08	0.18	0.69	(1, 716) = 1.14	p = 0.286	ns
Consonante				(7, 716) = 38.81	p < 0.001	***
β	-1.81	0.30	-6.05			
ð	-1.06	0.30	-3.56			
f	-0.04	0.30	-0.13			
m	-1.10	0.30	-3.68			
n	-0.03	0.28	-0.11			
p	-0.09	0.30	-0.31			
s	3.23	0.30	10.71			
t	(0.90)					
Unidad x Acento				(1, 716) = 0.01	p = 0.917	ns
Consonante x Unidad				(7, 716) = 4.14	p < 0.001	***
Consonante x Acento				(7, 716) = 3.62	p < 0.001	***

No se produce una interacción significativa entre la Unidad y el Acento; en cambio, sí contribuyen al modelo las interacciones bidireccionales entre el tipo de Consonante y cada uno de los dos factores prosódicos. Resulta especialmente interesante la interacción Consonante x Unidad (*vid.* Figura 6.11). En casi todas las consonantes la mandíbula alcanza una posición más elevada en posición inicial de palabra (en mayor o menor grado según la identidad consonántica), pero la [f] muestra un patrón en dirección contraria.

6.9. H₉: altura del ápice de la lengua

El M₂₁ (Tabla 6.42 y Tabla 6.43) pone a prueba la H₉ sobre la altura del ápice de la lengua, que predice una posición más elevada de dicho órgano cuando la consonante aparece en una posición fuerte.

Como es lógico, la identidad de la Consonante determina en buena medida la altura del ápice, que se eleva más en los casos en que es necesario un cierre completo de la salida del aire ([n] y [t]), algo menos cuando debe producirse el estrechamiento crítico propio de la fricativa [s], y menos aún en el caso de la espirante [ø].

Partiendo de esas diferencias iniciales, la Unidad tiene un efecto significativo sobre la magnitud del gesto apical: las consonantes en posición inicial de palabra presentan una mayor altura del ápice, tal y como predice la H_{9A}. El Acento no resulta significativo.

De entre las interacciones comprobadas, solo alcanza el nivel de significatividad requerido la que se produce entre la Consonante y la Unidad: existe un patrón cruzado tal que el ápice sube más en ω que en σ en todas las consonantes menos en la [n], en la que se obtiene una altura apical ligeramente menor en inicial de palabra.

Tabla 6.42. Media (y desviación estándar), por condiciones prosódicas, de la altura (en mm) del ápice de la lengua.

	Palabra	Sílaba	TOTAL	
Acento	-6.71 (3.27)	-7.43 (3.52)	-7.07 (3.41)	
No acento	-6.97 (3.33)	-7.26 (3.24)	-7.12 (3.28)	
TOTAL	-6.84 (3.29)	-7.35 (3.37)	-7.10 (3.34)	n = 380

Tabla 6.43. M₂₁: Altura del ápice de la lengua.

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	-7.26	0.97	-7.52		
Unidad	0.29	0.27	1.18 (1, 362) = 6.96	p < 0.01	**
Acento	-0.17	0.27	-0.36 (1, 362) = 0.22	p = 0.637	ns
Consonante			(3, 362) = 98.71	p < 0.001	***
ø	-2.80	0.29	-9.71		
n	1.92	0.28	6.90		
s	0.42	0.29	1.43		
t	(0.46)				
Unidad x Acento			(1, 362) = 0.97	p = 0.326	ns
Consonante x Unidad			(3, 362) = 5.21	p < 0.01	**
Consonante x Acento			(3, 362) = 2.26	p = 0.081	ns

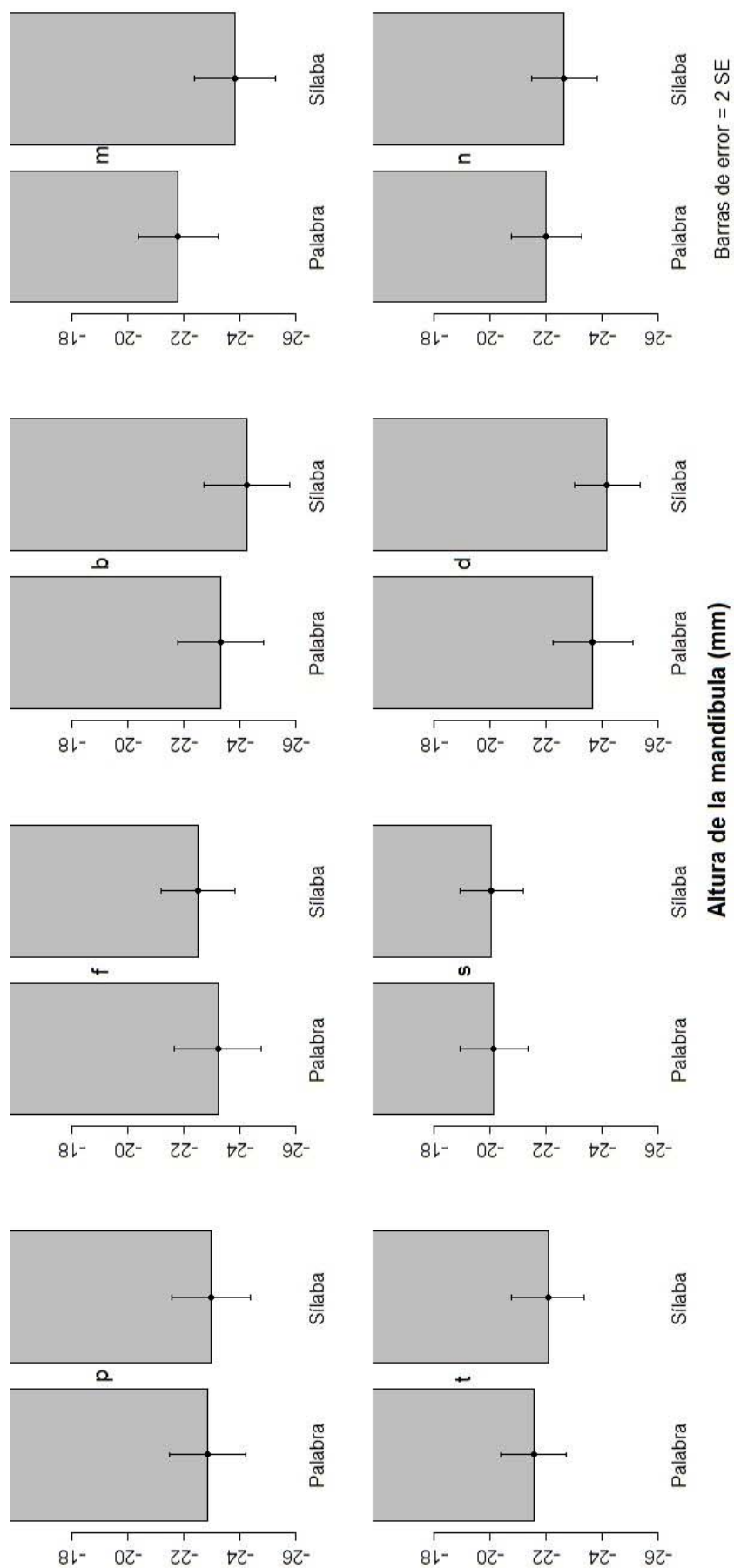


Figura 6.11. M_{20} : Interacción Consonante x Unidad en la altura de la mandíbula.

6.10.H₁₀: abertura labial

Siguiendo el mismo esquema que hasta ahora, la Tabla 6.44 recoge el resumen descriptivo de los datos de abertura labial en las distintas condiciones prosódicas estudiadas. La Tabla 6.45 expone los resultados del modelo mixto realizado para determinar los factores que afectan dicha abertura de manera significativa.

Por supuesto, existen diferencias entre los distintos tipos de Consonante: la [β] y la [f] se articulan con una mayor separación labial que la [m] y la [p]. Los valores (para posición inicial de sílaba no acentuada) de la nasal y la oclusiva son 13.93 – 0.49 y 13.93 – 1.27, respectivamente, como resulta de sumar los coeficientes β de cada consonante al término constante del modelo (ordenada en el origen), que es el nivel de referencia. Nótese que estos valores no son iguales a cero por la imposibilidad de superponer los dos sensores labiales del dispositivo de EMA. Sin embargo, se sobreentiende que los valores alcanzados por estas consonantes equivalen a un contacto total entre los labios, con una mayor compresión del tejido blando en el caso de la [p].

Las consonantes en posición inicial de palabra se producen con una separación labial significativamente mayor que las que se sitúan tan solo en inicio de sílaba (algo que es contrario a la predicción de la H_{10A}). En cuanto al Acento, no se observa ningún efecto significativo sobre la abertura labial.

A partir de la observación de los datos de la Tabla 6.44 se comprende que la interacción Unidad x Acento sea significativa, ya que el carácter prominente de la sílaba favorece una menor abertura labial en la posición inicial de palabra, mientras que aumenta dicha abertura en las consonantes en posición media. El resto de las interacciones no resultan significativas.

Tabla 6.44. Media (y desviación estándar), por condiciones prosódicas, de la abertura labial (distancia entre el labio superior y el labio inferior, en mm).

	Palabra	Sílaba	TOTAL	
Acento	14.29 (4.22)	14.23 (4.53)	14.26 (4.36)	
No acento	15.03 (4.59)	13.93 (4.40)	14.47 (4.52)	
TOTAL	14.67 (4.42)	14.08 (4.45)	14.37 (4.44)	n = 366

Tabla 6.45. M₂₂: Abertura labial.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	13.93	1.69	8.23			
Unidad	1.10	0.31	3.72	(1, 348) = 7.38	p < 0.01	**
Acento	0.30	0.31	0.80	(1, 348) = 1.84	p = 0.176	ns
Consonante				(3, 348) = 14.24	p < 0.001	***
β	1.02	0.33	3.12			
f	0.74	0.32	2.28			
m	-0.49	0.33	-1.50			
p	(-1.27)					
Unidad x Acento				(1, 348) = 6.19	p < 0.05	*
Consonante x Unidad				(3, 348) = 0.71	p = 0.544	ns
Consonante x Acento				(3, 348) = 2.02	p = 0.111	ns

6.11.H₁₁: flujo y presión nasal

Los datos de U_n , promediados por condiciones prosódicas en la Tabla 6.46, muestran un menor flujo de aire nasal de [m] y [n] en posición inicial de palabra que en inicial de sílaba; y también menor en sílaba acentuada que en sílaba no acentuada. El test F (ANOVA) practicado para determinar la significatividad de cada una de las variables prosódicas en el modelo M_{23} corrobora la influencia del efecto principal de la Unidad, pero no la del Acento. Por otro lado, se aprecia una diferencia significativa en el nivel de U_n de la [m] y de la [n] (es mayor en la consonante labial). Ninguna de las interacciones posibles de este modelo resulta significativa.

Los datos de P_n (M_{24}) aparecen en la Tabla 6.48 y la Tabla 6.49 y confirman los efectos observados para el flujo: existe una menor P_n cuando la consonante aparece en la frontera inicial de una palabra. Además, la [m] presenta una P_n media mayor que la [n]. Una posible explicación del mayor grado de nasalidad de la labial es que la mayor duración de este segmento permita que el gesto de descenso del velo se desarrolle también durante más tiempo y, por ello, se alcance un mayor pico de P_n . El resultado de la correlación de Pearson entre la duración acústica y el nivel de presión nasal podría apoyar esta idea, $r = 0.51$.

Tabla 6.46. Media (y desviación estándar), por condiciones prosódicas, del flujo de aire nasal (en ml/s).

	Palabra	Sílaba	TOTAL	
Acento	6.30 (1.77)	6.85 (1.62)	6.57 (1.70)	
No acento	6.65 (1.99)	7.51 (2.65)	7.08 (2.36)	
TOTAL	6.48 (1.88)	7.20 (2.23)	6.84 (2.08)	n = 123

Tabla 6.47. M_{23} : Flujo de aire nasal.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	7.51	0.78	9.64			
Unidad	-0.86	0.36	-2.74	(1, 113.1) = 8.67	p < 0.01	**
Acento	-0.66	0.37	-1.51	(1, 113.1) = 1.63	p = 0.204	ns
Consonante				(1, 113.1) = 25.45	p < 0.001	***
m	0.74	0.22	3.35			
n	(-0.74)					
Unidad x Acento				(1, 113.1) = 0.73	p = 0.394	ns
Consonante x Unidad				(1, 113.1) = 1.01	p = 0.317	ns
Consonante x Acento				(1, 113.1) = 2.79	p = 0.097	ns

Tabla 6.48. Media (y desviación estándar), por condiciones prosódicas, de la presión de aire nasal (en cm de H₂O).

	Palabra	Sílaba	TOTAL	
Acento	0.65 (0.35)	0.74 (0.45)	0.69 (0.40)	
No acento	0.56 (0.37)	0.72 (0.47)	0.64 (0.43)	
TOTAL	0.61 (0.36)	0.73 (0.46)	0.67 (0.42)	n = 122

Tabla 6.49. M_{24} : Presión de aire nasal.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.72	0.18	3.94			
Unidad	-0.16	0.06	-3.07	(1, 112) = 9.20	p < 0.01	**
Acento	0.02	0.06	0.36	(1, 112) = 3.03	p = 0.084	ns
Consonante				(1, 112) = 24.49	p < 0.001	***
m	0.12	0.04	3.13			
n	(-0.12)					
Unidad x Acento				(1, 112) = 1.55	p = 0.216	ns
Consonante x Unidad				(1, 112) = 0.01	p = 0.913	ns
Consonante x Acento				(1, 112) = 0.07	p = 0.794	ns

6.12.H₁₂: %t pico de magnitud, pico de U_n y pico de P_n

Si los apartados anteriores se centraban en el pico de magnitud alcanzado por los distintos gestos articulatorios, este analiza cuánto tiempo tarda en alcanzarse dicho pico de magnitud, medido en términos proporcionales sobre el total de la duración acústica de la consonante (algo que abreviaremos como %t).

Comencemos por el %t pico de magnitud del movimiento de la mandíbula (M_{25}), cuyos datos se muestran en la Tabla 6.50 y la Tabla 6.51. Hay diferencias significativas entre unas consonantes y otras, como recogen los distintos coeficientes y valores t de las diferentes consonantes en la Tabla 6.51. En general, puede decirse que las labiales (a excepción de la [p]) alcanzan más tarde (que las corales) el pico de magnitud del movimiento de ascenso de la mandíbula durante su fase de constricción. Además, las oclusivas alcanzan el pico antes que el resto de los modos de articulación de su misma serie (labial o coronal), para asegurar el cierre del tracto vocal (y, por consiguiente, la característica fase de silencio) cuanto antes. Además, esto permite generar el aumento de P_o necesario para que la explosión distensiva sea lo más intensa posible (no hay que olvidar que en esa distensión y en la inmediata transición hacia la vocal siguiente se hallan muchas claves fundamentales para la percepción, entre otras cosas, del punto de articulación de la consonante).

El efecto principal de la Unidad contribuye significativamente a modelar los datos, al igual que el del Acento. Ambos resultados son contrarios a las predicciones de la H₁₂, ya que muestran un pico de magnitud de la constricción más temprano en posición inicial de palabra que en inicial de sílaba, y también más adelantado en las sílabas prominentes que en las no prominentes. Sin embargo, hay una interacción entre Consonante y Acento debido a un patrón cruzado: en las labiales, especialmente en la [β], la condición prominente adelanta el momento en que se alcanza el pico del movimiento de ascenso mandibular, pero en todas las corales se retrasa ligeramente.

Dejando de lado la mandíbula, veamos ahora qué pasa con el gesto del ápice de la lengua (Tabla 6.52 y Tabla 6.53). De los datos del modelo M_{26} resulta evidente que ni la posición inicial de palabra ni la prominencia acentual modifican significativamente el momento en que se alcanza el pico de magnitud del gesto. Tampoco hay interacciones

significativas. En cambio, sí hay diferencias entre los distintos tipos de Consonante: de nuevo se alcanza antes el pico cuando debe establecerse un contacto completo ([t̪] y [n]).

Tabla 6.50. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del movimiento de ascenso de la mandíbula.

	Palabra	Sílaba	TOTAL	
Acento	0.46 (0.16)	0.54 (0.17)	0.50 (0.17)	
No acento	0.48 (0.17)	0.56 (0.18)	0.52 (0.18)	
TOTAL	0.47 (0.17)	0.55 (0.18)	0.51 (0.18)	n = 746

Tabla 6.51. M₂₅: %t pico de magnitud del movimiento de ascenso de la mandíbula.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.56	0.02	26.53			
Unidad	-0.08	0.02	-4.79	(1, 715.6) = 47.88	p < 0.001	***
Acento	-0.02	0.02	-1.61	(1, 715.6) = 6.08	p < 0.05	*
Consonante				(7, 715.6) = 8.70	p < 0.001	***
β	0.03	0.03	1.13			
ð	-0.01	0.03	-0.51			
f	0.10	0.03	3.86			
m	0.10	0.03	3.64			
n	-0.08	0.03	-3.35			
p	-0.01	0.03	-0.54			
s	-0.03	0.03	-1.16			
t̪	(-0.10)					
Unidad x Acento				(1, 715.6) = 0.04	p = 0.837	ns
Consonante x Unidad				(7, 715.6) = 1.40	p = 0.202	ns
Consonante x Acento				(7, 715.6) = 2.73	p < 0.01	**

Tabla 6.52. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del gesto de ascenso del ápice de la lengua.

	Palabra	Sílaba	TOTAL	
Acento	0.48 (0.17)	0.53 (0.18)	0.50 (0.18)	
No acento	0.51 (0.20)	0.50 (0.17)	0.50 (0.19)	
TOTAL	0.49 (0.19)	0.51 (0.18)	0.50 (0.18)	n = 380

Tabla 6.53. M₂₆: %t pico de magnitud del gesto de ascenso del ápice de la lengua.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.50	0.02	22.58			
Unidad	0.01	0.03	0.09	(1, 361.9) = 1.58	p = 0.210	ns
Acento	0.03	0.03	0.89	(1, 361.9) = 0.02	p = 0.898	ns
Consonante				(3, 361.9) = 5.67	p < 0.001	***
ð	0.07	0.03	2.44			
n	-0.05	0.03	-1.81			
s	0.03	0.03	1.01			
t̪	(-0.05)					
Unidad x Acento				(1, 361.9) = 1.92	p = 0.167	ns
Consonante x Unidad				(3, 361.9) = 0.50	p = 0.681	ns
Consonante x Acento				(3, 361.9) = 0.39	p = 0.758	ns

En el caso del gesto labial (Tabla 6.54 y Tabla 6.55), el M_{27} refleja que los efectos principales de Unidad y Acento no son significativos, aunque sí su interacción (hay un patrón cruzado, como se aprecia en la primera de esas tablas).

También es significativo el efecto principal de la Consonante, de modo que la oclusiva [p] es la que antes alcanza el pico de magnitud del gesto de constricción labial, seguida de la fricativa [f], la nasal [m] y la espirante [β], en ese orden, aunque solo los dos extremos —[p] y [β]— difieren significativamente de la media intergrupala.

Por lo que se refiere a los datos aerodinámicos, serán tomados como indicio indirecto del desarrollo del gesto velar. Técnicamente, puede haber un ligero retraso entre el momento en que se alcanza el pico de magnitud del gesto velar y el momento en que se registra, a la salida de la nariz, el pico de flujo y el pico de presión. Sin embargo, está claro que, si uno de los factores prosódicos ocasiona un retardo en la consecución del pico articulatorio, también se verá afectado el registro aerodinámico. Por tanto, esta medida es válida para comparar los efectos relativos de cada posición prosódica.

Los datos de flujo nasal se muestran en la Tabla 6.56 y la Tabla 6.57, mientras que los resultados de la presión nasal aparecen recogidos en la Tabla 6.58 y la Tabla 6.59. Ambos modelos (M_{28} y M_{29}) son bastante similares, aunque en el primero hay un efecto aleatorio del Sujeto que afecta a la pendiente de Acento, mientras que en el segundo es la pendiente de Unidad la afectada.

En efecto, la Figura 6.12 muestra que Pablo experimenta un significativo adelanto del pico de magnitud de U_n en $^1\sigma$, y Rebecca apunta en la misma dirección pero de manera mucho más tímida. En cambio, Adriana y Lourdes producen dicho pico de manera más tardía bajo el efecto de la prominencia acentual (recuérdese que esta sería la dirección del cambio prevista por la H_{12B}), pero sin llegar a un efecto significativo.

Tabla 6.54. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del gesto de constricción labial.

	Palabra	Sílaba	TOTAL	
Acento	0.57 (0.15)	0.55 (0.14)	0.56 (0.14)	
No acento	0.51 (0.16)	0.57 (0.13)	0.54 (0.15)	
TOTAL	0.54 (0.16)	0.56 (0.14)	0.55 (0.15)	n = 366

Tabla 6.55. M_{27} : %t pico de magnitud del gesto de constricción labial.

	β	SE	t (df) F	p (test F)	
Ordenada en el origen	0.57	0.03	19.55		
Unidad	-0.06	0.02	-2.81	(1, 347.8) = 0.94	p = 0.334 ns
Acento	-0.02	0.02	-1.24	(1, 347.8) = 1.50	p = 0.222 ns
Consonante				(3, 347.8) = 6.26	p < 0.001 ***
β	0.04	0.02	2.03		
f	-0.02	0.02	-0.93		
m	0.03	0.02	1.50		
p	(-0.05)				
Unidad x Acento				(1, 347.8) = 8.75	p < 0.01 **
Consonante x Unidad				(3, 347.8) = 1.97	p = 0.118 ns
Consonante x Acento				(3, 347.8) = 0.76	p = 0.517 ns

Tabla 6.56. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud del flujo nasal.

	Palabra	Sílaba	TOTAL	
Acento	0.43 (0.31)	0.53 (0.29)	0.48 (0.30)	
No acento	0.60 (0.31)	0.52 (0.36)	0.56 (0.33)	
TOTAL	0.52 (0.32)	0.52 (0.32)	0.52 (0.32)	n = 123

Tabla 6.57. M₂₈: %t pico de magnitud del flujo nasal.

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.52	0.09	5.48			
Unidad	0.08	0.06	1.35	(1, 110.6) = 0.01	p = 0.938	ns
Acento (Sujeto)	0.01	0.14	0.04	(1, 3) = 0.40	p = 0.574	ns
Consonante				(1, 110.6) = 43.83	p < 0.001	***
m	-0.21	0.04	-5.51			
n	(0.21)					
Unidad x Acento				(1, 110.6) = 3.77	p = 0.055	ns
Consonante x Unidad				(1, 110.6) = 3.36	p = 0.069	ns
Consonante x Acento				(1, 110.6) = 0.84	p = 0.361	ns

Tabla 6.58. Media (y desviación estándar), por condiciones prosódicas, del %t pico de magnitud de la presión nasal.

	Palabra	Sílaba	TOTAL	
Acento	0.57 (0.25)	0.62 (0.21)	0.60 (0.23)	
No acento	0.66 (0.28)	0.53 (0.25)	0.60 (0.27)	
TOTAL	0.62 (0.27)	0.57 (0.24)	0.60 (0.25)	n = 123

Tabla 6.59. M₂₉: %t pico de magnitud de la presión nasal.

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.53	0.03	15.71			
Unidad (Sujeto)	0.13	0.09	1.62	(1, 3.6) = 0.29	p = 0.622	ns
Acento	0.09	0.05	1.78	(1, 113.2) = 0.09	p = 0.759	ns
Consonante				(1, 113.2) = 57.24	p < 0.001	***
m	-0.17	0.03	-5.67			
n	(0.17)					
Unidad x Acento				(1, 113.2) = 8.01	p < 0.01	**
Consonante x Unidad				(1, 113.2) = 3.81	p = 0.053	ns
Consonante x Acento				(1, 113.2) = 0.03	p = 0.862	ns

Por su parte, la Figura 6.13 deja ver que, para Lourdes y Pablo, la ubicación de la consonante en posición inicial de palabra motiva un retraso en el pico de magnitud de P_n , pero para Rebecca y, sobre todo, para Adriana sucede justo lo contrario.

En ninguno de los dos modelos llega a alcanzarse un valor significativo para los efectos principales de Unidad y Acento, seguramente por la discrepancia entre los sujetos, aunque la tendencia de la Unidad apunta en la dirección esperada (el retraso del pico en ω).

Por otro lado, hay un efecto significativo de la Consonante, tal que el pico de magnitud —tanto de U_n como de P_n — se alcanza antes en la [m] que en la [n]. Por último, el %t (P_n) se ve afectado por una interacción Unidad x Acento (efecto cruzado).

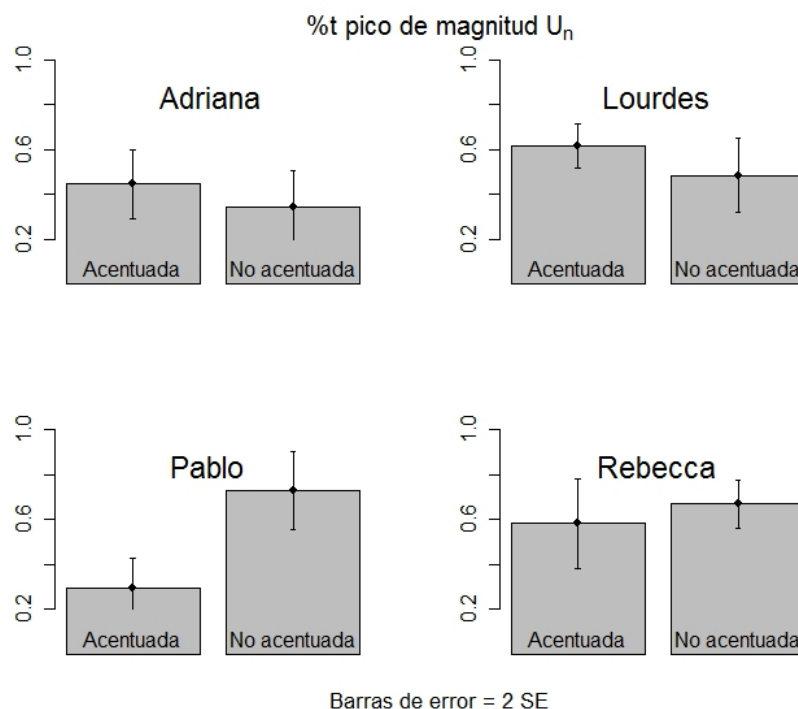


Figura 6.12. M_{28} : Efecto aleatorio del Sujeto sobre la pendiente de Acento en el %t pico de magnitud del flujo nasal.

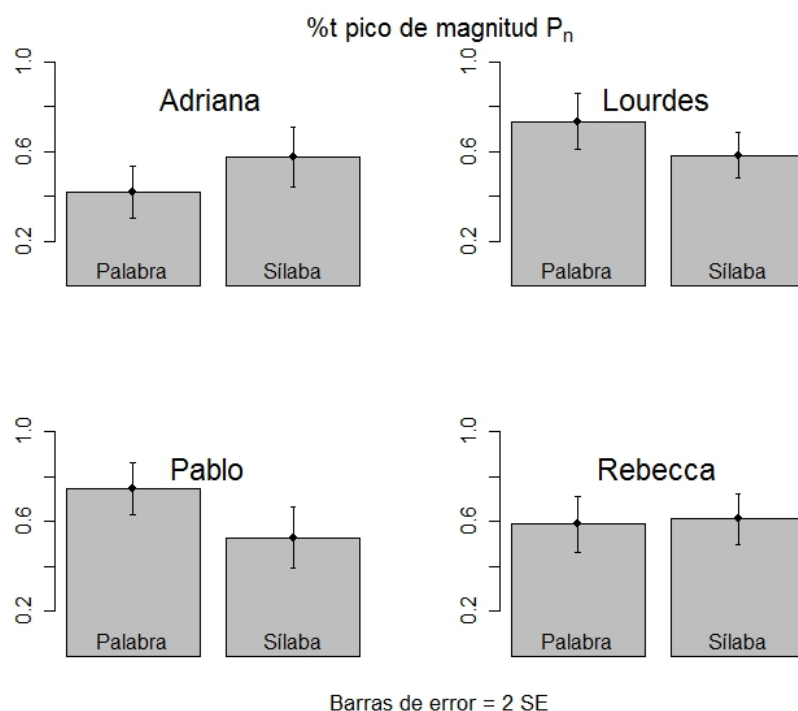


Figura 6.13. M_{29} : Efecto aleatorio del Sujeto sobre la pendiente de Unidad en el %t pico de magnitud de la presión nasal.

6.13.H₁₃: intensidad relativa

Por último, la H_{13A} predecía una menor intensidad relativa de las consonantes en posición inicial de palabra, excepto para las fricativas, mientras que la H_{13B} relaciona la prominencia acentual con una menor intensidad relativa en el caso de las espirantes, pero mayor (que en σ) en el caso de las fricativas y nasales. Veamos si los resultados de los modelos correspondientes confirman estas predicciones.

El M₃₀ se refiere a la intensidad de las fricativas (*vid.* Tabla 6.60 y Tabla 6.61), y presenta un efecto significativo de la Unidad, tal que las fricativas son más intensas en la condición ω que en σ (conforme a la H_{13A.a}). En cuanto al Acento, la dirección del cambio apunta una tendencia en dirección contraria, y frente a lo esperado. No obstante, la significatividad de este factor resulta controvertida, ya que es apoyada por el test F pero no por el test t. Lo mismo ocurre con la Consonante.

El comportamiento de las espirantes (M₃₁) se describe en la Tabla 6.62 y la Tabla 6.63. El único efecto significativo es el del Acento, y eso aun teniendo en cuenta que este modelo presenta un efecto aleatorio del Sujeto (por diferencias de magnitud) que afecta, precisamente, a la pendiente de dicha variable. Las espirantes en sílaba acentuada presentan una menor intensidad relativa que las situadas en sílaba no acentuada, algo que se puede interpretar como indicio de un mayor grado de constricción oral. Este resultado concuerda con lo esperado (H_{13B.b}).

Tabla 6.60. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las fricativas (proporción).

	Palabra	Sílaba	TOTAL	
Acento	0.70 (0.07)	0.70 (0.05)	0.70 (0.06)	
No acento	0.75 (0.06)	0.72 (0.06)	0.73 (0.06)	
TOTAL	0.72 (0.07)	0.71 (0.06)	0.71 (0.06)	n = 187

Tabla 6.61. M₃₀: Intensidad relativa de las fricativas.

	β	SE	t (df)	F	p (test F)	
Ordenada en el origen	0.72	0.01	50.40			
Unidad	0.03	0.01	2.79	(1, 174.7) = 4.12	p < 0.05	*
Acento	-0.02	0.01	-1.83	(1, 174.7) = 20.09	p < 0.001	***
Consonante				(1, 174.7) = 4.19	p < 0.05	*
f	0.01	0.01	1.26			
s	(-0.01)					
Unidad x Acento				(1, 174.7) = 3.62	p = 0.059	ns
Consonante x Unidad				(1, 174.7) = 0.34	p = 0.560	ns
Consonante x Acento				(1, 174.7) = 0.52	p = 0.473	ns

Tabla 6.62. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las espirantes (proporción).

	Palabra	Sílaba	TOTAL	
Acento	0.88 (0.08)	0.89 (0.05)	0.88 (0.07)	
No acento	0.94 (0.06)	0.93 (0.06)	0.94 (0.06)	
TOTAL	0.91 (0.08)	0.91 (0.06)	0.91 (0.07)	n = 183

Tabla 6.63. M₃₁: Intensidad relativa de las espirantes.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.93	0.02	50.59			
Unidad	0.01	0.01	0.49	(1, 165.6) = 0.03	p = 0.853	ns
Acento ([Sujeto])	-0.04	0.02	-2.40	(1, 4.9) = 9.25	p < 0.05	*
Consonante				(1, 165.6) = 1.38	p = 0.242	ns
β	-0.01	0.01	-1.11			
ϕ	(0.01)					
Unidad x Acento				(1, 166.1) = 0.77	p = 0.381	ns
Consonante x Unidad				(1, 165.6) = 0.22	p = 0.639	ns
Consonante x Acento				(1, 166.1) = 0.06	p = 0.811	ns

Tabla 6.64. Media (y desviación estándar), por condiciones prosódicas, de la intensidad relativa de las nasales (proporción).

	Palabra	Sílaba	TOTAL	
Acento	0.90 (0.05)	0.88 (0.05)	0.89 (0.05)	
No acento	0.91 (0.06)	0.94 (0.03)	0.93 (0.05)	
TOTAL	0.91 (0.05)	0.91 (0.05)	0.91 (0.05)	n = 190

Tabla 6.65. M₃₂: Intensidad relativa de las nasales.

	β	SE	t	(df) F	p (test F)	
Ordenada en el origen	0.94	0.01	77.53			
Unidad	-0.03	0.01	-2.60	(1, 177.6) = 0.02	p = 0.900	ns
Acento	-0.06	0.01	-6.72	(1, 177.6) = 35.36	p < 0.001	***
Consonante				(1, 177.6) = 16.68	p < 0.001	***
m	-0.01	0.00	-2.26			
n	(0.01)					
Unidad x Acento				(1, 177.6) = 11.96	p < 0.001	***
Consonante x Unidad				(1, 177.6) = 0.32	p = 0.571	ns
Consonante x Acento				(1, 177.8) = 0.08	p = 0.779	ns

Por último, la intensidad relativa de las nasales (M₃₂) se describe en la Tabla 6.64 y la Tabla 6.65. La [m] es significativamente menos intensa que la [n]. La ubicación en una sílaba prominente hace que disminuya la intensidad relativa de cualquiera de esos segmentos (un descubrimiento contrario al esperado, según la H_{13B:c}). En cuanto al efecto principal de la Unidad, vuelve a darse un caso de contradicción entre el resultado del test F y del test t (el primero considera que la aportación de la variable al modelo no es significativa, mientras que el segundo señala que el coeficiente estimado para la

condición ω es significativamente distinto de cero y, por tanto, contribuye al modelo). En cualquier caso, la tendencia presentada por esta variable consiste en una reducción de la intensidad en posición inicial de palabra (lo que coincide con la previsión de $H_{13A:c}$). Finalmente, hay una interacción Unidad x Acento, tal que la intensidad de las nasales es mayor en sílaba no prominente, pero aún mayor si esa sílaba es interior de palabra (*vid.* Figura 6.14).

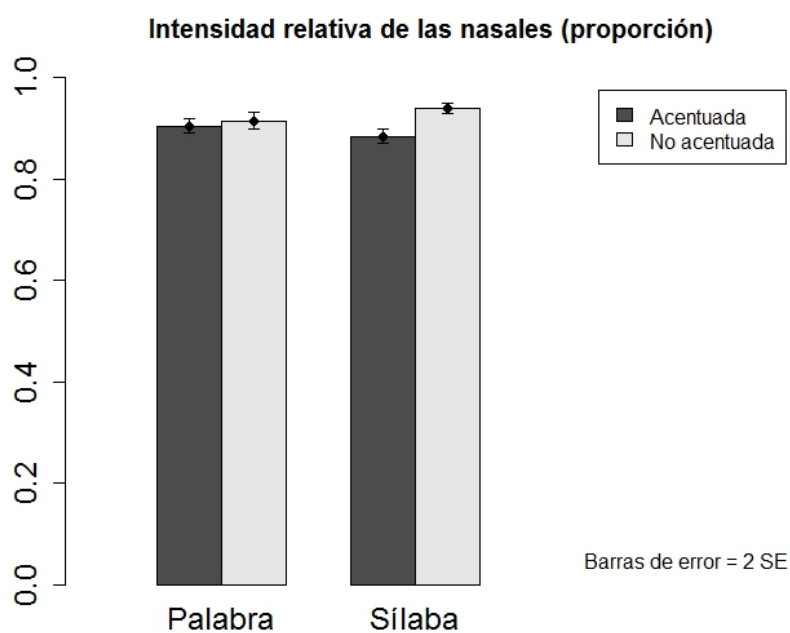


Figura 6.14. M_{32} : Interacción Unidad x Acento en la intensidad relativa de las nasales.

7. Discusión y conclusiones

La prosodia cumple un papel fundamental en la fonología de las lenguas, al proporcionar una organización de los elementos de contenido fonológico mediante una estructura jerárquica. Además, esta estructura prosódica hace las veces de interfaz entre el componente fonológico y otros componentes lingüísticos, como el morfosintáctico o el semántico, de modo que la identificación de dicha estructura puede ser de gran ayuda para la correcta interpretación del significado de un enunciado.

En esta tesis se han analizado los fenómenos de refuerzo consonántico como cambios en el desarrollo espaciotemporal de los gestos articulatorios provocados por las posiciones prosódicas en que estos aparecen. Los fenómenos de refuerzo (y de debilitamiento) se reflejan en marcas fonéticas que permiten a los usuarios de la lengua codificar y decodificar la organización jerárquica de los elementos fonológicos. En efecto, la dinámica gestual sufre modificaciones específicas en posiciones colindantes con una frontera entre dos constituyentes prosódicos, y las marcas fonéticas que de ello derivan pueden cumplir una función demarcativa al poner de manifiesto la presencia de dicha frontera. Por otro lado, hay fenómenos de refuerzo desencadenados por el carácter prominente de un constituyente frente a otros del mismo nivel, y sus marcas fonéticas asociadas desempeñan una función culminativa que permite reconocer el número de unidades que existen en el nivel jerárquico inmediatamente superior.

Ambos tipos de refuerzo (por causa de la posición y por causa de la prominencia) comparten algunas características pero difieren en otras. Por otro lado, los efectos de uno y otro presentan tendencias comunes en todas las lenguas que se han estudiado hasta ahora, pero también muestran diferencias de un idioma a otro. No en vano, el refuerzo afecta a los gestos articulatorios que están fonológicamente especificados en una determinada palabra, de modo que las diferencias fonológicas inherentes a los distintos sistemas lingüísticos justifican el distinto efecto que el refuerzo puede tener en uno u otro idioma.

Tal carácter idiosincrásico justifica que esta tesis haya abordado el estudio del refuerzo en una lengua como el español, que no se había investigado en profundidad desde esta perspectiva. Para ello se ha recopilado un corpus de habla de seis hablantes mexicanos¹⁶⁶, que incluye datos articulográficos, aerodinámicos y acústicos. Este estudio ha permitido identificar algunos de los efectos que presentan el refuerzo en posición inicial de palabra (frente a inicial de sílaba) y el refuerzo asociado a las sílabas prominentes (acentuadas). Por otro lado, algunos de los efectos que esperábamos encontrar no han resultado significativos. Esta ausencia puede deberse a que esas

¹⁶⁶ Aunque en las siguientes páginas se hable, en general, del «español», cabe precisar que no podemos estar seguros de si los mismos resultados se replicarían en otras variedades de esta lengua. No obstante, dado que los efectos del refuerzo (por posición o por prominencia) actúan sobre la base de los gestos articulatorios fonológicamente especificados en la forma léxica de las palabras, tampoco se esperarían grandes diferencias entre las distintas zonas hispanohablantes a menos que en una de ellas supusiéramos un sistema fonológico (parcialmente) distinto.

marcas no afloren hasta niveles superiores de la jerarquía, que faltarían por investigar. También puede ser que en español nunca lleguen a observarse dichas marcas.

En cualquier caso, los datos aquí obtenidos son compatibles con la idea —reiterada hace un momento— de que el refuerzo presenta puntos comunes y puntos diferenciales entre distintas lenguas, como se verá en la interpretación de conjunto que se ofrece a continuación¹⁶⁷.

Para facilitar una visión global de los distintos resultados obtenidos y abonar la subsiguiente discusión de los mismos, la Tabla 7.1 recoge el sentido de los distintos efectos significativos encontrados para la Unidad y el Acento. Una marca de verificación (✓) señala los casos que coinciden con la predicción de la respectiva hipótesis, mientras que una equis (✕) marca los resultados que van en sentido contrario al esperado. Los puntos suspensivos (...) identifican los casos que no obtuvieron un resultado significativo.

Aunque la H_1 no recogía ninguna predicción específica sobre el efecto de la prominencia, en la tabla se han registrado los dos efectos significativos que se han encontrado al respecto.

7.1. Efectos de la posición prosódica

El modelo de los gestos π de Byrd y Saltzman (2003) postula como causa primera de los efectos posicionales de refuerzo una ralentización local en la ejecución de los gestos articulatorios adyacentes a una frontera prosódica, ralentización que es tanto mayor cuanto más alto en la jerarquía se sitúe el constituyente prosódico en cuestión. En efecto, los datos del español concuerdan con esta propuesta, pues hemos visto que la velocidad de constricción correspondiente a los movimientos de la mandíbula y de los labios se ve sensiblemente reducida en la posición inicial de palabra, por comparación con la inicial de sílaba. No se puede decir lo mismo del movimiento del ápice, cuya deceleración no llega a ser significativa, aunque sí apunta en la dirección esperada. Es posible que la pérdida de velocidad en dicho movimiento sí aflore con efectos significativos en niveles superiores de la jerarquía prosódica (aunque la comprobación de si esto es cierto o no deberá quedar pendiente para futuros estudios). Por otro lado, debido al carácter local de la ralentización, el subsiguiente gesto de distensión no siempre se ve afectado en su velocidad: en el corpus analizado, solo el movimiento de la mandíbula (pero ya no el de los labios) continúa mostrando una modificación significativa.

Una consecuencia lógica de la ralentización debería ser el aumento de la duración gestual (*vid.* los ejemplos revisados en el capítulo 3). Sin embargo, hemos encontrado resultados contrarios a esta expectativa, al detectar un decremento temporal en el

¹⁶⁷ La discusión se centrará en los efectos principales de las variables Unidad y Acento, foco de interés de este estudio. Por ejemplo, no se entrará a analizar en detalle las diferencias de comportamiento entre las distintas consonantes que se han observado en el capítulo anterior, aunque allí se notificaran como resultados que son (recuérdese que el hecho de incluir la variable Consonante en los modelos respondía a la intención de controlar las variaciones inherentes a la identidad segmental para dar a la Unidad y el Acento su peso justo en la explicación de la varianza de los datos). Por supuesto, los efectos ajenos a las variables prosódicas Unidad y Acento son también interesantes en sí mismos y merecen una investigación detallada, pero exceden los propósitos de esta tesis.

Tabla 7.1. Recapitulación de los resultados del refuerzo por posición y por prominencia.

REFUERZO POR POSICIÓN			REFUERZO POR PROMINENCIA		
Pico de velocidad de la constricción					
✓	mandíbula	$\omega < \sigma$	${}^l\sigma < \sigma$	mandíbula	
	ápice	...	...	ápice	
✓	labios	$\omega < \sigma$	${}^l\sigma < \sigma$	labios	
Pico de velocidad de la distensión					
✓	mandíbula	$\omega < \sigma$	${}^l\sigma > \sigma$	mandíbula	✓
	ápice	...	...	ápice	
	labios	...	${}^l\sigma > \sigma$	labios	✓
Duración del gesto de constricción					
✗	mandíbula	$\omega < \sigma$	...	mandíbula	
✗	ápice	$\omega < \sigma$	...	ápice	
	labios	...	...	labios	
Duración del gesto de distensión					
	mandíbula	...	${}^l\sigma > \sigma$	mandíbula	✓
	ápice	...	${}^l\sigma > \sigma$	ápice	✓
	labios	...	${}^l\sigma > \sigma$	labios	✓
Solapamiento temporal entre los movimientos articulatorios de la constricción					
✓	mandíbula-ápice	$\omega < \sigma$	...	mandíbula-ápice	
✓	mandíbula-labios	$\omega < \sigma$	...	mandíbula-labios	
Duración acústica del segmento consonántico					
	orales	...	${}^l\sigma > \sigma$	orales	✓
	nasales	...	${}^l\sigma > \sigma$	nasales	✓
VOT					
✓		$\omega = \sigma$	${}^l\sigma = \sigma$		✓
Altura de la mandíbula					
✓		$\omega > \sigma$	...		
Altura del ápice					
✓		$\omega > \sigma$	...		
Abertura labial					
✗		$\omega > \sigma$	...		
Flujo y presión nasal					
✓		$\omega < \sigma$	...		
% t pico de magnitud, y pico U_n / P_n					
✗	mandíbula	$\omega < \sigma$	${}^l\sigma < \sigma$	mandíbula	✗
	ápice	...	...	ápice	
	labios	...	...	labios	
	U_n / P_n	...	...	U_n / P_n	
Intensidad relativa					
✓	fricativas	$\omega > \sigma$	${}^l\sigma < \sigma$	fricativas	✗
	espirantes	...	${}^l\sigma < \sigma$	espirantes	✓
	nasales	...	${}^l\sigma < \sigma$	nasales	✗

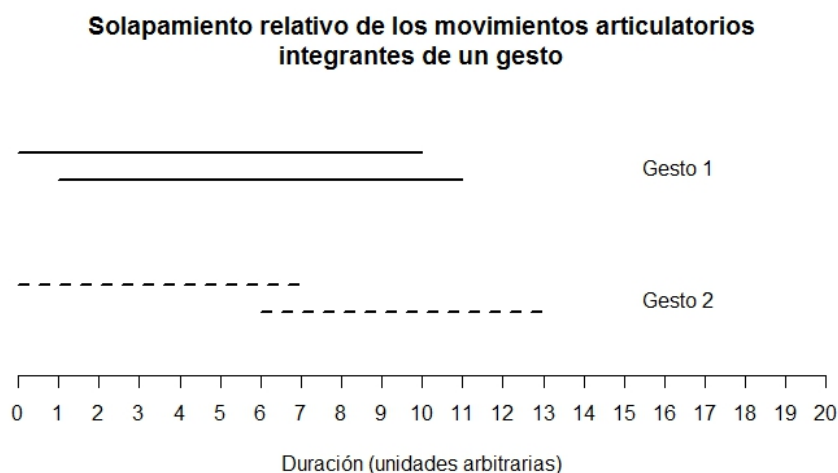


Figura 7.1. Solapamiento relativo de los movimientos articulatorios integrantes de un gesto.

movimiento de la mandíbula y en el del ápice de la lengua. Ahora bien, los gestos orales se caracterizan por el movimiento coordinado de la mandíbula con el de otros órganos, como la lengua o los labios. Las particularidades de dicha coordinación están sujetas a variación entre distintas lenguas y también entre hablantes de una misma lengua. Creo que el inesperado resultado de la reducción temporal del movimiento articulatorio de la mandíbula y del ápice podría verse iluminado si tenemos en cuenta otro de los resultados de este estudio, a saber, la disminución —siempre en inicial de palabra— del solapamiento temporal entre dichos movimientos. Es posible pensar que, aunque cada uno de estos movimientos dure menos en sí mismo, sí puede aumentar la duración total del gesto compuesto por ambos, como ilustra la Figura 7.1. En ella se representa de manera idealizada la extensión temporal de dos gestos, cada uno de los cuales está integrado por el movimiento coordinado de dos estructuras u órganos articulatorios. Si bien los movimientos que integran el primer gesto (en línea continua) son, cada uno de ellos, más largos que los movimientos del segundo de los gestos (en línea discontinua), el menor grado de solapamiento de estos últimos hace que el segundo gesto (contemplado en su totalidad) dure más que el primero. Aunque sería necesario hacer una investigación más detallada al respecto, puede que un enfoque integrado de estos dos resultados (*i. e.* la reducción temporal de los movimientos y la reducción temporal del solapamiento) permita alcanzar una explicación compatible con la expectativa de que la duración de los gestos debe aumentar en posición inicial de palabra (frente a la inicial de sílaba)¹⁶⁸.

Del mismo modo que disminuye el solapamiento entre la mandíbula y el ápice, también se ha constatado una reducción del tiempo de coactivación de la mandíbula y los labios. En el § 3.1 se expuso que uno de los efectos del refuerzo posicional es que, a través de una frontera prosódica, se reduce el nivel de coarticulación entre segmentos o

¹⁶⁸ Además, los futuros trabajos en esta línea podrían enriquecerse comparando distintas velocidades de habla, pues este factor puede modificar no solo las duraciones de los distintos constituyentes, sino también la coordinación temporal entre gestos o segmentos sucesivos, como sugieren Celata y Calamai (2011).

entre gestos contiguos. Aquí se ha comprobado que también existe una menor coarticulación entre los movimientos que pertenecen a un mismo gesto. Todo ello convierte a esta posición prosódica en una posición relativamente poco proclive a los procesos de asimilación. Esto no quiere decir que no puedan darse asimilaciones contextuales en frontera de palabra pero, si lo hacen, se puede establecer la implicación de que seguro que también actuarán en interior de palabra si las demás condiciones permanecen constantes.

Por otro lado, quizá también esta modificación en la coordinación intragestual pueda explicar por qué, frente a lo esperado, el pico de magnitud del movimiento de la mandíbula se alcanza antes en posición inicial de palabra que en inicial de sílaba. De hecho, este es el único movimiento articulatorio que ha mostrado un cambio significativo de dicho parámetro en tal posición. Los movimientos de la mandíbula presentan características peculiares frente a los demás órganos articulatorios, pues la mandíbula es una estructura cuyo desplazamiento acompaña al de otros órganos (el movimiento mandibular es el elemento común de todos los gestos orales). Por otro lado, en el § 3.5.2 vimos cómo los movimientos de la mandíbula presentan un notable grado de independencia con respecto a los otros movimientos articulatorios, pudiendo incluso presentar movimientos contrapuestos a estos en determinadas ocasiones. Por estos motivos, debemos estudiar más a fondo el comportamiento de la mandíbula en relación con el de los órganos con que se coordina, ya que ese tipo de investigación podría iluminar resultados como los aquí obtenidos.

Otra de las hipótesis preveía una modificación en la duración acústica de las consonantes en inicial de palabra, con un efecto diferente en el caso de los segmentos orales y en el de los nasales. Sin embargo, este es uno de los casos en los que no se ha encontrado ningún efecto significativo de la posición prosódica.

Tampoco hay diferencias significativas en el VOT, dato que concuerda con las expectativas. Teniendo en cuenta que el español es una lengua que basa las oposiciones laringeas de las obstruyentes en la presencia o ausencia de sonorización (y no de aspiración), se puede asumir que las oclusivas sordas son el elemento fonológicamente no marcado de la oposición y que, por tanto, carecen de especificación alguna en cuanto al gesto glotal —siempre desde el punto de vista fonológico—. Por ello, el refuerzo posicional no puede tomar como objetivo el gesto glotal, ya que este se halla ausente de la partitura de gestos correspondiente. Este es un ejemplo de cómo los efectos del refuerzo pueden ser distintos en lenguas que difieren en sus sistemas fonológicos, pues las lenguas de aspiración sí verían incrementado el VOT (*vid.* § 3.3.1).

Por su parte, las medidas de la magnitud gestual (correspondientes a la altura de la mandíbula, la altura del ápice, la abertura labial y el flujo y la presión nasales) muestran todas un efecto significativo de la posición inicial de palabra. Como se justificó en el § 3.1, la peculiar dinámica de los gestos implica que, dado un aumento temporal, estos pueden alcanzar una mayor magnitud cuando van seguidos de otro gesto correspondiente a la misma variable de tracto, pero de diferente valor. En cambio, la magnitud decrece cuando el gesto siguiente controla una variable de tracto distinta. En efecto, hemos observado un aumento de la magnitud en el movimiento de la mandíbula,

que alcanza una posición más alta en las consonantes situadas en inicio de palabra. Del mismo modo, el ápice de la lengua también experimenta este tipo de refuerzo espacial. Por su parte, como esperábamos, el descenso de U_n y P_n se interpreta como una menor bajada del velo del paladar durante la ejecución del gesto de descenso velar de las nasales en posición inicial de palabra. Es decir, la magnitud del gesto es menor de la habitual, hecho que se justifica en el carácter monovalente del gesto velar, gesto que, por tanto, nunca va seguido de otro gesto que modifique la posición del velo en un sentido diferente.

Ahora bien, los resultados correspondientes al gesto de constricción labial son contrarios a los propuestos por la hipótesis H_{10A} . En posición inicial de palabra se constata una mayor abertura labial, es decir, un menor desplazamiento de los labios en su movimiento de cierre. A la vista de este resultado, y en retrospectiva, creo que existe una explicación plausible para este comportamiento, similar al del gesto velar. En efecto, hay que tener en cuenta que las consonantes labiales del experimento (en realidad, todas las consonantes) aparecían seguidas de dos posibles vocales, la [a] y la [i], ninguna de las cuales presenta una especificación de un gesto labial. Por este motivo, el gesto de constricción labial de la consonante no va seguido de otro gesto labial, y esas son las circunstancias en las que se prevé una reducción de la magnitud en posición inicial de palabra, al igual que ocurre con el gesto velar. Esto explica, precisamente, el patrón observado de aumento de la distancia entre los labios en dicha posición prosódica. Habría que comprobar, en futuros estudios, si dicho comportamiento desaparece cuando la consonante va seguida de una [o] o una [u], vocales que sí incluyen un gesto labial (de redondeamiento) en su especificación fonológica.

Por último, hemos encontrado un aumento de la intensidad relativa de las fricativas (con respecto a la vocal a la que preceden) en la posición inicial de palabra, lo que seguramente responde a un mayor estrechamiento del canal de salida del aire, que provoca, a su vez, un mayor ruido de fricción. En el caso de la [s] está claro que el aumento de magnitud en el movimiento del ápice es el responsable de dicho estrechamiento. Por lo que respecta a la [f], podría sorprender este aumento en la intensidad del ruido de fricción a la luz del dato que hemos comentado de la mayor abertura labial. Sin embargo, hay que tener en cuenta que la turbulencia correspondiente a la [f] no se genera entre los dos labios, sino entre el labio inferior y los incisivos superiores. Por este motivo, una mayor separación interlabial no necesariamente implica un menor estrechamiento en el verdadero punto de articulación (labiodental), así que ambos resultados son perfectamente compatibles.

7.2. Efectos de la prominencia prosódica

En el caso del refuerzo por prominencia, la H_{1B} no preveía ningún cambio en la velocidad de la constricción, pero se ha encontrado el mismo efecto que en la posición inicial de palabra, es decir, la aminoración de dicha velocidad en el caso del movimiento de la mandíbula y el de los labios, aunque no hay un reflejo significativo en la duración (ni en la magnitud) de dicho gesto de constricción.

Esta dinámica contrasta con la del gesto correspondiente a la distensión consonántica (y formación de la vocal), que es donde radican las diferencias más importantes entre sílaba acentuada y no acentuada. La velocidad de la distensión es mayor en las sílabas prominentes: los resultados del español lo confirman en el caso del movimiento de la mandíbula y del de los labios.

Por otro lado, dicho gesto de distensión se desarrolla durante más tiempo (tanto en el caso de la mandíbula y de los labios como en el del ápice). Esto se interpreta como un retardo en el gesto de cierre subsiguiente a la vocal nuclear de la sílaba acentuada. Tiene sentido que, si la prominencia afecta a toda la unidad silábica y la sílaba es la unidad de coordinación gestual, la presencia del acento se manifieste mediante una reorganización de las relaciones temporales entre gestos sucesivos.

La extensión temporal del gesto de distensión se corresponde parcialmente con una porción acústica del final de la consonante prenuclear y con una porción acústica del inicio de la propia vocal. Dado el aumento de la duración articulatoria de dicho gesto, cabría esperar que también se vieran alargadas las correspondientes duraciones acústicas de ambos segmentos, como, en efecto, se ha comprobado. Los resultados del M_{13} (*vid.* Tabla 6.27 en la p. 220) confirman un alargamiento significativo de la vocal en condiciones de prominencia. Efectivamente, la mayor duración vocálica es una de las claves de las sílabas acentuadas, como se vio en el § 2.2.1. Además, de acuerdo con las expectativas de la H_{6B} , también aumenta la duración acústica de la consonante prenuclear. El alargamiento neto de la sílaba acentuada se manifiesta, pues, en dos segmentos: la vocal que ejerce de núcleo y la consonante inmediatamente anterior.

Tanto el aumento de velocidad de apertura como la mayor duración deben de permitir que el grado de abertura vocálica alcance una mayor magnitud, lo que, unido al propio alargamiento temporal, supone un incremento de la energía acústica radiada, como parece desprenderse de los resultados de la intensidad relativa. La H_{13B} esperaba un aumento de la intensidad relativa de las fricativas y las nasales con respecto a la vocal. Este incremento se debería, en el primer caso, a un mayor ruido de fricción ocasionado por el mayor estrechamiento del canal de salida del aire, causado por un hipotético aumento de magnitud en el gesto de constricción oral. Sin embargo, tal refuerzo espacial no se constata en los datos del corpus aquí analizado y, por el contrario, se observa una disminución de la intensidad relativa (en lugar de un aumento). En el caso de las nasales, también se esperaba que un aumento de magnitud del gesto velar (detectable como una subida en el valor de U_n y P_n) llevaría a una mayor intensidad relativa de dichos segmentos nasales. Frente a esas hipótesis, no hemos encontrado ninguna modificación significativa en los parámetros aerodinámicos, lo que se interpreta como una ausencia de cambio en la magnitud del gesto del velo del paladar. Del mismo modo que ocurre con las fricativas, también se constata una menor intensidad relativa de las nasales en sílaba acentuada, frente a lo esperado.

De estos datos podemos deducir que, si ha habido algún pequeño cambio de magnitud en los respectivos gestos articulatorios, no ha sido suficiente para ser detectado como significativo. Por otro lado, lo que sí ha aumentado ha sido la intensidad de la vocal, como consecuencia de las modificaciones en la velocidad y en la

duración del gesto de distensión. De este modo, al permanecer inalterada la intensidad de la consonante y aumentar la de la vocal, se puede entender que la intensidad relativa del segmento consonántico se haya reducido. Dicho de otra forma, los cambios en la vocal superan en magnitud a cualquier posible cambio menor en la consonante. Puede suponerse que esos cambios vocálicos, encaminados a un aumento de la duración y de la intensidad, son el objetivo de las modificaciones que la prominencia silábica introduce en la dinámica gestual. Siendo la vocal el núcleo de la sílaba, es comprensible que una mayor duración e intensidad vocálicas sirvan para poner de relieve toda la unidad prominente.

Por lo demás, las sílabas acentuadas no presentan ninguna modificación sensible del VOT, por los mismos motivos que los aducidos en el apartado anterior.

Tampoco se ha encontrado una diferencia significativa en el grado de solapamiento intragestual entre las sílabas acentuadas y no acentuadas. Efectivamente, no había ninguna hipótesis H_{5B} al respecto. Además, este resultado es coherente con algunas propuestas acerca de la clasificación tipológica de las lenguas en cuanto al ritmo. Por ejemplo, Bertinetto (1989) repasa algunos de los problemas que se han encontrado los investigadores cuando han intentado detectar una base empírica para la tradicional distinción entre las lenguas de ritmo silábico y las lenguas de ritmo acentual. En ese trabajo, Bertinetto propone que ambos tipos rítmicos no se oponen —como se ha creído hasta ahora— en que uno fundamente su temporización en una unidad (la sílaba) y otro en otra (el pie), pues —argumenta— es difícil pensar que una lengua pueda mantener cierta flexibilidad en las duraciones de una de esas unidades y, al mismo tiempo, ser más bien rígida en cuanto a las duraciones de la otra¹⁶⁹. Más bien las lenguas presentan un mayor o menor grado de flexibilidad general en cuanto a las duraciones de los segmentos, tanto desde una perspectiva intrasilábica como intersilábica (*i. e.* dentro de un pie). Las lenguas tradicionalmente clasificadas como de ritmo acentual tienden a compensar la duración de los segmentos (fundamentalmente, las vocales): si bien una sílaba acentuada dura más que una no acentuada, esa diferencia tiende a reducirse cuando un pie debe dar cabida a más de una sílaba no acentuada. Por otro lado, también en estas lenguas las sílabas no acentuadas sufren una mayor reducción temporal. En cambio, las lenguas tradicionalmente descritas como de ritmo silábico tienden a controlar la duración de los segmentos y a mantener esta más o menos constante (*vid.* también Bertinetto y Bertini, 2008, por ejemplo).

En las lenguas de compensación, como el inglés, la mayor reducción que se produce en sílaba no acentuada se manifiesta, entre otras cosas, como un mayor solapamiento de gestos. Esto vuelve más difusas las fronteras entre las sílabas: en efecto, estas lenguas suelen presentar casos de ambisilabicidad (recuérdese lo explicado en el § 2.1.1.4). Además, los sistemas métricos de estas lenguas —argumenta Bertinetto en su trabajo de 1989— tienden a contar acentos. Esto es así, por un lado, porque la diferencia entre las sílabas acentuadas y no acentuadas se siente como más prominente y, por otro, precisamente debido a esa difuminación de las fronteras silábicas, que hace más difícil

¹⁶⁹ Saltzman *et al.* (2008) proponen un sistema, basado en la dinámica de tareas, que modela la manera en que se puede alcanzar el equilibrio dadas esas tensiones contrapuestas.

un sistema basado en el cómputo de sílabas. En cambio, las lenguas de control de la duración, como el español, se caracterizan por una menor diferencia en el grado de solapamiento entre las sílabas acentuadas y las no acentuadas (que es precisamente lo que muestran los resultados aquí obtenidos, con una diferencia no significativa desde el punto de vista estadístico). El menor solapamiento hace que las fronteras entre sílabas sean más nítidas y propicia que los sistemas métricos de la poesía de estas lenguas se basen más bien en el recuento silábico (ya que, por otro lado, el contraste entre sílabas acentuadas y no acentuadas no es tan acusado)¹⁷⁰.

7.3. Valoración global y aplicaciones de este estudio

Los resultados del análisis del corpus ilustran las variaciones en el detalle fonético que sufren una serie de consonantes del español como consecuencia de su especificación fonológica en términos de gestos articulatorios, y de su inclusión en determinadas posiciones de la jerarquía prosódica.

Son numerosos los efectos del refuerzo que se han podido constatar, incluso a pesar de que los niveles estudiados —la sílaba y la palabra fonológica— se sitúan en la mitad inferior de dicha jerarquía. Dado el carácter acumulativo del refuerzo, es de esperar que en los niveles superiores estas variaciones se amplíen en magnitud del efecto y que también afloren otros efectos que aquí no han alcanzado el nivel de significatividad requerido. Esto puede ser objeto de una prometedora investigación en el futuro.

Aquí ha parecido interesante centrarse en los niveles léxicos por varios motivos. Por un lado, como se argumentó e ilustró en el capítulo 4, las variaciones subsegmentales detectadas instrumentalmente apuntan tendencias que, a la larga, pueden consolidarse como cambios segmentales que tomen carta de naturaleza en el sistema fonológico de la lengua, bien en la forma de alternancias alofónicas sincrónicas, o bien como cambios fonémicos desde una perspectiva histórica. Además, de este modo, los fenómenos de refuerzo condicionan qué tipos de segmentos aparecen habitualmente en unas posiciones y otras en el seno de las palabras. En concreto, hemos visto que algunas particularidades de los patrones fonotácticos exclusivos de los extremos de palabra pueden justificarse a la luz de los efectos fortitivos de la frontera prosódica.

Por otro lado, como se vio en el § 2.3.3.1, en español la sílaba juega un importante papel en los procesos psicolingüísticos de segmentación del continuo sonoro que apoyan el acceso al léxico. La identificación de las unidades del nivel de la palabra prosódica, por motivos obvios, también es vital en dicho acceso. En definitiva, las marcas fonéticas que permiten diferenciar estos dos niveles prosódicos suponen una pista muy valiosa acerca de qué claves pueden ser aquellas en las que se basen los oyentes a la hora de segmentar en palabras los enunciados que están tratando de decodificar. Naturalmente, la confirmación del uso perceptivo de estas claves debe ser

¹⁷⁰ No obstante, Piera (2001) —entre otros autores— sostiene lo siguiente: «In mainstream Spanish verse both the distribution of stresses and the number of syllables are relevant. [...] Spanish metrics is foot-based» (p. 206). En cualquier caso, es cierto que los patrones acentuales están menos restringidos que en la poesía de las lenguas de «ritmo acentual»; en cambio, en español —según se argumenta en ese artículo— hay factores entonativos que condicionan los esquemas métricos aceptados en el verso.

objeto de futuros estudios. En ellos se debería poner a prueba, de manera controlada, la capacidad de los sujetos de tomar decisiones sobre la estructura prosódica de un enunciado cuando se varían las marcas fonéticas relevantes. Algunos de los resultados de esta tesis apuntan directamente a claves acústicas, como la duración segmental o la intensidad relativa de las consonantes. Sin embargo, otros parámetros, de naturaleza articulatoria, también pueden traducirse en modificaciones acústicas cuyo potencial como claves perceptivas debe ser investigado. Por ejemplo, los cambios en la velocidad articulatoria pueden tener su reflejo en la velocidad de las transiciones espectrales, y el aumento de la magnitud de los gestos también puede traer consigo un desplazamiento en los valores de los formantes o en el centro de gravedad.

Los rasgos fonéticos propios del refuerzo también pueden utilizarse, en el ámbito de las tecnologías del habla, para que los sistemas de reconocimiento automático de habla detecten la estructura prosódica de los enunciados (no hay que olvidar que hay casos en los que dicha estructura puede condicionar la interpretación semántica de la oración). Asimismo, si se afinan los mecanismos de síntesis de habla para que incluyan las variaciones de detalle fonético adecuadas, eso conseguirá resultados más naturales y más informativos para los oyentes.

Otra línea de investigación que puede resultar muy interesante tiene que ver con el análisis de los efectos del refuerzo por causa prosódica en las lenguas de signos. La dinámica de tareas se puede aplicar no solo a los movimientos de los órganos del habla, sino también a los movimientos propios de las lenguas de signos (*vid.* p. ej. el estudio de Tyrone, Nam, Saltzman, Mathur y Goldstein, 2010, sobre la lengua de signos americana —estadounidense—). Falta mucho por investigar en este ámbito y, que yo sepa, no existe ningún estudio de esta índole sobre la lengua de signos española. Es de esperar que los efectos motores provocados por la presencia de una frontera prosódica sean los mismos en este caso, a saber, ralentización y alargamiento gestual, con una posible modificación de su magnitud. Por otro lado, dependiendo de si ese cambio espacial consiste en un aumento o en una disminución, se puede llegar a establecer cuáles son los equivalentes¹⁷¹ signados de las variables de tracto, pues recuérdese que solo se espera un refuerzo sustantivo del gesto si este va seguido por otro que controle la misma variable (en caso contrario, se espera un debilitamiento de la magnitud).

En efecto, a lo largo de esta tesis se ha insistido en la necesidad de diferenciar las variaciones fonéticas causadas por la posición prosódica (que se manifiestan como instancias de refuerzo o de debilitamiento) de aquellas motivadas por el contexto segmental circundante (que se suelen reflejar como coarticulaciones o asimilaciones, aunque también pueden tener un efecto disimilatorio). No obstante, también se ha visto que ambos tipos de fenómenos pueden interactuar y, de hecho, lo hacen a menudo. Por ejemplo, las posiciones débiles son más propicias a la coarticulación, pues en posiciones de refuerzo los cambios en la dinámica gestual conllevan un menor grado de solapamiento entre gestos sucesivos, y también entre movimientos integrantes de un

¹⁷¹ Gran parte de la investigación que se ha hecho sobre la fonología de las lenguas de signos consiste en establecer qué elementos de estas corresponden a los rasgos distintivos, los segmentos, las sílabas, etc. de las lenguas orales (*vid.* p. ej. Brentari, 1995).

mismo gesto. Además, el tipo de modificación sobre la magnitud de los gestos colindantes con una frontera prosódica depende de si el gesto siguiente afecta o no a la misma variable de tracto. Esto es, según el contexto siguiente, el gesto inicial de un constituyente puede ver aumentada o disminuida su magnitud. Aquí hemos visto que la disminución es, efectivamente, lo que ocurre en el caso del gesto velar, que alcanza una posición de descenso menos acusada, y se ha apuntado que quizá es lo que ocurre también con el gesto labial de una consonante cuando la vocal siguiente no presenta otro gesto que también afecte a los labios.

Teniendo en cuenta el peculiar comportamiento del gesto velar en posición inicial de constituyente, se pueden hacer ciertas predicciones acerca de las características aerodinámicas de las consonantes en esa posición. Si las nasales se vuelven menos nasales y las orales ven aumentada su magnitud de constricción oral, todas las consonantes que experimentan el refuerzo posicional deben caracterizarse por un aumento en la presión oral (P_o). Además, la posición más elevada de la mandíbula contribuye a que disminuya el volumen de la cavidad oral y, por tanto, a que aumente dicha presión.

De hecho, es posible establecer experimentalmente el peso de la contribución de cada gesto articulatorio a la P_o , mediante un modelo de regresión lineal basado en las mediciones simultáneas de los distintos movimientos articulatorios y de las variaciones de presión dentro de la cavidad oral. Una mayor abertura glotal, una menor abertura velofaríngea y una menor abertura oral conllevan un aumento de la P_o , y viceversa. Sin embargo, puede ser interesante saber cómo afectan los distintos gestos a la presión cuando esos gestos no colaboran todos en la misma dirección. Es de esperar que el gesto glotal tenga un mayor peso que los demás gestos, por ejemplo, pues se trata de la única vía de entrada de aire a la cavidad oral durante la producción de las consonantes pulmonares (otra cuestión, también de interés, sería el comportamiento aerodinámico de los chasquidos —en inglés, *clicks*—, las implosivas y las eyectivas en condiciones de refuerzo). McGowan y Saltzman (1995) proponen la necesidad de incorporar parámetros aerodinámicos al modelo de la dinámica de tareas para lograr predicciones más completas. En concreto, añaden la presión subglotal (P_s) y la presión transglotal (P_t). Sin embargo, si se investigaran, como propongo aquí, las correlaciones múltiples entre los gestos y la P_o , sería posible ahorrar el parámetro de la P_t (que es igual a la diferencia entre P_s y P_o), puesto que la P_s ya está incluida en el modelo y la P_o se podría deducir a partir los datos de los demás gestos, también incluidos en el modelo.

En definitiva, son muchas y muy prometedoras las investigaciones que aún se pueden hacer sobre las relaciones entre fonética y fonología en el ámbito de la prosodia de las lenguas, en general, y del español, en particular. Espero que esta tesis haga una aportación en este campo y estimule la aparición de nuevos estudios en esta línea.

8. General discussion and conclusions

Prosody plays a fundamental role in a language's phonology by organizing elements of phonological content using a hierarchical structure. Additionally, this prosodic structure sometimes acts as an interface between the phonological component and other linguistic components, such as morphosyntax or semantics; therefore, identifying this structure may be of great help in correctly interpreting the meaning of an utterance.

This dissertation analyzed the phenomena of consonant strengthening as changes in the spatiotemporal development of articulatory gestures caused by the prosodic positions in which said gestures appear. Strengthening (and weakening) phenomena are reflected in phonetic cues that allow a language's users to code and decode the hierarchical organization of phonological elements. Gestural dynamics undergo specific modifications in positions adjacent to a boundary between two prosodic components, and the phonetic cues that arise therefrom may fulfill a demarcative function as they make the presence of said boundary more noticeable. On the other hand, there are strengthening phenomena that are triggered by the prominent nature of one component compared to others on the same level, and the associated phonetic cues fulfill a culminative function that allows us to recognize the number of units that exist on the hierarchical level immediately above.

Both types of strengthening (due to position and prominence) share some characteristics, but differ on others. Furthermore, the effects of one or the other have common tendencies in all of the languages that have been studied until now, but they also show some differences from one language to another. Because strengthening affects articulatory gestures that are phonologically specified in a certain word, the phonological differences between different linguistic systems explain the varying effects that strengthening may have in one language or another.

This idiosyncratic nature justifies the fact that this dissertation studied reinforcement in a language like Spanish, which had not been the subject of in-depth research from this perspective. To perform the study, a speech corpus has been compiled with six Mexican speakers,¹⁷² which includes articulographic, aerodynamic, and acoustic data. This study has enabled the identification of some of the effects of strengthening in the word-initial position (compared to the syllable-initial position) and strengthening associated with prominent (stressed) syllables. Moreover, some of the effects that we expected to find have turned out not to be significant. This absence may be due to the fact that these cues do not appear until higher levels of the hierarchy, which remains to be researched. It may also be that in Spanish said cues are never observed.

¹⁷² Although in the following pages a reference is made to "Spanish" in general, it should be mentioned that we cannot be sure that the same results would be replicated in other varieties of this language. However, given that the effects of strengthening (due to position or prominence) act upon the basis of the articulatory gestures that are phonologically specified in the lexical form of words, we would not expect large differences between the different Spanish-speaking areas, unless one had a (partially) different phonological system.

In any case, the data obtained in this study are compatible with the idea, mentioned above, that strengthening has common points and differences from language to language, as will be seen in the below interpretation.¹⁷³

In order to provide a comprehensive vision of the different results obtained and lay the groundwork for their discussion, Table 8.1 presents the different significant effects found for Unit and Stress. A check mark (✓) indicates those cases that coincide with the prediction in the respective hypothesis, while an x (✗) marks unexpected results. An ellipsis (...) identifies those cases in which a significant result was not obtained.

Although H₁ did not make a specific prediction regarding the effect of prominence, the table has included the two significant effects that have been observed in this regard.

8.1. Effects of prosodic position

The π -gesture model of Byrd & Saltzman (2003) postulates that the main cause of the positional effects of strengthening is a local slowing of the execution of articulatory gestures adjacent to a prosodic boundary; this slowing is more pronounced the higher the prosodic component in question is located in the hierarchy. Indeed, the Spanish data support this proposal, as we have seen that the constriction speed corresponding to jaw and lip movements noticeably decreases in the word-initial position, compared to the syllable-initial position. The same cannot be said of tongue tip movement; its deceleration is not significant, although it does point in the expected direction. It is possible that the loss of speed in said movement would indeed be significant in higher levels of the prosodic hierarchy (although proving this to be right or wrong must be the work of future studies). Moreover, due to the local nature of the slowing, the speed of the subsequent release gesture is not always affected: in the corpus analyzed here, only jaw movement (but not lip movement) continues to show a significant modification.

A logical consequence of slowing should be an increase in the gesture's duration (see the examples reviewed in chapter 3). However, we have found results contrary to this expectation, and have detected a temporal decrease of jaw and tongue tip movement. Oral gestures are characterized by the coordinated movement of the jaw with other organs, such as the tongue or lips. The particular details of said coordination are subject to variations between different languages and also between speakers of the same language. I believe that the unexpected result of the durational decrease of the articulatory movement of the jaw and tongue tip could be explained if we keep in mind another result from this study, that is, the decrease of temporal overlap (always in the word-initial position) between said movements. It is possible that, although each of these movements alone lasts for less time, the total duration of the gesture made up of

¹⁷³ The discussion will center on the main effects of the variables Unit and Stress, the focus of this study. For example, a detailed analysis will not be provided for the behavioral differences between the different consonants that have been observed in the prior chapter, although they were identified there as results (remember that the variable Consonant was included in the models with the intention of controlling the variations inherent to segmental identity so as to give Unit and Stress the appropriate weight in the explanation of variance in the data). The effects not related to the prosodic variables Unit and Stress are, of course, interesting in and of themselves and deserve detailed study; however this analysis exceeds the goals of this dissertation.

Table 8.1 Summary of the results for strengthening due to position and prominence.

STRENGTHENING DUE TO POSITION			STRENGTHENING DUE TO PROMINENCE		
Constriction peak velocity					
✓	jaw	$\omega < \sigma$	${}^l\sigma < \sigma$	jaw	
	tongue tip	...	...	tongue tip	
✓	lips	$\omega < \sigma$	${}^l\sigma < \sigma$	lips	
Release peak velocity					
✓	jaw	$\omega < \sigma$	${}^l\sigma > \sigma$	jaw	✓
	tongue tip	...	...	tongue tip	
	lips	...	${}^l\sigma > \sigma$	lips	✓
Duration of the constriction gesture					
✗	jaw	$\omega < \sigma$	...	jaw	
✗	tongue tip	$\omega < \sigma$	...	tongue tip	
	lips	...	...	lips	
Duration of the release gesture					
	jaw	...	${}^l\sigma > \sigma$	jaw	✓
	tongue tip	...	${}^l\sigma > \sigma$	tongue tip	✓
	lips	...	${}^l\sigma > \sigma$	lips	✓
Overlap between the articulatory movements of the constriction					
✓	jaw-tongue tip	$\omega < \sigma$	...	jaw-tongue tip	
✓	jaw-lips	$\omega < \sigma$	...	jaw-lips	
Acoustic duration of the consonantal segment					
	orals	...	${}^l\sigma > \sigma$	orals	✓
	nasals	...	${}^l\sigma > \sigma$	nasals	✓
VOT					
✓		$\omega = \sigma$	${}^l\sigma = \sigma$		✓
Jaw height					
✓		$\omega > \sigma$	...		
Tongue-tip height					
✓		$\omega > \sigma$	...		
Lip aperture					
✗		$\omega > \sigma$	...		
Nasal flow and pressure					
✓		$\omega < \sigma$	...		
(Proportional) time of achievement of the magnitude peak, or the U_n / P_n peak					
✗	jaw	$\omega < \sigma$	${}^l\sigma < \sigma$	jaw	✗
	tongue	...	...	tongue tip	
	lips	...	...	lips	
	U_n / P_n	...	...	U_n / P_n	
Relative intensity					
✓	fricatives	$\omega > \sigma$	${}^l\sigma < \sigma$	fricatives	✗
	spirant approximants	...	${}^l\sigma < \sigma$	spirant approximants	✓
	nasals	...	${}^l\sigma < \sigma$	nasals	✗

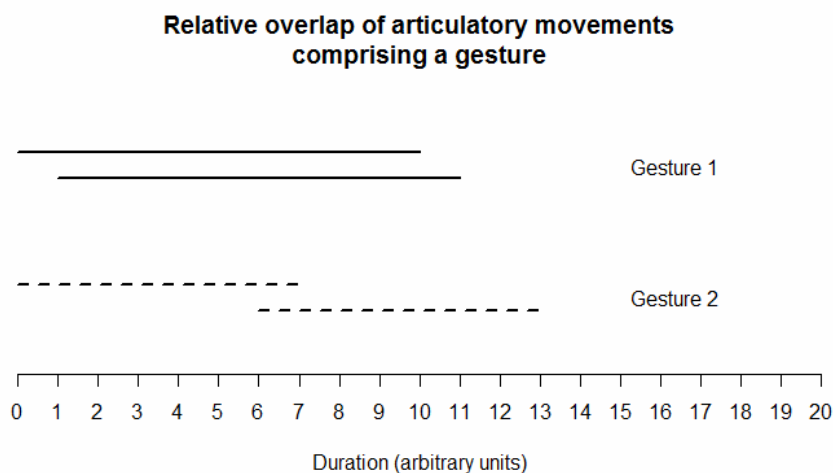


Figure 8.1. Relative overlap of articulatory movements comprising a gesture.

both could increase, as can be seen in Figure 8.1. This figure shows, in a simplified fashion, the duration of two gestures, each of which consists of the coordinated movement of two articulatory structures or organs. While the movements that make up the first gesture (solid line) are each longer than the movements in the second gesture (dashed line), the shorter overlap in the second gesture makes it last longer than the first (taken as a whole). Although a more detailed investigation into the subject is required, it may be that an approach using these two results (i.e. the decrease in movement duration and the decrease in overlap duration) will allow us to reach an explanation that is compatible with the hypothesis that the length of the gestures should increase at the word-initial position (compared to the syllable-initial position) ¹⁷⁴.

In the same way that the overlap between the jaw and tongue tip decreases, a reduction in the time of jaw and lip coactivation has also been observed. In § 3.1 it was stated that one of the effects of positional strengthening is that, through a prosodic boundary, the level of coarticulation between contiguous segments or gestures is decreased. Here it has also been proven that there is a lower level of coarticulation between movements that belong to the same gesture. All of this means that this prosodic position is not highly favorable to assimilation processes. This does not mean that contextual assimilations cannot occur on word boundaries but, if they do, they will most likely also act upon the medial part of the word if the rest of conditions remain constant.

Furthermore, this modification in intra-gestural coordination may also explain why, contrary to expectations, the magnitude peak of jaw movement is reached earlier in the word-initial position than the syllable-initial position. In fact, this is the only articulatory movement that has shown a significant change for said parameter in this position. Jaw movements present unique characteristics compared to other articulatory organs, as the jaw is a structure whose movement accompanies that of other organs (jaw

¹⁷⁴ Additionally, future work along these lines should be enriched by comparing different speech rates, provided that this factor can modify not only the duration of the different constituents, but also the temporal coordination between successive gestures or segments, as suggested by Celata & Calamai (2011).

movement is the common element of all oral gestures). Moreover, in § 3.5.2 we saw how jaw movements present a notable degree of independence with regard to other articulatory movements, and may even present movements opposed to other articulatory movements on certain occasions. For these reasons, we must study jaw behavior in relation to the organs with which it coordinates more in depth, as this type of study could shed some light on results like the ones obtained here.

Another hypothesis predicted a modification in the acoustic duration of consonants in the word-initial position, with a different effect for oral segments and nasal segments. However, this is one of the cases in which the prosodic position was found to have no significant effect.

There are also no significant differences in the VOT, which matches the hypothesis. Keeping in mind that Spanish is a language that bases its laryngeal oppositions for obstruents on the presence or absence of voicing (and not on aspiration), it can be assumed that voiceless stops are the phonologically unmarked element of the opposition and that, therefore, they lack specification in terms of the glottal gesture (always from the phonological point of view). For this reason, positional strengthening cannot target the glottal gesture, as said gesture is absent from the corresponding gestural score. This is an example of how the effects of strengthening can differ for languages that have different phonological systems, as aspirated languages would have an increase in VOT (see § 3.3.1).

As regards measurements of gestural magnitude (corresponding to the height of the jaw, the height of the tongue tip, lip aperture and nasal air flow and pressure), all are significantly affected by the word-initial position. As was explained in § 3.1, the unique dynamics of gestures imply that, with a time increase, they may reach a greater magnitude when they are followed by another gesture that corresponds to the same tract variable, but has a different value. On the contrary, the magnitude decreases when the following gesture controls a different tract variable. Indeed, we have observed an increase in the magnitude of jaw movement, which reaches a higher position for consonants in the word-initial position. Similarly, the tongue tip also experiences this type of spatial strengthening. Furthermore, as was expected, the decrease in U_n and P_n is interpreted as a decrease in the lowering of the soft palate during the velum lowering gesture for nasals in the word-initial position. That is, the magnitude of the gesture is more reduced than usual, which is explained by the monovalent nature of the velar gesture, a gesture which, therefore, is never followed by another gesture that modifies the position of the velum in a different direction.

The results for labial constriction are contrary to the predictions of hypothesis H_{10A} . In the word-initial position, there is greater labial opening, that is, a lower degree of lip displacement in their closing movement. Given this result and, retrospectively, I believe that there is a plausible explanation for this behavior, similar to that for the velar gesture. We must take into account that the labial consonants in the experiment (in fact, all of the consonants) are followed by two possible vowels, [a] and [i], neither of which specifies a labial gesture. Therefore, the consonant's gesture of labial constriction is not followed by another labial gesture; these are the circumstances under which a decrease

in magnitude is predicted in the word-initial position, as occurs with the velar gesture. This explains the pattern observed of an increase in the distance between the lips for said prosodic position. Future studies would have to confirm if said behavior disappears when the consonant is followed by an [o] or a [u], vowels that do include a labial gesture (lip rounding) in its phonological specification.

Lastly, we have found an increase in the relative intensity of fricatives (with respect to the preceding vowel) in the word-initial position, which is surely due to greater narrowing of the air output canal which, in turn, causes greater friction noise. In the case of [s] it is clear that an increase in tongue tip movement magnitude is responsible for said narrowing. With regard to [f], this increase in friction noise may be surprising given the data we have pointed out regarding greater lip aperture. However, it should be kept in mind that the friction corresponding to [f] is not created between the two lips, but rather between the lower lip and the upper incisors. For this reason, greater interlabial separation does not necessarily imply a lesser degree of narrowing in the actual articulation point (labiodental), and therefore both results are perfectly compatible.

8.2. Effects of prosodic prominence

In terms of strengthening due to prominence, H_{1B} did not predict any changes in constriction speed; however, the same effect was found as for word-initial position, that is, said speed decreases for jaw and lip movement, although there is no significant effect on the duration (or magnitude) of said constriction gesture.

This dynamic contrasts with that of the consonant release gesture (and vowel formation), which is where the most important differences between stressed and unstressed syllable arise. Release velocity is greater in prominent syllables: the Spanish results confirm this in the cases of jaw and lip movement.

Furthermore, said release gestures last longer (both for the jaw and lips as well as for the tongue tip). This is interpreted as a delay in the closing gesture following the nuclear vowel of the stressed syllable. It makes sense that if prominence affects the entire syllabic unit and the syllable is the unit of gestural coordination, then the presence of stress expresses itself through a reorganization of the temporal relationships between successive gestures.

The temporal extension of the release gesture partially corresponds to an acoustic portion of the end of the prenuclear consonant and an acoustic portion of the onset of the vowel itself. Given the increase of the articulatory duration of said gesture, it is to be expected that the corresponding acoustic durations of both segments would increase, as has indeed been observed. The results of M_{13} (see Table 6.27 on p. 220) confirm a significant lengthening of the vowel in conditions of prominence. In effect, longer vowel duration is one of the keys of stressed syllables, as was seen in § 2.2.1. Additionally, in accordance with the predictions of H_{6B} , the acoustic duration of the prenuclear consonant also increases. The net lengthening of the stressed syllable is seen, then, in two segments: the vowel that acts as the nucleus and the immediately preceding consonant.

Both the increase of opening speed as well as the longer duration should allow the degree of vowel openness to reach a greater magnitude which, together with the temporal increase, means an increase in the acoustic energy emitted, which can be deduced from the results regarding relative intensity. H_{13B} predicted an increase in the relative intensity of fricatives and nasals with regard to the vowel. This increase would be due, in the first case, to greater friction noise caused by increased narrowing of the air output canal, caused by a hypothetical increase in the magnitude of the oral constriction gesture. However, this spatial strengthening is not supported by the data collected from the corpus analyzed here and, on the contrary, a decrease in relative intensity was observed (instead of an increase). In the case of nasals, it was also expected that an increase in the magnitude of the velar gesture (detected through an increase in the values of U_n and P_n) would lead to greater relative intensity of said nasal segments. For this hypothesis, we have not found any significant modifications in the aerodynamic parameters, which is interpreted as the absence of a change in the magnitude of the velum gesture. As with the fricatives, a lower relative intensity was also observed for nasals in the stressed syllable, contrary to the hypothesis.

From these data, we can deduce that if there were a small change in the magnitude of the respective articulatory gestures, it was not sufficient to be detected as significant. Furthermore, what did increase was the intensity of the vowel, as a result of the modifications in the speed and duration of the release gesture. Thus, as the intensity of the consonant was unaltered and the vowel's intensity increased, it can be understood that the relative intensity of the consonant segment decreased. Put another way, the changes in the vowel surpass in magnitude any possible smaller change in the consonant. It can be assumed that these vowel changes, aimed at an increase in duration and intensity, are the target of the modifications that syllabic prominence introduces in the gestural dynamics. Given that the vowel is the nucleus of the syllable, it is understandable that greater vowel duration and intensity serve to highlight the entire prominent unit.

In terms of the other aspects analyzed, the stressed syllables show no significant modifications of the VOT, for the same reasons as those presented in the prior section.

Furthermore, no significant difference was found in the degree of intra-gestural overlap between stressed and unstressed syllables. In effect, there was no hypothesis H_{5B} in that regard. Additionally, this result is consistent with some ideas regarding the classification of languages in terms of rhythm. For example, Bertinetto (1989) reviews some of the problems that researchers have encountered when trying to detect an empirical basis for the traditional distinction between syllable-timed languages and stress-timed languages. In this work, Bertinetto postulates that both rhythmic types are not opposing—as was believed until now—in that one bases its timing on one unit (the syllable) and the other bases its timing on another (the foot). He argues that it is difficult to believe that one language can maintain a certain degree of flexibility regarding the duration of one of these units and, at the same time, be more rigid in terms of the

duration of the other.¹⁷⁵ It is more likely that languages present a greater or lesser degree of general flexibility in terms of segment duration, both from an intra-syllabic and inter-syllabic (i.e. within a foot) perspective. Languages traditionally classified as stress-timed tend to compensate the duration of the segments (mainly, the vowels): while a stressed syllable lasts more than an unstressed syllable, this difference tends to be lessened when a foot must fit more than one unstressed syllable. Moreover, in these languages the unstressed syllables also experience greater temporal reduction. On the contrary, languages traditionally described as syllable-timed tend to control the duration of segments and to keep said duration more or less constant (see also Bertinetto & Bertini, 2008, for example).

In compensating languages, like English, the larger reduction in the unstressed syllable expresses itself, among other ways, as greater gesture overlap. This makes the boundaries between syllables more diffuse: indeed, these languages generally present cases of ambisyllabicity (remember the explanation in § 2.1.1.4). Furthermore, Bertinetto argues in his 1989 paper, the metric systems of these languages tend to count stresses. This is so, on the one hand, because the difference between stressed and unstressed syllables is felt as more prominent and, on the other, precisely due to this blurring of syllabic boundaries, which makes a system based on syllable counting more difficult. On the other hand, duration-controlling languages, like Spanish, are characterized by a lesser difference in the degree of overlap between stressed and unstressed syllables (which is precisely what the results obtained here show, with a non-significant difference from the statistical point of view). This lower degree of overlapping means that boundaries between syllables are more defined and allows for the metric systems for poetry in these languages to be based more heavily on syllable counting (as the contrast between stressed and unstressed syllables is not as pronounced)¹⁷⁶.

8.3. Overall evaluation and applications of this study

The results of the analysis of the corpus illustrate the variations in phonetic details experienced by a series of consonants in Spanish as a result of their phonetic specification in terms of articulatory gestures, and of their inclusion in specific positions of the prosodic hierarchy.

The effects of strengthening observed here are numerous, despite the fact that the levels studied—the syllable and the phonological word—are located in the lower half of the hierarchy. Given the accumulative nature of strengthening, it is expected that in higher levels the effects of these variations would increase, and that other effects would

¹⁷⁵ Saltzman *et al.* (2008) propose a system based on task dynamics that models the way in which balance can be achieved given these opposing tensions.

¹⁷⁶ However, Piera (2001), among other authors, argues the following: “In mainstream Spanish verse both the distribution of stresses and the number of syllables are relevant. [...] Spanish metrics is foot-based” (p. 206). In any case, it is true that stress patterns are rather unconstrained with respect to those present in the poetry of “stress-timed” languages; instead, in Spanish, as argued in that paper, there are intonational factors that condition the metrical patterns accepted in verse.

appear that here did not reach the level of significance required. This could be the goal of a promising future research project.

Here we have focused on the lexical levels for several reasons. On the one hand, as was discussed and shown in chapter 4, subsegmental variations detected using instruments show trends that, over time, may be consolidated as segmental changes that become natural in the language's phonological system, either in the form of synchronic allophonic alternations, or as phonemic changes from a historical perspective. Furthermore, in this way, strengthening phenomena condition what types of segments usually appear in one position or another in words. Specifically, we have seen that some unique aspects of the phonotactic patterns exclusive to the borders of words may be explained by the fortitive effects of the prosodic boundary.

Moreover, as was seen in § 2.3.3.1, in Spanish the syllable plays an important role in the psycholinguistic processes of segmenting continuous sound, which supports lexical access. Identifying units at the prosodic word level is also vital for lexical access, for obvious reasons. Overall, the phonetic cues that allow us to differentiate these two prosodic levels are a very valuable clue regarding which keys listeners use to segment utterances that they are trying to decode into words. Naturally, confirming the perceptive use of these keys could be the object of future studies. These studies, in a controlled manner, should test the capacity of the subjects to make decisions regarding the prosodic structure of an utterance when the relevant phonetic cues are varied. Some of the results from this dissertation directly indicate acoustic cues, such as segment duration or the relative intensity of consonants. However, other parameters of an articulatory nature can also be translated into acoustic modifications whose potential as perceptive cues should be researched. For example, changes in articulatory speed may be reflected in the speed of spectral transitions, and the increase in the magnitude of the gestures may also involve a displacement in formant values or in the center of gravity.

The phonetic traits inherent to strengthening may also be used in the field of speech technology, in order for automatic speech recognition systems to detect the prosodic structure of utterances (we must not forget that there are cases in which said structure may condition the semantic interpretation of the phrase). Furthermore, if speech synthesis mechanisms are refined to include the appropriate variations in phonetic detail, this would help achieve more natural and informative results for listeners.

Another line of research that could be very interesting has to do with analyzing the effects of strengthening due to prosodic structure in sign languages. Task dynamics can be applied not only to speech organ movements, but also to movements in sign languages (see for example, the study by Tyrone *et al.*, 2010, on American Sign Language). There is still much research to be done in this field and, as far as I know, there is no study of this type on Spanish Sign Language. It is to be expected that the motor effects caused by the presence of a prosodic boundary would be the same as in this case, that is, gestural slowing and lengthening, with a possible modification in magnitude. Furthermore, depending on whether this spatial change consists of an

increase or a decrease, we could establish which signs are equivalent¹⁷⁷ to tract variables; remember that a substantial strengthening of the gesture is only expected if it is followed by another gesture that controls the same variable (otherwise, a decrease in magnitude is expected).

In effect, throughout this dissertation emphasis has been made on the importance of differentiating between phonetic variations caused by prosodic position (which express themselves as instances of strengthening or weakening) and those caused by the surrounding segmental context (which are usually reflected as coarticulations or assimilations, although they can also have a dissimilatory effect). However, we have also seen that both types of phenomena can interact and, in fact, they do so frequently. For example, weak positions are more prone to coarticulation, given that in positions of strengthening, changes in the gestural dynamics lead to a lower degree of overlap between successive gestures, and also between movements that make up the same gesture. Additionally, the type of modification on the magnitude of gestures adjacent to a prosodic boundary depends on whether the next gesture affects the same tract variable or not. That is, depending on the following context, the magnitude of the initial gesture of a component may increase or decrease. We have seen here that a decrease does indeed occur in the case of the velar gesture, which reaches a less pronounced lowering position, and it has been noted that this may also occur with a consonant's labial gestures when the following vowel does not present a gesture that also affects the lips.

Keeping in mind the unique behavior of the velar gesture in the initial position of a component, some predictions can be made regarding the aerodynamic characteristics of the consonants in that position. If nasals become less nasal and orals experience an increase in the magnitude of oral constriction, all consonants that experience positional strengthening should be characterized by an increase in oral pressure (P_o). Furthermore, the higher position of the jaw contributes to decreasing the volume of the oral cavity and, therefore, to increasing said pressure.

In fact, it is possible to experimentally establish the weight of each articulatory gesture's contribution to P_o , using a linear regression model based on simultaneous measurements of the different articulatory movements and of the pressure variations in the oral cavity. A larger glottal opening, smaller velo-pharyngeal opening and smaller oral opening lead to an increase in P_o , and vice-versa. However, it may be interesting to see how the different gestures affect pressure when these gestures are not all working in the same direction. For example, the glottal gesture is expected to have more weight than the other gestures, given that it is the only way for air to enter the oral cavity during the production of pulmonic consonants (another matter, also of interest, would be the aerodynamic behavior of clicks, implosives and ejectives in conditions of strengthening). McGowan & Saltzman (1995) propose the necessity of including aerodynamic parameters in the task dynamics model in order to achieve more complete predictions. Specifically, they add subglottal pressure (P_s) and transglottal pressure (P_t).

¹⁷⁷ A large part of the research that has been done on sign language phonology consists of establishing which elements correspond to distinctive features, segments, syllables, etc. as in oral languages (see, for example, Brentari, 1995).

However, if, as I propose here, the multiple correlations between gestures and P_o are studied, it would be possible to do away with the parameter of P_t (which is the difference between P_s and P_o), given that P_s is already included in the model and the P_o could be calculated using the data from the other gestures, which are also included in the model.

In conclusion, there are many promising areas of research still to be explored involving the relationships between phonetics and phonology in the field of the prosody of languages in general, and of Spanish in particular. I hope this dissertation provides a contribution to this field and encourages new studies along these lines.

Referencias bibliográficas

- Aitchison, J. (1987). *Words in the mind. An introduction to the mental lexicon*. Oxford: Blackwell.
- Alarcos Llorach, E. (1958). Fonología y Fonética (A propósito de las vocales andaluzas). *Archivum Oviedo*, 8, 193-205.
- Alarcos Llorach, E. (1965). *Fonología española* (4.^a ed.). Madrid: Gredos.
- Alcoba, S. (1999). La flexión verbal. En I. Bosque y V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (Vol. 3, pp. 4915-4992). Madrid: Espasa Calpe.
- Alonso, A. (1974). *Estudios lingüísticos: temas españoles*. Madrid: Gredos.
- Alvar, M. (Ed.). (1996). *Manual de dialectología hispánica. El español de España*. Barcelona: Ariel.
- Auer, E. T., Jr., Kim, S., Keating, P. A., Scarborough, R. A., Alwaan, A. y Bernstein, L. E. (2004). Optical phonetics and visual perception of lexical and phrasal boundaries in English. *Journal of the Acoustical Society of America*, 116(4), 2644.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baer, T. (1975). *Investigation of phonation using excised larynxes* (Tesis doctoral). MIT.
- Baken, R. J. (1996). *Clinical measurement of speech and voice*. San Diego, CA: Singular Publishing Group.
- Barbón Rodríguez, J. A. (1975). El rehilamiento. *Phonetica*, 31, 81-120.
- Barbón Rodríguez, J. A. (1978). El rehilamiento: descripción. *Phonetica*, 35, 185-215.
- Bates, D. (2005). Fitting linear mixed models in R. *R news*, 5(1), 27-30.
- Bates, D., Maechler, M. y Bolker, B. (2012). lme4: Linear mixed-effects models using «Eigen» and S4 (Versión 0.999999-0) [R package]. Recuperado a partir de <http://CRAN.R-project.org/package=lme4>
- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Beckman, M. E. y Edwards, J. (1994). Articulatory evidence for differentiating stress categories. En P. A. Keating (Ed.), *Phonological structure and phonetic form* (pp. 7-33). Cambridge: Cambridge University Press.
- Beckman, M. E., Edwards, J. y Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. En G. J. Docherty y D. R. Ladd (Eds.), *Gesture, segment, prosody* (pp. 68-86). Cambridge: Cambridge University Press.
- Beckman, M. E. y Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3(1), 255-309.
- Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental context. En M. K. Huffman y R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (pp. 63-85). New York: Academic Press.
- Bertinetto, P. M. (1989). Reflections on the dichotomy 'stress' vs. 'syllable-timing'. *Revue de Phonétique Appliquée*, 91(93), 99-130.

- Bertinetto, P. M., Agonigi, M., Cioni, L., García Lecumberri, M. L. y González Parra, E. (1998). *An experimental investigation on syllable geometry in Spanish (Expanded version)*. Recuperado a partir de <http://linguistica.sns.it/QLL/QLL98/spanish.pdf>
- Bertinetto, P. M. y Bertini, C. (2008). On modeling the rhythm of natural languages. En P. A. Barbosa, S. Madureira y C. Reiss (Eds.), *Proceedings of Speech Prosody 2008 (Campinas, Brazil)* (pp. 427-430).
- Blecua Falgueras, B. (2001). *Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos* (Tesis doctoral). Universitat Autònoma de Barcelona.
- Blevins, J. (1995). The syllable in phonological theory. En J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206-244). Oxford: Blackwell.
- Blevins, J. (2003). The independent nature of phonotactic constraints: An alternative to syllable-based approaches. En C. Féry y R. Van de Vijver (Eds.), *The syllable in optimality theory* (pp. 375-403). Cambridge: Cambridge University Press.
- Blevins, J. (2004). *Evolutionary phonology. The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bloch, B. (1950). Studies in Colloquial Japanese IV: Phonemics. *Language*, 26(1), 86-125.
- Boersma, P. (2006). Prototypicality judgments as inverted perception. En G. Fanselow, C. Féry, M. Schlesewsky y R. Vogel (Eds.), *Gradience in grammar: Generative perspectives* (pp. 167-184). Oxford: Oxford University Press. Recuperado a partir de <http://145.18.230.100/paul/papers/Prototypicality.pdf>
- Boersma, P. y Weenink, D. (2011). Praat: doing phonetics by computer (Versión 5.3.03).
- Bosch, A. R. K. (2011). Syllable-internal structure. En M. Van Oostendorp, C. J. Ewen, E. Hume y K. Rice (Eds.), *Companion to phonology* (Vol. 2, cap. 33). Oxford: Blackwell.
- Brandão de Carvalho, J., Scheer, T. y Ségéral, P. (2008). *Lenition and fortition*. Berlin: Mouton de Gruyter.
- Brentari, D. (1995). Sign language phonology: ASL. En J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 615-639). Oxford: Blackwell.
- Browman, C. P. y Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.
- Browman, C. P. y Goldstein, L. M. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45(2), 140-155.
- Browman, C. P. y Goldstein, L. M. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.
- Browman, C. P. y Goldstein, L. M. (1992). Articulatory phonology: an overview. *Phonetica*, 49(3-4), 155-180.
- Browman, C. P. y Goldstein, L. M. (1995). Gestural syllable position effects in American English. En R. Bell-Berti (Ed.), *Producing speech: A Festschrift for Katherine Safford Harris* (pp. 19-33). New York: American Institute of Physics.
- Buenafuentes de la Mata, C., Madrigal Montoya, N. y Garrido Almiñana, J. M. (2000). Análisis acústico de las variaciones micromelódicas en las curvas del f0 en español. *Español Actual*, 73, 65-77.

- Bustos Tovar, E. (1960). *Estudios sobre asimilación y disimilación en el ibero románico*. Madrid: CSIC.
- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- Byrd, D. (1995). C-Centers revisited. *Phonetica*, 52, 285-306.
- Byrd, D., Kaun, A., Narayanan, S. S. y Saltzman, E. L. (2000). Phrasal signatures in articulation. En M. B. Broe y J. B. Pierrehumbert (Eds.), *Acquisition and the lexicon* (pp. 70-87). Cambridge: Cambridge University Press.
- Byrd, D., Krivokapić, J. y Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America*, 120(3), 1589-1599.
- Byrd, D., Lee, S., Riggs, D. y Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America*, 118(6), 3860-3873.
- Byrd, D. y Saltzman, E. L. (2003). The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31(2), 149-180.
- Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Barcelona: Universitat de Barcelona.
- Carrasco, P., Hualde, J. I. y Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, 69(3), 149-179.
- Catford, J. C. (1977). *Fundamental problems in Phonetics*. Bloomington: Indiana University Press.
- Celata, C. y Calamai, S. (2011). Timing in Italian VNC sequences at different speech rates. En *Proceedings of Interspeech* (pp. 405-408).
- Chomsky, N. y Halle, M. (1976 [1968]). *Principios de fonología generativa*. Madrid: Fundamentos.
- Cho, T. (2003). Lexical stress, phrasal accent and prosodic boundary in the realization of domain-initial stops in Dutch. En M. J. Solé, D. Recasens y J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*.
- Cho, T. y Keating, P. A. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155-190.
- Cho, T. y Keating, P. A. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4), 466-485.
- Cho, T., McQueen, J. M. y Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, 35(2), 210-243.
- Cho, Y.-M. Y. y King, T. H. (2003). Semisyllables and universal syllabification. En C. Féry y R. Van de Vijver (Eds.), *The syllable in optimality theory* (pp. 183-212). Cambridge: Cambridge University Press.
- Christophe, A., Mehler, J. y Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, 2(3), 385-394.

- Christophe, A., Peperkamp, S., Pallier, C., Block, E. y Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language*, 51(4), 523-547.
- Clements, G. N. (1985). The geometry of phonological features. *Phonology*, 2(1), 225-252.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. En J. Kingston y M. E. Beckman (Eds.), *Between the grammar and physics of speech* (pp. 283-333). Cambridge: Cambridge University Press.
- Clements, G. N. y Keyser, S. J. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT.
- Colina, S. (2012). Syllable structure. En J. I. Hualde, A. Olarrea y E. O'Rourke (Eds.), *The Handbook of Hispanic Linguistics* (pp. 133-151). Oxford: Wiley-Blackwell.
- Condon, W. S. y Sander, L. W. (1974). Synchrony demonstrated between movement of the neonate and adult speech. *Child Development*, 45, 456-462.
- Cooper, A. (1991). *Glottal gestures and aspiration in English*. Yale University.
- Côté, M.-H. (2011). Final consonants. En M. Van Oostendorp, C. J. Ewen, E. Hume y K. Rice (Eds.), *Companion to phonology* (Vol. 2, cap. 36). Oxford: Blackwell.
- Cutler, A. y Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31(2), 218-236.
- Cutler, A. y Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2(3), 133-142.
- Cutler, A., Hawkins, J. A. y Gilligan, G. (1985). The suffixing preference: a processing explanation. *Linguistics*, 23(5), 723-758.
- Cutler, A. y Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21(1/2), 103-108.
- Cutler, A., Mehler, J., Norris, D. y Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
- Cutler, A., Mehler, J., Norris, D. y Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24(3), 381-410.
- Cutler, A. y Norris, D. (1988). The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121.
- Cutler, A. y Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824-844.
- De Brosches, C. (1765). *Traité de la formation mécanique des langues, et de principes physiques de l'étymologie*. Paris: Chez Saillant, Vincent, Desaint.
- Degenshein, R. (2004). Dholuo interdentals: Fricatives or affricates? Evidence from domain-initial strengthening. Presentado en From Sound to Sense, Boston, MA: MIT.
- Dejean de la Bâtie, B. (1993). *Word boundary ambiguity in spoken French* (Tesis doctoral). Monash University, Victoria, Australia.
- De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97(1), 491-504.

- Delais-Roussarie, E. y Fougeron, C. (2004). Les notions de domaine et de constituant en phonétique et en phonologie. En *Actes des 4èmes Journées d'Etudes Linguistiques de Nantes* (pp. 189-196).
- De Pijper, J. R. y Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America*, 96(4), 2037-2047.
- Di Cristo, A. (1982). *Prolégomènes à l'étude de l'intonation. Micromélorie*. Paris: CNRS.
- Dilley, L. C., Mattys, S. L. y Vinke, L. (2010). Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. *Journal of Memory and Language*, 63(3), 274 - 294.
- Dilley, L. C., Shattuck-Hufnagel, S. y Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, 24, 423-444.
- Dixit, R. P. (1991). Palatoglossus activity during VCV utterances containing oral and nasal consonants of Hindi. En *Proceedings of the 12th International Congress of Phonetic Sciences, Aix-en Provence* (Vol. 4, pp. 2-5).
- Dixit, R. P., Bell-Berti, F. y Harris, K. S. (1987). Palatoglossus activity during nasal / nonnasal vowels of Hindi. *Phonetica*, 44, 210-226.
- Donegan, P. J. (1978). *On the natural phonology of vowels* (Tesis doctoral). Ohio State University.
- Donegan, P. J. y Stampe, D. (1979). The study of natural phonology. En D. A. Dinnsen (Ed.), *Current approaches to phonological theory* (pp. 126-173). Bloomington: Indiana University Press.
- Dressler, W. U. (1984). Explaining Natural Phonology. *Phonology Yearbook*, 1, 29-51.
- Dziubalska-Kołodziejczyk, K. (1995). *Phonology without the syllable: A study in the natural framework*. Poznan: Motiwex.
- El Aissati, A., McQueen, J. M. y Cutler, A. (2012). Finding words in a language that allows words without vowels. *Cognition*, 124(1), 79-84.
- Féry, C. y Van de Vijver, R. (Eds.). (2003). *The syllable in Optimality Theory*. Cambridge: Cambridge University Press.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL quarterly*, 39(3), 399-423.
- Fletcher, J. (2010). The prosody of speech: Timing and rhythm. En W. J. Hardcastle, J. Laver y F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2.^a ed., pp. 523-602). Oxford: Wiley-Blackwell.
- Fougeron, C. (1999a). Articulation of prosodic phrasing in French. En J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville y A. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco* (pp. 675-678). Berkeley, CA: UC Berkeley.
- Fougeron, C. (1999b). Prosodically conditioned articulatory variations: A review. *UCLA Working Papers in Phonetics*, 97, 1-74.
- Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109-135.

- Fougeron, C. y Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101(3), 3728-3740.
- Fougeron, C. y Ridouane, R. (2008). On the phonetic implementation of syllabic consonants and vowel-less syllables in Tashlhiyt. *Estudios de Fonética Experimental*, 17, 140-175.
- Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, 99(3), 1730-1741.
- Fowler, C. A. y Galantucci, B. (2005). The relation of speech perception and speech production. En D. B. Pisoni y R. E. Remez (Eds.), *The handbook of speech perception* (pp. 633-652). Oxford: Blackwell.
- Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press.
- García Arias, X. L. (2003). *Gramática histórica de la lengua asturiana* (2.^a ed.). Oviedo: Academia de la Llingua Asturiana.
- García Santos, J. F. (2002). *Cambio fonético y fonética acústica*. Salamanca: Ediciones Universidad de Salamanca.
- Gelman, A. y Hill, J. (2007). *Data analysis using regression and multilevel / hierarchical models*. Cambridge: Cambridge University Press.
- Gick, B. (2003). Articulatory correlates of ambisyllabicity in English glides and liquids. En J. Local, R. Ogden y R. Temple (Eds.), *Phonetic interpretation* (pp. 222-236). Cambridge: Cambridge University Press.
- Gil Fernández, J. (2000a). *Panorama de la fonología española actual*. Madrid: Arco/Libros.
- Gil Fernández, J. (2000b). Un cuarto de siglo de fonología española. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 19-100). Madrid: Arco/Libros.
- Gil Fernández, J. (2006). Implicaciones fonológicas de la base de articulación. En *Filología y lingüística: Estudios ofrecidos a Antonio Quilis* (pp. 219-252). Madrid: CSIC.
- Gil Fernández, J. (2007). *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco/Libros.
- Gil Fernández, J. y Llisterri, J. (Eds.) (en prensa). *Fonética y fonología descriptivas de la lengua española*. Madrid: CSIC.
- Goad, H. (2011). The representation of sC clusters. En M. Van Oostendorp, C. J. Ewen, E. Hume y K. Rice (Eds.), *Companion to phonology* (Vol. 2, cap. 38). Oxford: Blackwell.
- Goldman, J.-P. (2011). EasyAlign: An automatic phonetic alignment tool under Praat. En *Proceedings of Interspeech*.
- Goldman, J.-P. y Schwab, S. (2014). EasyAlign Spanish: An (semi-)automatic segmentation tool under Praat. En Y. Congosto Martín, M. L. Montero Curiel y A. Salvador Plans (Eds.), *Fonética experimental, educación superior e investigación* (pp. 629-640). Madrid: Arco/Libros.
- Goldsmith, J. A. (1976). *Autosegmental phonology* (Tesis doctoral). MIT.

- Goldsmith, J. A. (1979). The aims of autosegmental phonology. En D. A. Dinnsen (Ed.), *Current approaches to phonological theory* (pp. 202-222). Bloomington: Indiana University Press.
- Goldsmith, J. A. (2000 [1981]). Las unidades subsegmentales en la fonología española: enfoque autosegmental. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 337-356). Madrid: Arco/Libros.
- Goldstein, L. M., Byrd, D. y Saltzman, E. L. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. En M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215-249). Cambridge: Cambridge University Press.
- Goman, R. D. (1979). *Consonants in natural phonology*. Ohio State University.
- Gómez Asencio, J. J. (2000 [1994]). Los fonemas consonánticos no líquidos orales del español. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 159-183). Madrid: Arco/Libros.
- Gordon, M. K. (1996). The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasals in Estonian. *UCLA Working Papers in Phonetics*, 92, 151-159.
- Gordon, M. K. (2004). Syllable weight. En B. Hayes, R. M. Kirchner y D. Steriade (Eds.), *Phonetically based phonology* (pp. 277-312). Cambridge: Cambridge University Press.
- Gósy, M. y Ringen, C. (2009). Everything you always wanted to know about VOT in Hungarian. Presentado en 9th International Conference on the Structure of Hungarian, Debrecen, Hungary. Recuperado a partir de http://www.icsh9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf
- Gout, A., Christophe, A. y Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51(4), 548-567.
- Gussenhoven, C. y Jacobs, H. (2011). *Understanding phonology* (3.^a ed.). London: Hodder Education.
- Halle, M. y Stevens, K. N. (1971). A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report*, 101, 198-213.
- Hardcastle, W. J. y Hewlett, N. (Eds.). (1999). *Coarticulation: Theory, data and techniques*. Cambridge: Cambridge University Press.
- Harrington, J. (2010). *Phonetic analysis of speech corpora*. Oxford: Wiley-Blackwell.
- Harrington, J., Fletcher, J. y Beckman, M. E. (2000). Manner and place conflicts in the articulation of Australian English consonants. En M. B. Broe y J. B. Pierrehumbert (Eds.), *Acquisition and the lexicon* (pp. 40-51). Cambridge: Cambridge University Press.
- Harrington, J., Fletcher, J. y Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics*, 23(3), 305-322.
- Harris, John (2006). The phonology of being understood: Further arguments against sonority. *Lingua*, 116, 1483-1494.

- Harris, John (2009). Why final obstruent devoicing is weakening. En K. Nasukawa y P. Backley (Eds.), *Strength relations in phonology* (pp. 9-45). Berlin: Walter de Gruyter.
- Harris, James W. (1969). *Spanish phonology*. Cambridge, MA: MIT.
- Harris, James W. (1991 [1983]). *La estructura silábica y el acento en español: Análisis no lineal*. Madrid: Visor.
- Harris, James W. (2000 [1989]). Lo que sabemos en la actualidad sobre la estructura silábica del español. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 485-510). Madrid: Arco/Libros.
- Harris, K. S., Gay, T., Sholes, G. N. y Liberman, P. (1968). Some stress effects on electromyographic measures of consonant articulations. *Haskins Laboratory Status Report*, 13/14, 137-151.
- Haugen, E. (1956). The syllable in linguistic description. En M. Halle (Ed.), *For Roman Jakobson* (pp. 213-221). The Hague: Mouton.
- Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic inquiry*, 20(2), 253-306.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: University of Chicago Press.
- Hayes, B. (2009). *Introductory phonology*. Oxford: Wiley-Blackwell.
- Hayes, J. R. y Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. En J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 221-234). New York: Wiley-Blackwell.
- Hertegard, S., Gauffin, J. y Lindestad, P.-A. (1995). A comparison of subglottal and intraoral pressure measurements during phonation. *Journal of Voice*, 9(2), 149-155.
- Hinton, V. A. (1996). Interlabial pressure during production of bilabial phones. *Journal of Phonetics*, 24, 337-349.
- Hirst, D. y Di Cristo, A. (1998). A survey of intonation systems. En D. Hirst y A. Di Cristo (Eds.), *Intonation systems. A survey of twenty languages* (pp. 1-44). Cambridge: Cambridge University Press.
- Hockett, C. F. (1955). *A manual of phonology*. Baltimore, MD: Waverly Press.
- Honeybone, P. (2008). Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. En J. Brandão de Carvalho, T. Scheer y P. Ségéral (Eds.), *Lenition and fortition* (pp. 9-92). Berlin: Mouton de Gruyter.
- Hooper, J. B. (1974). *Aspects of natural generative phonology*. Bloomington: Indiana University Linguistics Club.
- Hooper, J. B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.
- Horiguchi, S. y Bell-Berti, F. (1987). The Velotrace: A device for monitoring velar position. *Cleft Palate Journal*, 24, 104-111.
- Hyman, L. M. (1975). *Phonology: Theory and analysis*. New York: Holt, Rinehart and Winston.
- Hyman, L. M. (1985). *A theory of phonological weight*. Dordrecht: Foris.

- Hyman, L. M. (2011). In Defense of Prosodic Typology: A Response to Beckman y Venditti. *UC Berkeley Phonology Lab Annual Report*, 200-235.
- Inkelas, S. y Zec, D. (Eds.). (1990). *The phonology-syntax connection*. Chicago, IL: University of Chicago Press.
- Iskarous, K. (2010). Vowel constrictions are recoverable from formants. *Journal of Phonetics*, 38(3), 375-387.
- Iverson, G. K. y Salmons, J. C. (2007). Domains and directionality in the evolution of German final fortition. *Phonology*, 24, 121-145.
- James, A. L. (1940). *Speech signals in telephony*. London: Sir I. Pitman and Sons, Ltd.
- Jespersen, O. (1904). *Phonetische Grundfragen*. Leipzig: Teubner.
- Jessen, M. y Ringen, C. (2002). Laryngeal features in German. *Phonology*, 19(2), 189-218.
- Joglar, J. A., Nguyen, C., Garst, D. M. y Katz, W. F. (2009). Safety of Electromagnetic Articulography in Patients With Pacemakers and Implantable Cardioverter-Defibrillators. *Journal of Speech, Language and Hearing Research*, 52(4), 1082-1087.
- Johnson, K., Flemming, E. S. y Wright, R. A. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language*, 69, 505-528.
- Jun, S.-A. (2005). Prosodic typology. En S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430-458). Oxford: Oxford University Press.
- Jun, S.-A., Beckman, M. E. y Lee, H.-J. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics*, 96, 43-68.
- Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology*. Orlando, FL: Academic Press.
- Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M. y Goldstein, L. (2014). The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics*, 44, 62-82.
- Kavet, R. y Tell, R. A. (1991). VDTs: field levels, epidemiology, and laboratory studies. *Health Physics*, 61(1), 47-57.
- Kaye, J. D., Lowenstamm, J. y Vergnaud, J. R. (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook*, 2, 305-328.
- Keating, P. A. (1984). Aerodynamic modeling at UCLA. *UCLA Working Papers in Phonetics*, 59, 18-28.
- Keating, P. A. (1988). Palatals as complex segments: X-ray evidence. *UCLA Working Papers in Phonetics*, 69, 77-91.
- Keating, P. A. (2006). Phonetic encoding of prosodic structure. En J. Harrington y M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 167-186). New York: Psychology Press.

- Keating, P. A., Cho, T., Fougeron, C. y Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. En J. Local, R. Ogden y R. Temple (Eds.), *Phonetic interpretation* (pp. 145-163). Cambridge: Cambridge University Press.
- Keating, P. A., Wright, R. A. y Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, 97, 157-173.
- Kingston, J. (1985). *The phonetics and phonology of the timing of oral and glottal events* (Tesis doctoral). UC Berkeley.
- Kiparsky, P. (1973). «Elsewhere» in phonology. En S. R. Anderson y P. Kiparsky (Eds.), *Festschrift for Morris Halle* (pp. 93-106). New York: Holt, Rinehart and Winston.
- Kirchner, R. M. (2001). Phonological contrast and articulatory effort. En L. Lombardi (Ed.), *Segmental phonology in Optimality Theory: constraints and representations* (pp. 79-117). Cambridge: Cambridge University Press.
- Kodzasov, S. V. (1977). Fonetika Archinskogo Jazyka. En A. E. Kibrik, S. V. Kodzasov, I. P. Olyannikova y D. S. Samedov (Eds.), *Opyt Structurnogo Opisanija Archinskogo Jazyka (Sketch of a Structural Description of the Archi language)* (pp. 184-355). Moscow: Moscow University.
- Kodzasov, S. V. (1990). Fonetika. En A. E. Kibrik y S. V. Kodzasov (Eds.), *Sopostavitelnoe Izuchenie Dagestanskix Jazykov (Comparative Study of Dagestani Languages)* (pp. 338-341). Moscow: Moscow University.
- Kodzasov, S. V. y Muravjeva, I. A. (1982). Fonetika Tabasaranskogo jazyka. En *Tabasaranskije Etjudy* (pp. 6-16). Moscow: Moscow University.
- Koike, Y. y Hirano, M. (1973). Glottal-area time function and subglottal-pressure variation. *Journal of the Acoustical Society of America*, 54(6), 1618-1627.
- Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. En M. K. Huffman y R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (pp. 87-116). New York: Academic Press.
- Kuzla, C., Cho, T. y Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, 35(3), 301-320.
- Kuznetsova, A., Brockhoff, P. B. y Christensen, R. H. B. (2012). lmerTest: Tests in linear mixed effects models (Versión 1.2-0) [R package]. Recuperado a partir de <http://CRAN.R-project.org/package=lmerTest>
- Labov, W. (1996 [1994]). *Principios del cambio lingüístico. Vol. 1: Factores internos*. Madrid: Gredos.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. R. (2008). *Intonational phonology* (2.^a ed.). Cambridge: Cambridge University Press.
- Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago, IL: University of Chicago Press.
- Ladefoged, P. y Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Lahoz-Bengoechea, J. M. (en preparación). *La prosodia de los pronombres de objeto en español*.

- Lahoz-Bengoechea, J. M. (2015). ¿Qué aporta la fonética contrastiva a la didáctica de ELE? En Y. Morimoto, M. V. Pavón y R. Santamaría (Eds.), *La enseñanza de ELE centrada en el alumno* (pp. 49-61). Málaga: ASELE.
- Laks, B., Calderone, B. y Celata, C. (2014). French liaison and the lexical repository. En C. Celata y S. Calamai (Eds.), *Advances in Sociophonetics* (pp. 29-54). Amsterdam: John Benjamins.
- Lapesa, R. (1981). *Historia de la lengua española* (9.^a ed.). Madrid: Gredos.
- Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lavoie, L. M. (2001). *Consonant strength: phonological patterns and phonetic manifestations*. New York: Garland.
- Lee, S., Parrell, B. y Byrd, D. (2009). Computational modeling of juncture strength using articulatory synthesis of prosodic gestures. *Journal of the Acoustical Society of America*, 125(4), 2572.
- Lee, Y. (2011). Prosodic strengthening of /sCV/ sequences in English. En *Proceedings of 17th International Congress of Phonetic Sciences, Hong Kong* (pp. 1190-1193).
- Lieberman, A. M. y Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.
- Lieberman, M. (1975). *The intonational system of English* (Tesis doctoral). MIT.
- Lieberman, M. y Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, 8(2), 249-336.
- Lindblom, B. E. (1990). Explaining phonetic variation: a sketch of the H y H theory. En W. J. Hardcastle y A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403-439). Dordrecht: Kluwer.
- Lindblom, B. E., Agwuele, A., Sussman, H. M. y Cortes, E. E. (2007). The effect of emphatic stress on consonant vowel coarticulation. *The Journal of the Acoustical Society of America*, 121(6), 3802-3813.
- Lisker, L. y Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Llisterri, J., Machuca, M. J., De la Mota, C., Riera, M. y Ríos, A. (2003). The perception of lexical stress in Spanish. En M. J. Solé, D. Recasens y J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona* (pp. 2023-2026).
- Llisterri, J., Machuca, M. J., Ríos, A., De la Mota, C. y Riera, M. (2003). Algunas cuestiones en torno al desplazamiento acentual en español. En Z. Herrera y P. Martín Butragueño (Eds.), *La tonía: dimensiones fonéticas y fonológicas* (pp. 163-185). México, D. F.: Colegio de México.
- Lloyd, P. M. (1993). *Del latín al español. Vol. 1, Fonología y morfología históricas de la lengua española*. Madrid: Gredos.
- Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.
- Löfqvist, A. y Gracco, V. L. (1997). Lip and jaw kinematics in bilabial stop consonant production. *Journal of Speech, Language and Hearing Research*, 40(4), 877-893.
- López Morales, H. (1984). Desdoblamiento fonológico de las vocales en el andaluz oriental: Reexamen de la cuestión. *Revista Española de Lingüística*, 14(1), 85-97.

- Lozano, M. d C. (2000 [1978]). Un argumento a favor de las especificaciones parciales: pruebas del español. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 323-335). Madrid: Arco/Libros.
- Macchi, M. J. (1988). Labial articulation patterns associated with segmental features and syllable structure in English. *Phonetica*, 45, 109-121.
- Macken, M. A. y Barton, D. (1980). The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. *Journal of Child Language*, 7, 433-458.
- Maddieson, I. (2001). Typological patterns-geographical distribution and phonetic explanation. Presentado en Conference on the Phonetics-Phonology Interface, Berlin.
- Maddieson, I. (2006). In search of universals. En R. Mairal y J. Gil Fernández (Eds.), *Linguistic universals* (pp. 80-100). Cambridge: Cambridge University Press.
- Malécot, A. (1970). The lenis-fortis opposition: its physiological parameters. *Journal of the Acoustical Society of America*, 47(2), 1588-1592.
- Marin, S. y Goldstein, L. M. (2012). A gestural model of the temporal organization of vowel clusters in Romanian. En P. Hoole, M. Pouplier, L. Bombien, C. Mooshammer y B. Kühnert (Eds.), *Consonant clusters and structural complexity*. Berlin: Mouton de Gruyter.
- Marrero, V. (2001). *Fonética perceptiva. Addenda*. Madrid: UNED.
- Martinet, A. (1974 [1955]). *Economía de los cambios fonéticos*. Madrid: Gredos.
- Martínez Celdrán, E. (1985). ¿Hasta qué punto es importante la sonoridad en la discriminación auditiva de las obstruyentes mates del castellano? *Estudios de Fonética Experimental*, 1, 241-291.
- Martínez Celdrán, E. (1991). Tensión frente a sonoridad en las consonantes mates del castellano. En *Fonética experimental: teoría y práctica* (pp. 131-141). Madrid: Síntesis.
- Martínez Celdrán, E. (1993). Un mismo parámetro fonético en el fondo de la lenición protorromance: la duración. En J. A. Bartol, J. F. García Santos y J. De Santiago Guervós (Eds.), *Estudios filológicos en homenaje a Eugenio de Bustos Tovar* (Vol. 2, pp. 621-640). Salamanca: Ediciones Universidad de Salamanca.
- Martínez Celdrán, E. y Fernández Planas, A. M. (2013). *Manual de fonética española* (2.^a ed.). Barcelona: Ariel.
- Martínez-Gil, F. y Colina, S. (Eds.). (2006). *Optimality-theoretic studies in Spanish phonology*. Amsterdam: John Benjamins.
- Martínez-Paricio, V. (2012). Superfeet as recursion. En N. Arnett y R. Bennett (Eds.), *Proceedings of the 30th West Coast Conference on Formal Linguistics (WCCFL 30)* (pp. 259-269). Somerville, MA: Cascadilla Proceedings Project. Recuperado a partir de <http://www.lingref.com/cpp/wccfl/30/paper2823.pdf>
- McGlone, R. E. y Profitt, W. (1967). Lingual pressure associated with speaker consistency and syllable variations. *Phonetica*, 17, 178-183.
- McGlone, R. E., Profitt, W. y Christiansen, R. (1967). Lingual pressures associated with alveolar consonants. *Journal of Speech and Hearing Research*, 10, 606-661.

- McGowan, R. S. y Saltzman, E. L. (1995). Incorporating aerodynamic and laryngeal components into task dynamics. *Journal of Phonetics*, 23(1-2), 255-269.
- McMahon, A. (1967). Phonemes and phonemic units of Cora (Mexico). *International Journal of American Linguistics*, 33, 128-134.
- McQueen, J. M., Norris, D. y Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 621-638.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U. y Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 298-305.
- Méndez Dosuna, J. (1987). La aspiración de s como proceso condicionado por el contacto de sílabas. *Revista Española de Lingüística*, 17(1), 15-35.
- Mielke, J. (2005). Ambivalence and ambiguity in laterals and nasals. *Phonology*, 22(2), 169-204.
- Mielke, J. (2009). Segment Inventories. *Language and Linguistics Compass*, 3(2), 700-718.
- Mondéjar, J. (1979). Diacronía y sincronía en las hablas andaluzas. *Lingüística Española Actual*, 1(2), 375-402.
- Montagu, J. (2004). Les sons sous-jacents aux voyelles nasales en français parisien: indices perceptifs des changements. En *Journées d'Étude sur la Parole* (pp. 385-388).
- Moreno Cabrera, J. C. (2003). *El universo de las lenguas. Clasificación, denominación, situación, tipología, historia y bibliografía de las lenguas*. Madrid: Castalia.
- Nam, H. (2007). Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. En J. Cole y J. I. Hualde (Eds.), *Laboratory Phonology* (Vol. 9, pp. 483-506). Berlin: Mouton de Gruyter.
- Navarro Tomás, T. (1996 [1918]). *Manual de pronunciación española* (26.^a ed.). Madrid: CSIC.
- Nespor, M. y Vogel, I. (1994 [1986]). *La prosodia*. Madrid: Visor.
- Netsell, R. (1969). Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. *Phonetica*, 20(2-4), 68-73.
- Nittrouer, S. (1991). Phase relations of jaw and tongue tip movements in the production of VCV utterances. *Journal of the Acoustical Society of America*, 90(4), 1806-1815.
- Nittrouer, S., Munhall, K. G., Kelso, J. A. S., Tuller, B. y Harris, K. S. (1988). Patterns of interarticulator phasing and their relation to linguistic structure. *Journal of the Acoustical Society of America*, 84(5), 1653-1661.
- Nogita, A., Yamane, N. y Bird, S. (2013). The Japanese unrounded back vowel /u/ is in fact rounded central/front [ɯ-ɤ]. Presentado en Ultrafest VI, Edinburgh. Recuperado a partir de http://www.qmu.ac.uk/casl/conf/ultrafest_2013/docs/A%20Nogita%20Ultrafest.pdf
- Nooteboom, S. G. (1981). Lexical retrieval from fragments of spoken words: Beginnings vs. endings. *Journal of Phonetics*, 9(4), 407-424.
- Nunes, J. (2004). *Linearization of chains and sideward movement*. Cambridge, MA: MIT.

- Núñez Cedeño, R. A. (2000 [1984]). Teoría de la organización silábica. En J. Gil Fernández (Ed.), *Panorama de la fonología española actual* (pp. 455-474). Madrid: Arco/Libros.
- Núñez Cedeño, R. A. y Morales-Front, A. (1999). *Fonología generativa contemporánea de la lengua española*. Washington, DC: Georgetown University Press.
- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. En P. F. MacNeilage (Ed.), *The production of speech*. New York: Springer Verlag.
- Ohala, J. J. (1990a). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. En M. Ziolkowski, M. Noske y K. Deaton (Eds.), *Parasession on the syllable in phonetics and phonology. Papers from the 26th Annual Regional Meeting of the Chicago Linguistic Society, part 2* (pp. 319-338). Chicago, IL: Chicago Linguistic Society.
- Ohala, J. J. (1990b). Respiratory activity in speech. En W. J. Hardcastle y A. Marchal (Eds.), *Speech production and speech modelling* (pp. 23-53). Dordrecht: Kluwer.
- Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. *Journal of the Acoustical Society of America*, 99(3), 1718-1725.
- Ohala, J. J. y Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. En S. Eliasson y E. H. Jahr (Eds.), *Language and its ecology: Essays in memory of Einar Haugen* (pp. 343-365). Berlin: Mouton de Gruyter.
- Ohala, J. J. y Ohala, M. (1993). The phonetics of nasal phonology: theorems and data. En M. K. Huffman y R. A. Krakow (Eds.), *Nasals, nasalization, and the velum* (pp. 225-249). New York: Academic Press.
- Otake, T., Hatano, G., Cutler, A. y Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258-278.
- Parker, S. G. (2002). *Quantifying the sonority hierarchy* (Tesis doctoral). University of Massachusetts, Amherst.
- Parrell, B. (2014). *Dynamics of consonant reduction* (Tesis doctoral). University of Southern California.
- Pensado Ruiz, C. (1984). *Cronología relativa del castellano*. Salamanca: Ediciones Universidad de Salamanca.
- Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I. y Jackson, M. T. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, 92, 3078-3096.
- Perlmutter, D. (1995). Phonological quantity and multiple association. En J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 307-317). Oxford: Blackwell.
- Piera, C. (2001). Intonational factors in metrics. *Belgian Journal of Linguistics*, 15(1), 205-228.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Tesis doctoral). MIT.

- Pierrehumbert, J. B. (1999). Prosody and intonation. En R. A. Wilson y F. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 479-482). Cambridge, MA: MIT.
- Pierrehumbert, J. B. y Talkin, D. (1992). Lenition of /h/ and glottal stop. En G. J. Docherty y D. R. Ladd (Eds.), *Gesture, segment, prosody* (pp. 90-117). Cambridge: Cambridge University Press.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Piñeros, C. E. (en prensa). Nasales: descripción fonológica. En J. Gil Fernández y J. Llisterri (Eds.), *Fonética y fonología descriptivas de la lengua española*. Madrid: CSIC.
- Pons, C. (2008). The sonority scale: Categorical or gradient? Presentado en CUNY Conference on the Syllable. Recuperado a partir de <http://www.cunyphonologyforum.net/SYLLABSTRACTS/PonsPoster.pdf>
- Poplack, S. (1980). Deletion and disambiguation in Puerto Rican Spanish. *Language*, 56(2), 371-385.
- Prieto, P. y Roseano, P. (Eds.). (2010). *Transcription of intonation of the Spanish language*. München: Lincom Europa. Recuperado a partir de http://prosodia.upf.edu/home/arxiu/publicacions/prieto/transcription_intonation_spanish.php
- Prince, A. (1983). Relating to the grid. *Linguistic Inquiry*, 14, 19-100.
- Puigvert Ocal, A. (2015). Sobre la acomodación de la base articulatoria de la lengua materna del alumno a la del español como lengua extranjera. En Y. Morimoto, M. V. Pavón y R. Santamaría (Eds.), *La enseñanza de ELE centrada en el alumno*. Málaga: ASELE.
- Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Real Academia Española. (2011). *Nueva gramática de la lengua española. Fonética y fonología*. Madrid: Espasa Libros.
- Recasens, D. (1999). Lingual coarticulation. En W. J. Hardcastle y N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 80-104). Cambridge: Cambridge University Press.
- Recasens, D. (2002). Weakening and strengthening in Romance revisited. *Rivista di Linguistica*, 14(2), 327-373.
- Recasens, D. y Romero, J. (1997). An EMMA study of segmental complexity in alveolopalatals and palatalized alveolars. *Phonetica*, 54(1), 43-58.
- Renard, R. (1971). *Introduction à la méthode verbo-tonale de correction phonétique*. Paris: Didier.
- Rost Bagudanch, A. (2011). *Variación en los procesos de palatalización de yod segunda (o cómo la sincronía permite la explicación de la diacronía)* (Tesis doctoral). Universitat de Girona.
- Rost Bagudanch, A. (2014). El papel de la percepción en el yeísmo: Estudio preliminar con jueces catalanohablantes. *Loquens. Spanish Journal of Speech Sciences*, 1(2), e010. <http://dx.doi.org/10.3989/loquens.2014.010>

- Rothenberg, M. (1977). Measurement of airflow in speech. *Journal of Speech and Hearing Research*, 20, 155-176.
- Rothenberg, M. (1982). Interpolating subglottal pressure from oral pressure. *Journal of Speech and Hearing Disorders*, 47(2), 219-220.
- Rousselot, P. J. (1901). *Principes de phonétique expérimentale*. Paris: Welter.
- Ruch, H. y Harrington, J. (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics*, 45, 12-25.
- Saffran, J. R., Newport, E. L. y Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606-621.
- Saltzman, E. L. y Kelso, J. A. S. (1987). Skilled actions: a task dynamic approach. *Psychological Review*, 94, 84-106.
- Saltzman, E. L. y Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1(4), 333-382.
- Saltzman, E. L., Nam, H., Krivokapić, J. y Goldstein, L. M. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. En P. A. Barbosa, S. Madureira y C. Reiss (Eds.), *Proceedings of Speech Prosody 2008 (Campinas, Brazil)* (Vol. 4, pp. 175-184).
- Sanderman, A. A. y Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech*, 40(4), 391-409.
- Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.
- Schiering, R., Hildebrandt, K. A. y Bickel, B. (2007). *Cross-linguistic challenges for the prosodic hierarchy: Evidence from word domains*. University of Leipzig. Recuperado a partir de http://www.researchgate.net/profile/Balthasar_Bickel/publication/237302034_Cross-linguistic_Challenges_for_the_Prosodic_Hierarchy_Evidence_from_Word_Domains/links/53f47e020cf2888a7491008e.pdf
- Ségéral, P. y Scheer, T. (2008). Positional factors in lenition and fortition. En J. Brandão de Carvalho, T. Scheer y P. Ségéral (Eds.), *Lenition and fortition* (pp. 131-172). Berlin: Walter de Gruyter.
- Segui, J., Frauenfelder, U. y Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, 72(4), 471-477.
- Selkirk, E. O. (1978). On prosodic structure and its relation to syntactic structure. En T. Fretheim (Ed.), *Nordic prosody II* (pp. 111-140). Tapir.
- Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry*, 11(3), 563-605.
- Selkirk, E. O. (1981). On the nature of phonological representation. En T. Myers, J. Laver y J. M. Anderson (Eds.), *The cognitive representation of speech* (pp. 379-388). Amsterdam: North-Holland.
- Selkirk, E. O. (1982). The syllable. En H. Van der Hulst y N. Smith (Eds.), *The structure of phonological representations. Part II* (pp. 337-383). Dordrecht: Foris.
- Selkirk, E. O. (1984a). On the major class features and syllable theory. En M. Aronoff y R. Oehrle (Eds.), *Language Sound Structure* (pp. 107-136). Cambridge, MA: MIT.

- Selkirk, E. O. (1984b). *Phonology and syntax. The relation between sound and structure*. Cambridge, MA: MIT.
- Selkirk, E. O. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, 3, 371-405.
- Selkirk, E. O. (1996). The prosodic structure of function words. En K. Demuth y J. Morgan (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 187-214). Hillsdale, NJ: Lawrence Erlbaum.
- Selkirk, E. O. (2000). The interaction of constraints on prosodic phrasing. En M. Horne (Ed.), *Prosody: Theory and experiment. Studies presented to Gösta Bruce* (pp. 231-261). Dordrecht: Kluwer.
- Shattuck-Hufnagel, S. y Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.
- Sievers, E. (1893). *Grundzüge der Phonetik*. Leipzig: Breitkopf and Hartel.
- Silbert, N. y De Jong, K. (2008). Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production. *The Journal of the Acoustical Society of America*, 123(5), 2769-2779.
- Slis, I. H. (1971). Articulatory effort and its durational and electromyographic correlates. *Phonetica*, 23, 171-188.
- Smith, J. L. (2005). *Phonological augmentation in prominent positions*. New York: Routledge.
- Solé, M. J. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics*, 30, 655-688.
- Solé, M. J. (2010). Effects of syllable position on sound change: An aerodynamic study of final fricative weakening. *Journal of Phonetics*, 38(2), 289-305.
- Solomon, N. P., Ramanathan, P. y Makashay, M. J. (2007). Phonation threshold pressure across the pitch range: preliminary test of a model. *Journal of Voice*, 21(5), 541-550.
- Sosa, J. M. (1999). *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.
- Spinelli, E., Cutler, A. y McQueen, J. M. (2002). Resolution of liaison for lexical access in French. *Revue Française de Linguistique Appliquée*, 7(1), 83-96.
- Spinelli, E., Grimault, N., Meunier, F. y Welby, P. (2010). An intonational cue to word segmentation in phonemically identical sequences. *Attention, Perception, and Psychophysics*, 72(3), 775-787.
- Spinelli, E., McQueen, J. M. y Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, 48(2), 233 - 254.
- Stampe, D. (1973). *A dissertation on natural phonology* (Tesis doctoral). University of Chicago.
- Steriade, D. (1993). Orality and markedness. En *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 334-347).
- Steriade, D. (1999). Alternatives to syllable-based accounts of consonantal phonotactics. En O. Fujimura, B. D. Joseph y B. Palek (Eds.), *Item order in language and speech* (pp. 205-245). Prague: Karolinum Press.

- Stone, M. L. (1981). Evidence for a rhythm pattern in speech production: Observations of jaw movement. *Journal of Phonetics*, 9, 109-120.
- Stone, M. L. (1997). Laboratory techniques for investigating speech articulation. En W. J. Hardcastle y J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 11-32). Oxford: Blackwell.
- Stone, M. L. y Hamlet, S. (1982). Variations in jaw and tongue gestures observed during the production of unstressed /d/s and flaps. *Journal of Phonetics*, 10(4), 401-415.
- Stone, M. L. y Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw and palate contributions to speech production. *Journal of Phonetics*, 23, 81-100.
- Straka, G. (1963). La division de sons du langage en voyelles et consonnes peut-elle être justifiée? *Travaux de Linguistique et de Littérature, Université de Strasbourg*, 1, 17-99.
- Strycharczuk, P. y Kohlberger, M. (2014). Derived onsets in Spanish: an experimental study of resyllabification. Presentado en 14th Laboratory Phonology Conference.
- Szigetvári, P. (2008). What and where? En J. Brandão de Carvalho, T. Scheer y P. Ségéral (Eds.), *Lenition and fortition* (pp. 93-129). Berlin: Walter de Gruyter.
- Terrell, T. D. (1980). Teoría generativo-transformacional y dialectología castellana. En J. M. Guitart y J. Roy (Eds.), *La estructura fónica de la lengua castellana: fonología, morfología y dialectología* (pp. 203-246). Barcelona: Anagrama.
- 't Hart, J., Collier, R. y Cohen, A. (1990). *A perceptual study of intonation*. Cambridge: Cambridge University Press.
- Titze, I. R. (2000). *Principles of voice production*. Iowa City, IA: National Center for Voice and Speech.
- Topintzi, N. (2010). *Onsets: Suprasegmental and prosodic behaviour*. Cambridge: Cambridge University Press.
- Trubetzkoy, N. S. (1973 [1939]). *Principios de fonología*. Madrid: Cincel.
- Tuller, B., Harris, K. S. y Kelso, J. A. S. (1982). Stress and rate: Differential transformations of articulation. *Journal of the Acoustical society of America*, 71(6), 1534-1543.
- Tyrone, M. E., Nam, H., Saltzman, E., Mathur, G. y Goldstein, L. (2010). Prosody and movement in American Sign Language: A task-dynamics approach. En *Proceedings of Speech Prosody V*.
- Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica*, 45, 122-139.
- Vaissière, J. (2006). *La phonétique*. Paris: Presses Universitaires de France.
- Van der Hulst, H. y Ritter, N. A. (1999a). Theories of the syllable. En H. Van der Hulst y N. A. Ritter (Eds.), *The syllable: views and facts* (pp. 13-52). Berlin: Walter de Gruyter.
- Van der Hulst, H. y Ritter, N. A. (Eds.). (1999b). *The syllable: views and facts*. Berlin: Walter de Gruyter.
- Van der Hulst, H. y Van de Weijer, J. (1995). Vowel harmony. En J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 495-534). Oxford: Blackwell.

- Vatikiotis-Bateson, E. y Ostry, D. J. (1995). An analysis of the dimensionality of jaw motion in speech. *Journal of Phonetics*, 23, 101-117.
- Vayra, M. y Fowler, C. A. (1992). Declination of supralaryngeal gestures in spoken Italian. *Phonetica*, 49, 48-60.
- Vennemann, T. (1994). Universelle Nuklearphonologie mit epiphänomenaler Silbenstruktur. En K. H. Ramers, H. Vater y H. Wode (Eds.), *Universale phonologische Strukturen und Prozesse* (pp. 7-54). Tübingen: Max Niemeyer Verlag.
- Wang, M. y Cheng, C. (2008). Subsyllabic unit preference in young Chinese children. *Applied Psycholinguistics*, 29(2), 291-314.
- Westbury, J. R. y Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, 22(1), 145-166.
- Whitney, W. D. (1874). *Oriental and linguistic studies*. New York: Scribner, Armstrong, and Company.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. y Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91(3), 1707-1717.
- Wright, R. A. (2004). A review of perceptual cues and cue robustness. En B. Hayes, R. M. Kirchner y D. Steriade (Eds.), *Phonetically based phonology* (pp. 34-57). Cambridge: Cambridge University Press.
- Yoon, Y. B. y Derwing, B. L. (2001). A language without a rhyme: Syllable structure experiments in Korean. *The Canadian journal of linguistics*, 46(3-4), 187-237.
- Yoshida, S. (1996). *Phonological government in Japanese*. Canberra: Australian National University.
- Zamora Munné, J. C. y Guitart, J. (1982). *Dialectología hispanoamericana: Teoría-descripción-historia*. Salamanca: Almar.