



Universitat Autònoma de Barcelona

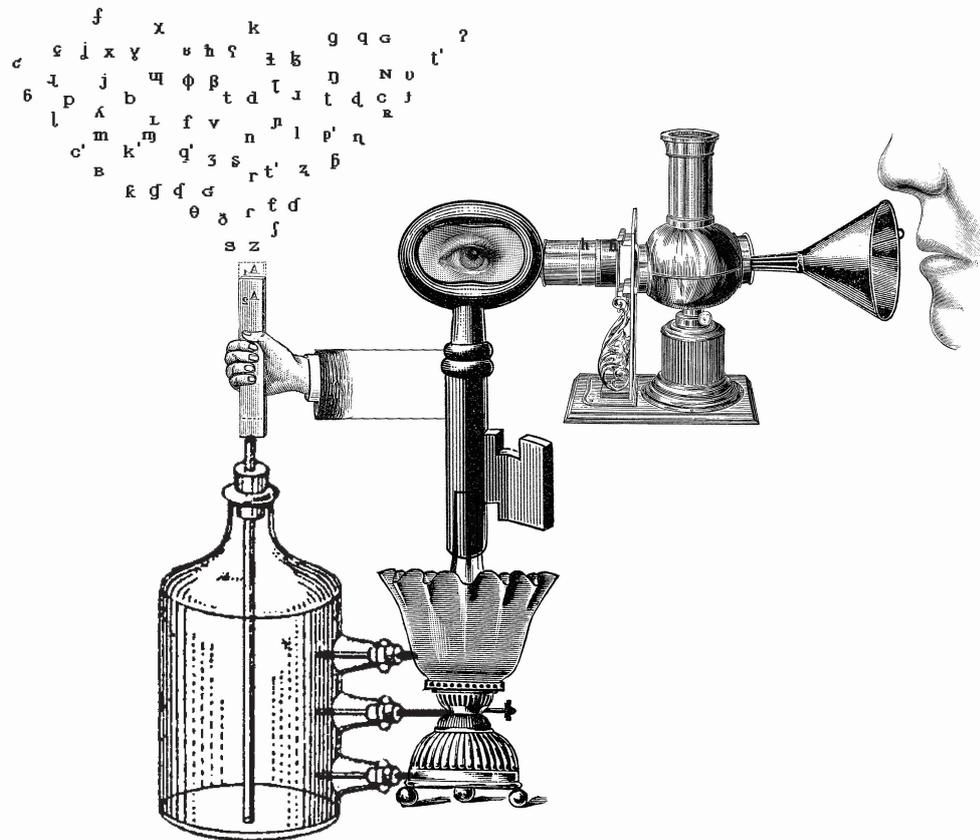
ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

Reducción fonética de las vocales del español de Bogotá (Colombia)

José Alejandro Correa Duarte





Universitat Autònoma de Barcelona
Departament de Filologia Espanyola
Facultat de Filosofia i Lletres
Programa de Doctorat en Filologia Espanyola

Reducción fonética de las vocales del español de Bogotá (Colombia)

Tesis doctoral

José Alejandro Correa Duarte

Dirigida por:

Dra. María J. Machuca Ayuso

Bellaterra, Barcelona, 2021

Tesis doctoral
Universidad Autónoma de Barcelona
Departamento de Filología Española
Facultad de Filosofía y Letras
Programa de Doctorado en Filología Española
Bellaterra, Barcelona, 2021

Imagen de portada: elaborada por ©Ixchel Estrada por comisión del autor
Diseño: José Alejandro Correa Duarte
Todas las figuras de esta tesis fueron elaboradas o adaptadas por el autor

Agradecimientos

Quiero agradecer a mi directora de tesis, María Machuca Ayuso, por su enorme paciencia, sus consejos oportunos y sus siempre acertadas recomendaciones: sin su dirección, guía, apoyo y comprensión nunca habría terminado la tesis.

Agradezco a Linda Carolina Rodríguez, quien ha sido mi principal interlocutora durante los últimos veinte años: gracias por las eternas discusiones lingüísticas, por los consejos, por el apoyo profesional y por su amor swedenborgiano. También quiero expresar mi gratitud a mi hijo, padres, hermanos y sobrinos por creer siempre en mí y darme aliento cuando lo he necesitado.

Quiero agradecer a los participantes en esta investigación y a mis estudiantes de la Maestría en Lingüística del Instituto Caro y Cuervo: como verán en las páginas que siguen, algunos temas incluidos en esta tesis los hemos discutido previamente en clase: ¡gracias por escucharme! Quiero que sepan que me siento honrado por haber sido su profesor de fonética.

Finalmente, agradezco a la Subdirección Académica del Instituto Caro y Cuervo por la financiación de esta investigación durante el periodo 2016-2020, y a la Facultad Seminario Andrés Bello de la misma institución por el apoyo económico para realizar mis estancias en Barcelona durante los años 2018 y 2019.

Resum

La reducció fonètica de les vocals és un procés de variació sincrònica per mitjà del qual les característiques temporals i espectrals de les vocals es modifiquen per la influència de factors lingüístics, orgànics, estilístics i expressius. Els principals efectes de la reducció fonètica són la disminució de la durada, la centralització dels formants i el minvament de l'espai acústic. El principal objectiu d'aquesta investigació és determinar la influència de factors com la velocitat de parla, la tonicitat, la variació contextual, el timbre, el tipus de paraula, l'estil de parla, el sexe dels parlants i la variació individual sobre la variabilitat de les vocals de l'espanyol de Bogotá. Per portar a terme aquest objectiu, es va analitzar un corpus de parla llegida a velocitat lenta, normal i ràpida, i un corpus de parla espontània tenint en compte paràmetres com la durada, les freqüències dels formants (F1 i F2), l'àrea d'espai vocàlic (AEV), la dispersió, la taxa de centralització dels formants (TCF), la diferència absoluta dels formants ($\Delta F1$ i $\Delta F2$), la velocitat de canvi dels formants (RF1 i RF2) i l'índex de centralització ∂ .

Els resultats mostren que les propietats temporals i espectrals de les vocals depenen de factors com l'entorn consonàntic, el timbre, la tonicitat, la velocitat de parla i el sexe dels parlants. No trobem efectes sobre els valors de freqüència dels formants relacionats amb el tipus de paraula i la posició de la vocal dins de la paraula i dins el grup fònic. L'àrea acústica és més gran en les vocals tòniques i en les mostres gravades a velocitat lenta, i menor en les vocals àtones i en els textos llegits a velocitat ràpida. Els resultats indiquen que la durada i l'àrea de l'espai vocàlic no van presentar diferències quan es va comparar la parla llegida a velocitat normal amb la parla espontània. Finalment, vam concloure que la variació fonètica individual exerceix una funció important sobre la reducció acústica de les vocals de l'espanyol de Bogotá.

Paraules clau: *reducció fonètica de les vocals, freqüències de formants, durada, dispersió, àrea de l'espai vocàlic, velocitat de parla, estil de parla, espanyol de Colòmbia*

Resumen

La reducción fonética de las vocales es un proceso de variación sincrónica por medio del cual las características temporales y espectrales de las vocales se modifican por la influencia de factores lingüísticos, orgánicos, estilísticos y expresivos. Los principales efectos de la reducción fonética son la disminución de la duración, la centralización de los formantes y la disminución del espacio acústico. El principal objetivo de esta investigación es determinar la influencia de factores como la velocidad de habla, la tonicidad, la variación contextual, el timbre, el tipo de palabra, el estilo de habla, el sexo del hablante y la variación individual sobre la variabilidad de las vocales del español hablado en Bogotá. Para cumplir con este objetivo, se analizó un corpus de habla leída a velocidad lenta, normal y rápida, y un corpus de habla espontánea teniendo en cuenta parámetros como la duración, las frecuencias de los formantes (F1 y F2), el área de espacio vocálico (AEV), la dispersión, la tasa de centralización de los formantes (TCF), la diferencia absoluta de los formantes ($\Delta F1$ y $\Delta F2$), la velocidad de cambio de los formantes ($rF1$ y $rF2$) y el índice de centralización ∂ .

Los resultados muestran que las propiedades temporales y espectrales de las vocales dependen de factores como el entorno consonántico, el timbre, la tonicidad, la velocidad de habla y el sexo del hablante. No encontramos efectos sobre los valores de frecuencia de los formantes relacionados con el tipo de palabra y la posición de la vocal dentro de la palabra y dentro del grupo fónico. El área acústica fue mayor en las vocales tónicas y en las muestras de habla leída a velocidad lenta, y menor en las vocales átonas y en los textos leídos a velocidad rápida. Los resultados indican que la duración y el área del espacio vocálico no presentaron diferencias cuando se comparó el habla leída a velocidad normal con el habla espontánea. Finalmente, concluimos que la variación fonética individual desempeña una función importante sobre la reducción acústica de las vocales del español de Bogotá.

Palabras clave: *reducción fonética de las vocales, frecuencias de formantes, duración, dispersión, área del espacio vocálico, velocidad de habla, estilo de habla, español de Colombia*

Abstract

Phonetic vowel reduction is a process of synchronic variation that changes the temporal and spectral characteristics of vowels owing to the influence of linguistic, organic, stylistic, and expressive factors. Phonetic reduction involves vowel shortening, centralized formant values and compression of the acoustic space. The main objective of the current research is to investigate the influence of speaking rate, lexical-stress, contextual variation, vowel quality, word category, speech style, sex, and inter-speaker differences on vowel variability in Bogota Spanish. A corpus of read speech produced at slow, normal, and fast rate, and a corpus spontaneous speech were analyzed by measuring duration, formant frequencies (F1-F2), vowel space area (VSA), dispersion, formant centralization rate (FCR), absolute formant frequency changes ($\Delta F1$ y $\Delta F2$), rates of formant frequency (rF1 y rF2) and centralization index (∂).

The results show that temporal and spectral properties of vowels depend on several factors such as consonantal context, vowel quality, lexical-stress, speech rate and speaker sex. We did not find clear effects of word category, vowel position within the word, and phrase context on formant frequencies. The vowel space area was largest for stressed vowels and slow speech, and smallest for unstressed vowels produced at fast rate. Results indicated that there were no differences in vowel duration and vowel space area between read speech at normal rate and spontaneous speech. Finally, it is argued that individual phonetic variation plays an important role on the acoustic reduction of vowels in Bogota Spanish.

Keywords: *phonetic vowel reduction, formant frequency, duration, dispersion, vowel space area, speech rate, speech style, Colombian Spanish*

Contenido

1.	Introducción	1
1.1	La reducción fonética.....	1
1.2	El español hablado en Bogotá.....	3
1.2.1	La pronunciación del español hablado en Bogotá	6
1.3	Motivación y estructura de la tesis	8
2.	La reducción fonética de las vocales	11
2.1	Definición	11
2.1.1	La centralización según el Alfabeto Fonético Internacional.....	14
2.1.2	La posición neutra del tracto vocal.....	17
2.2	Factores que determinan la reducción vocálica	19
2.2.1	Duración y entorno consonántico.....	19
2.2.2	Velocidad de elocución.....	28
2.2.3	El esfuerzo articulatorio	30
2.2.4	El acento	32
2.2.5	Propiedades léxicas.....	35
2.2.6	El estilo de habla.....	39
2.2.7	Diferencias entre hombres y mujeres	44
2.2.8	Variación dialectal y contacto de lenguas	47
2.3	Resumen	53
2.4	Preguntas y objetivos de la tesis.....	57
3.	Metodología	61
3.1	La constitución del corpus	61
3.1.1	Los estilos de habla.....	61
3.1.2	La velocidad de elocución.....	64
3.1.3	Entorno fonético de las vocales.....	65
3.1.4	Variación prosódica	67
3.1.5	Variación léxica	67
3.2	Los participantes	68
3.3	Las grabaciones.....	70

3.3.1	Transcripción y anotación	71
3.4	Parámetros acústicos	72
3.4.1	Mediciones estáticas de la duración y los formantes vocálicos	73
3.4.2	Mediciones dinámicas de los formantes	74
3.4.3	Medición del área del espacio vocálico	75
3.4.4	Medición de la centralización de las vocales	75
3.4.5	Normalización de las frecuencias formánticas	76
3.5	Tratamiento estadístico del corpus	77
3.5.1	Modelos lineales de efectos mixtos (MLM)	79
3.5.2	Variables consideradas en el análisis estadístico	80
4.	Resultados: Duración, formantes, espacio acústico y centralización de las vocales en tres velocidades de habla	83
4.1	Introducción	83
4.2	Procedimiento de análisis	85
4.3	Resultados según los parámetros analizados	86
4.3.1	Duración	86
4.3.2	Valores absolutos de F1 y F2	90
4.3.3	Valores normalizados de F1 y F2	97
4.3.4	Correlación entre los valores temporales y espectrales	99
4.3.5	Área del espacio vocálico (AEV)	101
4.3.6	Dispersión vocálica	105
4.3.7	Tasa de centralización de los formantes (TCF)	108
4.4	Conclusiones parciales	110
5.	Resultados: La influencia del entorno consonántico sobre los formantes vocálicos	115
5.1	Introducción	115
5.2	Procedimiento de análisis	117
5.3	Resultados según los parámetros analizados	117
5.3.1	Frecuencias de formantes en función del contexto consonántico	117
5.3.2	Variaciones dinámicas de las frecuencias de los formantes de /i/	123
5.3.3	Variaciones dinámicas de las frecuencias de los formantes de /e/	126
5.3.4	Variaciones dinámicas de las frecuencias de los formantes de /a/	131
5.3.5	Variaciones dinámicas de las frecuencias de los formantes de /o/	135
5.3.6	Variaciones dinámicas de las frecuencias de los formantes de /u/	138
5.4	Conclusiones parciales	141

6.	Resultados: Duración, formantes, espacio acústico y centralización de las vocales en habla leída y en habla espontánea.....	145
6.1	Introducción.....	145
6.2	Procedimiento de análisis.....	146
6.3	Análisis del habla leída en velocidad normal vs. habla espontánea.....	147
6.3.1	Duración.....	147
6.3.2	Valores de frecuencia de F1 y F2.....	150
6.4	Área vocálica, centralización y dispersión de las vocales.....	159
6.4.1	Área del espacio vocálico (AEV).....	159
6.4.2	Tasa de centralización de los formantes (TCF).....	162
6.4.3	Correlación entre el espacio vocálico y la tasa de centralización.....	163
6.4.4	Dispersión vocálica.....	164
6.4.5	Índice ∂	167
6.5	Conclusiones parciales.....	169
7.	Discusión y conclusiones.....	175
7.1	Discusión de los resultados.....	175
7.1.1	Características temporales y espectrales de las vocales bogotanas.....	175
7.1.2	Efectos de la velocidad de elocución sobre la duración y los valores de frecuencia de F1 y F2.....	176
7.1.3	La correlación entre la duración y los formantes vocálicos.....	177
7.1.4	La influencia de la tonicidad sobre la duración y los valores de frecuencia de F1 y F2.....	178
7.1.5	Factores que inciden sobre el área y la centralización del espacio acústico.....	179
7.1.6	La influencia del entorno consonántico.....	180
7.1.7	Factores léxicos.....	183
7.1.8	Posición de la vocal dentro de la palabra y el grupo fónico.....	184
7.1.9	Diferencias entre hombres y mujeres.....	185
7.1.10	El estilo de habla vs. velocidad de elocución.....	186
7.1.11	La variación individual.....	188
7.2	Conclusiones.....	189
	Bibliografía.....	197
A.	Anexo: Textos usados en la tarea de lectura.....	217
B.	Anexo: Temas de discusión utilizados para obtener habla conversacional.....	219
C.	Anexo: Valores medios y desviaciones estándar de la duración vocálica en función la tonicidad, la velocidad de elocución y el sexo del hablante.....	221

D.	Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (Hz) en función del timbre, la velocidad de elocución y el sexo del hablante.....	223
E.	Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (normalizadas con el método de Lobanov) en función del timbre, la velocidad de elocución y el sexo del hablante.....	225
F.	Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (Datos en hercios (Hz) y normalizados con el método de Lobanov) en función del timbre y el sexo del hablante en habla espontánea	227

Lista de figuras

Figura 1. División dialectal de Colombia y ubicación de Bogotá D. C.	4
Figura 2. Reducción fonológica en catalán, búlgaro y portugués de Brasil.....	12
Figura 3. Reducción fonética de las vocales de español de Bogotá	13
Figura 4. Las vocales cardinales	14
Figura 5. Vocales centrales después de la convención de Kiel (1989) y la corrección propuesta por J. C. Catford (1990)	15
Figura 6. Cuadro vocálico en 1993	16
Figura 7. Espectro y modelo de tubo de una <i>schwa</i> [ə].....	17
Figura 8. Transiciones en los entornos consonánticos de las sílabas [ba], [da], [ga].....	21
Figura 9. Efecto del contexto consonántico sobre las vocales del inglés	22
Figura 10. Ecuación de locus aplicada a los valores de los dos primeros formantes de la vocal /ə/ en función del entorno consonántico y la duración	24
Figura 11. Reducción fonética según van Bergem (1994, p. 57). Entorno consonántico velar [w_ɫ] (abajo) y alveolar [t_n].....	27
Figura 12. Reducción vocálica en inglés, francés, alemán y español.....	33
Figura 13. Distribución geográfica de la reducción fonética de las vocales en español	50
Figura 14. Factores que determinan la reducción fonética de las vocales.	56
Figura 15. Proporción de vocales analizadas según el punto de articulación de la consonante precedente C ₁ en habla leída.	66
Figura 16. Proporción de vocales analizadas según el punto de articulación de la consonante precedente C ₁ en habla espontánea.....	67
Figura 17. Bogotá Distrito Capital. Principales lugares de residencia de los 22 participantes en el estudio.....	69

Figura 18. Transcripción y anotación en Praat para el habla leída y el habla espontánea.	72
Figura 19. Transcripción ortográfica del habla espontánea en elan.	72
Figura 20. Segmentación de la vocal a partir de la forma de onda y del espectrograma.	73
Figura 21. Coeficientes de los factores fijos con relación al intercepto del modelo de la duración vocálica.	88
Figura 22. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo de la duración vocálica.	89
Figura 23. Coeficientes de los factores fijos incluidos con relación al intercepto del modelo de F1 (Hz).	91
Figura 24. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de habla para el modelo de F1 (Hz).	92
Figura 25. Coeficientes de los factores fijos con relación al intercepto del modelo de F2 (Hz).	94
Figura 26. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo del formante F2 en Hz.	95
Figura 27. Contorno bidimensional de F1 y F2 (Hz) para hombres y mujeres en tres velocidades de elocución.	96
Figura 28. Desviación típica de F1 y F2 para hombres y mujeres en tres velocidades de elocución.	96
Figura 29. Contorno bidimensional de F1 y F2 (normalizados) para las vocales tónicas y átonas producidas por hombres y mujeres en tres velocidades de elocución.	98
Figura 30. Correlaciones entre la duración y el primer formante para las vocales /i e a o u/ en tres velocidades de elocución.	100
Figura 31. Correlaciones entre la duración y el segundo formante para las vocales /i, e, a, o, u/ en tres velocidades de elocución.	101
Figura 32. Área del espacio vocálico (AEV) en Bark ² en función de la velocidad de habla, la tonicidad y el sexo del hablante.	102
Figura 33. Área del espacio vocálico (AEV) en Bark ² en función de la velocidad de habla, la tonicidad y el hablante.	104
Figura 34. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo del Área del espacio vocálico (AEV).	104

Figura 35. Coeficientes de los factores con relación al intercepto del modelo dispersión vocálica (Distancia euclídea).....	106
Figura 36. Dispersión vocálica en función de la velocidad de elocución, la tonicidad de la vocal y el timbre vocálico.....	107
Figura 37. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo de la dispersión vocálica.	108
Figura 38. Tasa de Centralización de los Formantes (TCF) en función de la velocidad de habla, la tonicidad y el hablante.	109
Figura 39. <i>K-medias</i> de los valores de frecuencia de los formantes de las vocales /i, e, a, o, u/ en función del entorno consonántico.....	119
Figura 40. Desviaciones típicas de los valores frecuencia de F1 y F2 (en hercios) en función del punto de articulación y el timbre.	121
Figura 41. Desviaciones típicas de F1 (en hercios) en función del hablante, el punto de articulación y el timbre vocálico.....	122
Figura 42. Desviaciones típicas de F2 (en hercios) en función del hablante, el punto de articulación y el timbre vocálico.....	122
Figura 43. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /i/ en función de la consonante precedente y la velocidad de elocución.	125
Figura 44. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /i/ en función de la consonante precedente y la velocidad de elocución.	126
Figura 45. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /e/ en función de la consonante precedente y la velocidad de elocución.	129
Figura 46. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /e/ en función de la consonante precedente y la velocidad de elocución.	130
Figura 47. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /a/ en función de la consonante precedente y la velocidad de elocución.	133
Figura 48. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /a/ en función de la consonante precedente y la velocidad de elocución.	134

Figura 49. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /o/ en función de la consonante precedente y la velocidad de elocución.	136
Figura 50. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /o/ en función de la consonante precedente y la velocidad de elocución.	137
Figura 51. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /u/ en función de la consonante precedente y la velocidad de elocución.	139
Figura 52. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /u/ en función de la consonante precedente y la velocidad de elocución.	140
Figura 53. Coeficientes de los factores fijos del modelo de la duración vocálica (ms) en habla leída y habla espontánea.	149
Figura 54. Interceptos aleatorios por hablante para el modelo de la duración en habla leída y habla espontánea.	150
Figura 55. Diagrama de puntos con los coeficientes de los factores fijos en el modelo de los valores de frecuencia de F1 (Hz) de las vocales producidas en habla leída y en habla espontánea.	152
Figura 56. Interceptos aleatorios por hablante para el modelo de los valores de frecuencia de F1.	153
Figura 57. Coeficientes de los factores fijos del modelo de los valores de frecuencia de F2 (Hz) de las vocales producidas en habla leída y en habla espontánea.	155
Figura 58. Interceptos aleatorios por hablante para el modelo de los valores de F2.	156
Figura 59. Valores medios y las desviaciones típicas de los valores de frecuencia de F1 y F2 (Hz) en función del sexo del hablante y el estilo de habla.	158
Figura 60. Diagrama con los valores medios y las desviaciones típicas de los valores de frecuencia de F1 y F2 (Lobanov) de las vocales tónicas y átonas producidas en habla leída y habla en espontánea.	158
Figura 61. Área del espacio vocálico (Bark) en función de la velocidad y el estilo de habla de 18 hablantes del español bogotano.	160
Figura 62. Área del espacio vocálico (AEV) en Bark ² por hablante y en función de la velocidad y el estilo de habla.	161

Figura 63. Diagrama de puntos con el Área del espacio vocálico (AEV) en Bark ² por hablante y en función de la velocidad y estilo de habla.	163
Figura 64. Correlación entre el área del espacio vocálico y tasa de centralización de los formantes.	164
Figura 65. Dispersión vocálica en función de la velocidad de elocución, el estilo y el timbre.	166
Figura 66. Índice ∂ para cuatro condiciones de elocución en función del timbre.	168

Lista de tablas

Tabla 1.	Lenguas con reducción fonética de las vocales y tareas de recolección de datos usadas para obtener habla espontánea	43
Tabla 2.	Frecuencia y porcentajes de aparición de las vocales en cada uno de los textos fonéticamente equilibrados utilizados para obtener las muestras de habla	63
Tabla 3.	Coefficientes de los efectos fijos y de los efectos aleatorios fijos para un modelo de la duración vocálica con un total de 6517 observaciones y 22 participantes	88
Tabla 4.	Coefficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F1 (Hz) con un total de 6517 observaciones y 22 participantes	91
Tabla 5.	Coefficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F2 (Hz) con un total de 6517 observaciones y 22 participantes.	94
Tabla 6.	Coefficientes de correlación y significativa de las correlaciones para las vocales /i, e, a, o, u/ en tres velocidades de elocución	99
Tabla 7.	Coefficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de Área del espacio vocálico (AEV) con un total de 132 observaciones y 22 participantes	102
Tabla 8.	Coefficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la dispersión vocálica (distancia euclídea) con 660 observaciones y 22 hablantes	106
Tabla 9.	Coefficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la Tasa de Centralización de los Formantes (TCF) con un total de 132 observaciones y 22 participantes	109
Tabla 10.	Desviaciones típicas de los valores frecuencia de F1 y F2 (en hercios) en función del punto de articulación y el timbre.	121
Tabla 11.	Coefficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de F2 (Hz) de la vocal /i/.....	125

Tabla 12.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /i/	126
Tabla 13.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de F2 (Hz) de la vocal /e/	129
Tabla 14.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /e/	130
Tabla 15.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de $\Delta F2$ (Hz) de la vocal /a/	133
Tabla 16.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /a/	134
Tabla 17.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de $\Delta F2$ (Hz) de la vocal /o/	136
Tabla 18.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /o/	137
Tabla 19.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de $\Delta F2$ (Hz) de la vocal /u/	139
Tabla 20.	Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /u/	140
Tabla 21.	Coeficientes de los efectos fijos y de los efectos aleatorios fijos para el modelo de la duración vocálica en habla leída y en habla espontánea con un total de 4223 observaciones y 18 participantes	149
Tabla 22.	Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F1 (Hz) con un total de 4071 observaciones y 18 participantes	152
Tabla 23.	Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F2 (Hz) con un total de 4071 observaciones y 18 participantes.	155
Tabla 24.	Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de Área del espacio vocálico (AEV) con un total de 144 observaciones y 18 participantes.	160
Tabla 25.	Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la Tasa de Centralización de los Formantes (TCF) con un total de 144 observaciones y 18 participantes	163

Tabla 26. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la dispersión vocálica (distancia euclídea) con 720 observaciones y 18 hablantes	166
Tabla 27. Índice $\hat{\sigma}$ para las vocales /i, e, a, o, u/ según la tonicidad en cuatro condiciones de elocución	168

1. Introducción

1.1 La reducción fonética

En todas las lenguas del mundo es común que en las interacciones comunicativas cotidianas los hablantes usen palabras con un menor número de sílabas, simplifiquen la articulación de las consonantes o modifiquen la duración de las unidades lingüísticas y el timbre de las vocales. Las realizaciones fonéticas resultantes llegan a ser incomprensibles para los hablantes no nativos o, incluso, para los hablantes nativos si la secuencia se extrae de su contexto lingüístico. Por ejemplo, como observan Hawkins y Smith (2001, p.170), el enunciado *I don't know* puede tener una realización fonética cercana a [ã̃n:əʊ] o [ǣ̃ǣ̃] en una conversación relajada entre familiares y ser perfectamente inteligible, pero en otra situación comunicativa más restringida estas secuencias no tienen significado. Este fenómeno es conocido como *reducción fonética* y, lejos de ocurrir de manera esporádica, se manifiesta de manera regular y sistemática en las lenguas en que se ha estudiado (Engstrand, 1992; Barry y Andreeva, 2001, entre otros).

La reducción fonética supone la existencia de realizaciones fonéticas normativas o pronunciaciones de referencia como las que derivamos del análisis de palabras aisladas (*citation form*). En este sentido, algunos autores interpretan el fenómeno como una simplificación de las unidades lingüísticas, como un conjunto de realizaciones fonéticas que se alejan de las formas canónicas (Wagner, 2019, p. 1) o como realizaciones con menos «substancia fonética» (Clopper y Turnbull, 2018, p. 25). También podemos definir la reducción fonética como una reorganización de los gestos articulatorios que resulta en palabras con menos sílabas, en procesos de lenición y elisión consonánticas, pero, paradójicamente, también puede manifestarse en realizaciones articulatoria y acústicamente más complejas (Simpson, 2001a; Correa y Rodríguez, 2018). Hawkins (2018, p. VII) utiliza una analogía muy acertada para explicar este punto: la reducción fonética, al igual que ocurre con la reducción de líquidos en el ámbito culinario, permite

la concentración de una sustancia, simplificar su volumen y forma, pero también resaltar los atributos de los ingredientes y lograr sabores más complejos.

El estudio de la reducción fonética no es nuevo, pues hay toda una tradición de investigaciones sobre habla conectada y fonética sintáctica inspirados en los fenómenos *sandhi* estudiados por los gramáticos de la India desde hace dos milenios (Kohler, 1996). Sin embargo, en los últimos 30 años las investigaciones han tomado vigor gracias al desarrollo de modelos teóricos orientados a explicar la variación fonética y el reconocimiento léxico. Así, en el plano teórico, es un hecho casi establecido que no hay invarianza en la señal acústica debido a que la producción del habla es adaptativa y está prospectivamente organizada, es decir, que los locutores ajustan su pronunciación a las demandas comunicativas y situacionales sin comprometer la discriminación de la información léxica (Lindblom, 1990; Kohler, 1996; Barry y Andreeva, 2001).

La reducción fonética es, entonces, un fenómeno intrínseco a la variabilidad fonético-acústica y es sensible a los factores lingüísticos y situacionales. Por esta razón, su estudio implica el uso de diferentes tareas de recolección de datos y requiere de diseños experimentales en los que se controle el grado de atención del hablante a su pronunciación (Warner, 2011). Por todo ello, el estudio del habla espontánea y el habla conversacional han resultado ser un terreno fértil para este campo de investigación. Este auge se ha visto beneficiado por la elaboración de grandes corpus orales y la aplicación de métodos cuantitativos para su análisis. Podemos afirmar que los estudios sobre la reducción fonética fomentan la pluralidad de estilística, permiten integrar el habla de laboratorio con otras tareas de recolección de datos y, de esta manera, aseguran la validez y fiabilidad de los resultados (Milder, 2013).

Cangemi *et al.* (2018), en un reciente volumen dedicado exclusivamente a la reducción fonética, notan que el término *speech reduction* tiene 58 ítems en Google Scholar para el periodo comprendido entre 1965 y 1990, cifra que se incrementó a 924 entre 1990 y 2018 y, según mis propias pesquisas, se publicaron 193 documentos nuevos entre 2018 y el 31 de diciembre de 2020. Estas cifras, a mi juicio, indican que se está consolidando una comunidad de investigadores que, como diría Peirce a propósito de su definición de ciencia, «se ayudan y estimulan unos a otros al comprender un conjunto particular de estudios como ningún extraño podría comprenderlos» (Peirce, 1905).

1.2 El español hablado en Bogotá

El español de Colombia se divide en dos superdialectos, uno costeño o meridional y otro denominado central o andino (Montes, 1982). A nivel fonético el primero se caracteriza por la aspiración y pérdida de /-s/, la neutralización y pérdida de las consonantes /r/ y /l/ en posición posvocálica y la velarización de /n/. El segundo se caracteriza por conservar la /s/ implosiva y mantener, en algunas zonas geográficas, la distinción entre /j/ y /ʎ/. De acuerdo con Montes (1982), este superdialecto andino se subdivide, a su vez, en dialecto centro-occidental, que incluye las variedades habladas desde la frontera con Ecuador hasta los departamentos cafeteros de Caldas y Antioquia, y en dialecto andino centro-oriental, que abarca las zonas montañosas de los departamentos de Tolima, Huila, Cundinamarca, Boyacá y Santander (Figura 1). El dialecto andino centro-oriental, al que pertenece la variante bogotana, se considera el más conservador del español colombiano (Canfield, 1962, p. 247).

El español hablado en Bogotá es una de las variedades lingüísticas más prestigiosas de Hispanoamérica y es la norma lingüística mejor valorada por los colombianos (Lipski, 1994, p. 207). Recientemente se realizó una investigación sobre las actitudes lingüísticas de los hablantes del español en veinte países, y los resultados muestran que el español hablado en España es el mejor valorado (41 %), seguido por la variedad hablada en el país de los encuestados (19 %) y, tercero, por el español hablado en Colombia (10 %). Los informantes mostraron preferencia por el español colombiano teniendo en cuenta rasgos como una pronunciación neutra y clara, la ausencia de elisión de fonemas y una velocidad de habla lenta (Bernal, Díaz y Munévar, 2018, p. 113). De acuerdo con los datos tomados en Colombia, el 73 % de las personas considera que el español bogotano es la mejor variedad del país (Bernal, 2016, pp. 246-258).

Como han observado varios investigadores, la idea según la cual el español de Colombia y, en particular, la variedad bogotana, tiene una pronunciación normativa y prestigiosa se debe, en parte, a la importancia de la filología y la corrección lingüística en la formación de la nación y la identidad cultural colombianas (Patiño Roselli, 2004; Espejo, 2012; Schwegler y Correa, 2018; Bernal, Díaz y Munévar, 2018, p. 113). La famosa obra de Rufino José Cuervo, *Apuntaciones críticas sobre el lenguaje bogotano* (1867), es un

ejemplo temprano del interés por la unidad de la lengua, los buenos usos idiomáticos y por «conformar nuestro lenguaje con el de Castilla» (Cuervo, 1867, p. VIII).



Figura 1. División dialectal de Colombia y ubicación de Bogotá D. C. Adaptado de Montes (1982).

Los resultados de las descripciones fonéticas del español hablado en Bogotá van en sentido contrario a lo que suponen los hablantes y a la tendencia conservadora que presuntamente lo caracteriza. De acuerdo con las caracterizaciones fonéticas existentes, el español hablado en Bogotá favorece la relajación articulatoria y la simplificación de las unidades lingüísticas a pesar de la presión social que impone su prestigio. Dos investigadores ya habían notado esta tendencia. El primero fue el dialectólogo colombiano Luis Flórez (1951), quien en su trabajo *Pronunciación del español en Bogotá*, escribió las siguientes observaciones fonéticas:

Son corrientes en la fonética bogotana [...] la poca tensión muscular y frecuente relajación en la articulación de las vocales y de las consonantes, hechos con los cuales se relacionan otros como el tono bajo, la poca abertura de la boca y la tendencia a suavizar la expresión en general (Flórez, 1951, pp. 326-327).

El abrir poco la boca produce la impresión de que todos los sonidos se articulan en la mitad anterior de la boca, hacia los dientes. Al mismo tiempo se nota una gran abreviación de las vocales (Flórez, 1951, p. 327).

El segundo autor fue Urbano González de la Calle, un filólogo español que estuvo en Bogotá en el año 1963 y escribió lo siguiente:

Ese tono apagado va muchas veces unido a una relajación articulatoria que suele engendrar equívocos o dificultades de comprensión en quienes no tenemos habituado el oído a percibir las inflexiones particularísimas del castellano de América. No hay que decir que las inflexiones aquí apuntadas quedan reducidas a términos tan restringidos en algunos casos, que apenas resultan perceptibles para un oído peninsular dotado de acuidad auditiva extraordinaria. (González de la Calle, 1963)

Estas afirmaciones son importantes por cuanto resaltan que la «relajación articulatoria», la «poca tensión muscular» y una menor «abertura de la boca» son la característica más sobresaliente del español bogotano. Segundo, se sugiere que esta tendencia comporta problemas perceptivos para algunos hablantes no nativos de la variedad. Y tercero, los testimonios tienen cerca de setenta y seis años (Flórez inició su trabajo en 1944), lo cual invita a pensar que estas tendencias de la pronunciación bogotana están bastante arraigadas desde hace varias generaciones. A continuación haré un resumen de las principales características fonéticas del español hablado en Bogotá y mostraré que la

evolución de su sistema consonántico refleja, en efecto, un proceso de debilitamiento más o menos estable a lo largo de las últimas décadas.

1.2.1 La pronunciación del español hablado en Bogotá

Como afirma Espejo (2012, pp. 200-201), no resulta fácil definir el español hablado en Bogotá debido a que «no se trata de una modalidad claramente diferenciada» y «en ella convergen todas las variedades del español de Colombia». Por tanto, algunos de los rasgos fonéticos también pueden encontrarse en otras variedades habladas en Colombia. Sin embargo, se han documentado algunos rasgos y procesos que podemos considerar, al menos, como hipótesis de trabajo, característicos del español hablado en la capital colombiana.

Actualmente, los fonemas /b/, /d/ y /g/ se realizan como oclusivas sonoras después de una consonante; En posición intervocálica se pronuncian como aproximantes [β, ð, ɣ] si los hablantes usan un estilo cuidado y enfático, pero si usan un estilo espontáneo la consonante /d/ y, en menor medida, /b/, se eliden. Hace medio siglo la tendencia al debilitamiento de las oclusivas sonoras estaba menos desarrollada. Por ejemplo, Canfield (1962, p. 247) anotaba que las oclusivas sonoras después de consonante eran tan marcadas que le daban al español bogotano «cierto efecto de staccato». De acuerdo con Flórez (1973), en posición intervocálica /d/ se elidía principalmente en las palabras terminadas en *-ado* y era común en habla espontánea y en hablantes de todos los niveles sociales. En cambio, el debilitamiento de /b/ y /g/ intervocálicas no era «regular ni uniforme» y ocurría «sobre todo en el hablar descuidado y, más aún, en la pronunciación vulgar» (Flórez, 1973, pp. 108-109). Dos décadas después, Montes *et al.* (1998, pp. 96-104) encontraron que la elisión de la oclusiva bilabial sonora /b/ en posición intervocálica era frecuente entre hablantes de un nivel socioeconómico bajo. De acuerdo con los datos presentados por estos autores, para finales de siglo XX la elisión de la oclusiva dental sonora intervocálica /d/ ya se había ampliado a un mayor número de contextos y estaba generalizada entre los hablantes de todos los niveles sociales.

Bogotá está ubicada en una zona dialectal que se caracteriza por conservar la fricativa sibilante /s/ en posición de coda silábica (Flórez, 1951, 1973; Canfield, 1962). De acuerdo con varios investigadores, en Bogotá predomina una /s/ alveolar, predorsal y sorda. Montes *et al.* (1998, p. 79) encontraron que la sibilante /s/ se conservaba en los

datos obtenidos a través de encuestas, sin embargo, en habla espontánea hallaron elisiones y aspiraciones en posición intervocálica y en coda silábica, lo cual fue más común entre los hablantes de un nivel socioeconómico bajo y medio. Como anotan Montes *et al.* (1998, p. 103), la pérdida de /s/ va en dirección contraria a las caracterizaciones fonéticas del español de Bogotá y sugiere, como propuso Canfield (1962), el desarrollo de tendencias de la fonética andaluza en las variedades andinas del centro de Colombia.

Uno de los rasgos que, según la bibliografía, distinguía al español hablado en Bogotá durante el siglo XX era la realización de /r/ final de sílaba como una vibrante asibilada [ř] (Flórez, 1951, 1973; Canfield, 1962). Este sonido dejó de usarse a lo largo de la segunda mitad del siglo a favor de una vibrante múltiple alveolar (Flórez, 1973; Montes *et al.*, 1998, p. 104), así que actualmente no es común encontrarla en hablantes jóvenes y solamente se oye entre algunos hablantes adultos (Espejo, 2012). En la misma línea, el español hablado en Bogotá conservaba, como otras variedades andinas, los fonemas /j/ y /ʎ/, pero esta distinción se perdió hace varias décadas (Montes *et al.*, 1998; Espejo, 2012).

A continuación, presentamos una transcripción fonética del *El viento norte y el sol*¹ (Coloma, 2005) ilustrando la pronunciación de una mujer de 25 años nacida en la ciudad de Bogotá y perteneciente a un nivel socioeconómico medio alto. La muestra, tomada del corpus de esta tesis, refleja los rasgos mencionados anteriormente, esto es, el debilitamiento de las oclusivas en posición intervocálica, la elisión de /s/ en posición posvocálica, yeísmo y, en general, una tendencia a la elisión de vocales y sílabas en posición final de grupo fónico.

[el 'bjɛnto 'nortɛ i el 'sol disku'tian soβre 'kwɔl'dejos era ɛl 'mas 'fwerte | 'kwɔndo
pa'so un eks'trajɔ βja'ħero m'bweltɔ en ñna 'ropa 'mwi aβri'ɣa: l kombi'njɛron ɛŋ
ke kjen 'aɲtes lo'ɣrara ɔβli'ɣaraɫ transe.ũnte ɔ ki'tarse la'βriɣo se'ria konsiɛ'raɔ

¹ «El viento norte y el sol discutían sobre cuál de ellos era el más fuerte, cuando pasó un extraño viajero envuelto en unas ropas muy abrigadas. Convinieron en que quien antes lograra obligar al transeúnte a quitarse el abrigo sería considerado más poderoso. El viento (...) sopló con gran furia, pero cuanto más soplabla, más se ceñía el hombre su ropa al cuerpo. Entonces se dio por vencido, y el sol empezó a brillar con mucha fuerza. Inmediatamente el viajero se despojó de su abrigo; y así ya quedó claro que el sol tenía superioridad respecto del viento».

'mas poɛ'ros || el bjeɲto so'plo kon 'mas 'furia | pero kwaɲto 'ma so'pla: | 'ma se se'nja e'lombre su 'ropal 'kwerpo | entonsej'o por ben'siðo | j̃el 'sol empe'so ɔ βri'jar ko'mutʃa 'fwersa || imej'atamentel bja'fhero se:spo'f̃o ðe swa'βriɣo | iɔ'si 'ja ke'ðo 'klaro kel 'soj̃ tenja superjori'a res'pekto ðel 'bjeɲto]

Como se ve, la mayoría de los estudios y las descripciones fonéticas se han centrado casi exclusivamente en el estudio de las consonantes. En lo que respecta al vocalismo, Flórez (1951, p. 327) escribió hace setenta años que las vocales eran breves y se producían con menor abertura que en el español peninsular. Este fenómeno, que se describirá más tarde, se conoce como *reducción fonética de las vocales* y ha recibido cierta atención en los estudios sobre el español hablado en las zonas altas de Ecuador, Perú y México (para mayor detalle véase la subsección 2.2.8). Sin embargo, a pesar de la claridad de los planteamientos de Flórez (1951, p. 327) y de la fecha de publicación de su trabajo, no contamos con investigaciones fonéticas sobre las vocales del español hablado en Bogotá y su relación con la tendencia del sistema consonántico hacia el debilitamiento articulatorio y, más específicamente, hacia la reducción fonética.

1.3 Motivación y estructura de la tesis

Tuve varias motivaciones para investigar el fenómeno de la reducción fonética y analizar las vocales del español hablado en Bogotá. Sin embargo, la principal fue el deseo de mejorar mis habilidades como investigador. Propuse una investigación de tipo cuantitativo para mejorar mi dominio de las técnicas y los métodos estadísticos que se usan en la actualidad para analizar un corpus fonético. El tema de investigación, la reducción fonética de las vocales, conjugaba mis intereses personales (la variabilidad vocálica) con mi trabajo como docente e investigador en el área de la fonética española. Finalmente, decidí estudiar la variedad de español hablada en Bogotá porque conozco la complejidad social de la ciudad y porque con este trabajo quiero dar continuidad a una tradición de lingüística que, como vimos, tiene cerca de 150 años.

Esta tesis está organizada en siete capítulos. Después de esta introducción, en el segundo capítulo se define la reducción fonética de las vocales, se mencionan las características que la distinguen de la reducción fonológica y se discute brevemente su representación en el AFI. El resto del capítulo se dedica a discutir los factores fonéticos, lingüísticos, situacionales y orgánicos que determinan la reducción fonética de las vocales. Entre los

factores fonéticos que discutimos está la duración vocálica y su relación con el entorno consonántico, el tempo y la velocidad articulatoria. Entre los factores lingüísticos consideramos la influencia del acento léxico y el contenido semántico de las palabras. Como se discute en este capítulo, una parte importante de los estudios publicados durante las últimas décadas ha estudiado la reducción fonética de las vocales en relación con el estilo de habla y, especialmente, cómo se modifica el área y la centralización del espacio vocálico en habla leída y en habla espontánea. Además, la variación individual y, concretamente, las diferencias anatómicas y sociofonéticas entre hombres y mujeres también han mostrado ser determinantes de la duración y el grado de centralización de las vocales. Finalmente, se discuten los estudios sobre la reducción vocálica en español y la influencia de la variedad lingüística y el contacto de lenguas sobre la variabilidad acústica de las vocales. El segundo capítulo cierra con un resumen de los factores que determinan la reducción acústica de las vocales y, con base en esta revisión, se definen las preguntas de investigación que se abordarán en el resto de la tesis.

En el tercer capítulo se describe la metodología de recolección y análisis de los datos. Primero, explicamos cómo se operacionalizaron las variables y los criterios que se tuvieron en cuenta para la constitución del corpus. Segundo, explicamos los parámetros acústicos y el procedimiento de análisis de la duración y de los formantes vocálicos. En particular, explicamos el método de normalización de las frecuencias de los formantes y las mediciones que usamos para analizar la reducción fonética de las vocales (formantes, velocidad de cambio, el área del espacio vocálico y tasa de centralización de los formantes, entre otros). Al final del capítulo se explica el procedimiento de anotación del corpus y los métodos de análisis estadístico que se adoptaron para responder a las preguntas e hipótesis de investigación.

En el cuarto capítulo se reportan los resultados del análisis de los valores de frecuencia de los formantes y de la duración vocálica de acuerdo con el análisis de un corpus de habla leída a tres velocidades de elocución. Para explicar los datos se construyeron cinco modelos de efectos mixtos con los dos primeros formantes, la duración, la dispersión vocálica y la tasa de centralización de los formantes como variables respuesta. Como factores incluimos el timbre, la tonicidad, la velocidad de elocución, el punto de articulación de la consonante precedente, el sexo del hablante y la posición de la vocal dentro de la palabra y dentro del grupo fónico. Gracias al número de variables incluidas,

en este capítulo logramos obtener una perspectiva general sobre la variación acústica de las vocales del español hablado en Bogotá y, adicionalmente, describir el comportamiento fonético individual de un gran número de hablantes.

En el quinto capítulo se presentan los resultados relacionados con la influencia del entorno consonántico sobre el timbre vocálico a partir del análisis del mismo corpus de habla leída estudiado en el cuarto capítulo. Para cuantificar el grado de coarticulación de las consonantes sobre las vocales y hacer inferencias sobre la velocidad de los movimientos articulatorios, incluimos mediciones espectrales obtenidas en la parte estable de la vocal y mediciones dinámicas de los valores de frecuencia de F1 y F2. Además del punto de articulación y los fonemas consonánticos, en el análisis cuantitativo también consideramos variables como la tonicidad, la velocidad de habla, el sexo del hablante y la posición de la vocal dentro de la frase. Este capítulo reporta un conjunto de resultados que refuerzan los hallazgos de investigaciones previas sobre la coarticulación al tiempo que presenta datos novedosos relacionados con la influencia del entorno consonántico sobre las vocales del español.

En el sexto capítulo se comparan los valores de la duración, los valores de frecuencia de los dos primeros formantes, y las mediciones del espacio acústico como el área, la tasa de centralización y la dispersión vocálica en habla leída y en habla espontánea. El objetivo es, de un lado, determinar si el estilo tiene un efecto independiente, o interactúa con otros factores como la tonicidad y, en ese sentido, incrementa o contrarresta los efectos de otras variables. Y de otro lado, en este capítulo buscamos determinar cómo organizan los hablantes el espacio acústico de las vocales. Resulta interesante saber, por ejemplo, si los bogotanos organizaron el área del espacio vocálico en un continuo organizado de menor a mayor grado de hipoarticulación determinado por el estilo de habla.

En el capítulo final se comparan los resultados obtenidos en nuestro estudio con los de otras investigaciones tomando como referencia las variables consideradas en el análisis cuantitativo. En el apartado de conclusiones se contestan las preguntas e hipótesis de investigación abordadas en este trabajo, además, se hace una síntesis de los resultados obtenidos en el análisis acústico de las vocales bogotanas.

2. La reducción fonética de las vocales

2.1 Definición

Las vocales se describen articulatoriamente teniendo en cuenta, al menos, tres parámetros: el grado de abertura linguomandibular, la posición de la lengua, y la disposición de los labios (Koenig, 2006, p. 50). Por esta razón distinguimos entre vocales cerradas, abiertas, anteriores, centrales, posteriores, redondeadas y no redondeadas. Esta categorización implica, en primer lugar, que cada vocal tiene una configuración articulatoria más o menos estable y, segundo, que las lenguas aprovechan estos tres parámetros para hacer distinciones fonológicas y, por tanto, léxicas. Cuando una vocal se realiza con una posición lingual y una postura de los labios específicas, se denomina *vocal plena (steady state)*, y su gesto prototípico se conoce como *objetivo articulatorio (target)*. En el análisis fonológico tradicional, este tipo de realización «estable» se considera la forma canónica de los fonemas vocálicos. No obstante, estas realizaciones pueden alterar sus características articulatorias en función de diferentes factores. En este sentido, la *reducción vocálica* ocurre cuando las vocales tienden a una posición articulatoria neutra y se alejan de su realización más «estable».

El tipo de reducción vocálica más conocida se denomina *reducción fonológica*. Ocurre cuando las vocales átonas tienen una duración más breve que las tónicas, de manera que los hablantes no alcanzan el objetivo articulatorio. Una consecuencia de esto es que las vocales átonas neutralizan las distinciones fonológicas basadas en el grado de altura lingual. De acuerdo con Flemming (2005) y Barnes (2006), este mecanismo es bastante frecuente en las lenguas del mundo, pero las vocales involucradas en la reducción y las realizaciones fonéticas resultantes dependen del inventario fonético y del sistema fonológico de la lengua. Por ejemplo, en inglés las vocales átonas se realizan como una vocal central o *schwa* [ə] en posición final, en posición intermedia e inicial de palabra p. ej. *sofa, emphasis, arise* ['sɒfə, 'ɛmfəsis, ə'reɪz] (Ladefoged y Johnson, 2010, p. 42).

Pero, como se ve en la Figura 2, en catalán oriental, búlgaro y portugués de Brasil el sistema vocálico átono está constituido por tres vocales y dos grados de abertura [i, ə, u].

Catalán oriental	Búlgaro oriental	Portugués de Brasil
/i/ → [i]	/i, e/ → [i]	/i, e, ε / → [i]
/e, ε, a/ → [ə]	/ə, a/ → [ə]	/a/ → [ə]
/u, o, ɔ/ → [u]	/u, o/ → [u]	/u, o, ɔ / → [u]

Figura 2.Reducción fonológica en catalán, búlgaro y portugués de Brasil. Fuente: catalán (Carbonell y Llisterra, 1992; Recasens, 2014), búlgaro y portugués (Barnes, 2006).

Existe un segundo tipo de reducción denominada *reducción fonética* o *reducción acústica* de las vocales. Esta ocurre cuando el sistema vocálico (o alguno de sus miembros) se centraliza debido a la influencia del entorno consonántico, a la duración, a la velocidad de elocución o al uso de un estilo de habla coloquial (Fourakis, 1991; Lindblom, 1963; Nord, 1986; van Bergen, 1993). En español, por ejemplo, las vocales se acercan a una vocal central media ([ɛ̃] o [ə]) en habla espontánea (Harmegnies y Poch, 1992). Además del estilo, el entorno consonántico y la clase gramatical de la palabra favorecen la aparición de vocales con un timbre cercano a *schwa*. En la Figura 3 se ilustra un caso en que la preposición ‘de’ se asimila con la sílaba átona precedente en el enunciado ‘el domingo de ramos’ [el: do'miŋɔ: rɒs], lo cual da como resultado una vocal central media alargada [ə:] con formantes característicos de una posición neutra del tracto vocal (500 Hz para el primer formante y 1500, para el segundo).

La aparición de *schwa* en contextos como el anterior no determina las distinciones léxicas en español, sino que es el resultado de un proceso de debilitamiento que afecta tanto a vocales como a consonantes. Los hablantes de una lengua con reducción léxica o fonológica se enfrentan a una situación diferente: la vocal central forma parte de la pronunciación normativa de las palabras y, por tanto, se produce intencionalmente en sílabas átonas, en posición intermedia de palabra u ocurre como cualquier otro fonema de la lengua (Oostendorp, 2014). Por esta razón, la reducción fonética se ha interpretado como un cambio en marcha y la reducción fonológica como su materialización en una comunidad de habla (van Bergen, 1995). Como veremos, los factores involucrados en la reducción fonológica son los mismos que impulsan la aparición de la reducción fonética en situaciones comunicativas concretas y su difusión entre un grupo de hablantes.

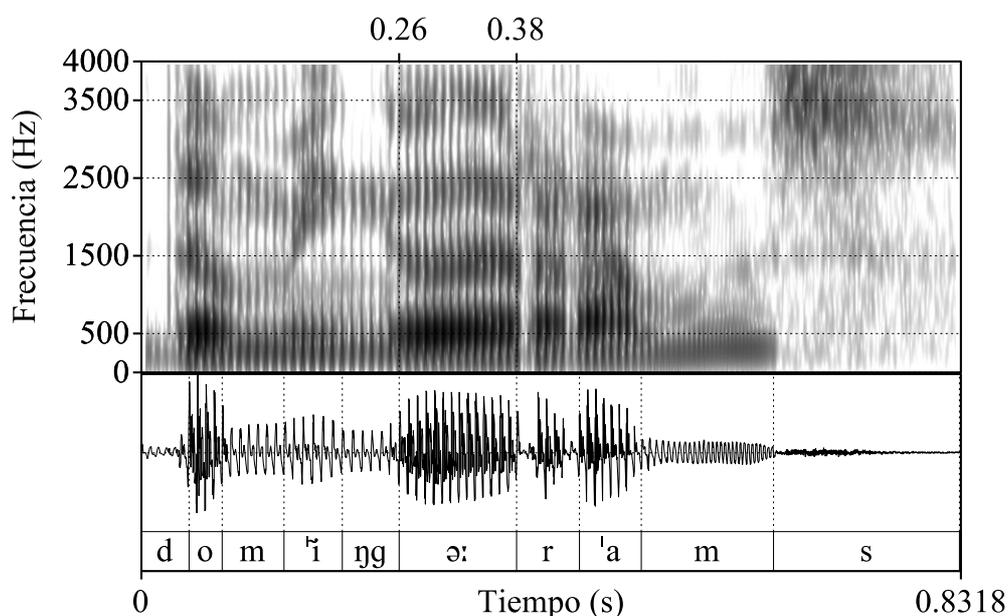


Figura 3.Reducción fonética de las vocales de español de Bogotá

Los trabajos sobre tipologías lingüísticas se han centrado exclusivamente en la reducción fonológica, pero concuerdan en que la vocal central media [ə] tiene un comportamiento particular en las lenguas del mundo. Por ejemplo, Crothers (1978) analizó 209 lenguas del *Stanford Phonology Archiving Project* y encontró que la mayoría tiene una vocal central media [ə] o alta [ɨ]. En la misma línea, Schwartz *et al.* (1997) estudiaron 317 lenguas de la base de datos UPSID (Maddieson, 1984) y observaron que *schwa* tiene una alta frecuencia en todos los sistemas vocálicos y su aparición es dos veces más alta que [ɨ]. Los autores aseguran que la vocal central media no interactúa con otras vocales en los sistemas analizados: se denomina a este comportamiento *regla de transparencia* y significa que la presencia o ausencia de una *schwa* no modifica la estructura de un sistema vocálico (Schwartz *et al.*, 1997, pp. 248-251).

Para Schwartz *et al.* (1997), la vocal central media [ə] constituye por sí misma un «sistema paralelo» que existe por la acción de principios diferentes a los que tienen las demás vocales. Por ejemplo, la existencia de una vocal redondeada como /y/ implica la existencia de su contraparte no redondeada /i/. Estos autores proponen que la aparición de *schwa* en las lenguas del mundo no exige este tipo de implicaciones porque ocurre independientemente del timbre y el número de vocales de una lengua. Según su planteamiento, cuando esta vocal aparece en una lengua, se debe a un proceso

sistemático de relajamiento articulatorio y, por tanto, su origen debe explicarse por medio del estudio de fenómenos fonéticos como puede ser el fenómeno de la coarticulación.

2.1.1 La centralización según el Alfabeto Fonético Internacional

Hasta 1988 en el *Alfabeto Fonético Internacional* (AFI) se caracterizaron las vocales siguiendo el sistema de representación creado por Daniel Jones (1922). Este sistema, conocido como *vocales cardinales* (Figura 4), fue creado para establecer puntos de referencia que permitieran la descripción articulatoria y la identificación auditiva de los timbres vocálicos de cualquier lengua. El sistema organiza estos sonidos según la altura de la lengua y su posición en el plano medio sagital del tracto vocal. Jones distingue entre vocales cardinales primarias y vocales cardinales secundarias según la relación entre el redondeamiento y la altura lingual: el esquema de las vocales cardinales se inicia con la representación de una vocal anterior cerrada no redondeada [i], la cual es seguida por tres vocales diferenciadas únicamente por el grado de abertura [e, ε, a]; en la quinta vocal [ɑ] la raíz de la lengua se desplaza hacia la pared faríngea, y en las vocales [ɔ, o, u] el dorso se eleva y se incrementa el redondeamiento de los labios. En las vocales cardinales secundarias la relación es inversa: las vocales anteriores [y, ø, œ] son redondeadas y las vocales posteriores [ʌ, ɤ, ɯ], a excepción de [ɐ], son no redondeadas.

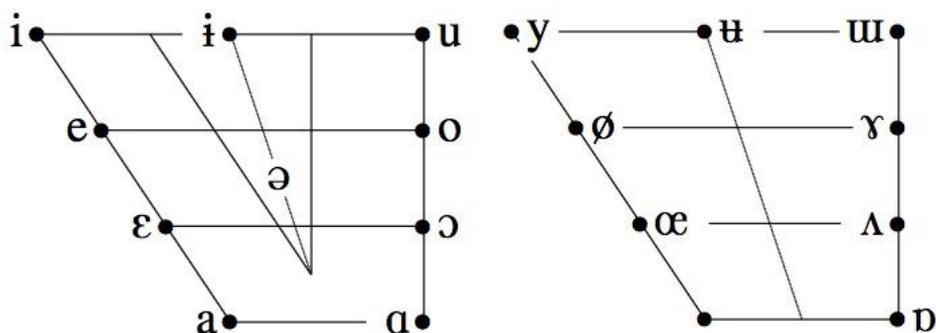


Figura 4. Las vocales cardinales (International Phonetic Association, 1949)

Como se ve en la Figura 4, el diagrama que aparece en *The Principles of International Phonetic Association* (1949) incluye las centrales [ɪ, ə] en las vocales primarias y [ɥ] entre las secundarias. Nótese, además, que la altura de la lengua es equidistante entre las vocales cardinales, sin embargo, el sonido [ə] puede ubicarse en un espacio articulatorio

y auditivo amplio e indeterminado. Cabe decir que el décimo octavo principio de la asociación sugiere:

The association also recommends that the letter ə be employed to denote any unrounded vowel situated in the interior triangle. If a language contains two unrounded central vowels, it is recommended that ə denotes the closer one and ɜ for the opener one. Occasionally ɜ may be employed to represent another variety of central vowel (p. 7).

En la convención de Kiel, celebrada en 1989, la asociación incluyó las vocales centrales abiertas [ɐ] y una vocal central media redondeada [ə] en el trapecio vocálico del *Alfabeto Fonético Internacional* (Figura 5). Catford (1990) reaccionó contra este cambio, pues consideraba que las vocales centrales debían tener realizaciones semicerradas y semiabiertas que sirvieran como referencia para la identificación de estos sonidos a través de la propiocepción de los movimientos articulatorios. Como se ilustra en la Figura 5, el autor propuso solucionar este problema con la inclusión de los símbolos [ə, ɜ] para las centrales semicerradas y [ɜ, ə] para las centrales semiabiertas.

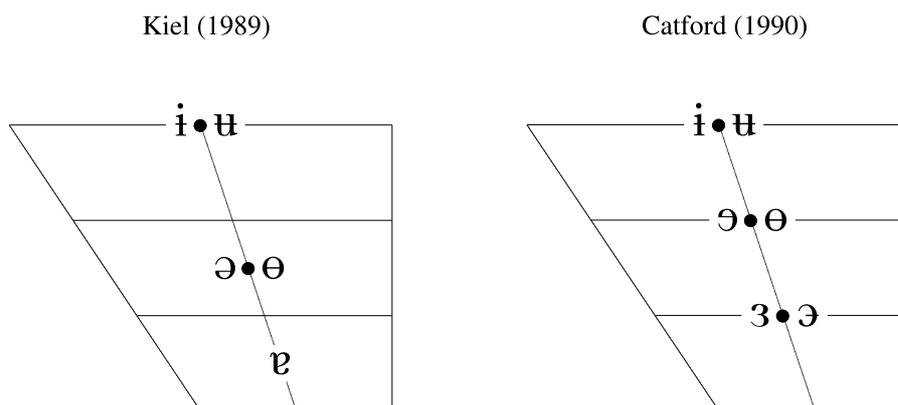


Figura 5. Vocales centrales después de la convención de Kiel (1989) y la corrección propuesta por J.C. Catford (1990)

En 1993 la asociación integró las vocales cardinales primarias y secundarias en un solo diagrama, y adoptó la propuesta de Catford (1990) al incluir los símbolos [ə, ɜ] y [ɜ, ə] para ajustar las vocales centrales a los grados de abertura existentes (Figura 6). Las vocales [ə] y [ɐ] conservaron su posición como central media y central abierta respectivamente. Para Pfitzinger y Niebuhr (2011) estas modificaciones significan un

retroceso porque introducen vocales de referencia innecesarias, y el sistema de vocales cardinales pierde la simplicidad que ganaba al integrar el redondeamiento de los labios con la posición de la lengua.

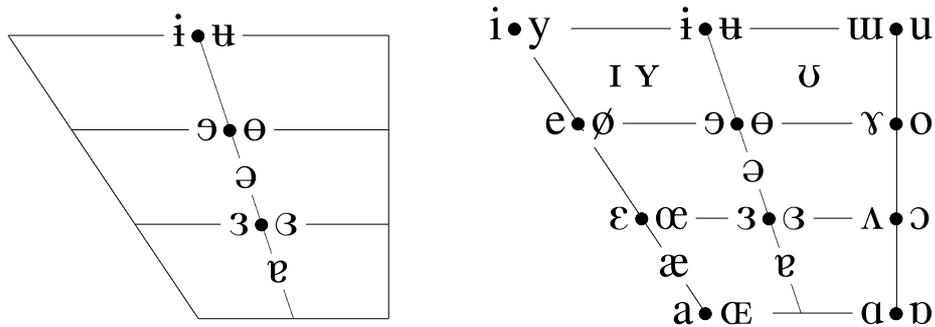


Figura 6. Cuadro vocálico en 1993 (International Phonetic Association, 1993)

Como se ve, en las diferentes revisiones del AFI la vocal [ə] no tiene una altura lingual específica. En la versión de 1949 (Figura 4), por ejemplo, esta vocal tiene una altura lingual indeterminada y se usa para abarcar una amplia variedad de realizaciones fonéticas. De acuerdo con las determinaciones tomadas en 1993 (IPA, 1993, p.33), se sugiere conservar el símbolo [ə] como una vocal central de carácter general. Ahora bien, como se aclara en el *Handbook of the International Phonetic Association* (1999), cuando una vocal se centraliza debido a la influencia de factores fonéticos, es conveniente usar diacríticos para indicar la naturaleza de la reducción fonética. Por ejemplo, cuando /e/ se *centraliza*, puede transcribirse como [ë] (equivalente a [ẽ]) y cuando la vocal /a/ se cierra hacia una altura media se sugiere usar [ą] (equivalente a [ɐ]). La transcripción de la centralización usando diacríticos tiene la ventaja de mostrar la direccionalidad del desplazamiento vocálico y, por tanto, es una representación útil de los cambios sincrónicos de los sistemas vocálicos. En las lenguas con reducción fonológica, los lingüistas usan, en cambio, símbolos menos específicos, como [ə]. Por esta razón, concordamos con Lass (2009, p.55) en que [ə] no es una vocal con un timbre específico, sino una manera de representar una gran variedad de vocales centrales cuya realización fonética es específica de la lengua o variedad lingüística.

2.1.2 La posición neutra del tracto vocal

La *teoría acústica de la producción del habla* interpreta la vocal central media [ə] como la posición neutra del tracto vocal (Chiba y Kajiyama, 1958; Fant, 1970; Stevens, 1998). Esta articulación tiene características acústicas similares a las de un tubo uniforme cerrado en un extremo (la glotis) y abierto en el otro (los labios) y, por tanto, se dice que sus frecuencias de resonancia o formantes son múltiplos impares de la primera resonancia.

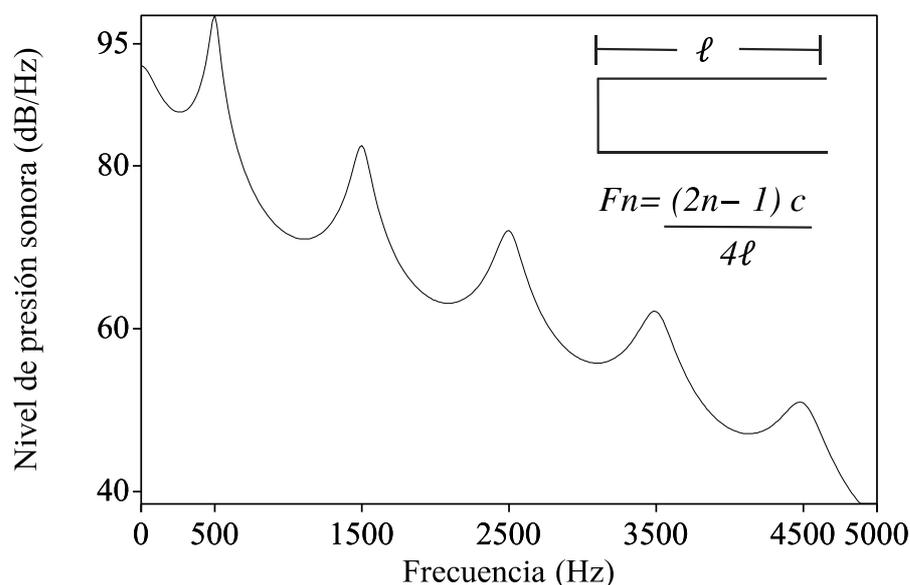


Figura 7. Espectro y modelo de tubo de una *schwa* [ə]. Elaborado a partir de la información reportada en Stevens (1998, p.138-139) y Johnson (2003, p. 94).

La Figura 7 muestra el espectro y un modelo de tubo de una *schwa*. Nótese, por ejemplo, que el segundo formante es tres veces más alto que el primero y el tercer formante es cinco veces más alto (Johnson, 2003, p. 94). Esta diferencia es igual a $(2n-1)$, donde n es el número del formante. Como indica la ecuación 1, si multiplicamos este valor por la velocidad del sonido (35 000 cm/seg) y dividimos el resultado por 4 veces la longitud (ℓ) del tracto vocal (17.5 cm), los cinco primeros formantes (F_n) tendrán frecuencias de 500 Hz, 1500 Hz, 2500 Hz, 3500 Hz y 4500 Hz.

$$F_n = \frac{(2n-1)c}{4\ell} \quad (1)$$

La noción de vocal neutra tiene importancia teórica, al menos, por dos razones. Primero, las posiciones articulatorias de las vocales que se ubican en los extremos del cuadro

vocálico o vocales periféricas ([i, e, a, o, u]) se han modelado exitosamente uniendo dos o más tubos uniformes, de manera que sus características acústicas se interpretan como desplazamientos o modificaciones con respecto a una posición neutra (Stevens, 1998, p. 142-294). Y segundo, algunas distinciones fonológicas clásicas están asociadas o dependen crucialmente de este concepto. Por ejemplo, en el inventario de rasgos distintivos propuesto por Jakobson, Fant y Halle (1969) se afirma que las vocales tensas se caracterizan por presentar una mayor desviación o alejamiento de la posición neutra del tracto vocal que las vocales laxas. Como ha anotado Catford (1977), esto implica la existencia de una posición neutra de carácter universal y bien definida. Sin embargo, este no es el caso, pues en la comunidad académica no hay acuerdo en cuanto a la altura y posición lingual que debería tener esta vocal, y tampoco es claro si debería transcribirse como una central media [ə] o debería usarse otro símbolo. Por ejemplo, Jakobson, Fant y Halle (1969) usan [æ], Chomsky y Halle (1968, p. 300) consideran que en inglés la articulación neutra del tracto vocal está cercana a [e], mientras que otros autores más recientes como Hayward (2000), siguen las convenciones del inglés británico y prefieren usar [ɜ] para representar una *schwa*. Debido a la vaguedad de este concepto, Catford (1977, p. 206-208) sugiere rechazar la distinción entre vocales tensas y laxas para explicar las vocales.

Como se ve, el concepto de vocal neutra ha sido útil para la teoría fonética porque ha permitido crear modelos cuantitativos de las frecuencias de resonancia del tracto vocal, y le ha permitido a la teoría lingüística explicar las distinciones léxicas y crear categorías fonológicas usando principios acústicos (por ejemplo, Jakobson, Fant y Halle, 1969). Pero, si se revisa con detenimiento, debemos admitir que su definición es vaga e, incluso, contradictoria. Resulta vaga porque, como vimos anteriormente, la interpretación articulatoria y las convenciones varían de autor a autor. También es contradictoria porque, como ya se mencionó, las vocales centrales no tienen una articulación definida, aunque, cuando se habla sobre la vocal *schwa* se alude a los valores de sus formantes (500, 1500 y 2500 Hz, etcétera) y se describe su articulación como si correspondieran a una vocal plena. Además, los valores citados en artículos y libros de texto casi siempre corresponden a un tracto vocal masculino, lo cual significa que el concepto de vocal neutra es androcéntrico (véase la sección 2.2.7).

2.2 Factores que determinan la reducción vocálica

Como ya se ha mencionado, la reducción fonética es un fenómeno que ocurre cuando las vocales de una lengua se centralizan debido a que el sonido es breve, al incremento de la velocidad de elocución y/o al uso de un estilo de habla coloquial. En otras lenguas, la reducción es fonológica, lo cual significa que las vocales centrales aparecen como realización fonética de las vocales átonas y/o son el resultado de la neutralización de las oposiciones fonológicas basadas en la altura lingual. Estos tipos de reducción vocálica no son aislados, sino que representan dos puntos en un continuo, en otras palabras, la reducción fonética es un cambio en marcha que se completa con la reducción fonológica (van Bergen, 1995). Por esta razón, este capítulo busca determinar qué factores motivan la reducción fonética de las vocales y, específicamente, cómo se modifican la duración y los valores de frecuencia de los dos primeros formantes de las vocales. La respuesta nos permitirá identificar similitudes y diferencias con otras lenguas y variedades lingüísticas del español, pero también entender cómo se inicia la neutralización de las vocales átonas o reducción fonológica.

El objetivo principal del capítulo es entender las condiciones en que surge la reducción fonética de las vocales. Un segundo objetivo es identificar los factores que debemos tener en cuenta para estudiar el fenómeno en la variedad de español de Bogotá (Colombia). La exposición trata de seguir un orden cronológico y resume —sin pretender abarcarlo todo— un periodo de cincuenta y seis años. Esto nos permitirá conocer el énfasis o interés de cada década y mostrarle al lector la vigencia de propuestas clásicas como pueden ser las de Stevens y House (1963) y Lindblom (1963).

2.2.1 Duración y entorno consonántico

En el apartado 2.1.2 vimos que la vocal *schwa* se caracteriza por presentar frecuencias de resonancia o formantes con valores de 500, 1500 y 2500 Hz. En el caso de las vocales no centrales o periféricas, se ha demostrado desde hace varias décadas que los dos primeros formantes están relacionados con los tres parámetros articulatorios que definen el timbre vocálico: la altura de la lengua, su posición horizontal y el redondeamiento de los labios. El valor de frecuencia de F1 es más alto cuanto mayor sea la abertura de la boca, por lo tanto, el valor de F1 será bajo en las vocales cerradas, cuando los labios

están redondeados o cuando la articulación es anterior, como en una [i] (Stevens y House, 1955). El valor del F2 es alto en las vocales anteriores y desciende a medida que la lengua se desplaza hacia la parte posterior del tracto vocal (Stevens y House, 1955; Lindblom y Sundberg, 1971).

Los formantes, sin embargo, no son un parámetro acústico estático, sino que cambian de manera continua como un reflejo de la sucesión de gestos articulatorios de la cadena de habla. Como se observa en la Figura 8, al inicio y al final de las vocales, los formantes realizan movimientos de ascenso o descenso según el punto de articulación de la consonante. Estos movimientos, llamados *transiciones*, son fundamentales para la percepción de las consonantes oclusivas y nasales. Esto se sabe desde las décadas de los cincuenta y sesenta, cuando los fonetistas de los Laboratorios Haskins encontraron que, si las transiciones del F2 se dirigían hacia un punto fijo o *locus*, se mejoraba la calidad de las oclusivas del habla sintética. Estos investigadores señalaron que la vocal que sigue a una oclusiva bilabial debía tener un F2 con una transición a los 720 Hz, mientras las consonantes dentales, que tienen una articulación anterior, debían tener una transición a los 1800 Hz (Delattre, Liberman y Cooper, 1955). En estudios posteriores se demostró que las consonantes velares requieren que el F2 y el F3 se aproximen hacia los 3000 Hz en el contexto de vocales anteriores; en el caso de las posteriores redondeadas la barra de explosión es el principal estímulo auditivo de las velares (Harris, Hoffman, Liberman, Delattre y Cooper, 1958; Delattre, 1969b). En cuanto a los valores de F1, los autores consideraban que un valor de 240 Hz arroja buenos resultados en la identificación auditiva de los diferentes puntos de articulación. Para estos autores, el *locus* se ubica 50 ms antes de la explosión de la oclusión y, en ese sentido, se entiende como un concepto abstracto que no se puede identificar en habla no sintética.

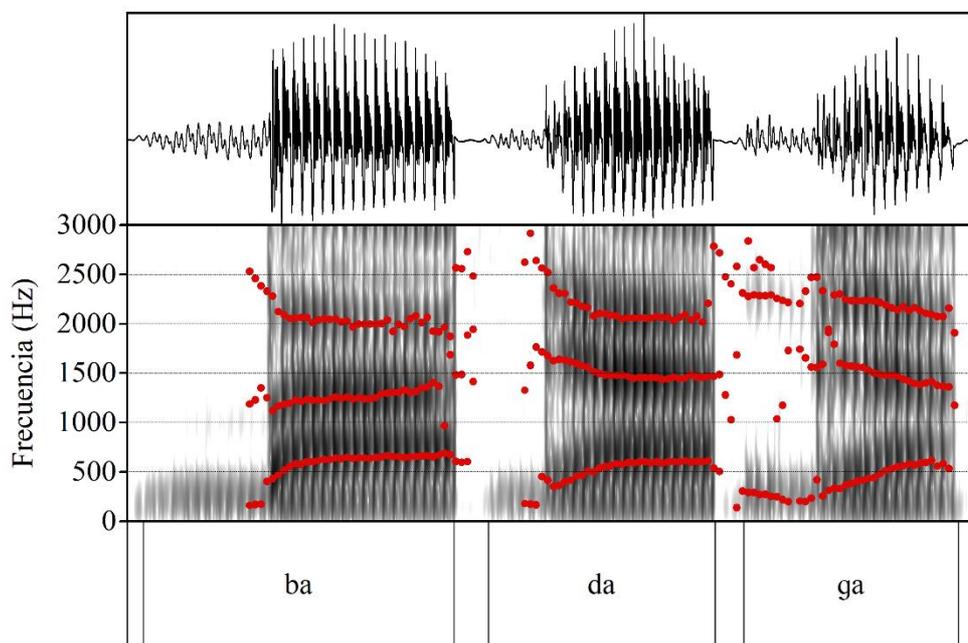


Figura 8. Transiciones en los entornos consonánticos de las sílabas [ba], [da], [ga]

Como se ve en la Figura 8, las transiciones del F2 en las sílabas [ba], [da], [ga] se dirigen hacia un valor común ubicado hacia los 1500 Hz, mientras que el F1 mantiene un valor cercano a los 600 Hz. En el caso de [ba], 1500 Hz corresponde al valor máximo del F2, mientras que en el caso de [ga] y [da] el *locus* concuerda con la parte estable de la vocal. Algunos autores llaman a este valor *objetivo articulatorio* o *target*; en este trabajo, siguiendo a Krull (1989, p. 87), le llamaremos simplemente *núcleo acústico* para diferenciarlo de su dimensión articulatoria y evitar el carácter invariante que el término *target* tiene en los estudios fonológicos. Por consiguiente, definimos núcleo como el valor máximo o mínimo de un formante o, en el caso de contornos estables, el valor del formante en el centro de la vocal.

Llegados a este punto, podemos decir que el tipo más común de reducción fonética ocurre cuando una vocal tiene una duración breve y el núcleo acústico del formante tiene valores de frecuencia cercanos al *locus* consonántico. Joos (1948) caracterizó este tipo de reducción fonética en su clásica monografía: «*the average effect of consonants upon vowels has been found to be a centralizing effect, both vertically and horizontally: all vowels are on the average shifted toward [a] in contact with consonants*». Esta idea fue retomada posteriormente por Stevens y House (1963), quienes compararon los formantes de las vocales del inglés /i, ɪ, ε, æ, u, ʊ, ʌ, ɑ/ en sílabas /hə'cvc/ con los valores de los

formantes de las mismas vocales producidas en el contexto nulo /h_d/. Sus resultados mostraron que el contexto consonántico tiene un efecto poco significativo sobre F1, pero causa un desplazamiento del F2 hacia el centro del espacio vocálico.

Como se ve en la Figura 9, en el trabajo de Stevens y House, las vocales anteriores experimentan un descenso del segundo formante y las vocales posteriores un ascenso, hecho que resulta especialmente notorio en el caso de /u/ y /ʊ/. La figura indica, además, que no todas las consonantes provocan el mismo grado de centralización: las vocales anteriores experimentan mayores cambios bajo la influencia de las consonantes labiales /p, b, f, v/ y coronales /θ, ð, s, z, t, d, ʃ, ʒ/, mientras que las vocales posteriores se modifican por la presencia de consonantes coronales. Los autores demuestran, además, que la duración es una variable determinante de la reducción, pues las vocales largas /i/, /æ/ y /ɑ/ sufren menos cambios que las breves /ɪ/, /ɛ/, /ʌ/ y /ʊ/.

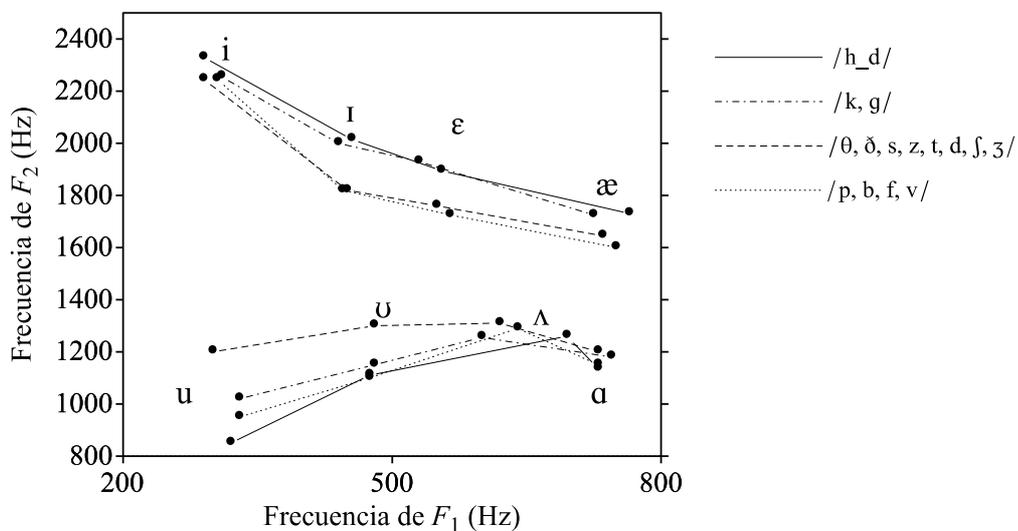


Figura 9. Efecto del contexto consonántico sobre las vocales del inglés. Adaptado de Stevens y House (1963, p. 119).

Para Stevens y House (1963), los efectos del contexto consonántico sobre los valores de frecuencia de F1 y F2 son consecuencia de la superposición de gestos articulatorios. En sílabas CVC, los articuladores deben asumir la configuración asociada a la primera consonante, luego efectuar el gesto vocálico y realizar sin discontinuidad la constricción de la consonante siguiente. De acuerdo con su interpretación, si el punto de articulación de estos sonidos está apartado, la vocal no alcanzará su posición plena (*target*), y esto

ocurrirá con mayor probabilidad si la duración de la vocal es breve. Según los autores, este proceso se conoce como *vowel undershoot* e involucra los mecanismos de *centralización* y *coarticulación*. El primero provoca un desplazamiento de F1 y F2 hacia los valores de una *schwa* (i.e., 500 y 1500 Hz en un hablante masculino), mientras que el segundo mecanismo hace que las frecuencias de los formantes se dirijan hacia el *locus* de las consonantes adyacentes (Stevens y House, 1963, p. 125). Sin embargo, es necesario aclarar que en este estudio la centralización es el resultado de promediar los datos de cada hablante y, por tanto, debe interpretarse tan solo como una tendencia. De hecho, como se ve en la Figura 9, ninguna de las vocales estudiadas alcanza el valor prototípico de frecuencias de los formantes de una *schwa*.

En un trabajo publicado el mismo año, Lindblom (1963) evaluó el papel de la centralización y la coarticulación en el proceso de reducción de las vocales. En su estudio analizó palabras del tipo CVC en las que se analizaban las vocales del sueco /i, ε, y, æ, ø, ɔ, u/ en los entornos consonánticos /b_b/, /d_d/ y /g_g/. Las sílabas se ubicaron en posición inicial y final de frase con el objetivo de variar la duración y el acento e.g., *CVC är det så/ så är det 'CVC*. Para asegurar una velocidad constante, la participante en el estudio oyó una señal periódica y proyectó el patrón rítmico en la lectura de 120 frases. La duración promedio de las vocales se ubicó entre 80 y 300 ms, y los tres primeros formantes sufrieron cambios a medida que se redujo la duración. Sin embargo, el F2 mostró mayor variación, lo cual confirma que la reducción vocálica es altamente dependiente del entorno consonántico.

A partir de la publicación de este trabajo, en las investigaciones sobre fonética acústica el valor de frecuencia del segundo formante en el centro de una vocal ($F2_0$) se explica mediante la siguiente función:

$$F2_0 = \kappa(F2_i - F2_t)e^{-\beta\tau} + F2_t \quad (2)$$

donde $F2_i$ es el valor del formante al inicio de la vocal y varía de acuerdo con la consonante y el timbre vocálico; $F2_t$ es el valor invariante del segundo formante (*target*), y es independiente de la consonante y de la duración; τ es la duración vocálica, mientras que κ y β son constantes que dependen del entorno consonántico. La expresión nos dice que el valor de $F2_0$ depende de la diferencia de frecuencia que hay entre $F2_i$, inicio de la

vocal, y $F2_t$, núcleo acústico de la vocal. En general, $F2_0$ se aleja de su valor invariante ($F2_i$) a medida que la duración vocálica (τ) decrece. Para el caso del primer formante ($F1_0$), el modelo establece un valor crítico de 375 Hz. Esto significa que, si el valor invariante del primer formante ($F1_t$) es menor a 375 Hz, el valor de $F1_0$ será idéntico a $F1_t$. Cuando $F1_t$ es mayor a 375 Hz, $F1_0$ se calcula aplicando la siguiente fórmula:

$$F1_0 = \kappa(375 - F1_t)e^{-\beta\tau} + F1_t \quad (3)$$

La Figura 10 refleja la aplicación de las ecuaciones 2 y 3, conocidas en conjunto como la *ecuación de locus*, al cálculo de los dos primeros formantes de la vocal /ə/ en función del entorno consonántico alveolar para sonidos con una duración entre 100 y 300 ms. Como se puede observar, los formantes se modifican a medida que la duración decrece, pero este patrón es especialmente significativo en el caso del valor de F2. Nótese que, a partir de los 200 ms, $F2_0$ se acerca a $F2_i$, lo cual indica que a menor duración habrá mayor influencia del entorno consonántico inmediato.

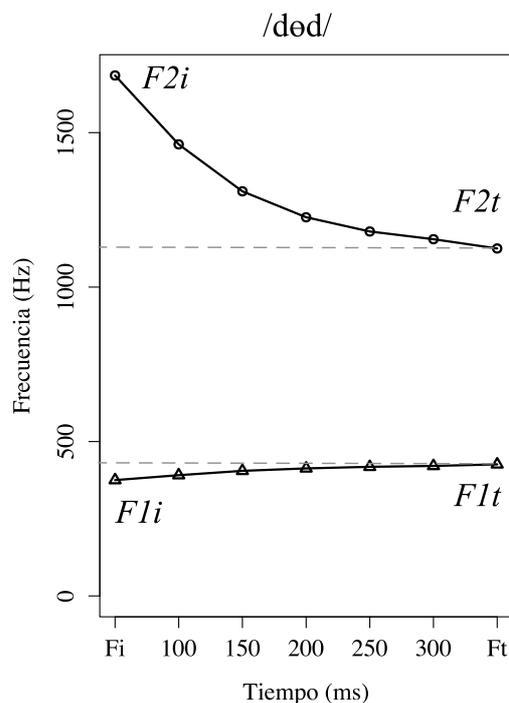


Figura 10. Ecuación de locus aplicada a los valores de los dos primeros formantes de la vocal /ə/ en función del entorno consonántico y la duración. Elaborado con las ecuaciones 1.1 y 1.2 y con los coeficientes κ y β reportados en Lindblom (1963).

Lindblom concluye que la reducción fonética de las vocales es el resultado de dos factores: la *duración* (independientemente de que sea el resultado de la tonicidad de la vocal o de la velocidad de elocución) y la *coarticulación* de las consonantes sobre las vocales, expresada como la distancia entre el valor del formante al inicio de la transición y en el centro de la vocal $|F2_i - F2_t|$. Este modelo confirma las observaciones hechas por Stevens y House (1963), pero difiere en un aspecto importante: la centralización se considera un efecto secundario de la coarticulación y no un proceso independiente. Para Lindblom (1963, p. 1780), «*an articulatory schwa component will be assimilated by a vowel in a dynamic situation only insofar as its immediate context contains schwa elements*». Por ejemplo, la vocal central media en la frase ‘el domingo de ramos’ [el: do' mĩŋgə: rɪms], que ya hemos mencionado antes (Figura 3), debe interpretarse como un efecto de fusión de vocales medias /o/ y /e/ a la consonante aproximante [ð], más no como el producto de una regla de realización fonética que cambie sistemáticamente las sílabas átonas por una *schwa*. En consecuencia, se recomienda usar el término *centralización* para describir el desplazamiento de los formantes en el plano formado por F1 vs. F2, pero no para referirse a un mecanismo de producción de las vocales diferente a la coarticulación (Lindblom, 1963).

A partir de los trabajos pioneros de Lindblom (1963) y Stevens y House (1963), la mayoría de los autores concuerdan en que la reducción vocálica es producto de la influencia del entorno consonántico y la duración sobre las vocales. Sin embargo, la importancia de la duración ha sido puesta en duda en varias publicaciones. Por ejemplo, en una investigación sobre el sueco, Nord (1986) estudió las vocales /a, i, e, ʌ/ en el contexto de consonantes dentales utilizando palabras pronunciadas de manera aislada. Las vocales tónicas no se modificaron por efecto del entorno consonántico en posición inicial y final de palabra, mientras que las vocales átonas se coarticulaban con el contexto consonántico en posición inicial y alcanzaron una posición neutra del tracto vocal en posición final de palabra. Dado que la duración de las vocales tónicas y átonas alcanzaron valores semejantes, Nord (1986) concluye que en sueco la reducción vocálica depende más del entorno consonántico y de la posición de la sílaba que de la duración.

Van Bergem (1994) desarrolla la propuesta de Lindblom (1963) y Stevens y House (1963) y afirma que *schwa* es una vocal que no está asociada con un gesto articulatorio determinado y, por tanto, carece de núcleo acústico. Para demostrar esta hipótesis

propone un análisis de la trayectoria de los formantes que tiene en cuenta el inicio, el centro y el final del formante. El autor analizó la vocal anterior media /e/ del neerlandés en posición átona, la cual suele realizarse como una [ə]. Dado que en esta lengua la reducción no es completamente predecible, el autor construyó palabras con las estructuras /Ce'CV/ y /'VCeC/, donde C corresponde a las consonantes /p, t, k, f, s, χ, m, n, ŋ, r, l, j, v/ y V corresponde a las vocales /i a: u/. La trayectoria de los formantes se modeló con un polinomio de segundo grado:

$$F(t) = C_0 + C_1t + C_2t^2, -1 \leq t \leq 1 \quad (4)$$

donde F es un formante que cambia en función del tiempo (*t*), C_2 es el coeficiente que mide la curvatura, C_0 el centro de la parábola y C_1 el punto final de desplazamiento; el valor temporal de cada movimiento está normalizado dentro del intervalo [-1, 1]. Para el caso del F2, la fórmula nos dice que la trayectoria del formante cambia de acuerdo con el timbre y los valores formánticos asociados a las consonantes adyacentes. En general, las vocales anteriores tienen una forma cóncava descendente \cap y las posteriores una forma cóncava ascendente \cup (van Bergem, 1993). Esto puede apreciarse en la Figura 11, que ilustra la magnitud del movimiento de los formantes en función del entorno consonántico. Por ejemplo, la vocal /ɔ/ se asimila al contexto velar /w_l/ y, por tanto, su formante tiene menos curvatura, mientras que, entre consonantes alveolares /t_n/, la distancia articulatoria y acústica entre la vocal y las consonantes se refleja en una mayor curvatura. En el caso de la vocal anterior /ɪ/, la relación entre curvatura y entorno consonántico es inversa a la vocal posterior /ɔ/.

Ahora bien, nótese en la Figura 11 que el formante de *schwa* [ə] no tiene un núcleo acústico (curvatura) sino que su valor depende del punto de articulación de las consonantes inmediatas. Por esta razón, van Bergem (1994, p.158) afirma que *schwa* es una vocal «sin una posición articulatoria definida que se asimila por completo a su contexto fonológico».

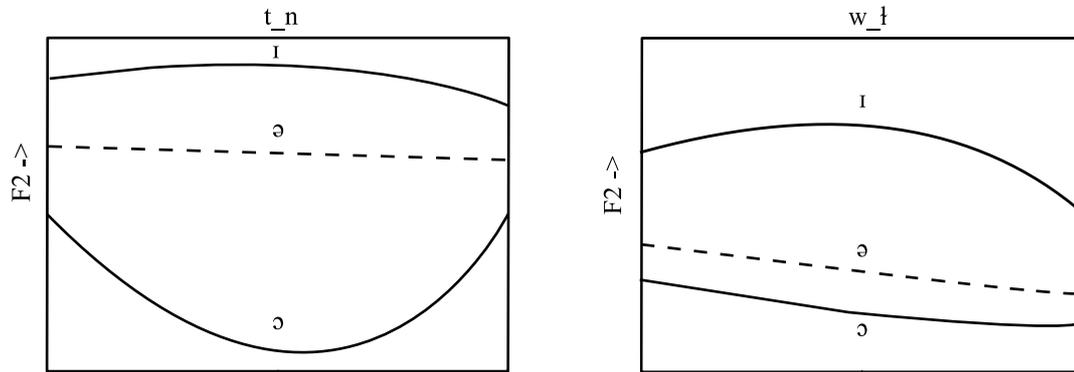


Figura 11. Reducción fonética según van Bergem (1994, p. 57). Entorno consonántico velar [w_t] (abajo) y alveolar [t_n] (arriba).

La hipótesis de van Bergem (1994) ha sido adoptada parcialmente por algunos autores y rechazada por otros. Por ejemplo, en un artículo sobre el inglés británico, Kondo (1994) adopta esta hipótesis para analizar secuencias del tipo [VCəCV]. Sus resultados indican que, durante la producción de *schwa*, F1 se mantiene estable, pero F2 es altamente variable debido a efectos coarticulatorios. Kondo concluye que *schwa* mantiene su identidad como vocal en el primer formante, pero no tiene especificación sobre el grado de anterioridad o posterioridad en el segundo. Flemming (2009) replicó el experimento en un estudio sobre el inglés americano y encontró tendencias similares: F2 presentó valores bajos de frecuencia en el contexto de consonantes posteriores y altos en el contexto de consonantes anteriores. También se encontró que los valores del primer formante de *schwa* son más altos en las vocales altas en secuencias [biCəCi], lo cual muestra una mayor estabilidad que en los valores de F2. No obstante, Flemming (2009) argumenta que la variación de los valores de F2 no significa que la *schwa* del inglés carezca de estabilidad articulatoria. Para afirmar esto, asegura, se debe construir un modelo de la coarticulación que sea aplicable a todos los contextos y a todas las vocales de la lengua.

La idea de la *schwa* [ə] como una vocal sin una posición articulatoria definida es similar a su representación en la versión del *Alfabeto Fonético Internacional* de 1949, pues esta admitía diferentes grados de abertura lingual y no sugería una posición definida. Barry (1998) considera poco viable esta hipótesis. Para probar su punto de vista, analizó la realización de *schwa* en alemán en contextos [o:dədo:], con velocidad de habla lenta, normal y rápida. Sus resultados mostraron que la influencia del contexto depende de la

velocidad de elocución y de la duración: cuando los hablantes producen las frases con velocidad rápida, el valor del F2 depende del contexto, pero, cuando se producen con velocidad lenta, este valor se aproxima a una vocal central media. Para Barry estos resultados significan que si el hablante cuenta con mayor tiempo para pronunciar una *schwa*, habrá mayor posibilidad de que la produzca con un punto de articulación estable (el espectrograma de la Figura 3 ilustra un ejemplo del español).

2.2.2 Velocidad de elocución

La velocidad de elocución o tempo hace referencia al número de fonemas, sílabas o palabras producidas por un hablante en una unidad de tiempo (Llisterri, 1991, p. 81; Gil, 2007, p. 548). Si bien algunos autores han tratado de establecer medidas y parámetros de referencia para caracterizar las diferentes velocidades, el tempo varía enormemente a causa de las preferencias individuales y la influencia de factores como la situación comunicativa, el tipo de discurso, las emociones, el sexo o el dialecto del hablante (Trouvain, 2003). Como ha expresado Lacheret-Dujour (1991, pp. 38), el habla lenta de una persona puede ser el habla rápida de otra.

A pesar de la variabilidad y de las dificultades para caracterizar la velocidad de elocución, se sabe que el habla rápida provoca una disminución de la duración de los elementos fónicos (consonantes, vocales, sílabas y pausas). Según la hipótesis de Lindblom (1963), a menor duración mayor coarticulación de la vocal con el entorno consonántico, pues la postura lingual no se logra en un intervalo breve de tiempo. Como vimos, esto se refleja acústicamente en valores de frecuencia de los formantes cercanos a las transiciones.

Varias investigaciones han estudiado si los cambios en la velocidad de elocución provocan una reorganización de los movimientos articulatorios con correlatos identificables en la señal acústica. En un trabajo sobre el español y el catalán, Nadeu (2014) encontró que, cuando los hablantes de estas lenguas utilizan habla rápida, producen las vocales en menor tiempo, lo cual produce un desplazamiento de los formantes hacia el centro del espacio vocálico. Jaworski (2009) presenta datos sobre diferentes procesos de debilitamiento vocálico y consonántico en español, polaco, ruso e inglés, y encuentra que el espacio acústico de las vocales de estas lenguas se contrae

en velocidad de habla rápida (i.e., hay mayor centralización), si bien el grado de reducción varía de lengua a lengua.

Gay (1977; 1978) analizó la duración y los tres primeros formantes de las vocales del inglés en sílabas CVC pronunciadas a velocidad normal y rápida. La parte estable de las vocales presentó una diferencia de 15 a 30 ms cuando se compararon ambas velocidades. Los valores de los formantes, por el contrario, no se alteraron por el cambio de velocidad ni por la duración. En los datos de Gay, las vocales átonas presentaron una ligera tendencia a la reducción, lo cual es un resultado esperable dado que en esta lengua hay reducción fonológica y, por tanto, las átonas se centralizan de manera sistemática.

En otro estudio sobre el inglés americano, Fourakis (1991) examinó acústicamente las vocales /i, ɪ, ε, æ, α, ɔ, ʊ, u, ʌ/ en los contextos /h_d/ y /b_d/. El autor encontró que el cambio de velocidad lenta a velocidad rápida reduce el 29,5 % de la duración de las sílabas, y las vocales átonas tienen un 34,5 % menos de duración que las tónicas. En este estudio, el entorno consonántico modificó el valor de los formantes vocálicos, pero no se produjeron cambios significativos asociados con la velocidad de elocución, la duración o la tonicidad de la vocal. El espacio vocálico resultó más amplio en las vocales tónicas producidas con velocidad lenta y más estrecho en las átonas producidas con velocidad rápida. Esto indica que en inglés la velocidad de elocución modifica las propiedades acústicas del sistema vocálico en conjunto, pero no los valores de frecuencia de los formantes de las vocales.

Los análisis de otras lenguas muestran resultados semejantes. Por ejemplo, en una serie de experimentos sobre el neerlandés, van Son y Pols (1990, 1992) estudiaron las frecuencias, los movimientos de los formantes y la duración de las vocales /i, y, u, o, a, α, ε/ en un texto de 850 palabras leído por un locutor profesional con velocidad de habla normal y rápida. Las vocales producidas con velocidad de habla rápida tuvieron menor duración que las producidas con velocidad normal, lo cual resultó más sobresaliente en las vocales largas /o, a/. Como en inglés, la velocidad de elocución no afectó los valores de los formantes, y la correlación entre los valores de frecuencia de los formantes y la duración fue marginal. Uno de los resultados más interesantes es que el hablante produjo un primer formante con una frecuencia más alta en velocidad rápida, esto es, realizó una mayor abertura linguomandibular para evitar la reducción vocálica.

Esta adaptación articulatoria a la velocidad de elocución se explica por el entrenamiento del locutor y corrobora la hipótesis de Kuehn y Moll (1976), según la cual los efectos de la velocidad de elocución varían de hablante a hablante.

Como hemos visto, si bien la velocidad de elocución modifica la duración, no necesariamente produce la reducción sistemática de las vocales (Engstrand, 1988; Brzezicha y Kul, 2014; Kul, 2015). Para producir las vocales en velocidad rápida, los hablantes pueden usar estrategias como aumentar el grado de abertura del tracto vocal o, como veremos en la próxima sección, aumentar la velocidad de los movimientos articulatorios (Harrington y Cassidy, 1999). Si bien la duración es un factor determinante de la reducción fonética de las vocales como se afirma en la teoría de Lindblom (1963), se puede contrarrestar aumentando el esfuerzo articulatorio.

2.2.3 El esfuerzo articulatorio

Moon y Lindblom (1994) revisaron el modelo propuesto por Lindblom (1963) con el fin de explicar por qué otras investigaciones no encontraron una relación entre duración y reducción fonética (e.g., Nord, 1986; Fourakis, 1991; van Son y Pols, 1990; van Son y Pols, 1992). Para cumplir con este objetivo, realizaron un estudio en el que controlaron cuidadosamente las variables relacionadas con la reducción (duración y entorno consonántico) en habla hiperarticulada (*clear speech*) y en la lectura normal de palabras aisladas (*citation form*). Los autores analizaron las vocales /i, ɪ, ε, eɪ/ entre una aproximante labiovelar y una lateral velarizada /'w_l/ para obtener desplazamientos amplios de las transiciones de los formantes, mientras que las variaciones de duración se lograron colocando sílabas átonas a la derecha, alargando así las palabras que formaban parte del corpus, como *will*, *willing* y *willingham*.

Los resultados de Moon y Lindblom indican, en primer lugar, que la reducción fonética resultó más sobresaliente en la lectura normal que en habla hiperarticulada. En la mayoría de los casos, una mayor duración significó un incremento de F2 e, inversamente, una duración breve trajo consigo el descenso del valor de frecuencia de este formante. Cuando los valores de la duración alcanzaron valores equivalentes en ambos estilos, las frecuencias de F2 resultaron más altas en habla hiperarticulada, lo cual significa que no es la duración sino el tipo de lectura el determinante de la reducción de las vocales

estudiadas. Para explicar este resultado, Moon y Lindblom introducen, como nuevo parámetro de medición, la velocidad de cambio de F2 (*F2 rate of change*), definida como la diferencia entre el valor del formante al inicio de la transición ($F2_L$) y el valor del formante en su parte estable ($F2_T$) dividida por la duración de la transición:

$$\frac{F2_T - F2_L}{\Delta T} \quad (5)$$

La velocidad de cambio de F2 crece progresivamente en las vocales /ε/, /ɪ/, /i/, esto es, a medida que el formante se aleja del punto de articulación. Esta tendencia fue especialmente notoria en el habla hiperarticulada, lo cual significa que, a mayor velocidad de los articuladores, hay menor influencia del contexto. En este estudio, la velocidad de cambio del F2 se interpreta como un índice del esfuerzo articulatorio. Para validar esta idea, se compara el mecanismo motor de la producción del habla con un sistema mecánico de segundo orden en el que el desplazamiento articulatorio depende de la duración, una fuerza de entrada (*input force*) y la rigidez (*stiffnes*) del sistema. La reducción vocálica puede evitarse incrementando la fuerza de entrada, aumentando la duración o disminuyendo la rigidez (Moon y Lindblom, 1994, p. 52-53). En esta analogía la distancia entre el inicio de la transición y el núcleo acústico (*locus-target distance*) corresponden al desplazamiento, la velocidad de cambio de F2 corresponde a la velocidad de respuesta del sistema, y la duración de la fuerza de entrada tiene relación con la duración vocálica.

El trabajo de Moon y Lindblom (1994) demuestra, entonces, que la reducción se puede compensar aumentando la duración de la vocal o aumentando la velocidad de los movimientos articulatorios involucrados en su producción. Vale aclarar que el grado de compensación articulatoria reportado por estos autores es dependiente del hablante como propusieron Kuehn y Moll (1976). Por ejemplo, uno de los participantes del estudio produjo el habla hiperarticulada y el habla leída con la misma velocidad de cambio de F2, mientras que otros participantes (hablantes W y S) presentaron grandes diferencias (Moon y Lindblom, 1994, p. 50).

En resumen, el trabajo de Moon y Lindblom (1994) sugiere que la reducción fonética depende de al menos tres factores: la *duración*, la *coarticulación* de las consonantes

sobre las vocales (*locus-target distance*) y la velocidad de cambio de F2. Por tanto, los resultados obtenidos por Fourakis (1991), van Son y Pols (1990, 1992), discutidos en la sección anterior, no deben interpretarse como prueba de que no hay relación entre reducción vocálica y duración. En el caso de Fourakis (1991), la ausencia de reducción se debe a que los contextos analizados (/h_d/ y /b_d/) no inducen a movimientos significativos de los formantes. En los análisis de van Son y Pols (1990, 1992) se analizaron datos leídos por un locutor profesional, así que el hablante pudo compensar la reducción aumentando la velocidad de cambio de F2 o realizando una mayor abertura del tracto vocal para evitar la reducción vocálica.

2.2.4 El acento

El acento es una propiedad prosódica de la palabra que permite a los usuarios de algunas lenguas delimitar e identificar la información léxica. Las sílabas tónicas o acentuadas son prominentes y se pronuncian con precisión articuladora porque funcionan como punto de anclaje para el reconocimiento de las palabras (van Bergen, 1993; de Jong, 1995), mientras que las sílabas átonas se producen con menor esfuerzo articulador y, por tanto, favorecen la aparición de procesos de reducción vocálica y consonántica (Barry y Andreeva, 2001).

Debido a su importancia funcional y fonética, algunos autores consideran que el acento es el principal determinante de la reducción vocálica (Harris, 1978; Nord, 1986; Engstrand, 1988). En uno de sus últimos trabajos, el fonetista Pierre Delattre realizó un análisis comparativo de las características articulatorias y acústicas de los sistemas vocálicos de cuatro lenguas (inglés, francés, alemán y español) y concluyó que el timbre de las vocales átonas era «menos distintivo» y «más oscuro» que el de las tónicas. Según Delattre esto se debe a que las vocales átonas se reducen física y fisiológicamente (1969b, p. 323). Sin embargo, advierte que, el grado y la direccionalidad de la reducción varían de lengua a lengua.

La Figura 12, elaborada a partir de los datos reportados por Delattre (1969a, pp. 303-306), permite entender su hipótesis. En inglés las vocales átonas se reducen sistemáticamente, lo cual se manifiesta en un desplazamiento del F2 hacia el centro del espacio acústico y en el descenso del F1 en las vocales no cerradas. En francés, el valor de F1 desciende principalmente en las vocales /a, ε, ɔ/, pero el del segundo formante

varía en menor proporción que en inglés. En alemán y en francés los formantes se desplazan ligeramente hacia el centro del espacio vocálico, lo cual es más sobresaliente en las vocales abiertas /a, ɔ/ y medio abiertas /ɛ, ɔ/. Finalmente, de acuerdo con los datos de Delattre, el español presenta menor diferencia entre tónicas y átonas, pues tan solo se observa una ligera centralización de las vocales medias /e, o/ y de la vocal abierta /a/.

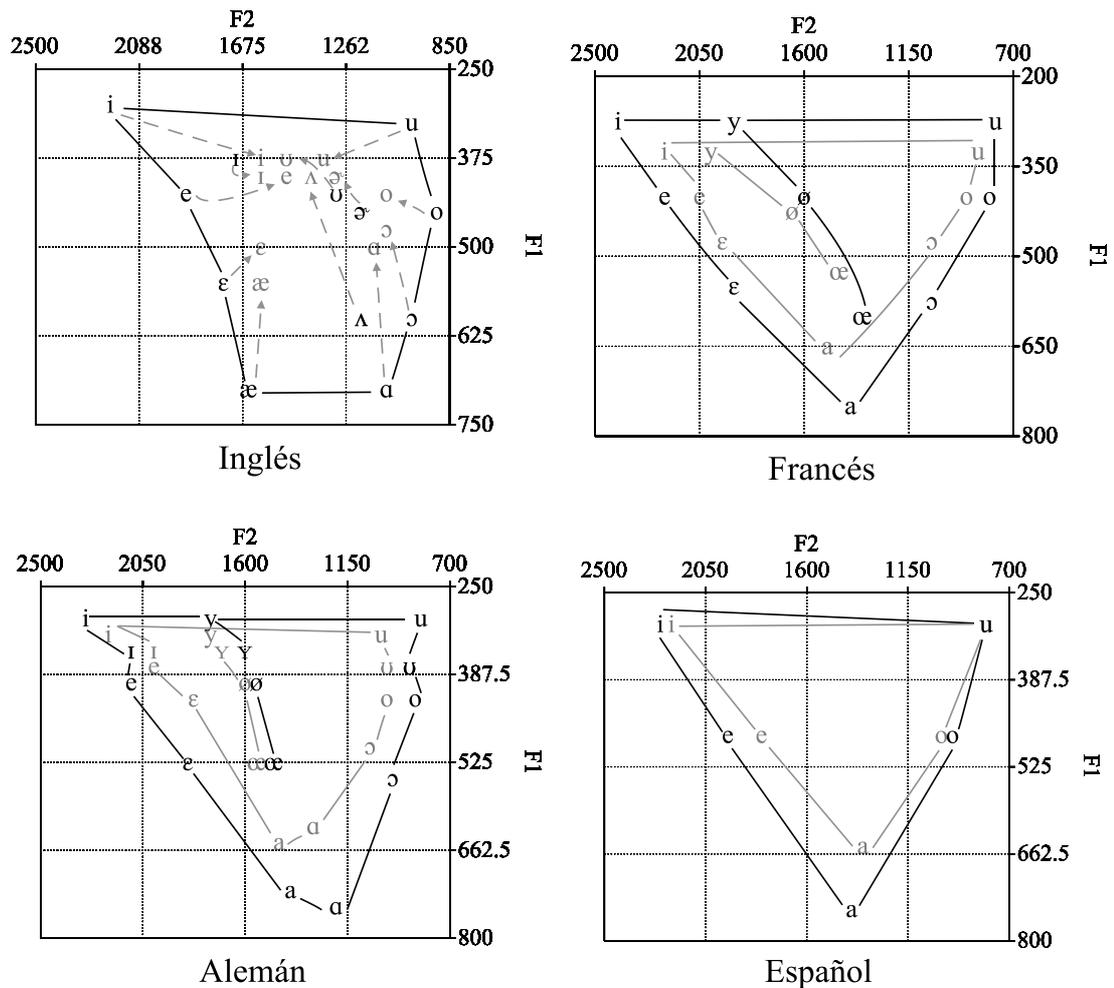


Figura 12. Reducción vocálica en inglés, francés, alemán y español. Elaborado a partir de los datos reportados en Delattre (1969a).

Ahora bien, ¿a qué se deben las diferencias en el grado y direccionalidad de la reducción fonética? Para responder a esta pregunta, es necesario aclarar que Delattre no distingue entre la reducción fonológica (en la cual las vocales átonas se centralizan de manera sistemática) y la reducción fonética (que es la centralización de las vocales debido a la influencia del entorno consonántico, a una breve duración de la vocal o, como veremos más adelante, al estilo de habla). La diferencia principal entre los dos tipos de reducción

está relacionada, sin embargo, con los correlatos fonéticos del acento. En una lengua con reducción fonológica como el inglés, el timbre vocálico es un correlato del acento tan importante como la duración, la frecuencia fundamental (f_0) y la intensidad. En lenguas con reducción fonética como el español, la tonicidad y percepción de una sílaba como acentuada tan solo involucra variaciones de la duración, la f_0 y la intensidad (Llisterri *et al.*, 2003), pero no diferencias en el timbre de las vocales (Cantero, 2002, p. 71).

Entre las cuatro lenguas que se muestran en la Figura 12, el inglés presenta mayor diferencia en la duración de tónicas y átonas, seguida por el alemán y, en menor proporción, por el francés y el español (Mairano *et al.* 2015; Gendrot y Adda-Decker, 2007). Por tanto, las diferencias observadas por Delattre pueden interpretarse como un reflejo de la importancia de la duración para marcar el acento en estas lenguas y no como un aspecto impredecible de la variación vocálica. Como anota Flemming (2005, pp. 6- 7), no es el acento *per se*, sino sus correlatos fonéticos los responsables de la reducción vocálica, de manera que una reducción sistemática y categórica solo puede observarse en lenguas en las cuales la duración y el timbre vocálico son los principales determinantes de la tonicidad de las vocales. Además del inglés, esto sucede en lenguas como el catalán oriental, el portugués de Brasil, el búlgaro, el italiano estándar, los dialectos sureños del ruso, entre otras (Barnes, 2006, pp. 21-22). Por el contrario, en las lenguas en que el timbre vocálico no está asociado con la tonicidad de las vocales, la reducción de las átonas se manifiesta principalmente en una contracción del espacio acústico, como se ha observado en francés, alemán, español, griego, checo y polaco (Barry y Andreeva, 2011).

Dado que en las lenguas con reducción fonética el timbre vocálico no está relacionado con la tonicidad, el grado de reducción varía de hablante a hablante y, además, las tónicas también pueden desplazarse hacia el centro del espacio acústico. Por ejemplo, Nadeu (2014) analizó veinte hablantes de español peninsular y encontró que nueve hablantes centralizaban la vocal tónica /u/ y tenían una pronunciación periférica de la /u/ átona, mientras que los 11 hablantes restantes se comportaban de manera inversa. En tanto, para los veinte hablantes de catalán analizados, la vocal tónica /u/ se realizó como una vocal posterior cerrada y la átona /u/ tuvo una pronunciación centralizada. Por esta razón, la autora concluye que los hablantes del español no reducen las vocales siguiendo un patrón específico, mientras que los hablantes del catalán centralizan de manera consistente las

vocales átonas. En la misma línea de trabajo, Correa (2017) analizó cinco hablantes masculinos en habla narrada y habla leída, y encontró que, en el español hablado en Bogotá (Colombia), la vocal /a/ se puede realizar como una vocal abierta [a], como una vocal media abierta [ɐ] o con un timbre muy cercano a *schwa* [ə]. Las realizaciones centralizadas no estuvieron asociadas exclusivamente con el grado de tonicidad sino también con las preferencias de cada hablante.

Las características acústicas de las vocales tónicas y átonas se modifican en posiciones prosódicamente destacadas, como son los límites de las frases entonativas, las sílabas nucleares o las sílabas focalizadas. En griego, una lengua que tiene un sistema vocálico muy parecido al español, la duración alcanza sus valores más altos cuando la vocal es tónica, focalizada y se encuentra en posición final de frase, y alcanza sus valores mínimos cuando la vocal es átona y no está asociada con una sílaba entonativamente prominente (Nicoladis, 2003). En español se han observado tendencias semejantes, pero, además, se ha encontrado que la posición de la vocal dentro de la palabra y la presencia de una pausa también modifican la duración y los formantes vocálicos. Gendrot, Adda-Decker y Santiago (2019) encontraron que en español y francés las sílabas iniciales de palabra presentan una duración menor y una f_0 más baja que las tónicas finales de palabra. En ambas lenguas, la presencia de una pausa después de la vocal supone valores de frecuencia más altos de los formantes y un espacio acústico periférico. Estos estudios muestran claramente que la articulación de vocales y consonantes se ven reforzados e interactúan con factores prosódicos de mayor nivel. Como sugieren Keating *et al.* (2004, p. 16), la estructura lingüística se proyecta en los detalles fonéticos y los constituyentes prosódicos pueden manifestarse en detalles articulatorios como la articulación de vocales y consonantes.

2.2.5 Propiedades léxicas

Las teorías fonológicas actuales han demostrado que hay una relación estrecha entre el contenido semántico del léxico, su predictibilidad en la lengua y los fenómenos de variación y cambio fonético. Un fenómeno bien documentado es la tendencia de las *palabras funcionales* al debilitamiento de sus fonemas y sílabas. La variabilidad de estas palabras se explica porque no están asociadas al contenido léxico, tienen funciones sintácticas, alta frecuencia de aparición y, por tanto, son unidades predecibles dentro de

la cadena de habla. En cambio, las *palabras de contenido léxico*, «asociadas a conceptos concretos e ideas que pueden ser evocadas» (Bosque, 1990, p.30), se pronuncian con mayor precisión, tienen baja frecuencia de aparición y son menos predecibles (Zipf, 1929; Koopmans-van Beinum y Harder, 1982; Rehor y Pätzold, 1996; Meunier y Espesser, 2011). De acuerdo con Lieberman (1963), los correlatos acústicos de las palabras funcionales son secundarios porque, aun cuando la reducción fonética sea extrema, el oyente puede identificar la información con ayuda del contexto (p.184). En cambio, las palabras con contenido léxico, y su sílaba tónica en particular, deben articularse claramente para que el oyente identifique fácilmente la información léxica (van Bergen, 1993; de Jong, 1995).

Los estudios fonéticos que han comparado palabras funcionales y palabras de contenido léxico concuerdan en que la duración y los formantes vocálicos están determinados por el tipo de palabra o la categoría morfológica a la que pertenece el sonido. Por ejemplo, Engstrand y Krull (1988, pp. 40-42) compararon la secuencia /fɔr/ del sueco cuando hace parte de una preposición y cuando hace parte de un prefijo, y observaron que, en el prefijo, la vocal /ø/ tiene una duración más breve y el segundo formante es más centralizado que cuando la vocal aparece en la preposición. En la misma línea, pero con un resultado contrario, Aguilar, Machuca y Martínez (1991) estudiaron la secuencia /de/ del español como preposición y como sílaba átona de una palabra de contenido. Las autoras encontraron diferencias significativas en la duración de los sonidos /d/ y /e/ asociadas al tipo de palabra, y un F2 más bajo (más centralización) cuando /e/ pertenecía a la preposición. En algunas lenguas, el contenido semántico de la palabra tiene mayor influencia en el timbre vocálico que el acento léxico: en neerlandés, por ejemplo, las palabras funcionales producidas con acento primario presentan mayor reducción fonética que las vocales átonas de las palabras con contenido léxico (van Bergen, 1993, p. 20), un comportamiento opuesto a los datos discutidos en la sección 2.2.4.

El estudio de Gósy y Horváth (2010) resulta de interés para nuestra discusión porque ilustra un caso de reducción fonética de las vocales asociado al tipo de la palabra durante un proceso de cambio lingüístico. Las autoras analizaron un corpus de habla espontánea del húngaro y encontraron que la conjunción *tehát* ('esto es', 'consecuentemente') y el pronombre *ilyen* ('tal', 'como este') también se usan como pausas llenas y marcadores discursivos. De acuerdo con su análisis, la conjunción *tehát* [tat] no presentó variaciones

en la duración asociadas a la función de la palabra, sin embargo, [a] tuvo una realización similar a una vocal abierta en la conjunción y a la de una vocal centralizada en la pausa llena.

En cuanto a la forma *ilyen*, las autoras reportan mayor duración en la pausa llena que en el pronombre; la articulación de la vocal átona /ɛ/ se realiza como una schwa [ə] en la pausa llena y como una vocal anterior medio abierta [ɛ] cuando *ilyen* es parte de un pronombre. Estos resultados muestran que la reducción fonética de las vocales, expresada en la duración y en la centralización de los valores de frecuencia de F1 y/o F2, está relacionada con el tipo de palabra o con su función discursiva.

La fonología basada en el uso (Bybee, 2001) sostiene que la frecuencia de aparición es el principal factor determinante de la reducción fonética y del cambio léxico. Según su hipótesis principal, los sonidos de las lenguas cambian debido a la automatización de las producciones de las palabras y frases más frecuentes (Bybee, 2001, p. 58; Bybee, 2002, pp. 268-269). La pronunciación de unidades que poseen una alta frecuencia de aparición generalmente está acompañada de la fusión de sílabas e, incluso, del debilitamiento de las sílabas tónicas. En el español de Colombia, por ejemplo, el marcador discursivo *o sea* tiene una mayor frecuencia que el subjuntivo *sea*, por esta razón, su realización fonética como marcador tiene una duración más breve y su forma se simplifica modificando la vocal tónica /'se.a/ > ['sa] (Bybee, File-Muriel y Napoleão de Souza, 2016).

La frecuencia de aparición no modifica de manera homogénea las unidades lingüísticas de una categoría gramatical. Pluymaekers, Ernestus y Baayen (2005a) analizaron los efectos de la frecuencia sobre la duración de los prefijos *ge-*, *ont-*, *ver-* y el sufijo *-lijk* del neerlandés, y encontraron que la variación es específica de cada afijo: el prefijo *ge-* tiene menor duración en las palabras con alta frecuencia; particularmente cuando está seguido de varias consonantes. En el caso de *ont-*, la frecuencia de aparición no tiene ningún efecto sobre los datos aportados por los hablantes jóvenes; en el prefijo *ver-* la frecuencia no tiene efectos significativos sobre la duración y, finalmente, la duración del prefijo *-lijk* está asociada a la posición del sufijo dentro de la frase. Los autores consideran que *ont-* y *ver-* no son sensibles a la frecuencia porque contribuyen al

significado léxico, mientras que *ge-* y *-lijk* tienen contenido gramatical y, por tanto, presentan una mayor tendencia a la reducción acústica.

Es importante tener presente que la frecuencia de aparición y el tipo de palabra pueden producir cambios independientes en la duración y en los valores de frecuencia de los formantes vocálicos, de manera que en una lengua un factor puede predominar sobre el otro. Dos estudios ilustran este punto. En el primero, Bell, Breiner, Gregory, y Jurafsky (2009) investigaron la duración vocálica, la frecuencia y el tipo de palabra en un corpus de 13 190 palabras del inglés. Como es de esperar, la duración de las palabras funcionales fue menor a la duración de las palabras con contenido léxico, sin embargo, la frecuencia de aparición no tuvo ningún efecto significativo sobre la duración de las vocales. En un estudio de Meunier y Espesser (2011) sobre la reducción fonética de las vocales francesas en habla conversacional, obtuvieron los mismos resultados: las palabras con contenido léxico se produjeron con mayor duración y tuvieron un valor de F1 más alto (mayor abertura), mientras que las palabras funcionales presentaron menor duración y un valor de F1 más bajo (mayor centralización). Sin embargo, la frecuencia no tuvo ningún efecto sobre la duración y, por tanto, no resultó un factor determinante de la centralización de las vocales.

Dado que los resultados aportados por el análisis de la frecuencia léxica pueden tener limitaciones, algunos investigadores consideran necesario complementarlo con medidas de similitud fonológica y predictibilidad. Desde la perspectiva de Wright (2004), la inteligibilidad de las palabras y la articulación de las vocales está determinada por, al menos, dos factores léxicos: la frecuencia y la similitud fonológica con el entorno léxico o densidad de la vecindad léxica (*neighborhood density*), definida como el número de palabras fonológicamente similares de una lengua (Wright, 2004, p. 77). De esta manera, Wright distingue, siguiendo a Luce y Pisoni (1998), entre «palabras difíciles» (aquellas con alta semejanza con el entorno léxico y baja frecuencia) y «palabras fáciles» (aquellas con poca semejanza con el entorno léxico y alta frecuencia). El autor presenta el análisis de un corpus de palabras aisladas del inglés y muestra que las palabras difíciles se produjeron con vocales hiperarticuladas y un espacio acústico amplio, en tanto, las vocales de las palabras fáciles tuvieron menor campo de dispersión y, por tanto, se redujeron fonéticamente. En otro estudio, Munson y Solomon (2004) ampliaron estas observaciones y lograron determinar que las palabras fáciles tienen un espacio acústico

más reducido tanto en palabras con alta frecuencia como en palabras con baja frecuencia de aparición. Esto sugiere que los efectos de la vecindad léxica (*neighborhood density*) sobre el timbre vocálico son independientes de los efectos de la frecuencia de aparición.

Si bien el tipo de palabra, la frecuencia de aparición y la densidad de la vecindad léxica son factores independientes, permiten hacer una distinción entre unidades léxicas predecibles o redundantes y unidades que tienen información que no puede ser extraída del contexto comunicativo. Como muestran los estudios citados, la reducción de la duración y la centralización de las vocales están relacionadas con la aparición de unidades lingüísticas redundantes, por esta razón, se ha propuesto que hay una relación directa entre la redundancia lingüística y la reducción fonética (van Son, Bolotonova, Lennes y Pols, 2004). Según la hipótesis de Aylett y Turk (2004, 2006), esto ocurre porque los hablantes buscan una comunicación efectiva y, para lograrlo, producen frases cuyos elementos tienen la misma probabilidad de aparición, de manera que la redundancia conduce a la reducción fonética de las unidades léxicas o sus fonemas para nivelar o equilibrar la señal. Según van Son *et al.* (2004, p.1280), la relación entre reducción fonética y redundancia es universal.

2.2.6 El estilo de habla

Por muchas décadas se ha considerado que el procedimiento más apropiado para estudiar los sonidos de una lengua es usar palabras pronunciadas de manera aislada o frases marco del tipo *Say ___ again*, o en español *diga ___ dos veces* (Jones, 1944, pp.127-128; Ladefoged, 2003, p. 9). Este método es completamente válido y conveniente, pues permite al investigador equilibrar el número de casos y controlar factores como el tipo de sílaba, la tonicidad, la velocidad de elocución, la categoría léxica de la palabra, entre otros (Llisterri, 1991, pp. 71-76). A pesar de las ventajas de esta técnica de recolección de datos, su uso comporta, por lo menos, dos problemas serios. Primero, los fenómenos de variación fonética más importantes desde un punto de vista sincrónico y diacrónico se manifiestan en interacciones cotidianas cuando el hablante no está prestando atención a su pronunciación, lo cual se conoce en la sociolingüística variacionista como el *principio vernacular* (Labov, 1972, p. 112). Y segundo, para demostrar la validez de las hipótesis desarrolladas en el laboratorio o a partir del análisis de datos controlados, es

necesario confrontar nuestros resultados con datos más heterogéneos y variables como los que ofrece el habla cotidiana (Rischel, 1992, p. 380).

A principios de la década de los noventa, el interés por explicar la variabilidad e invariabilidad fonéticas, sumado al desarrollo de nuevos modelos de la producción y de percepción del habla (Stevens, 1989; Lindblom, 1990), motivaron estudios teóricos y experimentales sobre la variación fonética condicionada por el estilo de habla. Como ejemplo de este creciente interés, tenemos, entre otros, los trabajos reunidos en los números especiales de las revistas *Speech Communication* (Llisterri y Poch-Olivé, 1992), *Journal of the International Phonetic Association* (Kohler, 2001) y *Journal of Phonetics* (Ernestus y Warner, 2011) y el reciente volumen de Cangemi *et al.* (2018). En la actualidad, la mayoría de los fonetistas es consciente de la necesidad de estudiar habla no leída, y cada vez son más comunes los estudios que comparan varios estilos para descubrir y explicar fenómenos de variación sincrónica, evaluar los modelos de la producción y la percepción del habla (Wagner y Trouvain, 2015) y determinar cómo influyen los métodos de recolección de datos sobre los resultados obtenidos (Niebuhr y Michaud, 2015).

En este trabajo definimos el concepto de estilo de habla como la variación fonética asociada a la interacción entre el emisor, su(s) interlocutor(es), y el grado de planificación del mensaje en una situación comunicativa determinada (Aguilar y Machuca, 1994). Los términos que se usan en la disciplina para designar los diferentes estilos varían enormemente, pero, generalmente, están relacionados con el grado de planificación y atención del hablante a su pronunciación. Jones (1944) distingue solamente entre la lectura de palabras aisladas (*citation form*) y *habla conectada* (unidades iguales o mayores a la frase que pueden ser leídas o no leídas). Otros autores llaman *habla de laboratorio* a las muestras obtenidas de manera controlada (Xu, 2010, p.329) y *habla espontánea* al habla no leída que se produce en tiempo real en respuesta a necesidades inmediatas (Simpson, 2013, p.155). El término habla espontánea es quizá el menos preciso de los mencionados, pues se usa para hacer referencia a monólogos, diálogos estructurados por medio de tareas de mapas, entrevistas (interacción entre investigador y participante), conversaciones entre amigos e, incluso, a diálogos extraídos de diferentes formatos periodísticos (Llisterri, 1992; Warner, 2011).

Ahora bien, en los estudios sobre reducción fonética de las vocales se han comparado los estilos más extremos del continuo: lectura de palabras vs. habla espontánea. La hipótesis de trabajo que subyace a la mayoría de los trabajos es que el espacio acústico, definido como el área acústica que forman los valores de frecuencia de F1 y F2 (Amir y Amir, 2007), se centraliza a menor grado de formalidad del estilo de habla. Koopmans-van Beinum (1983) comparó, por ejemplo, vocales y palabras producidas de manera aislada con habla conversacional en neerlandés, japonés e italiano. La autora encontró que las vocales extraídas de las conversaciones presentaron una menor dispersión que las vocales producidas aisladamente, y las palabras leídas se ubicaron en un espacio intermedio. Uno de los resultados más interesantes de este trabajo es que, independientemente del hablante y de la lengua, el valor de frecuencia de F2 fue tres veces más alto que el de F1, lo cual concuerda con la distribución de los valores formánticos cuando el tracto vocal asume una posición «neutra» (Fant, 1970). Para Koopmans-van Beinum, este resultado demuestra que, a pesar de la variabilidad, los sistemas vocálicos se mantienen equilibrados y organizados alrededor de un centroide.

Harmegnies y Poch Olivé (1992), por su parte, corroboraron los resultados obtenidos por Koopmans-van Beinum (1983) en un estudio sobre el español. En su análisis se compararon vocales extraídas de una lectura de palabras con vocales producidas en una conversación informal. El sistema vocálico del español que, como vimos en la sección anterior, es bastante simétrico y estable, resultó altamente variable en habla espontánea: todas las vocales presentaron una desviación típica alta y se solaparon alrededor del centro del espacio acústico. En un trabajo posterior sobre la variedad de español hablado en la Ciudad de México, Martín-Butragueño (2014) muestra que la centralización también es característica del español de América. En su corpus de entrevistas sociolingüísticas, las vocales medias /e, o/ son más cerradas y se solapan con las vocales cerradas /i, u/. La vocal /a/ es central medio abierta [a], y el efecto es más sobresaliente en el habla masculina (sobre este efecto, véase también la sección 2.3.7). Correa (2017) analizó narraciones y frases leídas por hombres de Bogotá y encontró que el español de Colombia sigue el mismo patrón del español mexicano. Sin embargo, en este trabajo la centralización de las vocales aparece tanto en habla leída como en habla no leída, lo cual hace suponer que en esta variedad el fenómeno está bastante arraigado.

En un estudio más reciente, Romanelli y Menegotto (2018) estudiaron las vocales /a, e, o/ del español de Argentina en dos estilos de habla (lectura de palabras y habla conversacional) teniendo en cuenta la tonicidad y el contexto consonántico. Las conclusiones de este trabajo reafirman las tendencias observadas en otras variedades del español. Por ejemplo, la vocal /a/ tuvo una realización abierta en las tónicas pertenecientes al habla leída y el valor de F1 fue menor en habla conversacional. Al igual que en el español mexicano y colombiano, las vocales medias /e, o/ del español rioplatense tuvieron una realización más cerrada en habla conversacional. Uno de los resultados más interesantes de este trabajo es que el contexto consonántico produjo cambios significativos sobre los valores de frecuencia de los formantes, mostrando así que factores situacionales como el estilo de habla interactúan con factores prosódicos y segmentales.

Koopmans-van Beinum (1983) sostiene que la reducción de las vocales en función del estilo de habla y su organización alrededor de un centroide acústico es una tendencia universal de las lenguas humanas. Si bien es cierto que la mayoría de los datos provienen de lenguas europeas, a medida que aumentan las publicaciones sobre el tema, esta hipótesis cobra mayor fuerza: Nicoladis (2003) y Lengeris (2012) han demostrado que las cinco vocales del griego son más breves y centralizadas en habla conversacional que en habla leída; Gendrot y Adda-Decker (2007) analizaron un amplio corpus de noticias de radio y de televisión de siete lenguas (inglés, francés, alemán, chino mandarín, portugués y español) y encontraron que a menor duración vocálica, mayor grado de centralización vocálica; Barry y Andreeva (2001) analizaron el checo, italiano, griego, bulgaro y polaco, y, en todos los casos el uso de habla espontánea puso de manifiesto la reducción fonética de las vocales; Amir y Amir (2007) usaron varias mediciones para analizar el hebreo, y reportan una gran variabilidad y centralización de las vocales en habla espontánea, especialmente aquellas producidas por hombres; Meunier y Espesser (2011), y Audibert, Fougeron, Gendrot y Adda-Decker (2015) replicaron estos resultados en el análisis del vocales del francés y, recientemente, DiCanio *et al.* (2015) también documentaron la reducción vocálica condicionada por el estilo de habla en el mixteco de Yoloxóchilt.

Tabla 1. Lenguas con reducción fonética de las vocales y tareas de recolección de datos usadas para obtener habla espontánea

Lenguas	Habla espontánea	Fuente
Inglés, portugués, alemán, francés, italiano, mandarín, español	Diálogos de radio y TV	Gendrot y Adda-Decker (2007)
Neerlandés	Conversaciones	Koopmans-van Beinum (1983), Swerts, Kloots, Gillis, y Schutter (2003)
Italiano	Conversaciones	Koopmans-van Beinum (1983) (Calamai, 2003)
Griego	Entrevista	Nicoladis (2003), Lengeris (2012)
Español	Entrevista	Harmegnies y Poch, (1992), Martín-Butragueño (2014), (Romanelli & Menegotto, 2108)
Francés	Diálogos de radio y TV, conversaciones	Audibert et al. (2015), Meunier y Espesser (2011)
Búlgaro, checo, polaco, italiano, griego, ruso	Tarea de mapas	Barry y Andreeva, (2001)
Hebreo	Conversaciones	Amir y Amir (2007)
Mixteco	Monólogos, narraciones	DiCanio et al. (2015)
Japonés	Conversaciones	Koopmans-van Beinum (1983)

Como se ve en la Tabla 1, en la que se recogen los trabajos que tratan sobre la variabilidad fonética y el estilo de habla, el tema que nos ocupa, la reducción fonética de las vocales es bastante común en habla no leída o en tareas de recolección de datos que involucran la interacción entre dos o más locutores (entrevista, conversación y diálogos). El concepto de *habla espontánea* agrupa todas las producciones lingüísticas no planificadas, pero no tiene relación con una situación comunicativa específica como conversar con un amigo o entablar un diálogo en un medio de comunicación. Esta imprecisión terminológica, que también se extiende a términos como *habla natural* o *habla casual*, no invalida ni resta importancia a los resultados obtenidos en estos estudios. Según Lindblom (1990), la producción y percepción del habla son adaptativas, de manera que los hablantes/oyentes ajustan su pronunciación a la situación comunicativa. En situaciones que requieren máxima inteligibilidad, el hablante hiperarticulará los sonidos de su lengua, mientras que en las situaciones en las que el mensaje requiere menos planificación y claridad, será más común la reducción fonética o hipoarticulación. Las tareas de recolección de datos listadas en la Tabla 1 cumplen esta condición, de manera que la reducción fonética es predecible a pesar de las diferencias en el uso del término *habla espontánea*.

2.2.7 Diferencias entre hombres y mujeres

Es bien sabido que existen variaciones individuales en la producción de los sonidos vocálicos que están relacionadas con la anatomía del tracto vocal y con patrones de variación sociolingüística (Ladefoged, 1967; Johnson, Ladefoged y Lindau, 1993). Las variaciones del primer tipo son características orgánicas o intrínsecas que no están bajo el control del hablante, mientras que las segundas, denominadas variaciones extrínsecas, son comportamientos aprendidos por los miembros de una comunidad lingüística (Llisterri, 1991, pp. 107-108). En una categoría intermedia podemos incluir las configuraciones articulatorias y ajustes musculares idiosincrásicos (sección 2.3.3.) que, si bien aportan información social y psicológica del hablante, están limitados por las características anatómicas del tracto vocal (Laver, 1996, pp. 241-247).

En las investigaciones sobre la reducción vocálica, el tipo de variación individual más mencionado es el sexo, un factor que integra información orgánica (e.g., la longitud del tracto vocal y las dimensiones laríngeas), variaciones cuasipermanentes (e.g., el grado de abertura oral individual) y variación sociofonética. Los estudios muestran que las vocales producidas por las mujeres tienen mayor duración que las producidas por los hombres, lo cual parece ser el resultado de usar una velocidad de elocución lenta o diferentes estrategias articulatorias para distinguir entre vocales tónicas y átonas (Ericsson y Ericsson, 2001; Weirich y Simpson, 2014, 2015). En consecuencia, se dice que las mujeres tienen la tendencia a pronunciar los sonidos de su lengua con un menor grado de reducción vocálica que los hombres (Simpson, 2009, pp. 635-636). Estos resultados, provenientes principalmente de trabajos sobre el inglés y el alemán, se han interpretado como una prueba de que las mujeres prefieren usar un estilo de habla claro y cuidadoso (Byrd, 1994; Whiteside, 1996; Simpson, 2009).

Para entender por qué el sexo del hablante condiciona el grado de reducción fonética de las vocales, debemos entender las diferencias anatómicas y sus consecuencias articulatorias y acústicas. De acuerdo con Fant (1966), los valores de los formantes de las vocales producidas por mujeres son, en promedio, 18 % más altos que las producidas por hombres, pero la diferencia no es uniforme, pues cada formante cambia dependiendo de la vocal. En su estudio, basado en datos del inglés y del sueco, los valores de los dos primeros formantes de las vocales posteriores redondeadas [o, ɔ, u] y del primer

formante de la vocal anterior cerrada [y] presentan poca diferencia entre sexos. En cambio, el valor del F1 de las vocales abiertas [æ, a] y el del F2 de las vocales anteriores [i, ɪ, ε] arrojaron una diferencia significativa. Fant, siguiendo a Chiba y Kajiyama (1958), consideraba que las diferencias se explican porque: (1) los hombres tienen una cavidad faríngea más grande que las mujeres y (2) los formantes de las vocales posteriores dependen más del redondeamiento de los labios que de la longitud del tracto vocal. Este planteamiento explica, por ejemplo, por qué en varias lenguas las vocales abiertas tienen un valor de F1 más alto en habla femenina y, por tanto, son menos centralizadas que las producidas por los hombres (Fant, 1975; Amir y Amir, 2007; Simpson, 2009; de Carvalho, 2010; Chládková, Escudero y Boersma, 2011; Nadeu, 2014; DiCanio *et al.*, 2015).

Simpson (2001b, 2002, 2009) afirma que las mujeres no solo tienen los valores de frecuencia del F1 altos sino un espacio acústico más amplio. Esto ocurre, según el autor, porque las dimensiones del tracto vocal femenino permiten alcanzar la postura lingual de las vocales recorriendo menor distancia articulatoria que la que recorren los articuladores de los hombres. Por esta razón, el grado de abertura oral es mayor en habla femenina, y el habla masculina muestra una fuerte tendencia a la reducción de las vocales abiertas. Simpson (2002) evaluó esta hipótesis mediante el análisis de datos acústicos y rayos X del diptongo /ai/ y la secuencia /ei a:/. Los resultados reportados indican que la forma y la trayectoria de los movimientos de la lengua son equivalentes entre hablantes. Sin embargo, la diferencia entre los valores de frecuencia de F1 procedentes de la voz femenina y los de la voz masculina aumentó con la abertura vocálica y, además, los hombres realizaron una mayor constricción y un desplazamiento dorso-palatal más amplios. Como la trayectoria articulatoria fue mayor para los hombres, estos aumentaron la velocidad de los movimientos articulatorios para alcanzar la posición plena de la vocal y compensar así las diferencias articulatorias con las mujeres.

Las diferencias no uniformes entre los valores de frecuencia de los formantes masculinos y femeninos también se han explicado asumiendo una relación directa entre el timbre vocálico y la frecuencia fundamental (f_0). De acuerdo con la formulación original de esta hipótesis (Ryalls y Lieberman, 1982), la inteligibilidad de las vocales disminuye cuando los valores de f_0 se incrementan. Esto implica que la voz femenina, que se caracteriza por tener una f_0 más alta que la masculina, tendrá armónicos dispersos y valores de

formantes con una resolución espectral pobre. En consecuencia, las mujeres usarían un espacio acústico amplio para evitar posibles problemas de identificación asociados con la f_0 (Simpson, 2009, pp. 630-632).

Diehl, Lindblom, Hoemeke y Fahey (1996) realizaron un estudio para determinar la relación entre f_0 y la percepción vocálica. Los autores presentaron a un grupo de oyentes versiones sintetizadas de [i] y [u] incrementando los valores de la f_0 a intervalos de 60 Hz. Como predice la hipótesis, la tasa de identificación disminuyó con el incremento de f_0 , pero la magnitud de este efecto fue mayor para [u] que para [i]. Según ellos, esto se explica porque los hablantes asocian un valor de frecuencia de f_0 alto con un timbre agudo como el de [i]. Si bien los datos analizados en este trabajo son bastante limitados, Diehl *et al.* (1996) llaman la atención sobre la existencia de posibles interacciones entre el valor de la f_0 , la percepción vocálica y el habla femenina.

Simpson, junto con otros autores, también ha evaluado la hipótesis de Ryalls y Lieberman (1982) con datos del alemán, pero sus resultados no son concluyentes. En Simpson y Ericsson (2007), por ejemplo, se analizaron los valores de la f_0 y de los formantes de las vocales /i: e: a: o: u:/ producidas por 17 hombres y 70 mujeres, sin embargo, el análisis estadístico no arrojó una correlación significativa entre el área del espacio acústico y el valor de f_0 . En un estudio posterior, Weirich y Simpson (2013) ampliaron el corpus, incluyeron participantes con diferentes rangos de f_0 y usaron nuevas mediciones, pero, a pesar de los cambios en el diseño experimental, tampoco se observó dicha correlación. Para Weirich y Simpson (2013, p. 2973), las características de las vocales de los hombres y de las mujeres dependen de otros factores orgánicos y sociolingüísticos diferentes al valor de f_0 .

Para resumir, los estudios revisados anteriormente afirman que: (1) las mujeres producen las vocales con mayor duración y menor grado de reducción fonética, por lo tanto, (2) tienen un espacio acústico más amplio que los hombres. El menor «grado» de reducción fonética se refleja articulatoriamente en una mayor abertura vocálica y acústicamente en el incremento del primer formante. Como vimos, estas diferencias se han explicado apelando a las dimensiones del tracto vocal, a la distancia recorrida por los movimientos articulatorios y a la percepción vocálica. Desde un punto de vista sociolingüístico, se ha

afirmado que la poca frecuencia de aparición de reducciones fonéticas en el habla femenina indica que estas prefieren una pronunciación cuidadosa.

Es necesario tener en cuenta que en la investigación fonética predomina el androcentrismo, es decir, trabajos que tratan sobre habla masculina y son llevados a cabo por hombres (Simpson, 2009, p. 622). La principal consecuencia de esta situación es que contamos con pocas comparaciones sistemáticas entre mujeres y hombres en diferentes lenguas. Por esta razón, es necesario recolectar y analizar más datos para determinar si ciertos rasgos como la abertura vocálica y el área del espacio acústico, son el resultado de una diferencia anatómica, o si también contribuyen factores sociolingüísticos como pertenecer a sociedades urbanas, como es el caso de los participantes de todos los estudios discutidos en este apartado.

2.2.8 Variación dialectal y contacto de lenguas

Hemos visto a lo largo de este capítulo que factores fonéticos como la duración y el entorno consonántico desplazan los valores de los dos primeros formantes hacia el centro del espacio vocálico, y que esto ocurre especialmente en sílabas átonas, en palabras funcionales con alta frecuencia de aparición y cuando el hablante usa habla espontánea. Según vimos, los estudios sociofonéticos también sugieren que la reducción fonética de las vocales resulta más común entre hombres que entre mujeres debido a diferencias orgánicas y a una preferencia marcada del habla femenina por la pronunciación clara o hiperarticulada (Byrd, 1994; Whiteside, 1996; Simpson, 2009). Adicionalmente, la variación diatópica y las situaciones de contacto lingüístico son factores fundamentales para entender la reducción vocálica. Una de las principales enseñanzas de los estudios sobre la variación dialectal es que la reducción fonológica y la reducción fonética de las vocales pueden coexistir en una misma lengua. Sirva como ejemplo el búlgaro, una lengua con seis fonemas vocálicos /i, e, a, â, o, u/, de los cuales /a/ y /â/ se realizan como una *schwa* [ə] en el habla espontánea de los dialectos occidentales. Sin embargo, en los dialectos orientales, la reducción vocálica ha avanzado, de manera que hay una neutralización sistemática del grado de abertura de las átonas: las anteriores /i, e/ se realizan [i], las posteriores /o, u/ convergen en [u] y las vocales /a/ y /â/ se resuelven a favor de [ə] (Barnes, 2006, pp. 32-35).

En otras lenguas, la mayoría de los hablantes reducen fonéticamente las vocales como parte de un fenómeno de variación sincrónica, debido, seguramente, a que el proceso se ha logrado incorporar a la norma de pronunciación sin mayores consecuencias fonológicas. El neerlandés ilustra este caso, pues las sílabas átonas se realizan como [ə] en la variedad normativa de Holanda (Gussenhoven, 1992), en la zona fronteriza con Bélgica y Francia (Heijmans y Gussenhoven, 1998; Peters, 2010) y en el neerlandés belga estándar (Verhoeven, 2005). En alemán, por el contrario, la reducción tiene una distribución diatópica limitada: las variedades estándar del norte de Alemania y de Suiza usan [ə] en las sílabas átonas (Kohler, 1999; Fleischer y Schmid, 2006), pero en la variedad estándar de Austria se realiza como una vocal central medio-abierta [ɐ] o una vocal plena (Piroth y Skupinski, 2011; Moosmüller, Schmid y Brandstätter, 2015).

El español tiene un sistema de cinco vocales /i, e, a, o, u/ que, como se sabe, es el más común en las lenguas del mundo (Crothers, 1978; Maddieson, 1984; Schwartz, Boë, Vallèe y Abry, 1997). De acuerdo con lo discutido en la sección 2.2.4, el timbre vocálico no está asociado con la tonicidad como en las lenguas germánicas, lo cual significa que no hay una relación marcada con la realización fonética del acento léxico (Delattre, 1969a; Quilis y Esgueva, 1983; Cantero, 2002; Hualde y Colina, 2014). Sin embargo, desde hace varias décadas se reconoce que, cuando los hablantes usan un estilo de habla coloquial o espontáneo, las átonas pueden sufrir cierto grado de reducción. Quizá fue Navarro Tomás (1918) el primero en llamar la atención sobre este fenómeno:

El timbre de nuestras vocales inacentuadas depende, especialmente, del esmero o descuido con que se habla [...] pero en lenguaje familiar y corriente reducen y relajan la articulación obedeciendo a diversas influencias y presentando matices diferentes. Toda vocal débil es, pues, una vocal relajada cuya imprecisión aumenta a medida que disminuye su intensidad; pero dentro de su relajación, la vocal débil española mantiene siempre su timbre en una relación suficientemente clara con el tipo normal a que corresponde, sin llegar, por consiguiente, al caso de las vocales indistintas que aparecen en otros idiomas (p.38).

Según Navarro Tomás, las vocales medias /e, o, a/ son más cerradas y las vocales /i, u/ son más abiertas cuando se producen con relajamiento articulatorio. Para este autor, todas las vocales del sistema pueden reducirse, pero, advierte, nunca hasta el punto de neutralizarse o perder su timbre. Los trabajos posteriores confirman que el sistema

vocálico del español permite cierto grado de variación: las tónicas son más abiertas que las átonas, y esta diferencia es particularmente significativa en el caso de las vocales medias cerradas /e/ y /o/ (Quilis y Esgueva, 1983, p. 243; Álvarez González, 1989; Martínez Celdrán y Fernandez Planas, 2007, p. 190; Albalá *et al.*, 2008, p.4; Torreira y Ernestus, 2011, p. 347). En habla espontánea, las realizaciones se solapan en el centro del espacio acústico, sin embargo, los estudios realizados por Harmegnies, Poch-Olivé y otros autores (por ejemplo, Harmegnies y Poch-Olivé, 1992 y Poch-Olivé, Huet, y Harmegnies, 2001) muestran que la vocal central media /e/ es la única que logra acercarse al timbre de una *schwa*.

Por su parte, el vocalismo del español de América presenta dos tipos de reducción fonética: la centralización de los formantes y el ensordecimiento de las átonas, que explicaremos más adelante. Varias investigaciones coinciden en que la centralización se manifiesta con un valor de frecuencia de F1 bajo o una articulación más cerrada en la vocal central abierta /a/. Los datos acústicos y las transcripciones reportadas sugieren que la vocal resultante se ubica entre una vocal central medio abierta [a̠] y una central media [ə]. Como se ve en la Figura 13, esta realización ha sido reportada en el español de Chile (Sadowsky, 2012), Perú (Chládkova, Escudero y Boersma, 2011; Chládková y Escudero, 2012), El Salvador (Azcúnaga, 2010), Nicaragua, (Rosales, 2010), México (Poch-Olivé, Harmegnies y Martín Butragueño, 2008; Martín-Butragueño, 2014) y Colombia (Correa, 2017).

Las vocales cerradas /i/ y /u/ son más variables en el español americano que en el peninsular. En algunas variedades como en la hablada en Bogotá, la vocal anterior /i/ es resistente a la reducción y tiene una realización fonética similar al español peninsular (Correa, 2017), mientras que /u/ normalmente se centraliza. En otras variedades como las habladas en México, Chile, Perú y Argentina, tanto /i/ como /u/ se centralizan debido a cambios en los valores de frecuencia del segundo formante (Godínez, 1978; Sadowsky, 2012). Consideramos que la direccionalidad de la centralización en español se debe indicar transcribiendo [ē] (equivalente a [ə] y [ẽ]) cuando procede de una vocal semicerrada, y [a̠] (equivalente a [ɐ] o [ə]) cuando procede de una realización de la vocal central abierta /a/. Para las vocales cerradas /i, u/ sugerimos usar [ĩ, ũ] (equivalentes a [i] y [u]) y para la vocal posterior media, [ö] (equivalente a [õ]).

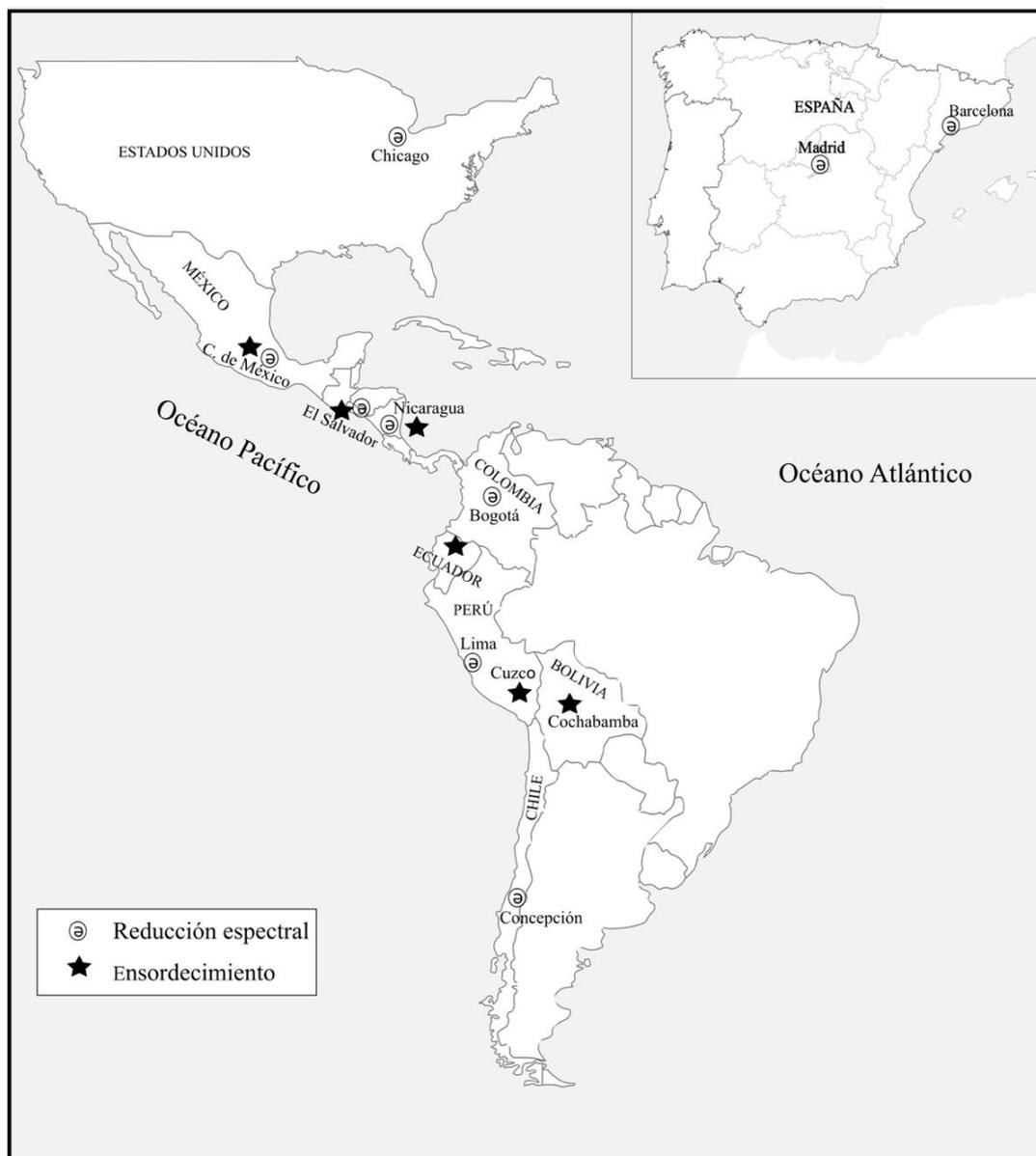


Figura 13. Distribución geográfica de la reducción fonética de las vocales en español

El concepto de reducción vocálica de las átonas (*unstressed vowel reduction*) ha sido usado en los estudios sobre el español de América con un sentido bien diferente al adoptado aquí (sección 2.1). No hace referencia a la neutralización y a la centralización de las vocales átonas, sino a un proceso de debilitamiento fonético que involucra el ensordecimiento o la elisión de las vocales átonas ante /s/ p. ej. *antes* ['ãntɛs] y *estos* ['estɔs]. En las variedades de español hablado en México y en Ecuador, el ensordecimiento ocurre principalmente en las vocales medias /e/ y /o/ (Lope Blanch, 1963; Lipski, 1990). Los datos reportados por Delforge (2008) para el español de Cuzco (Perú) presentan un patrón de variación semejante, sin embargo, las vocales no solo se

debilitan antes de /s/, sino en el contexto de consonantes sordas en general. Como se ilustra en el mapa de la Figura 13, este rasgo se ha documentado en otras zonas del Perú (Sessarego, 2012a), en Bolivia (Sessarego, 2012b), y en países centroamericanos como El Salvador (Azcúnaga, 2010) y Nicaragua (Rosales, 2010). Evidentemente, se trata de una reducción fonética porque modifica la duración y el timbre vocálico, pero concordamos con Delforge (2008) en que resulta más apropiado llamarle *ensordecimiento*, y reservamos el término *reducción vocálica* para hacer alusión a la centralización de los dos primeros formantes de las vocales.

Otra situación bastante común pero menos estudiada es la influencia del contacto lingüístico y el bilingüismo sobre el vocalismo del español. En el caso de las lenguas que comparten el inventario o algunas vocales, interesa saber si estas ocupan el mismo espacio acústico, si las características de una lengua predominan sobre la otra o, por el contrario, si los sistemas se mantienen diferenciados fonéticamente. El quechua, por ejemplo, tiene tres vocales /i, a, u/ y dos vocales semicerradas [e, o] como alófonos de las vocales cerradas. Si bien los dos sistemas son bastante centralizados y tienen una distribución similar, el estudio de Holliday y Martin (2017), realizado en Cochabamba (Bolivia), muestra que los bilingües mantienen las dos lenguas claramente diferenciadas: en quechua [i, e] son más abiertas (valores de frecuencia más altos para F1) y las vocales posteriores [o, u] son más centralizadas (valores más altos de F2) que en español. En Perú, los bilingües producen las vocales con el mismo patrón, sin embargo, como ha mostrado O'Rourke (2010), las diferencias entre dialectos predominan sobre el grado de bilingüismo: los bilingües de Cuzco realizan las vocales [i, e, a, o, u] con una pronunciación adelantada y abierta, pero las vocales realizadas por los hablantes monolingües en español originarios de Lima, presentan un espacio acústico más estrecho y centralizado.

Machuca y Poch-Olivé (2016, p. 153) estudiaron cómo influye el porcentaje de uso de una lengua en las características acústicas de las vocales. Las autoras analizaron dos informantes de dominancia catalana, dos informantes que usan el español en la mayoría de sus interacciones diarias, y dos informantes que usan las dos lenguas con la misma frecuencia. En general, los hablantes en los que predomina el uso del español produjeron formantes con mayor variación, desorganización del espacio acústico y desviaciones típicas más altas, mientras que los hablantes de dominancia catalana produjeron vocales

con mayor organización y estructura. Así, las vocales /e/ y /o/, pronunciadas por los hablantes de dominancia española, presentaron realizaciones dispersas en el espacio acústico, mientras que los hablantes de dominancia catalana pronunciaron /e/ y /o/ en un área definida y distinta de la que ocupan las vocales /ɛ/ y /ɔ/ del catalán (Machuca y Poch-Olivé, 2016, pp. 167-168). Los hablantes que hacen un uso similar de las dos lenguas presentaron un comportamiento variable y más difícil de generalizar.

Los resultados de las investigaciones sobre el contacto catalán-castellano y sobre el contacto quechua-español demuestran de manera consistente que la pronunciación de las vocales puede estar determinada por el grado de bilingüismo y por el porcentaje de uso de las lenguas implicadas. En otras situaciones, el grado de bilingüismo es más difícil de establecer, y la reducción integra tendencias fonéticas de ambos sistemas vocálicos. Esto ocurre en el español como lengua de herencia, la variedad hablada por aquellos hispanohablantes criados en los Estados Unidos que usan el inglés en la mayoría de sus interacciones diarias y han tenido acceso limitado o tardío a la educación en español. Ronquest (2013) encontró que los hablantes de origen mexicano y puertorriqueño del área de Chicago (Figura 13) producen las vocales átonas con un valor de F1 más bajo que las tónicas, pero esta diferencia es mayor a la de los hablantes monolingües de español. En el caso del F2, los valores de los formantes se desplazaron hacia el centro del espacio vocálico, pero no se centralizaron con la misma sistematicidad que en inglés (Ronquest, 2013, p. 165). En otras palabras, el español de herencia conserva las tendencias de reducción de ambas lenguas, pero, al mismo tiempo, tiene un patrón de variación característico.

Una posible explicación de los escenarios descritos es que cada variedad lingüística (superdialecto o dialecto) tiene su propio conjunto de hábitos articulatorios o *base de articulación* para producir las vocales. Se dice que dos lenguas tienen diferente base de articulación cuando sus hablantes producen un mismo fonema con diferencias sistemáticas en la postura de los articuladores (Gil, 2007, pp. 190-192). Así, las vocales /i/, /e/, /o/ y /u/ del inglés americano tienen una postura lingual más adelantada que en español (Bladlow, 1995), y el español americano tiene una articulación más cerrada (valores más bajos de F1) que el español peninsular (Poch-Olivé, Harmegnies y Martín Butragueño, 2008; Sadowsky, 2012; Correa, 2017). El contacto y el bilingüismo quechua-español favorecen la utilización de una misma base de articulación en el español

peruano (O'Rourke, 2010), pero esta se mantiene diferenciada en el caso del español y el quechua bolivianos (Holliday y Martin, 2017).

Para finalizar, permítaseme decir que, a pesar de la cantidad de estudios sobre el vocalismo del español, las investigaciones sobre la reducción fonética están en sus inicios. Sabemos que el habla espontánea lleva al relajamiento y a la centralización, pero sabemos poco sobre la influencia de factores como el entorno consonántico, la velocidad de elocución, la tonicidad, el sexo del hablante y el contenido léxico. Una excepción a este estado de la cuestión son los trabajos de Chládková y Escudero (2012), Sadowsky (2012) y Nadeu (2014), quienes analizan la variabilidad vocálica en diferentes condiciones experimentales y aportan análisis estadísticos llevados a cabo sobre datos acústicos que permiten la comparación interdialectal. En cualquier caso, ninguno de los estudios publicados ha tenido en cuenta todos los factores fonéticos, lingüísticos y extralingüísticos analizados en este capítulo.

2.3 Resumen

En este capítulo se examinaron los factores que determinan la reducción fonética de las vocales. La hipótesis más difundida y desarrollada sostiene que este fenómeno es causado por la duración y el entorno consonántico: cuando la vocal es breve, la lengua no alcanza su objetivo articulatorio o *target* y, por este motivo, los formantes resultantes tienen valores de frecuencia que dependen del punto de articulación de las consonantes adyacentes (Stevens y House, 1963; Lindblom, 1963; van Bergen, 1994; Kondo, 1994; Flemming, 2009; Mooshammer y Geng, 2008). Desde esta perspectiva, la reducción fonética se debe a la influencia de las consonantes sobre las vocales (*vowel undershoot*). Este efecto puede ser contrarrestado aumentando el esfuerzo articulatorio (Moon y Lindblom, 1994) o incrementando la abertura linguomandibular (van Son y Pols, 1990; van Son y Pols, 1992).

Otros autores como Delattre (1969a) argumentan que el acento léxico es la causa principal de la reducción de las vocales y que la duración es un factor secundario. Según esta hipótesis, las vocales tónicas se producen como vocales plenas para destacar semánticamente las palabras, mientras que las átonas tienen un timbre «menos distintivo» y «oscuro». Para el autor el grado de reducción depende de cada lengua, de

tal suerte que es extrema en lenguas como el inglés y apenas se manifiesta en lenguas como el español. Como vimos, el principal problema de esta propuesta podría ser que posiblemente se está confundiendo la reducción fonológica y la fonética.

Para aclarar este punto, debemos decir que en las lenguas con reducción fonológica la duración y el timbre vocálicos son correlatos fonéticos del acento, y por esta razón las oposiciones fonológicas formadas por las átonas se neutralizan de manera sistemática y categórica. Esta neutralización elimina las distinciones basadas en el grado de abertura lingual, lo cual se manifiesta acústicamente en un descenso de los valores de frecuencia de F1 (Flemming, 2005, p. 20). Dado que la duración es breve, las vocales se coarticulan por efecto del entorno consonántico (*vowel undershoot*), lo cual acerca, a su vez, el F2 a los valores de frecuencia del inicio de la transición o *locus*. Aunque este comportamiento es bastante estable en las lenguas con reducción fonológica, las vocales involucradas en el proceso y los timbres vocálicos resultantes varían de lengua a lengua (Figura 1). En lenguas como el español, en que el timbre vocálico no es un correlato fonético del acento, las vocales átonas también tienden a ser más breves que las tónicas, y los valores de los formantes se desplazan hacia el centro del espacio vocálico (Barry y Andreeva, 2001). En todo caso, estos cambios son asistemáticos, no siguen un patrón específico e, incluso, pueden modificar el timbre de las vocales tónicas (Poch, Huet y Harmegnies, 2001; Nadeu, 2014). Adicionalmente, el grado de reducción fonética varía entre hablantes e, incluso, entre los dialectos de la misma lengua (sección 2.3.8).

Existen otros factores lingüísticos que provocan la reducción fonética de las vocales, entre los que se destaca el contenido semántico de las palabras y su predictibilidad en la cadena de habla. En concreto, las palabras funcionales tienden al debilitamiento fonético de sus fonemas y/o sílabas, mientras que los mismos fonemas y sílabas mantienen su cualidad fonética cuando pertenecen a palabras con contenido léxico. La fonología basada en el uso (Bybee, 2001) argumenta que, más que la categoría de las palabras es su frecuencia de aparición, de tal suerte que las unidades que tienen una alta frecuencia están acompañadas de fenómenos de reducción fonética debido a que su pronunciación se somete a un proceso permanente de automatización. Ahora bien, varios de los estudios discutidos muestran que la frecuencia de aparición y el tipo de palabra producen cambios independientes en la duración y en los valores de frecuencia de los formantes, de manera

que, dependiendo de las características de la lengua, un factor puede predominar sobre el otro (Bell *et al.*, 2009; Meunier y Espesser, 2011).

Wright (2004) y Munson y Solomon (2004) aseguran que la similitud fonológica y la predictibilidad permiten explicar mejor la relación entre los factores léxicos y la reducción fonética. Así, las palabras que se asemejan al entorno léxico y tienen una baja frecuencia de aparición se pronuncian con vocales plenas, mientras que las palabras que tienen alta frecuencia de aparición y fonemas diferentes a las palabras vecinas tienen mayor probabilidad de reducirse fonéticamente. En todo caso, las palabras funcionales, las palabras con alta frecuencia de aparición y aquellas que tienen menor similitud fonética con el entorno léxico son unidades lingüísticas predecibles o redundantes en la cadena del habla, y su debilitamiento fonético se explica porque contienen información que puede ser extraída del contexto lingüístico y comunicativo (Lieberman, 1963, p. 184; Aylett y Turk, 2004, 2006). Dado que esta relación entre redundancia y reducción fonética es independiente de la fonología de la lengua, se considera un universal lingüístico (Van Son *et al.*, 2004).

El estilo de habla es otro factor que determina la reducción fonética y también puede considerarse un universal lingüístico. En este caso, la reducción fonética se entiende como el desplazamiento de los valores de los formantes hacia el centro del espacio acústico y/o como un estrechamiento del área acústica formada por todas las vocales del sistema. Los estudios sobre diferentes lenguas permiten concluir que las vocales producidas en habla espontánea tienen una duración más breve y un mayor grado de centralización que las vocales producidas de manera aislada o en habla leída. De acuerdo con Lindblom (1990), esto se debe a que la producción y la percepción del habla son adaptativas y, por este motivo, en las situaciones comunicativas que requieren mayor inteligibilidad de la información, el hablante hiperarticula los sonidos de su lengua, mientras que, en aquellos contextos que requieren menor planificación, es común la hipoarticulación de las vocales y de las consonantes.

La variación individual y, concretamente, las diferencias anatómicas y sociofonéticas entre hombres y mujeres, también determinan la duración y el grado de centralización de las vocales. Los estudios revisados muestran una tendencia acusada de las mujeres a producir vocales más largas que los hombres, lo cual lleva a que las tónicas y átonas

tengan una menor diferencia temporal (Ericsson y Ericsson, 2001; Weirich y Simpson, 2014). Una consecuencia de esto es que las vocales femeninas tienen menor influencia del entorno consonántico y, en general, el espacio vocálico es más amplio que el masculino. Este patrón se ha explicado apelando a las diferencias en la longitud del tracto vocal y, desde un punto de vista sociolingüístico, a la preferencia de las mujeres por una pronunciación cuidadosa (Byrd, 1994; Whiteside, 1996; Simpson, 2009).

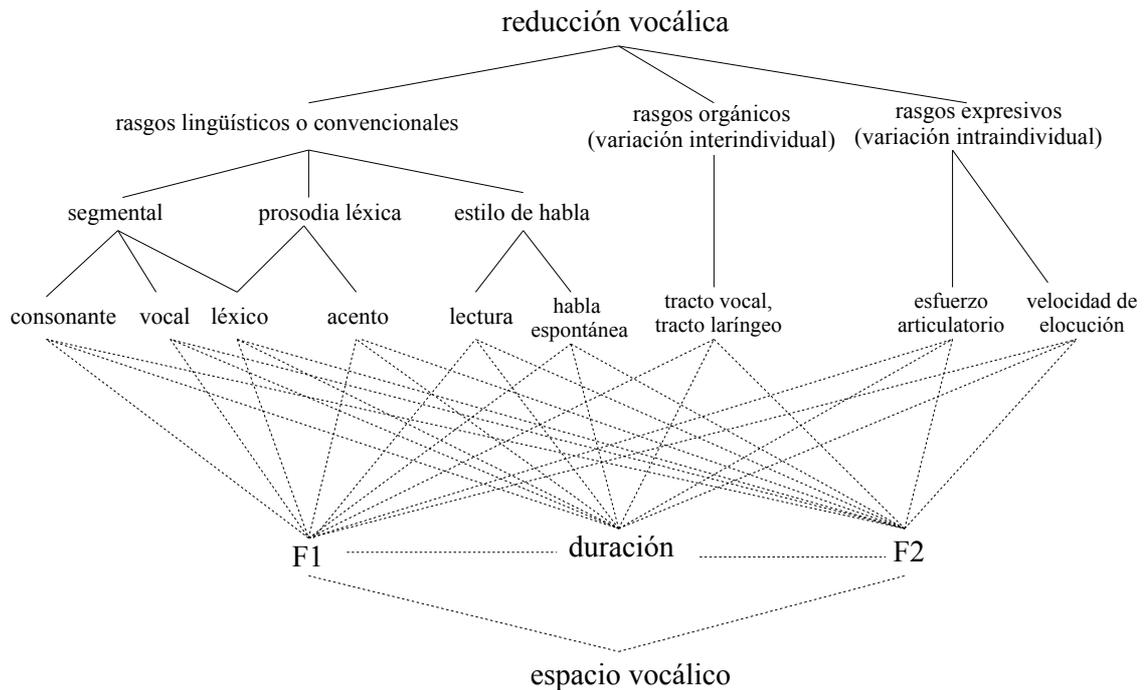


Figura 14. Factores que determinan la reducción fonética de las vocales.

La Figura 14 resume los factores que determinan la reducción fonética de las vocales. Siguiendo a Traunmüller (2005), los hemos agrupado según el tipo de información que transmite la señal de habla. En primer lugar, tenemos los factores lingüísticos de tipo segmental, esto es, el entorno consonántico, que ya hemos discutido previamente, y el tipo de vocal. Como habrá advertido el lector a lo largo del capítulo, la vocal abierta [a] y las vocales medias [e, o] son las más susceptibles a la reducción fonética, mientras que las vocales cerradas como [i, u] parecen ser resistentes a la centralización y a la coarticulación con los sonidos adyacentes (Fant, 1975; Amir y Amir, 2007; Simpson, 2009; de Carvalho, 2010; Chládkova, Escudero y Boersma, 2011; Nadeu, 2014; DiCano *et al.*, 2015, Correa, 2017, entre otros). Así, pues, desde el punto de vista del cambio fónico, la reducción fonética se inicia con el cierre de la vocal más abierta del sistema, y

posteriormente, las demás vocales pierden las oposiciones basadas en el grado de abertura lingual (sobre este punto véase Flemming, 2005 y Barnes, 2006).

Otro constituyente de los factores lingüísticos es el acento pues, como vimos, en la mayoría de lenguas las vocales átonas admiten diferentes grados de reducción vocálica y consonántica. En un punto intermedio entre la fonética segmental y la prosodia léxica ubicamos el tipo de palabra y demás propiedades léxicas. Aquí queremos resaltar el hecho de que las palabras funcionales, que generalmente son átonas, tienen una alta frecuencia de aparición, son precepciones del contexto lingüístico y, por tanto, es común el relajamiento o debilitamiento de sus fonemas.

El estilo de habla es un factor lingüístico en tanto la lingüística tradicional ha determinado la norma de pronunciación de los fonemas de una lengua o dialecto con base en el análisis de palabras aisladas (*citation form*). En este sentido, los fenómenos de variación sincrónica, como la reducción fonética, se interpretan en la mayoría de los estudios como un alejamiento de las formas canónicas y, por tanto, de una pronunciación no correcta desde el punto de vista normativo. Dentro de los rasgos orgánicos, se encuentran las diferencias en las dimensiones del tracto vocal y de la cavidad laríngea, que permiten explicar las diferencias de pronunciación entre hombres y mujeres. Finalmente, dentro de los rasgos expresivos, encontramos el esfuerzo articulatorio, la velocidad de elocución, que pueden contrarrestar o incrementar el grado de reducción fonética. Nótese, finalmente, que todos los factores discutidos en este capítulo pueden modificar los valores de la duración, los valores de frecuencia de F1 o de F2, los cuales, a su vez, determinan el grado de expansión o contracción del espacio vocálico.

2.4 Preguntas y objetivos de la tesis

De acuerdo con la bibliografía revisada, los factores que determinan la reducción fonética de las vocales, según quedan recogidos en la Figura 14, son la duración, el entorno consonántico, el timbre, la tonicidad, la velocidad de elocución, el estilo de habla, la categoría morfológica de las palabras (palabra de función vs. palabra con contenido léxico), la posición de la vocal dentro del grupo fónico y el sexo del hablante. Como vimos, la mayoría de los estudios publicados se centran en lenguas germánicas. En cambio, los estudios sobre la variabilidad acústica de las vocales del español y, en

especial, sobre las variedades lingüísticas habladas en América son relativamente recientes, aunque se destacan los trabajos de Poch-Olivé, Harmegnies y Martín Butragueño (2008), Martín-Butragueño (2014), Sadowsky (2012), Chládkova, Escudero y Boersma (2011), Chládková y Escudero (2012), Correa (2017), Romanelli y Menegotto (2018). Por tanto, el objetivo de esta investigación es contribuir al estudio del vocalismo del español de América teniendo en cuenta variables lingüísticas, orgánicas y expresivas. En esta tesis se estudia, en primer lugar, la relación entre la velocidad de elocución, la duración y las características espectrales de las vocales, esto es, los valores de frecuencia de los dos primeros formantes, el área del espacio acústico, la dispersión y la centralización de los formantes. Al respecto, nos planteamos las siguientes preguntas que van de cuestiones más generales a más concretas:

- 1) ¿Cómo afecta la velocidad de elocución en las características temporales y espectrales de las vocales considerando factores como el timbre, la tonicidad, el entorno consonántico, el tipo de palabra (palabras funcionales vs. palabras con contenido léxico), la posición de la vocal dentro la palabra, dentro del grupo fónico y el sexo del hablante?
- 2) ¿Cómo afecta la duración (que se supone que se reduce en las diferentes velocidades) en los valores de los formantes considerando el timbre de la vocal y la velocidad de habla?

En segundo lugar, se presenta un análisis de la influencia del entorno consonántico sobre las frecuencias de los formantes en la parte estable de las vocales y sobre la trayectoria de los formantes. En concreto, se busca dar respuesta a las siguientes preguntas:

- 3) ¿Cómo afecta el entorno consonántico en las variaciones espectrales de los dos primeros formantes de las vocales del español bogotano considerando el timbre, la tonicidad, el tipo de palabra, la velocidad de elocución, la posición de la vocal dentro de la palabra y dentro del grupo fónico y el sexo del hablante?

En tercer lugar, estudiamos la relación entre el estilo de habla, la duración, la frecuencia de los dos primeros formantes, el área del espacio acústico, la dispersión y la centralización de los dos primeros formantes de las vocales:

- 4) ¿Cómo afecta el estilo de habla en las características temporales y espectrales de las vocales considerando el timbre, la tonicidad, el tipo de palabra, la posición de la vocal dentro del grupo fónico y el sexo del hablante?

Finalmente, nos interesa describir la variación individual en cuanto nos aporta información sobre las estrategias articulatorias que usan los hablantes en las diferentes condiciones estudiadas. Por consiguiente, con la última pregunta se busca saber:

- 5) ¿Cómo afectan las características fonéticas individuales en los resultados obtenidos para cada uno de los parámetros acústicos analizados?

3. Metodología

3.1 La constitución del corpus

Las investigaciones discutidas en el capítulo anterior indican, en conjunto, que hay diferentes factores lingüísticos y extralingüísticos que están asociados a la reducción fonética de las vocales. Cuando se trata el fenómeno en alguna lengua germánica, predominan los estudios centrados en el entorno consonántico, el acento y, en las últimas dos décadas, los relacionados con el estilo de habla. En las lenguas romances el fenómeno se asocia principalmente al estilo de habla, por lo cual disponemos de pocos datos para entender la influencia de las consonantes sobre los valores de frecuencia de los formantes, la importancia de la velocidad de elocución o la variación individual. Por este motivo, el corpus utilizado en esta investigación se construyó con el objetivo de obtener información que permita entender la influencia de factores lingüísticos, orgánicos y expresivos sobre las variables dependientes de nuestro interés: la duración y las frecuencias de los dos primeros formantes de las vocales. Para seleccionar el tipo de corpus más apropiado tuvimos en cuenta los criterios que discutimos a continuación.

3.1.1 Los estilos de habla

Los estudios que tratan de la influencia del estilo de habla sobre las frecuencias de los formantes, comparan vocales producidas en palabras aisladas con vocales extraídas de habla espontánea (Koopmans-van Beinum, 1983; Koopmans-van Beinum y Harder, 1982; Harmegnies y Poch Olivé, 1992; Nicoladis, 2003; Calamai, 2003; Poch Olivé, Harmegnies, y Martín Butragueño, 2008, entre otros). En estos trabajos los resultados coinciden en mostrar que el área acústica se expande en la lectura de palabras y se contrae o desplaza hacia el centro del espacio vocálico en habla espontánea. El primer problema de esta hipótesis es que la categoría *habla espontánea* no es una categoría homogénea y

se obtiene por medio de tareas de recolección de datos tan diversas como una entrevista con el investigador, narraciones, tareas de mapas y diálogos entre amigos (Niebuhr y Michaud, 2015; Llisterri, 1992; Warner, 2011). Y segundo, puede argumentarse que comparar palabras aisladas con habla espontánea se corre el riesgo del sesgo de *reactividad*, en el cual la relación causal (la reducción vocálica como el resultado de un cambio de estilo de habla) es producto del diseño metodológico y no se observa en situaciones fuera del laboratorio (Milder, 2013, p.33). En un estudio preliminar (Correa, 2017), analizamos las vocales producidas por 5 hombres bogotanos en una narración y en la lectura de la transliteración de sus producciones, pero las diferencias entre las frecuencias de los formantes de las vocales no resultaron estadísticamente significativas. Esto nos hace afirmar que, contrario a lo que se piensa, otros factores diferentes al estilo pueden determinar la reducción fonética de las vocales del español.

No ponemos en duda la naturalidad de las muestras consideradas como habla espontánea en los trabajos previos, ni dudamos de la utilidad de analizar palabras aisladas para entender la reducción vocálica. Sin embargo, para mejorar la aplicabilidad de los resultados fuera del laboratorio (*validez ecológica*) y evitar los problemas señalados, adoptamos los criterios propuestos por Niebuhr y Michaud (2015, p. 11-12) para seleccionar las tareas de recolección de los datos:

- 1) Grado de control de las variables: con la tarea se debe obtener un control de las variables dependientes e independientes.
- 2) Densidad de casos: con la tarea se debe obtener un número de casos representativo. Se mide en número de casos por unidad de tiempo.
- 3) Expresividad: con la tarea se debe lograr emotividad e informalidad de las producciones lingüísticas.
- 4) Intención comunicativa: los participantes deben tener como objetivo transmitir un mensaje y una intención clara.
- 5) Homogeneidad de comportamiento: la tarea debe ser lo suficientemente clara como para poder comparar el comportamiento de diferentes participantes.

Teniendo en cuenta estos parámetros, decidimos analizar habla leída y habla conversacional obtenida en condiciones de laboratorio. Para el caso del habla leída seleccionamos tres textos (Anexo A) que han sido usados para estudiar el español americano (Coloma, 2005) y peninsular (Bruyninckx, Harmegnies, Llisterri y Poch, 1994; Ortega, González y Marrero, 2000) y, además, diseñamos un cuarto texto (texto *ad hoc*) para incrementar el número de casos de la vocal posterior cerrada /u/, pues su frecuencia de aparición no permitía llevar a cabo un análisis estadístico de esos datos que fuera fiable. La Tabla 2 ilustra la frecuencia y los porcentajes de aparición de las vocales en cada uno de los textos utilizados para obtener las muestras de habla.

Tabla 2. Frecuencia y porcentajes de aparición de las vocales en cada uno de los textos fonéticamente equilibrados utilizados para obtener las muestras de habla

Texto	Vocal	Frecuencia	Porcentaje
Texto <i>ad hoc</i>	a	265	14.34
	e	197	10.66
	i	66	3.57
	o	722	39.06
	u	598	32.35
	Total	1848	100
Ortega, González y Marrero (2000)	a	860	38.32
	e	526	23.440
	i	0	0
	o	792	35.29
	u	66	2.94
	Total	2244	100
Coloma (2005)	a	196	19.79
	e	328	33.13
	i	131	13.23
	o	201	20.30
	u	134	13.53
	Total	990	100
Bruyninckx et al. (1994)	a	729	28.32
	e	790	30.69
	i	462	17.94
	o	527	20.47
	u	66	2.56
	Total	2574	100
Total		7.656	100

Con los textos aseguramos el control de las variables, la densidad de casos analizados y la homogeneidad en la realización de la tarea. Para lograr datos lingüísticos con expresividad e informalidad grabamos conversaciones siguiendo el procedimiento del *Nijmegen Corpus of Casual Speech* (Torreira y Ernestus, 2010), el cual consiste en grabar diálogos entre amigos que discuten temas de actualidad y sobre los cuales los

participantes deben llegar a un acuerdo. Elegimos esta tarea para crear una situación comunicativa que requiriera la participación activa de los informantes y, de esta forma, olvidaran el lugar en el que estaban. La familiaridad entre los dos participantes de esta tarea favorecía, además, el grado de relajación de los hablantes y, por tanto, la obtención de un tipo de habla verdaderamente espontánea.

3.1.2 La velocidad de elocución

Para medir el tiempo se suele medir la *velocidad de articulación*, que es el número de sonidos por unidad de tiempo sin considerar las pausas, o la *velocidad de habla*, que abarca el número de sonidos y de pausas de una emisión o turno (Laver, 1994, p.158). En los trabajos sobre la reducción fonética de las vocales, la velocidad de articulación es la medición predominante, debido a que la atención de estos estudios se centra en sonidos o transiciones entre dos sonidos (como en las sílabas CVC) y, adicionalmente, el análisis se basa en frases marco, lo cual imposibilita el análisis de las pausas. En esta tesis analizamos el *tempo subjetivo*, es decir, la velocidad planeada por el hablante cuando se le solicita que incremente su velocidad, lo disminuya o use un tempo «normal». Para lograr esto, cada participante leyó los tres textos a velocidad normal, después a velocidad lenta y, finalmente, a velocidad rápida. En adelante, al tempo subjetivo le denominaremos simplemente *velocidad de elocución*.

Existen al menos dos razones para usar el *tempo subjetivo*. En primer lugar, nuestro objetivo no es medir la velocidad de elocución sino determinar los cambios que provoca esta variable en la articulación y en las propiedades acústicas de las vocales. Por esta razón no optamos, como sugiere la bibliografía, por métricas más populares y objetivas como las sílabas fonéticas y los fonos por segundo (Trouvain, 2003, p. 46). En segundo lugar, en esta investigación nos interesa la variación fonética individual, de manera que las variaciones temporales características de cada locutor no resultan una limitación para el análisis. Para asegurar el éxito de la tarea de lectura en las diferentes velocidades de habla, mantuvimos controlado el nivel educativo, así que todos los participantes en el estudio contaban, en el momento de la grabación, con formación universitaria.

3.1.3 Entorno fonético de las vocales

Los trabajos dedicados a la influencia del entorno consonántico sobre las vocales analizan secuencias CV\$C, donde C es una de las consonantes obstruyentes de la lengua. Para facilitar el análisis acústico, C tiene el mismo punto y modo de articulación en ambas apariciones, lo que se conoce como un contexto consonántico simétrico. En el español hay una baja aparición de palabras con estas características, lo cual obliga a usar logatomas o palabras poco comunes. Chládkova *et al.* (2011), por ejemplo, construyeron palabras bisílabas CV₁CV₂, donde C era una obstruyente /p, t, k, f, s/, V₁ una de las cinco vocales del español, y V₂ las vocales medias /e, o/. De esta manera, los autores construyeron un corpus con frases como *en fife y fifo tenemos i*, con tres apariciones de cada vocal por frase. Nadeu (2013, 2014) usó un procedimiento semejante, pero las palabras seleccionadas, si bien son poco comunes, fueron juzgadas como palabras reales por algunos de sus informantes.

El uso de logatomas y frases marco para estudiar la influencia del entorno consonántico sobre el timbre es un procedimiento bastante difundido, pero lo consideramos poco apropiado para investigar el fenómeno de nuestro interés por dos razones. Primero, como vimos en el capítulo anterior, las palabras poco frecuentes tienden a pronunciarse con mayor precisión articulatoria, mientras que la reducción fonética es común en la pronunciación de las unidades lingüísticas más redundantes de la lengua (van Son *et al.*, 2004). Esto significa que, durante la lectura de frases marco con logatomas, los hablantes pronunciarán las palabras poco conocidas de manera enfática (por ejemplo, en *diga pupe dos veces* o *diga fife dos veces*) y evitarán la reducción articulatoria y acústica. Segundo, en español predominan los contextos asimétricos. Desde un punto de vista metodológico, esto no resulta un problema para el análisis de los datos pues, como han mostrado diferentes autores, estos contextos permiten estudiar la reducción fonética en términos de la diferencia entre la frecuencia de los formantes en el locus consonántico y núcleo vocálico, o bien en términos de la curvatura de la trayectoria formántica (Moon y Lindblom, 1994; van Bergem, 1994; Sussman, Dalson y Gumbert, 1998).

Por estas dos razones, para el diseño del corpus optamos por el análisis de entornos consonánticos C₁VC₂, donde C₁ son las consonantes del español en ataque de sílaba clasificadas según el punto de articulación (labial, dentoalveolar y velar), V es una de las

cinco vocales /i, e, a, o, u/ y C_2 —el ataque de la sílaba siguiente— corresponde a las consonantes dentoalveolares (/l, r, ʀ, s, t, ʀ/). Se eligió este punto de articulación dado que varios estudios han mostrado que las consonantes alveolares son más resistentes a la coarticulación que, por ejemplo, las labiales (Recasens y Espinosa, 2009, p. 2289; Iskarous, y otros, 2013, p. 1272).

La Figura 15 presenta la distribución de las vocales analizadas en función del entorno consonántico del corpus de habla leída. Como puede verse, el contexto más común para C_1 es el *dentoalveolar*, seguido del *labial* y del *velar*. Las cinco vocales aparecen en los tres entornos consonánticos, a excepción de /i/ que no reporta casos para el contexto velar. En la Figura 16 se ilustran los entornos consonánticos analizados en habla espontánea, los cuales mantienen una proporción cercana al habla leída en cuanto al número de vocales en los tres entornos consonánticos. Como puede verse, los criterios adoptados y el número de casos que conforman el corpus permiten la comparación entre los dos estilos (lectura vs. habla espontánea).

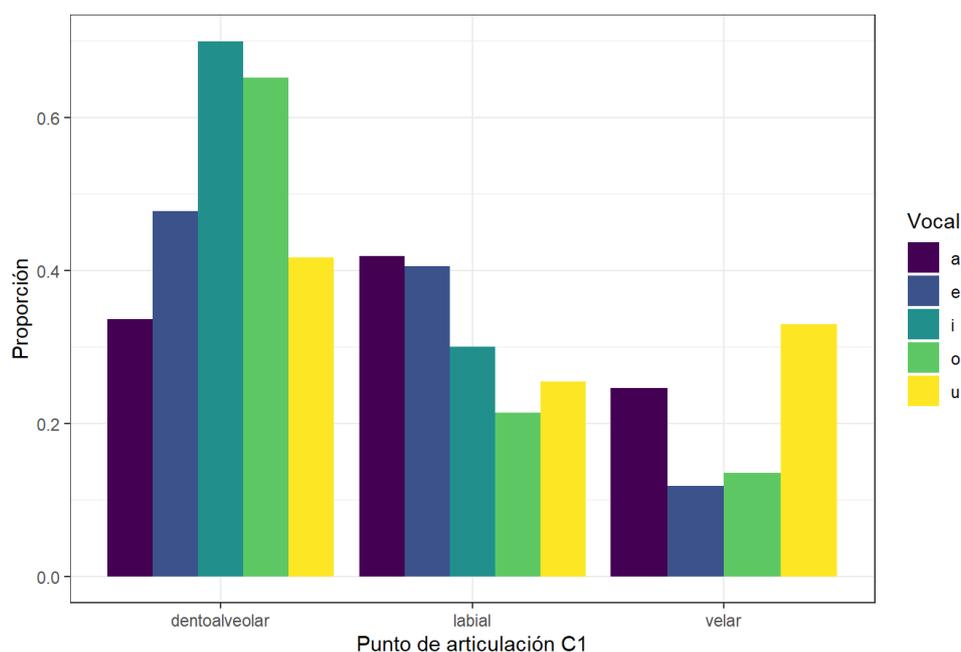


Figura 15. Proporción de vocales analizadas según el punto de articulación de la consonante precedente C_1 en habla leída.

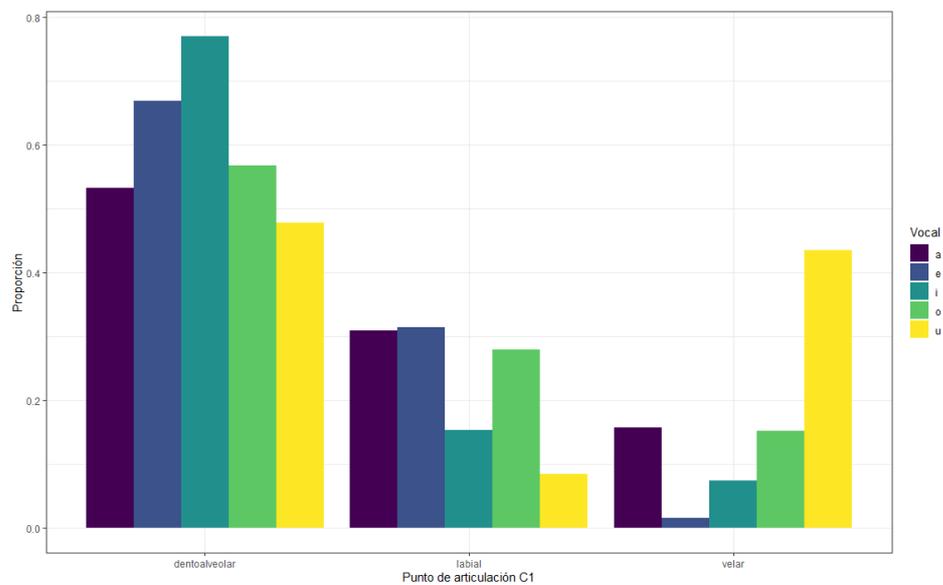


Figura 16. Proporción de vocales analizadas según el punto de articulación de la consonante precedente C_1 en habla espontánea

3.1.4 Variación prosódica

En cuanto a la tonicidad, distinguimos entre vocales tónicas y átonas tomando como referencia el acento léxico. También tenemos en cuenta las vocales ubicadas en posición inicial, intermedia y final de frase, con el objetivo de identificar cambios potenciales en la duración y el timbre relacionados con la prosodia de la frase. Por ejemplo, las tónicas y átonas iniciales pueden modificarse a causa de fenómenos como el reajuste de f_0 o la modalidad del enunciado, las vocales intermedias pueden reforzar las variaciones melódicas de las sílabas prenucleares, y en posición final las vocales son constituyentes del núcleo o tonema de la frase entonativa. Finalmente, incluiremos la posición de la vocal dentro de la palabra, dado que estudios recientes han mostrado que se trata de un factor determinante de la duración y los valores de frecuencia de los formantes (Gendrot, Adda-Decker y Santiago, 2019).

3.1.5 Variación léxica

Como mencionamos en el capítulo anterior, una de las hipótesis más sólidas sobre la relación entre propiedades léxicas y la reducción fonética sostiene que las palabras con contenido léxico se pronuncian con mayor precisión, tienen baja frecuencia de aparición y son menos predecibles, mientras que las palabras funcionales tienden al debilitamiento de sus fonemas y sílabas (Zipf, 1929; Koopmans-van Beinum y Harder, 1982; Rehor y

Pätzold, 1996; Meunier y Espesser, 2011). Por esta razón, en nuestro diseño estadístico incluimos como variable el tipo de palabra, distinguiendo entre aquellos ítems que tienen contenido léxico de los que no (palabras funcionales). Como veremos más adelante, el tipo de modelo estadístico que implementaremos (los modelos lineales de efectos mixtos) permite controlar el grado de varianza introducida por los ítems léxicos, de manera que la variabilidad que no pueda ser explicada por medio del contenido semántico de la palabra (palabra funcional vs. palabra con contenido léxico) será considerada dentro del análisis cuantitativo de los diferentes parámetros acústicos.

3.2 Los participantes

En este estudio participaron 11 hombres y 11 mujeres nacidos en la ciudad de Bogotá D. C. (Colombia) y su área metropolitana (Figura 17). En el momento de la obtención del corpus (2017-2018) los hombres y mujeres tenían una edad media de 28.3 (± 7.1) años. Los hablantes eran licenciados o estudiantes de posgrado y ninguno reportó problemas de producción o percepción del habla. Mantuvimos el nivel educativo controlado para asegurar el éxito de la tarea de lectura en las tres velocidades de habla.

Las sesiones de grabación consistieron en una tarea de lectura en solitario y una conversación. Las sesiones de conversación se realizaron en parejas: uno de los participantes era una persona conocida por el investigador (cómplice) y su interlocutor era un amigo o un familiar cercano. Esta estrategia, como ya se ha mencionado, la adaptamos de Torreira y Ernestus (2010) pero, a diferencia del *Nijmegen Corpus*, grabamos conversaciones en pareja y no entre el cómplice y dos informantes. Esta decisión se tomó después de realizar una prueba piloto y comprobar que dos informantes son suficientes para producir habla conversacional. Otra diferencia entre nuestro procedimiento y el utilizado por Torreira y Ernestus (2010) es que no restringimos la conversación a informantes del mismo sexo, sino que dimos mayor importancia al grado de confianza entre los interlocutores. De esta manera grabamos a personas del mismo sexo o a parejas conformadas por un hombre y una mujer siempre que, a juicio de ellos, pudieran mantener una conversación fluida sobre diferentes temas.

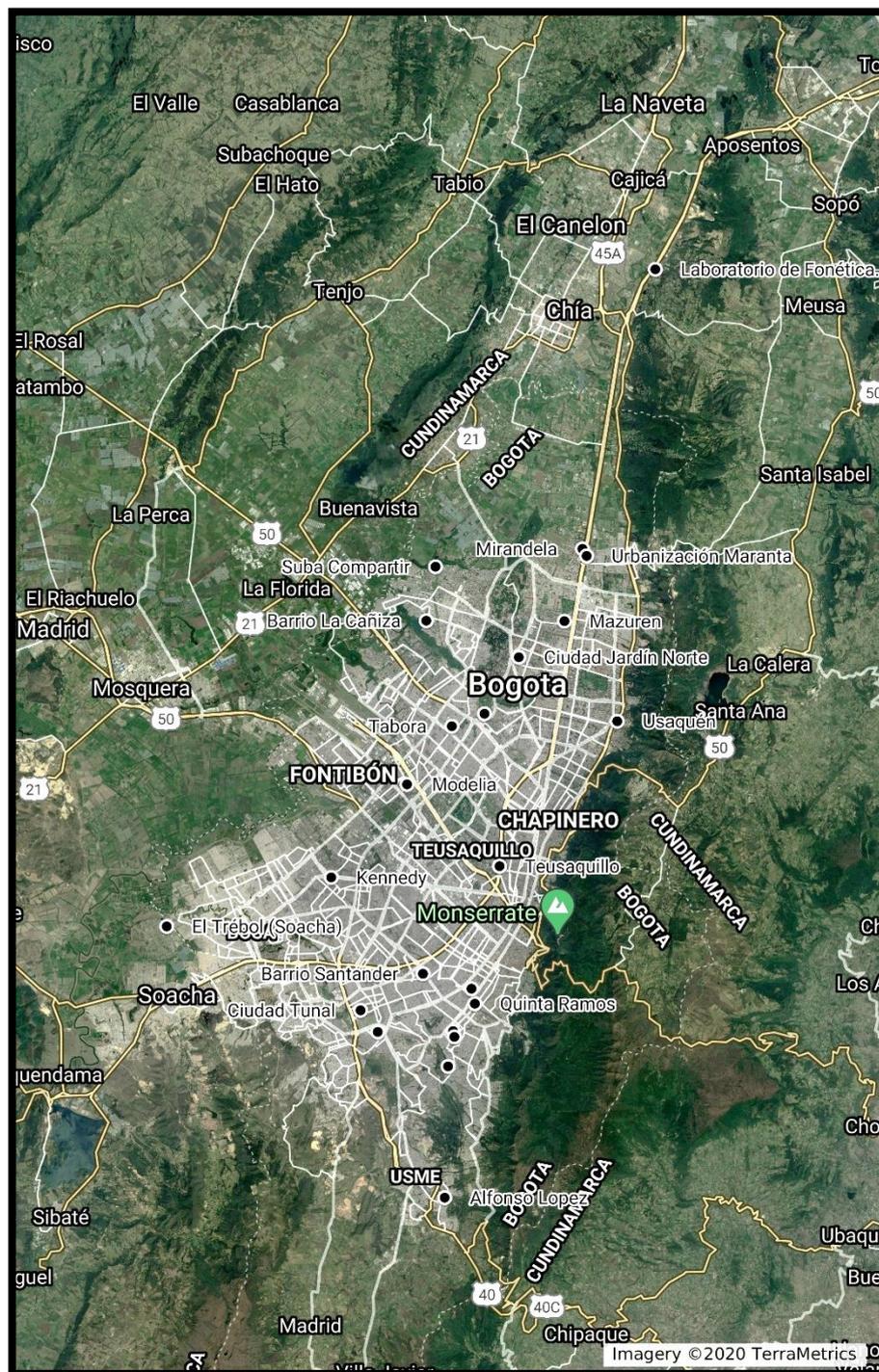


Figura 17. Bogotá Distrito Capital. Principales lugares de residencia de los 22 participantes en el estudio.

Los participantes en esta investigación han nacido en barrios del norte, sur, occidente y centro de la ciudad y han residido la mayor parte de sus vidas en estos lugares. La Figura 17 presenta un mapa de Bogotá con los barrios en que vivían los hablantes en el momento del estudio. En el punto más extremo al norte se encuentra el Laboratorio de Fonética

Experimental del Instituto Caro y Cuervo, ubicado a 7 km de la ciudad, mientras que en la parte inferior está la localidad de Usme, a un poco más de 50 km al sur. Al suroccidente de la ciudad se encuentra el municipio de Soacha, un importante núcleo poblacional que está completamente integrado a la zona urbana. Esta distribución muestra que los informantes aportaron datos de diferentes puntos de la ciudad y, si bien tienen un nivel educativo homogéneo, podemos decir que son representativos de la población joven y de los diferentes grupos sociales que residen en la capital de Colombia.

3.3 Las grabaciones

Las sesiones de grabación se llevaron a cabo en la cabina insonorizada del Laboratorio de Fonética Experimental del Instituto Caro y Cuervo. La lectura se realizó de manera individual y la conversación se realizó cara a cara sin presencia del investigador. Las grabaciones se realizaron con una grabadora digital TASCAM DR-100 a 44 100 Hz y 16 bits. Se utilizaron micrófonos de diadema Audio-Technica PRO 92cW, los cuales tienen patrón polar omnidireccional, y una respuesta de frecuencia entre 20-20 000 Hz. Los micrófonos se fijaron a la oreja y se ubicaron a 5 cm de los labios aproximadamente.

Cada sesión de grabación se dividió en dos partes. En la primera, cada participante leyó en solitario los textos a velocidad normal, después a velocidad lenta y, finalmente, a velocidad rápida. Después de la lectura, que duró en promedio treinta minutos, se realizó una pausa de veinte minutos durante la cual los participantes tuvieron la oportunidad de descansar y conversar entre sí. Siguiendo el procedimiento utilizado en la recolección de datos del *Nijmegen Corpus of Casual Speech* (Torreira y Ernestus, 2010), la conversación se obtuvo por medio de la discusión de temas de actualidad (Anexo B). Por ejemplo, se preguntó sobre las últimas cifras de inversión en la Plaza de Toros de Bogotá, tema que dio lugar a discusiones sobre diferentes cuestiones, como los derechos de los animales, la inversión del dinero público o la tauromaquia.

A cada pareja se le dieron las siguientes instrucciones: 1) leer cada situación de actualidad (dos o tres líneas de textos) y manifestar su posición sobre cada tema, 2) argumentar su posición y contraargumentar, 3) tratar de llegar a un acuerdo al final de las intervenciones sobre cada tema, 4) respetar el turno del interlocutor y 5) no intentar dominar la conversación. Las grabaciones resultantes tuvieron una duración de

aproximadamente 25 minutos. Durante este tiempo, el investigador supervisó la grabación, pero se mantuvo fuera del campo visual de los informantes.

3.3.1 Transcripción y anotación

La transcripción y la anotación de los archivos de audio se realizaron mediante el programa Praat (Boersma y Weenink, 2021). Cada texto se anotó de manera semiautomática con *Easyalign* (Goldman, 2011) en palabras, sílabas y fonos. Luego se añadieron tiras (TextGrid) para delimitar las vocales de nuestro interés e incluir información sobre la posición de la vocal dentro de cada grupo fónico (inicial, intermedia, final). Para realizar la segmentación de las vocales se tuvo en cuenta el grado de sonoridad, la periodicidad y la amplitud de la forma de onda, tal como se ilustra en la Figura 18. La tonicidad de la vocal se determinó a partir del acento léxico de cada palabra.

Las vocales se clasificaron como *ensordecidas* cuando la vocal no tuvo amplitud de onda ni una duración definida o cuando no fueron claros los contornos de los formantes. Esto ocurrió entre consonantes sordas y al final de los grupos fónicos, por lo cual las vocales ensordecidas tienen una forma de onda con ruido. Las vocales se clasificaron como *plenas* cuando el contorno de formantes estaba claramente definido y fue posible segmentarlas y diferenciarlas del contexto consonántico. En nuestro análisis incluimos solamente vocales plenas debido a que necesitábamos mediciones precisas de la duración y los valores de frecuencia de los formantes para determinar la reducción fonética.

En el caso de las conversaciones, se realizó una transcripción ortográfica y se dividieron los turnos de cada participante en grupos fónicos mediante el programa ELAN (Brugman y Russel, 2004), como se ilustra en la Figura 18. Con ayuda de esta transcripción se seleccionaron las vocales teniendo cuenta los porcentajes de aparición de cada vocal y los entornos consonánticos de los textos fonéticamente equilibrados. El objetivo de este procedimiento fue delimitar el número de casos y facilitar la comparación entre el habla conversacional y el habla leída. Después de seleccionar los casos en cada conversación, se organizaron los ficheros por hablante y se anotaron los audios con el programa Praat (Boersma y Weenink, 2021) y la extensión *Easyalign* (Goldman, 2011) tal como se ha descrito anteriormente.

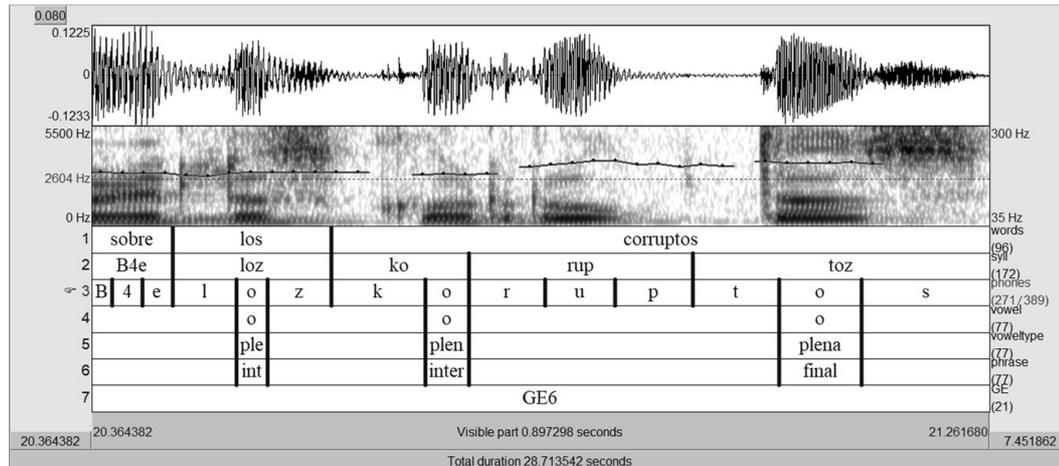


Figura 18. Transcripción y anotación en Praat para el habla leída y el habla espontánea.



Figura 19. Transcripción ortográfica del habla espontánea en ELAN.

3.4 Parámetros acústicos

Para responder a las preguntas de investigación, se seleccionaron mediciones para estimar la duración vocálica, los valores de frecuencia de los dos primeros formantes, la trayectoria de los formantes, el área vocálica y su grado de centralización. Estos parámetros acústicos los podemos clasificar en cuatro tipos:

1. Mediciones estáticas: duración, F1 y F2 en el centro de la vocal.
2. Mediciones dinámicas: diferencia absoluta ($\Delta F1/\Delta F2$) y velocidad de cambio ($rF1/rF2$).

3. Mediciones del espacio vocálico: Área del espacio vocálico (AEV).
4. Mediciones de la centralización: Tasa de centralización de los formantes (TCF), distancia euclídea e índice δ .

Los valores de frecuencia de F1 y F2 se analizaron en hercios y se normalizaron de acuerdo con el método de Lobanov (1971). El AEV, la TCF y la dispersión vocálica se midieron en escala Bark. A continuación, explicamos cada uno de los parámetros acústicos y las técnicas de normalización implementadas.

3.4.1 Mediciones estáticas de la duración y los formantes vocálicos

La duración se anotó manualmente con ayuda de la forma de onda y del espectrograma. Cada vocal se delimitó en el primer y último pulso glotal, teniendo en cuenta solamente variaciones cuasiperiódicas, es decir, aquellos pulsos que presentaron una amplitud y duración regulares. No incluimos, por ejemplo, aquellas variaciones acústicas que se encuentran en los límites entre sonidos y no tienen presencia de formantes. La Figura 20 ilustra el criterio de segmentación de las vocales en el entorno de consonantes obstruyentes. La parte sombreada indica los límites entre fonos (verde) y la duración vocálica (rojo).

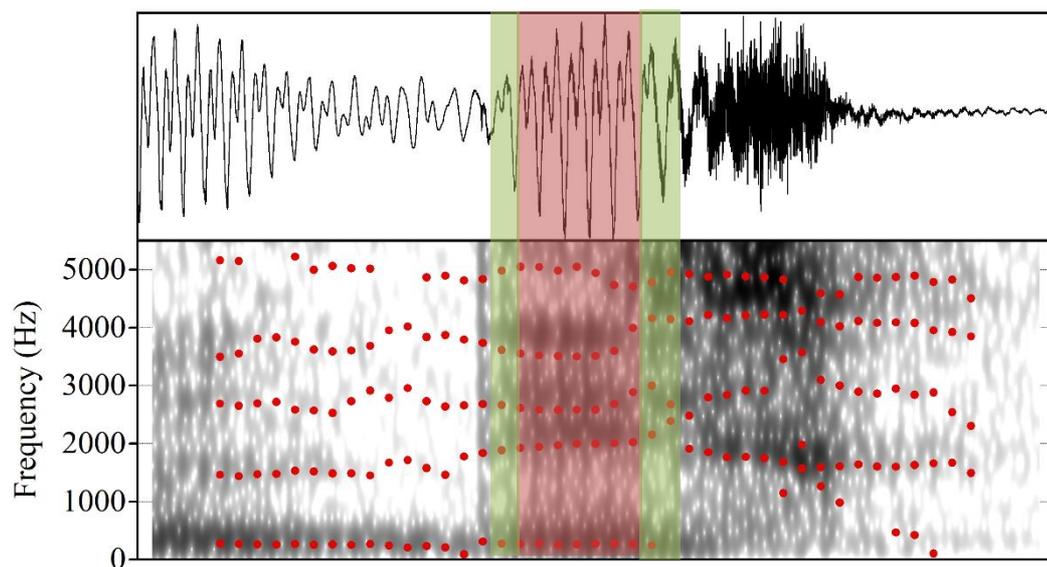


Figura 20. Segmentación de la vocal a partir de la forma de onda y del espectrograma.

Las frecuencias de los dos primeros formantes de las vocales del español bogotano se midieron en el primer pulso de la vocal para obtener el valor de la transición

consonántica. El núcleo acústico se midió en el centro de la vocal, pues los autores que han usado otros criterios, como el promedio del formante, el momento de mayor intensidad o el valor máximo/mínimo del formante, han encontrado que la diferencia entre los diferentes procedimientos no es significativa (van Son y Pols, 1990, 1992). En ese sentido, concordamos con Pätzold y Simpson (1997) en que el centro de la vocal es una elección adecuada para el análisis acústico, ya que en esta parte de la vocal hay menor influencia del contexto (sobre este punto véase también Stevens y House, 1963).

Los formantes fueron calculados con el algoritmo Burg de Praat (Boersma y Weenink, 2021), con una ventana de 25 ms, una ventana gaussiana de 50 ms y un formante máximo de 5000 Hz para los hombres y de 5500 para las mujeres. Dentro de estos parámetros incluimos cinco formantes, sin embargo, posteriormente se hizo una revisión manual de los datos y tuvimos que aumentar o reducir este cuando el programa identificó más o menos formantes de los esperados. Por ejemplo, en la vocal [u] del habla femenina suele aparecer un F2 falso que se corrigió ajustando el valor a 4.5 o 4 formantes (Correa, 2014, p. 41). Estos parámetros se aplicaron a los datos por medio de un *script* de Praat.

3.4.2 Mediciones dinámicas de los formantes

Para medir los cambios dinámicos en los valores de frecuencia de los formantes, calculamos, en primer lugar, el cambio absoluto de la frecuencia de los dos primeros formantes ($\Delta F1$ y $\Delta F2$). Este parámetro es la diferencia entre los valores de frecuencia al inicio Fn_i y el centro de la vocal Fn_c . De acuerdo con Herrmann, Cunningham y Whiteside (2014, p. 60), cuando el cambio absoluto de frecuencia del segundo formante tiene un valor bajo, hay mayor grado de coarticulación:

$$\Delta Fn = |Fn_c - Fn_i| \quad (6)$$

La segunda medición dinámica es la velocidad de cambio de los dos primeros formantes ($rF1$ y $rF2$), la cual permite caracterizar la tasa de cambio de la transición en función del tiempo y, en particular, permite hacer inferencias sobre las diferencias individuales en el grado de esfuerzo articulatorio involucrado en la producción de las vocales. La velocidad de cambio de F2, propuesta originalmente en Gay (1977) y Moon y Lindblom (1994), es la diferencia entre el centro de la vocal y el locus ($F2_T - F2_L$) sobre la duración de la

transición vocálica (ΔT). En nuestro análisis implementamos la velocidad de cambio con la ecuación propuesta por Herrmann, Cunningham y Whiteside (2014, p. 60) para los dos primeros formantes:

$$rFn = \frac{Fn\Delta}{0.5 \times t_{vocal}} \quad (7)$$

3.4.3 Medición del área del espacio vocálico

Las mediciones del espacio vocálico se usan para comparar variedades dialectales, estilos de habla y, en los ámbitos de las patologías del habla, para evaluar tratamientos en pacientes con disartria y tartamudeo. En esta investigación nos interesa calcular el área vocálica para evaluar la importancia de factores como el estilo, la velocidad de elocución, la tonicidad y el sexo del hablante. En este trabajo calcularemos el Área del espacio vocálico (AEV) diagramando F1 en función del F2 en un triángulo acústico formado por las vocales /i, a, u/. De acuerdo con Blomgren, Robb y Chen (1998, p. 1045), el AEV es igual a:

$$\frac{(F2u+F2i)(F1u-F1i)-(F2a+F2u)(F1a-F1u)-(F2a+F2i)(F1a-F1i)}{2} \quad (10)$$

3.4.4 Medición de la centralización de las vocales

Para medir la centralización de las vocales se usan dos métricas. La primera es la Tasa de centralización de los formantes (TCF), que sirve para determinar el grado de centralización del espacio acústico y, en ese sentido, es complementaria al AEV. Esta tasa se expresa como:

$$\frac{(F2u+F2a+F1i+F1u)}{(F2i+F1a)} \quad (12)$$

De acuerdo con Sapir, Raming y Spielman, (2010, p. 116) esta medida es independiente del sexo del hablante, es sensible a la articulación en habla normal y patológica y está diseñada para que las frecuencias de los formantes del numerador aumenten y las frecuencias del denominador disminuyan con la centralización. En el análisis que

presentan Sapir *et al.* (2010), esta métrica resultó ser más efectiva que el área vocálica para diferenciar las características fonéticas de la disartria vs. el habla normal.

Para determinar la distancia entre las vocales y el centroide del espacio acústico se calculará la distancia euclídea (Koopmans-van Beinum, 1983). En este estudio, el centroide es el valor medio de $\bar{F}1$ y $\bar{F}2$ de todas vocales de un hablante, y la distancia se calcula como la raíz cuadrada de la suma de las diferencias entre cada formante y su valor medio al cuadrado:

$$de = \sqrt{(F1 - \bar{F}1)^2 + (F2 - \bar{F}2)^2} \quad (11)$$

La distancia euclídea se ha usado con éxito en numerosos estudios para comparar estilos de habla y estudiar la variación dialectal porque arroja información precisa sobre la distancia entre cada vocal y el centro acústico (por ejemplo, Harmegnies y Poch Olivé, 1992; Recasens y Espinosa, 2006; Lengeris, 2012; Rathcke, Suart-Smith, Torsney y Harrington, 2017, entre otros). Por esa razón, en el capítulo 6 implementamos, además, el *índice* δ el cual permite establecer, para cada timbre vocálico, la diferencia entre las distancias euclídeas de dos estilos de habla (Harmegnies y Poch-Olivé, 1992; Poch-Olivé, Dhainaut y Harmegnies, 2003, p. 523). Los valores cercanos a cero indican que no hay diferencia entre los dos estilos; mientras que si el índice δ se incrementa, se debe a que hay una diferencia entre los estilos comparados. Para analizar nuestros datos, calculamos la diferencia entre la distancia euclídea del habla lenta (la pronunciación más hiperarticulada de nuestro corpus) y la distancia euclídea de las vocales pronunciadas en habla leída a velocidad rápida, normal y en habla espontánea.

3.4.5 Normalización de las frecuencias formánticas

Los valores en hercios fueron transformados a la escala Bark para analizar y representar los datos en una escala auditiva y, especialmente, para calcular la dispersión, el área y la tasa de centralización vocálica. La conversión la realizamos con la ecuación de Traunmüller (1990, 1997) que arroja las diferencias en unidades tipificadas o *z scores*:

$$z = \frac{26.81 \text{ Hz}}{\left(\frac{1960}{f} + 1\right)} - .53 \text{ Bk} \quad (8)$$

Las frecuencias de los formantes, expresadas en hercios y barks, retienen una gran cantidad de variabilidad acústica asociada a diferencias orgánicas, como el tamaño del tracto vocal y las dimensiones del tracto laríngeo. Para minimizar los efectos anatómicos e individuales, las frecuencias formánticas se normalizaron para 1) eliminar las diferencias entre hablantes, 2) conservar la variación sociolingüística y dialectal, y 3) conservar las distinciones fonológicas.

Los métodos de normalización se pueden aplicar sobre cada hablante (intrínseco al hablante) o sobre varios (extrínseco al hablante); es posible normalizar las vocales según el timbre (intrínseco a la vocal) o teniendo en cuenta el sistema vocálico en conjunto (extrínseco a la vocal). También es posible normalizar un formante a la vez (intrínseco al formante) o comparando entre los valores de frecuencia de F1, F2 y F3 (Wissing y Wikus, 2014, p. 99). Para esta investigación usamos el método de Lobanov (1971), el cual es intrínseco al hablante, extrínseco a la vocal e intrínseco al formante. De acuerdo con la adaptación de Adank *et al.* (2004), el método de Lobanov (1971), basado en unidades tipificadas (*z scores*), se calcula por medio de la siguiente ecuación:

$$F_{ti}^{Lobanov} = \frac{F_{ti} - F\mu_{ti}}{\delta_{ti}} \quad (9)$$

De acuerdo con este procedimiento, el valor de un formante (F_{ti}) es la distancia entre cada medición y la media de la frecuencia del formante ($F\mu_{ti}$) dividido por la desviación típica (δ_{ti}) de la frecuencia del formante ($F\mu_{ti}$). Este procedimiento, a pesar de su antigüedad, ha demostrado ser superior a los demás en diferentes análisis comparativos (Adank, *et al.*, 2004; Schützler, 2015).

3.5 Tratamiento estadístico del corpus

Durante las últimas décadas, la fonética ha consolidado su carácter cuantitativo gracias a la disponibilidad de corpus orales y de herramientas informáticas para analizar los datos. Por ejemplo, mientras que en el *International Congress of Phonetic Sciences* (ICPhS) de 1967 solo el 20 % de las ponencias reportó algún tipo de análisis estadístico, en 2011 el número de estudios cuantitativos alcanzó el 75 % (Šturm, 2015). Este crecimiento también se refleja en los test empleados: mientras que en 1967 predominó el t-test, las correlaciones y el análisis de varianza (ANOVA), en 2011 se observa la

aparición de modelos lineales de efectos mixtos, además de otras pruebas estadísticas diferentes a las mencionadas.

Si bien el uso creciente de análisis cuantitativos ha significado un gran avance en nuestra comprensión de los mecanismos involucrados en la producción y percepción del habla, también es fácil cometer errores en el diseño experimental. Esto ocurre, por ejemplo, cuando el investigador toma decisiones deliberadas como ajustar el análisis para obtener el resultado (*sesgo confirmatorio*) o tratar un estudio exploratorio como un estudio confirmatorio (Wagenmakers, Wetzels, Borsboom, van der Maas y Kievit, 2012). En el ámbito de la fonética, Roettger (2018) ha señalado que uno de los errores más comunes es el error Tipo I, esto es, falsar de manera errónea una hipótesis nula.

Uno de los factores que incrementa el error Tipo I es el uso de ítems repetidos dentro del conjunto de datos. Para probar esto, Winter (2011) simuló 5000 palabras con un VOT de 60 ms y una desviación típica de 10 ms. Si bien se trataba de un solo conjunto de datos, el autor dividió las observaciones en dos grupos (laxas y aspiradas) para la simulación estadística. Winter observó que la posibilidad de encontrar una diferencia estadísticamente significativa de manera errónea se incrementó cuando se promediaron los ítems y cuando se aumentó el número de repeticiones. Repetir ítems conlleva al problema de la pseudoreplicación, que consiste en analizar datos interrelacionados o correlacionados como si fueran datos independientes (Winter, 2011). Este problema es bastante serio porque la condición de independencia es una condición necesaria para aplicar la mayoría de métodos paramétricos de estadística inferencial. Una de las maneras de evitar los efectos de las repeticiones y algunos de los sesgos mencionados, es el uso de modelos de regresión de efectos mixtos.

Los datos de análisis obtenidos del corpus de esta tesis se enfrentan a varios desafíos desde un punto de vista estadístico. Primero, incluyen repeticiones de ítems en la lectura a tres velocidades y, por tanto, no se cumple la condición de independencia. Segundo, el diseño experimental incluye muestras no equilibradas (en habla espontánea no se puede controlar el número de casos) y, para responder a las preguntas de investigación, es necesario explicar la influencia de la variación individual sobre los valores de duración y de frecuencia de los sonidos vocálicos. Como veremos a continuación, los *modelos*

lineales de efectos mixtos (MLM) permiten resolver los desafíos metodológicos mencionados y hacer un tratamiento estadístico apropiado de los datos.

3.5.1 Modelos lineales de efectos mixtos (MLM)

Los modelos lineales de efectos mixtos (MLM) son un método estadístico que combina factores fijos con factores aleatorios. Los factores fijos son aquellos en que los niveles agotan los valores de una variable. Por ejemplo, la variable *sexo* es un factor fijo porque sus dos niveles (*masculino* y *femenino*) agotan los valores que se pueden observar. Los factores aleatorios introducen variación a las tendencias de la población (Baayen, 2008, p. 15) y sus valores cambian de respuesta a respuesta. Para el caso de los datos lingüísticos, los factores aleatorios aportan información sobre la influencia de las características individuales de los participantes y sobre la variabilidad introducida por las propiedades léxicas.

Si tomamos como ejemplo un análisis de la duración en función de los factores fijos *tonicidad* (tónica vs. átona) y *velocidad de elocución* (velocidad normal vs. velocidad rápida), la duración es la variable dependiente o respuesta (Y_K), mientras que la tonicidad (X_1) y la velocidad de habla (X_2) son los predictores o variables independientes. En el modelo (13), el intercepto o nivel de referencia, β_0 , es el valor medio de las átonas en velocidad normal, β_1 es la pendiente de la tonicidad, calculada como la diferencia entre duración media de las vocales átonas y tónicas, β_2 es la pendiente de la variable velocidad de elocución, y ϵ_K el error residual.

$$Y_K = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon, \quad (13)$$

$$K = 1, 2 \dots 5$$

Una de las ventajas de los modelos mixtos es que permite incluir factores aleatorios para explicar, por ejemplo, la variación asociada a las diferencias entre participantes o entre ítems léxicos. Estos ajustes se conocen como interceptos aleatorios y, cuando se aplican a la pendiente, se conocen como pendientes aleatorias (Baayen, 2011, p. 671). De acuerdo con Baayen (2013, pp. 351-352), la inclusión de interceptos aleatorios y pendientes aleatorias para los participantes e ítems mejora sustancialmente los modelos. Este autor añade que los estudios que no los incluyen pueden incluir valores de p más

bajos de lo esperado y, por tanto, el analista puede creer erróneamente en la validez de un efecto no significativo. Gries (2015), Baayen (2011) y Winter (2011) destacan las siguientes ventajas de los modelos mixtos frente a los métodos tradicionales de la estadística inferencial:

1. Permiten trabajar con diseños experimentales que tienen muestras no equilibradas (Gries, 2015) como, por ejemplo, las extraídas de habla espontánea.
2. Si el diseño experimental contiene repeticiones, es posible incluirlas en el modelo como un factor fijo, lo cual permite controlar las interacciones con el fenómeno de interés (Winter, 2011, p. 2140).
3. Permiten modelar los efectos aleatorios y entender las diferencias individuales en los resultados de los experimentos (Baayen, 2011; Gries, 2015).
4. Los modelos mixtos son, en general, mejores detectando efectos significativos (Baayen, 2011, p.673).

Como se ve, los modelos de regresión de efectos mixtos solucionan los problemas de tratamiento estadístico mencionados. Los modelos presentados en esta tesis se implementaron con el entorno de programación R (R Core Team, 2014) y, en particular, con el paquete *lm4* (Bates, Bolker y Walker, 2015). Los valores de *p* para los coeficientes fueron calculados con el paquete *afex* (Singmann, Bolker, Westfall y Aust, 2016), y la significatividad de los factores se determinó con la prueba de razón de verosimilitud (*likelihood ratio test*), como se explica en Winter (2020). Finalmente, la visualización se realizó con los paquetes *ggplot2* (Wickham, 2016) y *vowels* (Kendall y Thomas, 2010). La representación de los modelos mixtos y los diagramas estadísticos fueron elaborados siguiendo los procedimientos y recomendaciones de Healy (2019).

3.5.2 Variables consideradas en el análisis estadístico

La Tabla 2 resume los factores fijos, sus niveles, y los factores aleatorios incluidos en los modelos estadísticos. Entre los factores segmentales tuvimos en cuenta el entorno consonántico y el timbre de la vocal (/i, e, a, o, u/). Entre los factores léxicos incluimos la categoría de la palabra y en los modelos mixtos los ítems léxicos se incluyeron como un factor aleatorio.

Tabla 3. Variables independientes consideradas en esta tesis

Descripción	Niveles
Punto de articulación C ₁	Labial, dentoalveolar, velar
Consonante C ₁	/f, p, b, d, r, l, r, t, s, k, g/
Timbre vocálico	/i, e, a, o, u/
Categoría léxica	Palabras de contenido vs. palabras funcionales
Tonicidad de la vocal	Tónica, átona
Posición de la vocal dentro de la frase	Inicial, intermedia, final
Posición de la vocal dentro de la palabra	Inicial, intermedia, final
Velocidad de elocución	Habla lenta, normal y rápida
Estilo de habla	Habla leída, habla conversacional
Sexo del hablante	Mujeres, hombres
Variable aleatoria	Ítems léxicos
Variable aleatoria	Participantes

Para estudiar las influencias de la prosodia léxica y la prosodia de la frase, las vocales se clasificaron como tónicas y átonas. Las vocales también se diferenciaron según su aparición al inicio, intermedio o final del grupo fónico. En cuanto a los factores relacionados con la elocución, analizamos habla leída y habla conversacional. Para estudiar los efectos de la velocidad de elocución los datos fueron leídos en tres velocidades (la velocidad subjetiva del hablante). Finalmente, la variación individual la estudiamos incluyendo interceptos y pendientes aleatorias en los modelos mixtos.

4. Resultados: Duración, formantes, espacio acústico y centralización de las vocales en tres velocidades de habla

4.1 Introducción

En este capítulo se reportan los resultados del análisis de las frecuencias de F1 y F2 y la duración vocálica en tres velocidades de elocución. Como se mencionó en el capítulo anterior, 22 bogotanos leyeron cuatro textos a velocidad normal, lenta y rápida. Con este procedimiento buscamos obtener vocales con un amplio rango de duraciones e investigar si los hablantes modifican las frecuencias de los formantes cuando aumentan y/o disminuyen su velocidad de elocución. En particular, en este capítulo se busca dar respuesta a las siguientes preguntas que nos planteábamos al principio de este trabajo:

- 1) ¿Cómo afectan las tres velocidades de elocución en las características temporales y espectrales de las vocales considerando el timbre, la tonicidad, el entorno consonántico, la posición de la vocal dentro de la palabra y dentro del grupo fónico y el sexo del hablante?
- 2) ¿Cómo afecta la duración en los valores de los formantes considerando el timbre de la vocal y la velocidad de habla?
- 3) Por último, nos interesa describir cómo afectan las características fonéticas individuales en los resultados obtenidos para cada uno de los parámetros acústicos estudiados.

Dentro de las características temporales analizamos principalmente la duración, mientras que dentro de las las características espectrales incluimos los valores de frecuencia de los dos primeros formantes, el área del espacio acústico, la dispersión y la centralización de los formantes. Con relación a la primera pregunta, se espera que el incremento de la velocidad de elocución provoque una disminución de la duración e, inversamente, que la lectura en velocidad lenta incremente la duración vocálica. En cuanto a los formantes,

las investigaciones previas sugieren dos hipótesis opuestas. De un lado se afirma que, si la duración vocálica es breve como, por ejemplo, en velocidad de habla rápida, los formantes vocálicos no alcanzarán sus valores canónicos y terminarán por acercarse a los valores de las transiciones consonánticas (Lindblom, 1963; Lindblom, Sussman y Agwuele, 2009). Por otro lado, varios estudios sugieren que la velocidad de elocución modifica la duración, pero no produce cambios significativos en los valores de frecuencia de los formantes (Engstrand, 1988; Van Son y Pols, 1990, 1992; Fourakis, 1991; Brzezicha y Kul, 2014, Kul, 2015).

Se predice que existen otros factores que modifican la duración y las frecuencias formánticas e interactúan con la velocidad de elocución. Dentro estos factores están el timbre, el entorno consonántico, la tonicidad de la vocal, el tipo de palabra, la posición de la vocal dentro de la frase y el sexo del hablante. En cuanto al timbre vocálico se espera, por ejemplo, una mayor duración de la vocal abierta /a/, una menor duración para la vocal cerrada /i/ y valores intermedios en las vocales /e, o, u/ (Gendrot y Adda-Decker, 2007, p. 1419; Correa, 2017, p.75). Como vimos en la sección 2.2.4, las vocales átonas tienen menor duración y un valor de F1 más bajo que las tónicas. Así que se espera encontrar formantes más centralizados en las átonas producidas en velocidad rápida que en las tónicas producidas en velocidad normal y lenta.

En la misma dirección, se espera que las vocales de las palabras con contenido léxico tengan mayor duración y un F1 más alto, mientras que, según la hipótesis más aceptada (sección 2.2.5), las palabras funcionales tendrán menor duración y mayor centralización. La posición de la vocal dentro del grupo fónico no puede ignorarse, pues es bien sabido que las posiciones prosódicamente prominentes como el inicio de frase y el núcleo entonativo están asociadas con una mayor duración y, por tanto, predecimos que tienen vocales más periféricas y un menor grado de centralización. Finalmente, como se discutió en 2.2.7, se estima que las mujeres producirán vocales más largas y menos centralizadas que los hombres.

En cuanto al espacio acústico de las vocales, se predice que los hablantes modificarán el área de su espacio vocálico, el grado de dispersión de las vocales (entendida como la distancia euclídea entre cada timbre y el centro del espacio acústico del hablante) y la tasa de centralización en función de factores como la velocidad de elocución, la tonicidad

de la vocal y el sexo del hablante. De acuerdo con investigaciones recientes (Nadeu, 2014), se espera, por ejemplo, que en habla rápida el espacio vocálico tenga menor área y las vocales se desplacen hacia el centro del espacio acústico ($F1 \times F2$), mientras que en habla lenta se espera encontrar vocales más periféricas y un menor grado de centralización. Estas tendencias, de acuerdo con la bibliografía estudiada, deben ser más sobresalientes en el caso de las vocales átonas que en el de las tónicas y, desde un punto de vista sociofonético, estas modificaciones del espacio acústico deben ser más comunes en el habla masculina.

La segunda pregunta busca descubrir si hay una correlación entre la duración y los formantes vocálicos. No hay una hipótesis de partida sobre lo que ocurre en el español hablado en Bogotá. Sospechamos, a partir de la bibliografía (Nadeu, 2014), que puede haber una correlación positiva entre la duración y el primer formante en el caso de la vocal /a/, y una correlación negativa entre la duración y el segundo formante de las vocales posteriores. No obstante, también es posible que el análisis estadístico no arroje correlaciones significativas para ninguna de las vocales.

Finalmente, uno de los objetivos de la tesis es describir la variación fonética individual, en cuanto nos dará nuevos datos sobre las estrategias articulatorias que utilizan los hablantes para pronunciar las vocales del español bogotano en las diferentes condiciones estudiadas. En ese sentido, la implementación de los modelos de efectos mixtos y, en particular, la posibilidad de incluir efectos aleatorios, permite describir cómo afecta la variación fonética individual sobre cada uno de los parámetros acústicos.

4.2 Procedimiento de análisis

Se construyeron modelos lineales de efectos mixtos con los dos primeros formantes y la duración vocálica como variables de respuesta. Como factores fijos, se incluyeron el timbre (/i, e, a, o, u/), la tonicidad (tónica vs. átona), la velocidad de elocución (normal, lenta y rápida), el punto de articulación de la consonante precedente C_1 , el tipo de palabra (contenido vs. función), el sexo del hablante, la posición de la vocal dentro de la palabra y dentro del grupo fónico. Los modelos incluyeron interceptos aleatorios para los ítems léxicos y para los hablantes, además de pendientes aleatorias para la velocidad de habla. Se realizaron modelos para los dos primeros formantes teniendo en cuenta los valores

absolutos en hercios, y modelos complementarios con los valores de frecuencia normalizados con el método de Lobanov (1971).

Para estudiar el espacio acústico y la centralización, se transformaron las frecuencias de los formantes a la escala Bark y se construyeron modelos con el área del espacio vocálico (AEV), la dispersión y la tasa de centralización de los formantes (TCF) como variables respuesta o variables dependientes. Como factores fijos se incluyeron la tonicidad, la velocidad de habla y el sexo del hablante. El modelo también consideró interceptos para los hablantes y pendientes aleatorias para la velocidad de habla.

4.3 Resultados según los parámetros analizados

A lo largo de este y los capítulos siguientes, la visualización de los coeficientes de los modelos mixtos se hará por medio de diagramas de puntos. En la Figura 21, por ejemplo, el intercepto del modelo tiene un valor de cero y una línea vertical en el centro de la abscisa. Cada variable independiente se representa con un punto, y la magnitud del efecto depende de la proximidad de cada punto al intercepto: las variables que se solapan con el intercepto no son significativas, y las que se encuentran a mayor distancia (positiva o negativa) producen más cambios sobre la variable respuesta. Además, cada modelo está acompañado de una tabla que resume los valores estimados para cada factor, los estadísticos (el error estándar, el valor de t , y el valor de p) y los valores correspondientes a los factores aleatorios y a los residuos del modelo. Finalmente, las tablas incluyen la varianza explicada (R^2) por los factores fijos (varianza marginal) y la varianza explicada por el modelo de acuerdo según los planteamientos de Nakagawa y Schielzeth (2013).

4.3.1 Duración

El modelo de la duración tiene un intercepto de 89.73 ms, que corresponde a la duración de una /a/ átona, ubicada al final del grupo fónico, en contexto dentoalveolar, y producida en velocidad lenta por el grupo de mujeres. De acuerdo con la Tabla 3 y la Figura 21, la duración está influida por la velocidad de elocución ($\chi^2(2)=40.44$, $p < .001$). De acuerdo con los coeficientes de los factores fijos, el paso de velocidad lenta a velocidad normal produce un cambio de -12.54 ms (-14%), mientras que la diferencia entre la velocidad lenta y la rápida es de -17.47 ms (-19.4%). La tonicidad tiene un efecto altamente significativo sobre la duración ($\chi^2(1)=37.72$, $p < .001$) y una interacción con la velocidad

de elocución ($\chi^2(2)=235.17, p < .001$). En consecuencia, las tónicas son 17.55 ms (19.5 %) más largas que las átonas en velocidad lenta y en velocidad normal la diferencia disminuye a 11.55 ms (13 %). Las vocales cerradas /i, u/ presentan duraciones semejantes para las tónicas y átonas producidas en velocidad rápida. En el caso de /e, a, o/, la duración vocálica tiene diferencias significativas en todas las velocidades. La posición de la vocal dentro de la palabra no es estadísticamente significativa.

El timbre contribuye de manera altamente significativa al modelo de la duración ($\chi^2(4)=136.83, p < .001$). La vocal abierta /a/ es más larga en todas las velocidades, pues tiene una diferencia aproximada de -12 ms con respecto a las vocales /i, e, u/ y una diferencia de -8.7 ms con respecto a /o/. En el Anexo C, que presenta las mediciones discriminadas por velocidad, tonicidad y sexo, puede observarse que en el grupo de las tónicas la duración está relacionada con el grado de abertura: la vocal anterior cerrada /i/ es breve, la vocal abierta /a/ es larga, y las vocales /o, e, u/ tienen duraciones intermedias y con valores semejantes. En cambio, la duración de las átonas depende fundamentalmente de la velocidad de elocución. Como es de esperar, la posición de la vocal dentro de la frase también produce cambios significativos sobre la duración ($\chi^2(2)=148.98, p < .001$). En concreto, la duración en posición intermedia se reduce -9.59 ms respecto a posición final de frase, y -5.37 ms respecto a la posición inicial de palabra.

Las palabras funcionales se incrementan 8.94 ms con respecto a las palabras de contenido. Este resultado es justamente el opuesto al esperado. Así que se verificamos los interceptos aleatorios para los ítems léxicos, y comprobamos que las palabras funcionales tienen, como se espera, menor duración que las palabras con contenido léxico. Este resultado se debe, entonces, a que el intercepto del modelo corresponde a la vocal más larga /a/, a que dos hablantes produjeron vocales muy largas y, en general, a la alta variación interindividual en velocidad lenta. Las vocales son 6 ms más breves cuando la consonante precedente es una labial o una velar ($\chi^2(2)=31.77, p < .001$). Esto quiere decir que las vocales tienen mayor duración cuando la consonante precedente tiene un punto de articulación dentoalveolar. El sexo es una variable muy significativa dentro del modelo de la duración ($\chi^2(1)=9.33, p = .002$), en cuanto las mujeres producen vocales 5.85 ms más largas que los hombres.

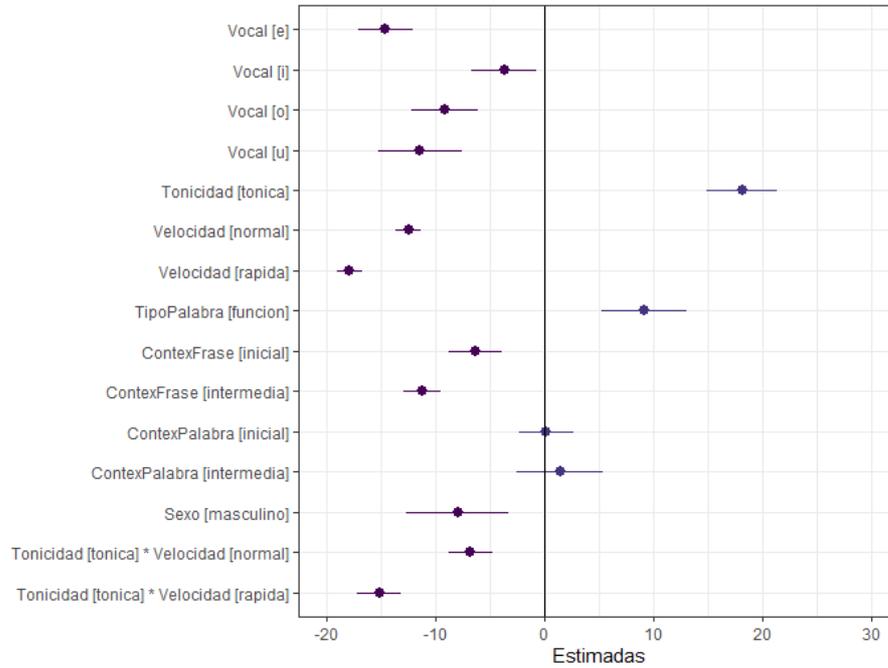


Figura 21. Coeficientes de los factores fijos con relación al intercepto del modelo de la duración vocálica.

Tabla 3. Coeficientes de los efectos fijos y de los efectos aleatorios fijos para un modelo de la duración vocálica con un total de 6517 observaciones y 22 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	89.73	3.60	28.26	<.001
/e/	-12.93	1.22	-10.52	<.001
/i/	-11.65	1.66	-7.00	<.001
/o/	-8.7	1.68	-5.19	<.001
/u/	-12.28	2.02	-6.06	<.001
Tonicidad (tónica)	17.55	1.06	10.39	<.001
Palabra (función)	8.94	1.88	4.74	<.001
Contexto Frase (inicial)	-5.37	1.17	-4.58	<.001
Contexto Frase (intermedia)	-9.59	0.81	-11.76	<.001
Velocidad (habla normal)	-12.54	1.91	-6.53	<.001
Velocidad (habla rápida)	-17.47	2.55	-6.83	<.001
Sexo (Hombres)	-5.85	1.66	-6.83	.001
PuntoC1labial	-6.00	1.71	-3.49	<.001
PuntoC1velar	-6.65	1.26	-5.23	<.001
Tonicidad(tónica): velocidad normal	-6.00	0.97	-6.16	<.001
Tonicidad(tónica): velocidad rápida	-15.06	0.98	-15.36	<.001
Factores aleatorios		Varianza	Std.dev	
Palabra		120.27	10.96	
Participante		133.98	11.57	
Velocidad lenta		22.56	4.75	
Velocidad rápida		14.27	3.77	
Residuos		215.07	14.66	
R ² Marginal = 0.32		R ² Condicional = 0.63		

Los factores aleatorios indican que la velocidad de elocución y los ítems léxicos dan cuenta de más del 50 % de la varianza de los datos. La Figura 22 ilustra los interceptos aleatorios por hablante y las pendientes aleatorias por velocidad. Es necesario aclarar que el diagrama no muestra los valores exactos de los interceptos sino la diferencia entre coeficientes del modelo y los coeficientes individuales en las tres velocidades. Como se ve, en velocidad normal y en velocidad rápida hay menor intervariabilidad en la duración vocálica. Los hablantes usan las duraciones en la velocidad lenta de manera diferente: un grupo de hablantes tiene preferencia por las vocales largas (hablantes BOG08, BOG10, BOG12 y BOG21), otros tienen preferencia por las vocales breves (hablantes BOG04, BOG06, BOG07, BOG22, BOG20) y, finalmente, un tercer grupo tienen duraciones con valores más cercanos al intercepto (hablantes BOG05, BOG09, BOG13-BOG19).

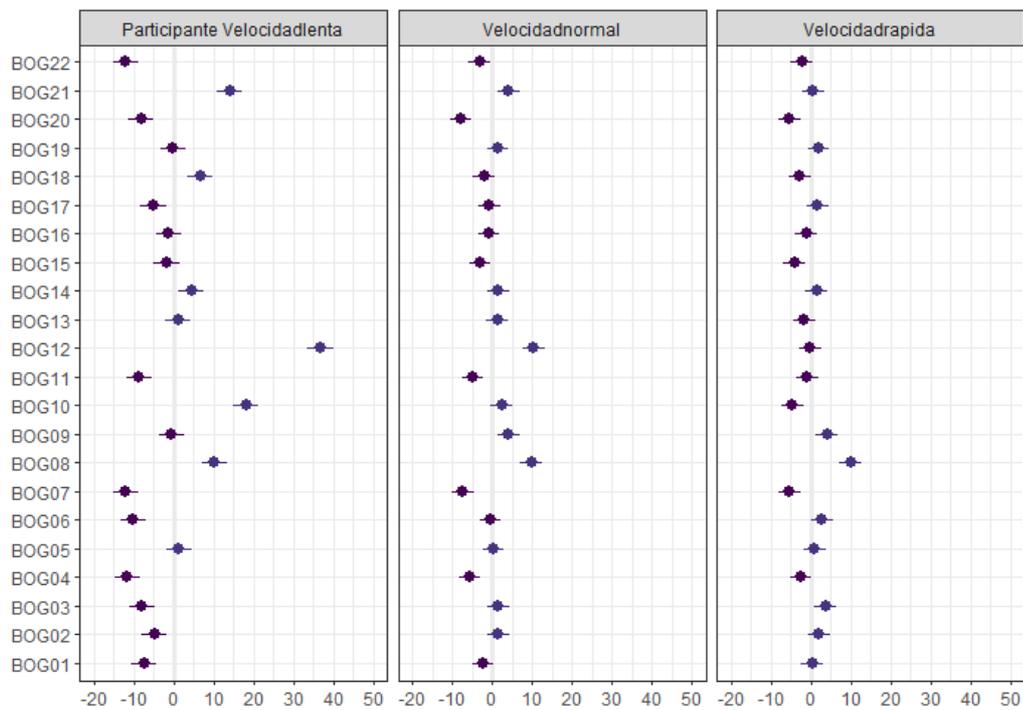


Figura 22. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo de la duración vocálica.

En este modelo para la duración vocálica la varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 32 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) es del 63 %.

4.3.2 Valores absolutos de F1 y F2

La Tabla 4 presenta los coeficientes de los factores fijos y de los efectos aleatorios del modelo del primer formante. La Figura 23 recoge los efectos de los factores fijos con relación al intercepto, el cual corresponde, a su vez, al valor de frecuencia de F1 de la vocal /a/ átona, ubicada al final del grupo fónico, en contexto dentoalveolar, y producida en velocidad lenta por el grupo de mujeres. Como podemos observar, el timbre vocálico es el factor que produce mayores cambios en los valores de frecuencia de F1 ($\chi^2(4)=3311$, $p < .001$). Las vocales medias /e, o/ tienen una diferencia de 160 Hz respecto a la vocal abierta /a/, mientras que /i/ es la vocal más cerrada con una diferencia de -296.53 Hz respecto al intercepto, seguida de /u/, con una diferencia de -256.71 Hz.

La tonicidad es un factor altamente significativo para el modelo del primer formante ($\chi^2(1)=24.51$, $p < .001$), lo cual se refleja en un incremento de 28.13 ± 4.30 Hz en las vocales tónicas. Como es de esperar, los valores absolutos del primer formante presentan diferencias muy significativas asociadas al sexo del hablante ($\chi^2(1)=30.37$, $p < .001$). Los hombres tienen un F1 que desciende -98.51 Hz respecto al primer formante de las mujeres, lo cual equivale a una diferencia entre sexos del 15.7 %. Encontramos una interacción entre la tonicidad y el sexo del participante que provoca una disminución del primer formante masculino en las vocales tónicas.

La velocidad de elocución tiene un efecto altamente significativo dentro del modelo ($\chi^2(2)=16.06$, $p < .001$). Si bien los cambios provocados en el valor de frecuencia de F1 son bajos comparados con los producidos por otros factores (Figura 23), resulta bastante interesante que no hay una diferencia significativa entre el valor de frecuencia de la velocidad lenta y la velocidad rápida. Esta similitud puede verse de manera más clara en los valores medios reportados en el Anexo D. En el modelo del primer formante también se observa un efecto significativo de la posición de la vocal dentro de la palabra ($\chi^2(2)=194.87$, $p < .001$), que produce un incremento de 50 Hz en posición inicial e intermedia de palabra o, lo que es igual, en posición final de palabra, las vocales se centralizan debido a un descenso de 50 Hz en el valor de frecuencia de F1.

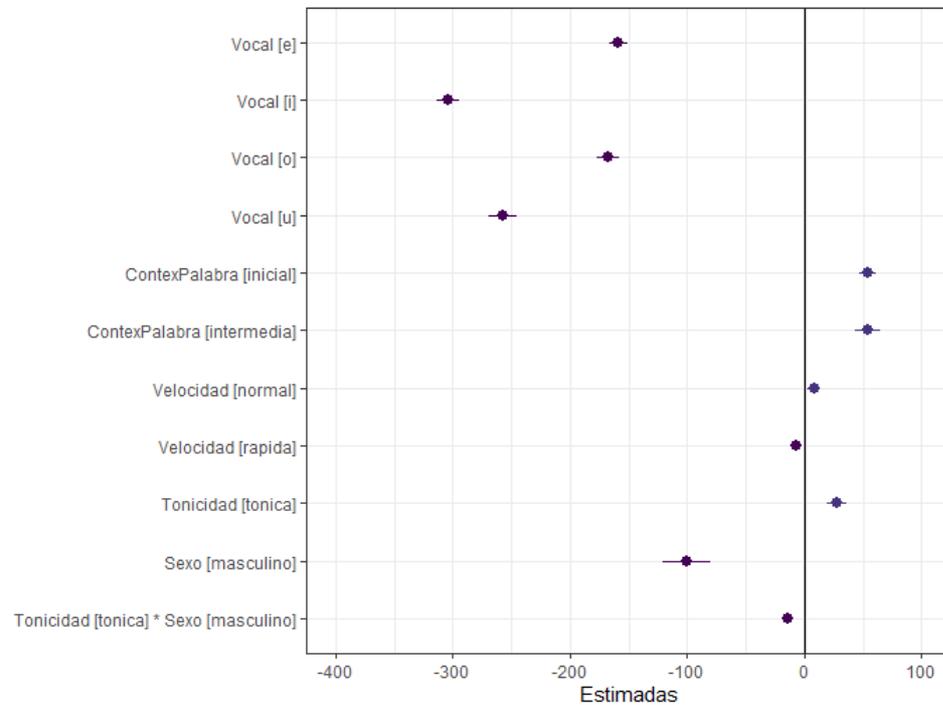


Figura 23. Coeficientes de los factores fijos incluidos con relación al intercepto del modelo de F1 (Hz).

Tabla 4. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F1 (Hz) con un total de 6517 observaciones y 22 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	623.47	9.35	66.62	<.001
/e/	-159.80	3.86	-41.44	<.001
/o/	-161.40	4.65	-34.66	<.001
/i/	-296.53	4.64	-63.88	<.001
/u/	-256.71	5.67	-45.24	<.001
Tonicidad (tónica)	28.13	4.30	-4.30	<.001
Contexto Palabra (inicial)	50.71	3.61	14.04	<.001
Contexto Palabra (intermedia)	49.19	5.58	8.81	<.001
Velocidad (habla normal)	9.20	2.75	3.33	.003
Velocidad (habla rápida)	-4.57	2.43	-1.87	.073
Sexo (masculino)	-98.51	10.24	-9.61	<.001
Tonicidad: Sexo (masculino)	-13.38	2.55	-5.24	<.001
Factores aleatorios		Varianza	Std.dev	
Palabra	791.9	28.14		
Participante Velocidad lenta	855.0	29.24		
Velocidad normal	583.6	24.16		
Velocidad rápida	916.0	30.27		
Residuos	2503	50.03		
$R^2_m = 0.74$		$R^2_c = 0.84$		

Ahora bien, el modelo incluyó interceptos aleatorios por participante y pendientes aleatorias por velocidad de elocución. Como se ve en la Figura 24, algunos hablantes presentan coeficientes menores a los estimados (BOG07, BOG12, BOG14, BOG18, BOG20, BOG22), un segundo grupo —la mayoría de los hablantes— produce un F1 con valores cercanos a los estimados, y un tercero tiene un primer formante más alto (vocales más abiertas) que el promedio (BOG02, BOG04, BOG13, BOG19, BOG21). Si bien hay una disminución de los coeficientes cuando se compara la velocidad lenta con la rápida, la mayoría de los hablantes produce vocales con valores de F1 similares en las tres velocidades. Finalmente, los interceptos aleatorios para el ítem léxico y los hablantes, y las pendientes aleatorias para velocidad de habla explican gran parte de la varianza observada (ver Tabla 4). La varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 74 %, mientras que la varianza explicada por el modelo (R^2_c) es del 84 %.

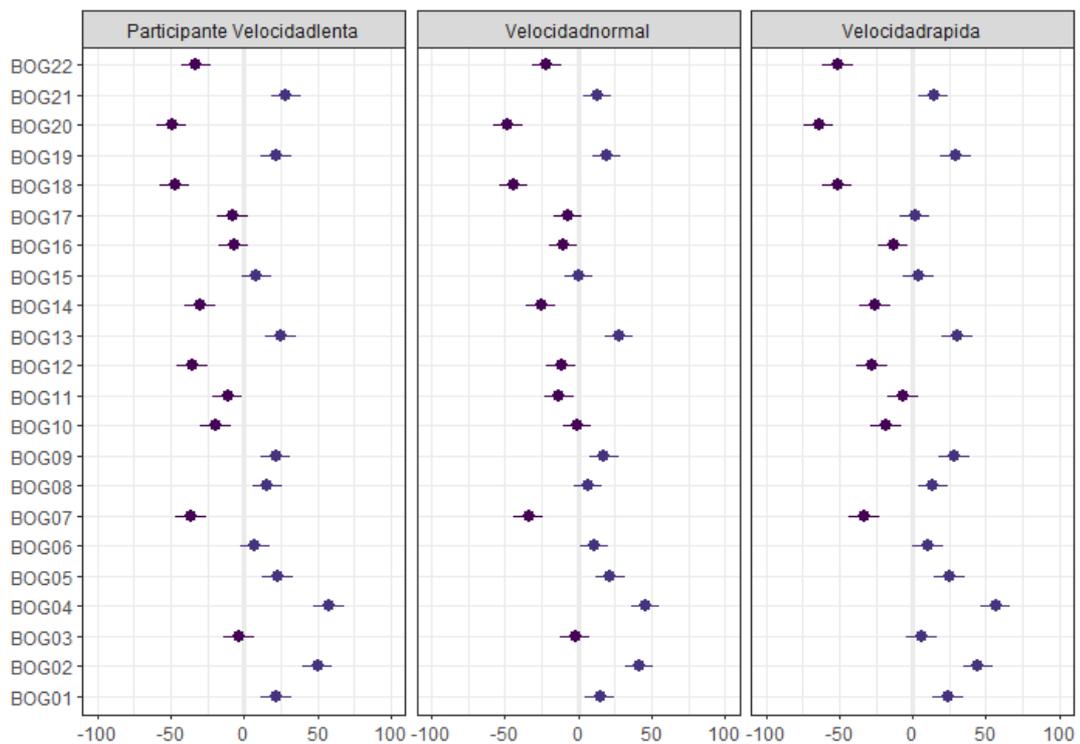


Figura 24. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de habla para el modelo de F1 (Hz).

El análisis del segundo formante pone de manifiesto que los valores de frecuencia están influidos por el timbre vocálico ($\chi^2(4)=3340$, $p < .001$). De acuerdo con los coeficientes del modelo (Figura 25 y Tabla 5), el intercepto corresponde al F2 de una /a/ átona,

ubicada al final del grupo fónico, en contexto dentoalveolar, y producida en velocidad lenta por el grupo de mujeres (1731.36 ± 23.25). De acuerdo con los coeficientes estimados, las vocales /e/, /o/ y /u/ tienen un F2 con valores relativamente cercanos al centro acústico (/e/ = 289.14 Hz, /o/ = -270.86 Hz, u = -384.31 Hz), mientras que /i/ tiene un segundo formante con 618.78 ± 14.42 Hz por encima del intercepto. Como es de esperar de los valores absolutos en hercios, el F2 tiene un valor más bajo en habla masculina (-263.20 Hz) y la diferencia entre hombres y mujeres es altamente significativa ($\chi^2(1)=41.33$, $p < .001$).

El punto de articulación de C₁ tiene un efecto altamente significativo sobre el F2 ($\chi^2(2)=188.71$, $p < .001$). Las consonantes labiales están asociadas a un F2 más bajo (-136.83 Hz) que las dentoalveolares, pero los cambios producidos por las consonantes velares no son diferentes a los característicos de las dentoalveolares ($M=-5.05$ Hz, $p = 0.66$). Como puede verse en la Figura 25 y en la Tabla 5, el tipo de palabra ($\chi^2(1)=6.04$, $p = .014$) y la tonicidad de la vocal ($\chi^2(1)=21.01$, $p < .001$) son efectos significativos, pero no producen grandes cambios en los valores de frecuencia de F2.

La Figura 26 muestra los interceptos aleatorios por hablante y pendientes aleatorias por velocidad de habla para el modelo de F2. Llama la atención que, si bien es predecible que algunos hablantes tengan valores de F2 por encima y por debajo del intercepto del modelo, los datos no se agrupan homogéneamente de acuerdo con el sexo del hablante: En el caso de la vocal /a/, por ejemplo, algunos hombres presentan valores por encima del intercepto (BOG21), y varias mujeres tienen interceptos con valores por debajo de las estimadas del modelo (BOG06, BOG14, BOG16).

El intercepto para los ítems léxicos explica parte de la varianza de los factores aleatorios (Tabla 5) y está asociado principalmente a palabras que incluyen vocales en el entorno de consonantes sordas o de una vibrante /r, r/. La varianza explicada por los factores fijos es del 75 %, mientras que la varianza explicada por el modelo es del 86 %.

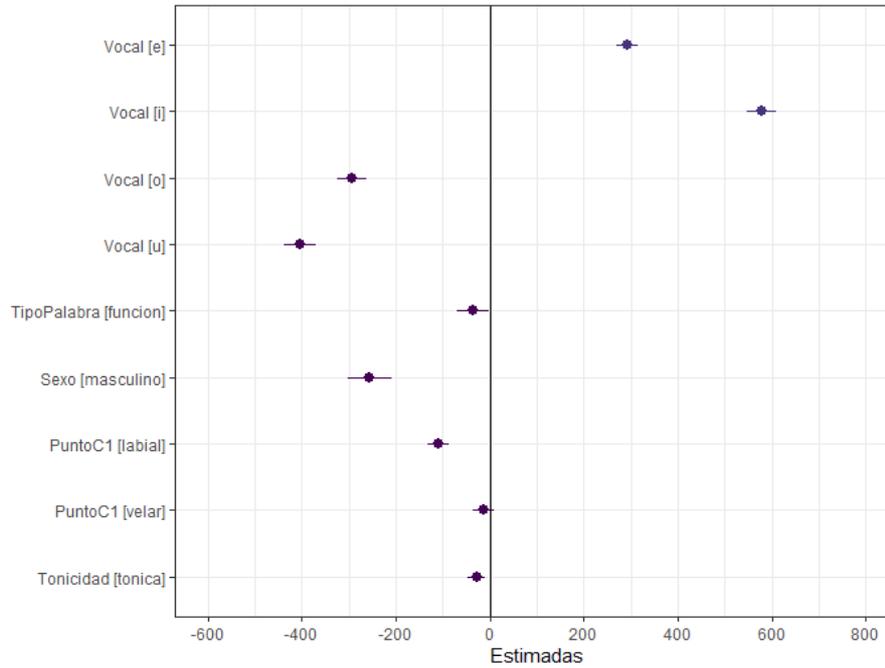


Figura 25. Coeficientes de los factores fijos con relación al intercepto del modelo de F2 (Hz).

Tabla 5. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F2 (Hz) con un total de 6517 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	1731.36	23.25	74.44	<.001
/e/	289.14	11.74	24.62	<.001
/o/	-270.86	14.58	-18.60	<.001
/i/	618.78	14.42	42.88	<.001
/u/	-384.31	15.96	-24.07	<.001
Tonicidad (tónica)	-42.15	9.21	-4.57	<.001
Palabra (función)	-42.56	17.32	-2.45	.014
Punto art. C1 (labial)	-136.83	9.97	-13.71	<.001
Punto art. C1 (velar)	-5.05	11.74	-0.43	0.66
Sexo (masculino)	-236.20	22.55	-11.66	<.001
Factores aleatorios		Varianza	Std.dev	
Palabra	1300	114.04		
Participante	2914	53.98		
Velocidad lenta	2883	53.13		
Velocidad normal	3583	59.86		
Velocidad rápida				
Residuos	21120	145.33		
$R^2_m = 0.75$		$R^2_c = 0.86$		

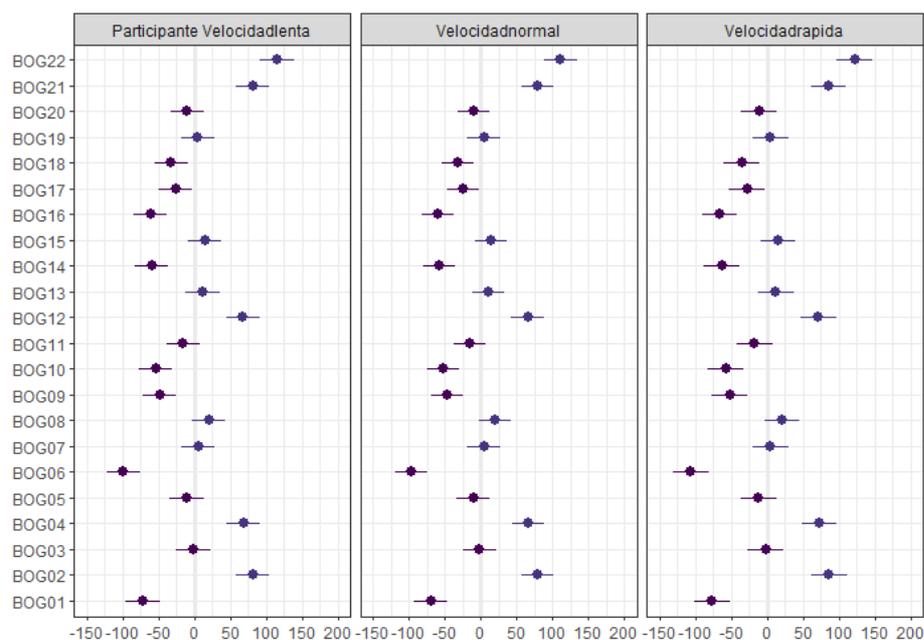


Figura 26. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo del formante F2 en Hz.

La distribución de los valores de frecuencia de F1 y F2 refuerza los resultados de los modelos de efectos mixtos y refleja tendencias adicionales (Figura 27 y Figura 28). En primer lugar, nótese que el valor medio de F1 de /u/ es bastante cercano al primer formante de la vocal media /o/ en todos los hablantes y en las tres velocidades. Estas vocales tienen desviaciones típicas altas para F2 (Figura 28), lo cual sugiere una fuerte influencia del entorno consonántico. En segundo lugar, la vocal abierta /a/ tiene realizaciones abiertas, medio abiertas y centralizadas [a, a̟, ə] en la mayoría de los hablantes. De hecho, en la Figura 27 se observa que el valor de frecuencia de F1 tiene dos centros de gravedad en habla femenina. En cambio, los hombres tienen una preferencia por una realización más centralizada de /a/. Esto también se refleja en la distribución de las desviaciones típicas. La Figura 28 ilustra, por ejemplo, que las desviaciones de los valores de F1 de /a/ son más altas y ocupan un mayor rango de valores en el habla femenina. En los hombres predomina, a juzgar por la distribución de los datos, una realización centralizada de /a/. La vocal /i/ tiene una realización periférica y varía a lo largo del segundo formante (Figura 28). En velocidad lenta y en velocidad normal, la vocal cerrada /i/ se mantiene completamente diferenciada de la vocal media /e/; en velocidad rápida tienen realizaciones muy cercanas.

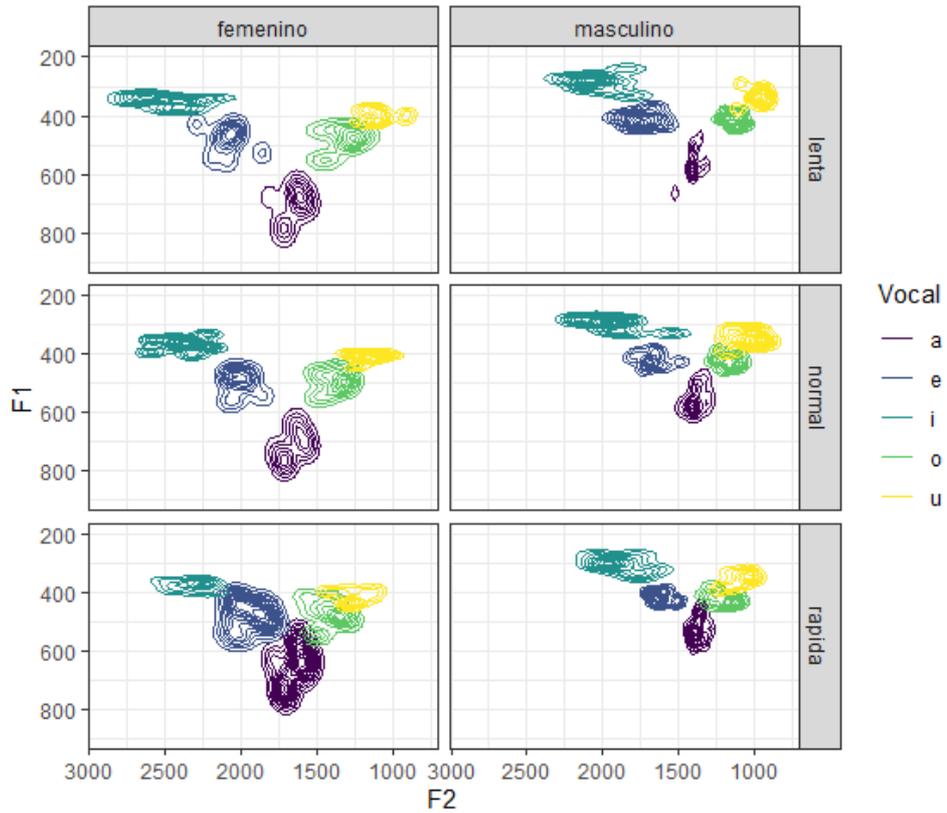


Figura 27. Contorno bidimensional de F1 y F2 (Hz) para hombres y mujeres en tres velocidades de elocución.

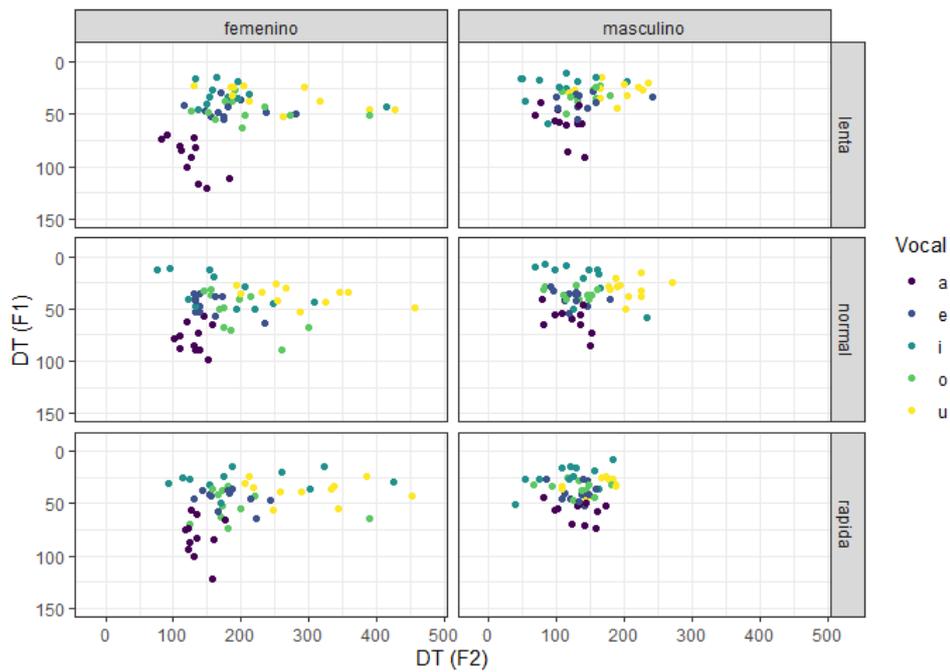


Figura 28. Desviación típica de F1 y F2 para hombres y mujeres en tres velocidades de elocución.

4.3.3 Valores normalizados de F1 y F2

Se construyeron modelos estadísticos con los valores normalizados de F1 y F2. Los coeficientes del F1 indican que el timbre tiene un efecto altamente significativo sobre la variable dependiente ($\chi^2(4)=3814.53$, $p < .001$). Al igual que los datos en hercios (Hz), hay una diferencia significativa entre el F1 de la velocidad lenta y la velocidad normal ($\chi^2(2)=19.87$, $p < .001$). El análisis también pone de manifiesto que el F1 está influido por la tonicidad de la vocal ($\chi^2(1)=17.41$, $p < .001$) y su posición dentro de la palabra ($\chi^2(2)=267.46$, $p < .001$). Esto confirma que las vocales átonas y las vocales en posición final de palabra tienen un valor más bajo de F1 y, en general, son más centralizadas.

La varianza del modelo está asociada a los interceptos aleatorios para los participantes e ítems léxicos, a las pendientes aleatorias para la velocidad de elocución y a los residuos. La varianza explicada por los factores fijos o varianza marginal (R^2_m) fue del 76 %, mientras que la varianza explicada por el modelo (R^2_c) fue del 83 %. Esto quiere decir que los factores que modifican el valor de F1 son los mismos para los dos tipos de datos, si exceptuamos el sexo del hablante, que está controlado en los datos normalizados.

En el modelo del segundo formante, el timbre tiene un efecto altamente significativo ($\chi^2(4)=3685.28$, $p < .001$), la velocidad de habla no modifica los valores de frecuencia de F2 ($\chi^2(2)=4.82$, $p = .090$) y la tonicidad ($\chi^2(1)=6.46$, $p = .011$) y el tipo de palabra ($\chi^2(2)=7.57$, $p = .006$) tienen efectos significativos. El contexto consonántico modifica los valores de F2 ($\chi^2(2)=61.31$, $p < .001$), pues las vocales en contexto labial se caracterizan por tener un valor más bajo que el dentoalveolar. Las palabras en posición inicial presentan un segundo formante bajo al inicio de palabra, y este efecto es altamente significativo para el modelo ($\chi^2(2)=27.67$, $p < .001$).

La varianza del segundo formante está asociada a los interceptos aleatorios para los participantes e ítems léxicos, a las pendientes aleatorias para la velocidad de elocución y a los residuos. La varianza explicada por los factores fijos (R^2_m) es del 76 %, mientras que la varianza explicada por el modelo (R^2_c) es del 87 %. Los factores que modifican el valor de F2 son los mismos para los dos modelos, si exceptuamos el sexo del hablante, que está controlado en los datos normalizados, y la posición de la vocal dentro de la palabra, que solamente es significativo en el modelo con valores normalizados.

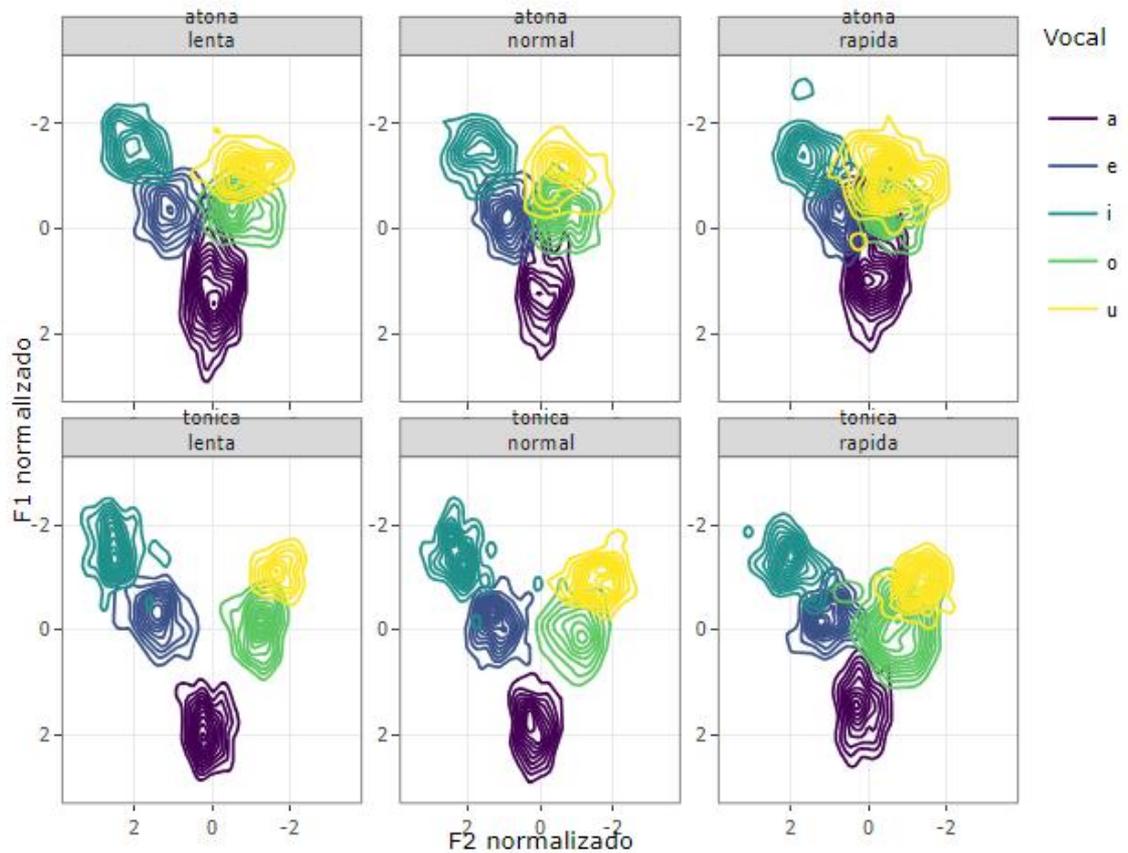


Figura 29. Contorno bidimensional de F1 y F2 (normalizados) para las vocales tónicas y átonas producidas por hombres y mujeres en tres velocidades de elocución.

La Figura 29 presenta un contorno bidimensional de los valores normalizados de F1 y F2 para las vocales tónicas y átonas producidas por hombres y mujeres en tres velocidades de elocución. Este diagrama se observa que, desde un punto de vista acústico, las vocales tónicas están más separadas y definidas, especialmente, en velocidad normal y lenta. Las vocales átonas se desplazan hacia el centro del espacio vocálico, al punto que las vocales /o, u/ se superponen. Las consecuencias de estas variables las podemos observar, de un lado, en las vocales tónicas en velocidad lenta que representan las realizaciones más alejadas del centroide acústico, mientras que las vocales átonas en velocidad rápida están altamente centralizadas y los valores de frecuencia y los márgenes de dispersión de las vocales se solapan.

4.3.4 Correlación entre los valores temporales y espectrales

Con el objetivo de estudiar la relación entre las características espectrales y las características temporales de las vocales del español de Bogotá, se realizaron correlaciones entre la duración y los dos primeros formantes para cada timbre en las tres velocidades. Se verificaron las condiciones de aplicación de los diferentes coeficientes de correlación (Levshina, 2015, p. 126) y se optó por una prueba no paramétrica (el coeficiente ρ (*rho*) de *Spearman*) debido a que en algunas vocales hay una relación lineal pero no monótonica. La Tabla 6 recoge los coeficientes y la significatividad de las correlaciones entre la duración y los dos primeros formantes para las vocales del español en tres velocidades de habla.

Tabla 6. Coeficientes de correlación y significativa de las correlaciones para las vocales /i, e, a, o, u/ en tres velocidades de elocución

	V. lenta		V. normal		V. rápida	
Vocal	F1					
/a/	$r=0.46$	$p < .001$	$r=0.47$	$p < .001$	$r=0.42$	$p < .001$
/e/	$r= 0.11$	$p = .004$	$r= 0.24$	$p < .001$	$r= 0.19$	$p < .001$
/o/	$r= 0.14$	$p < .00$	$r= 0.26$	$p < .001$	$r= 0.16$	$p < .001$
/i/	$r= 0.11$	$p = .09$	$r= 0.18$	$p = .01$	$r= 0.20$	$p = .01$
/u/	$r= 0.21$	$p < .00$	$r= 0.08$	$p = .08$	$r= 0.14$	$p = .00$
	V. lenta		V. normal		V. rápida	
Vocal	F2					
/a/	$r= 0.27$	$p < .001$	$r= 0.30$	$p < .001$	$r= 0.32$	$p < .001$
/e/	$r= 0.35$	$p < .001$	$r= 0.33$	$p = 1$	$r= 0.27$	$p < .001$
/o/	$r=-0.19$	$p = 1$	$r= -0.04$	$p = .2$	$r= 0.03$	$p = .2$
/i/	$r= 0.46$	$p < .001$	$r= 0.40$	$p < .001$	$r= 0.43$	$p < .001$
/u/	$r= -0.49$	$p = 1$	$r= -0.57$	$p = 1$	$r= -0.43$	$p = 1$

Como puede verse en la Tabla 6 y en la Figura 30, las correlaciones entre la duración y el primer formante son, en general, débiles. La excepción es la vocal /a/ que tiene una correlación positiva moderada y estadísticamente significativa en las tres velocidades. Esto quiere decir que la frecuencia de F1 se incrementa —es más abierta y menos centralizada— cuando la duración aumenta. La correlación no es perfecta porque, como se sabe, las correlaciones son sensibles a los valores atípicos (Levshina, 2015, p. 124), y nuestros resultados muestran precisamente que los valores de los formantes y la duración de /a/ tienen un amplio rango de valores y altas desviaciones típicas en las tres velocidades.

Las correlaciones entre la duración y el segundo formante (Tabla 6 y Figura 31) son débiles o no significativas para las vocales /e/, /a/ y /o/. Los valores de F2 y de la duración presentan correlaciones negativas para /u/, pero ninguna es estadísticamente significativa. En cambio, para la vocal /i/ se obtuvieron correlaciones moderadas y estadísticamente significativas en las tres velocidades. Nuevamente, la correlación no tiene un coeficiente más alto debido a los valores atípicos, al amplio rango de valores y a las altas desviaciones típicas que arrojan los datos de /i/ en las tres velocidades de habla. Sin embargo, la correlación indica claramente que el valor de frecuencia de F2 y, por tanto, la anterioridad de /i/ se incrementa con la duración.

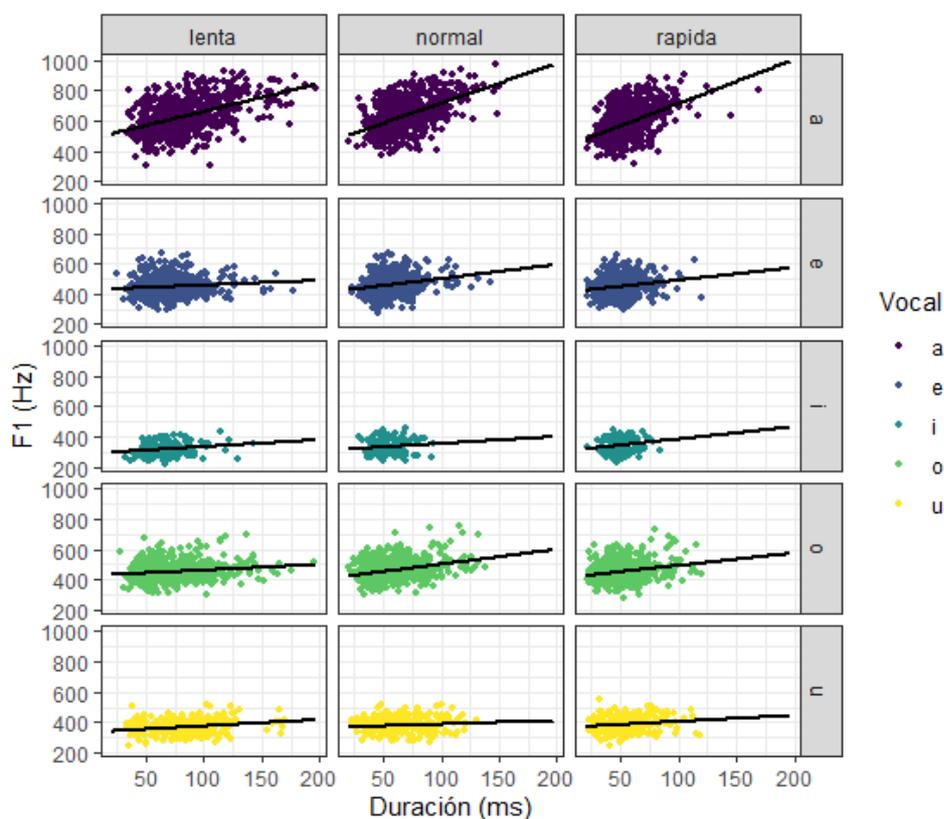


Figura 30. Correlaciones entre la duración y el primer formante para las vocales /i/ e /a o u/ en tres velocidades de elocución.

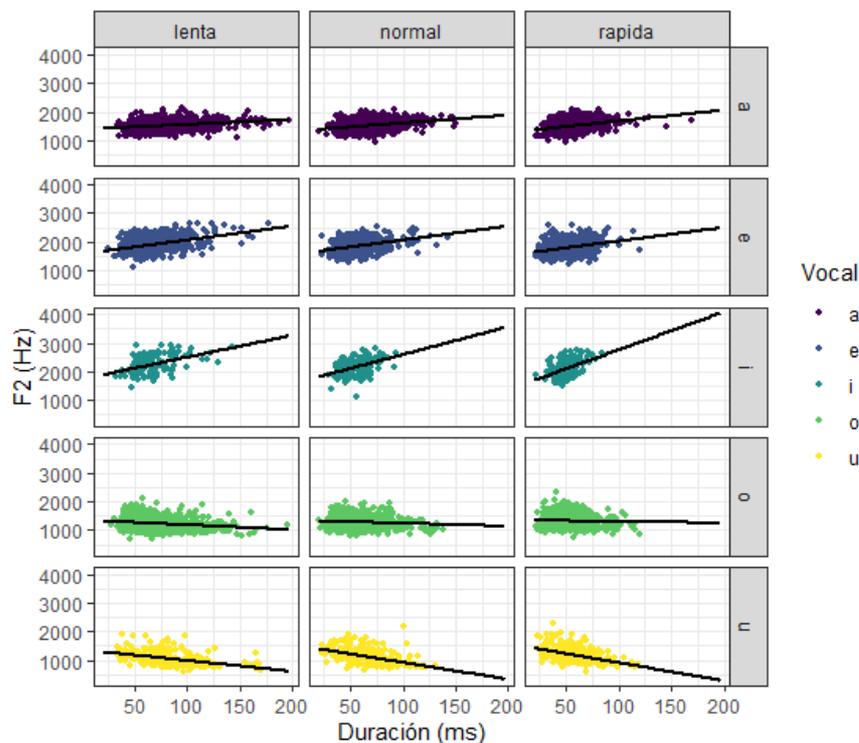


Figura 31. Correlaciones entre la duración y el segundo formante para las vocales /i, e, a, o, u/ en tres velocidades de elocución.

4.3.5 Área del espacio vocálico (AEV)

Se construyó un modelo mixto con el área del espacio vocálico (AEV) de cada hablante como variable dependiente (Tabla 8 y la Figura 32). Como factores fijos, se incluyó la velocidad de elocución, la tonicidad y el sexo del hablante. Además, se incluyeron interceptos aleatorios por hablante, pendientes aleatorias por velocidad y una interacción entre la tonicidad y la velocidad de elocución. El intercepto corresponde al AEV de las vocales átonas producidas con velocidad lenta por las mujeres. La tonicidad produce cambios altamente significativos en el AEV ($\chi^2(1)=179.58$, $p < .001$), ya que el AEV de las tónicas se incrementa 4.51 Bark^2 respecto del AEV de las vocales átonas. La velocidad de elocución también produce cambios muy significativos ($\chi^2(2)=37.20$, $p < .001$), que se reflejan en un área amplia para la velocidad lenta, un área intermedia para la velocidad normal (-1 Bark^2) y una menor área acústica con respecto a la velocidad rápida (-2 Bark^2). La interacción entre velocidad y tonicidad es muy significativa ($\chi^2(2)=9.36$, $p = .009$), lo cual significa que la diferencia entre el AEV de las átonas y las tónicas es mayor en velocidad rápida que en velocidad lenta.

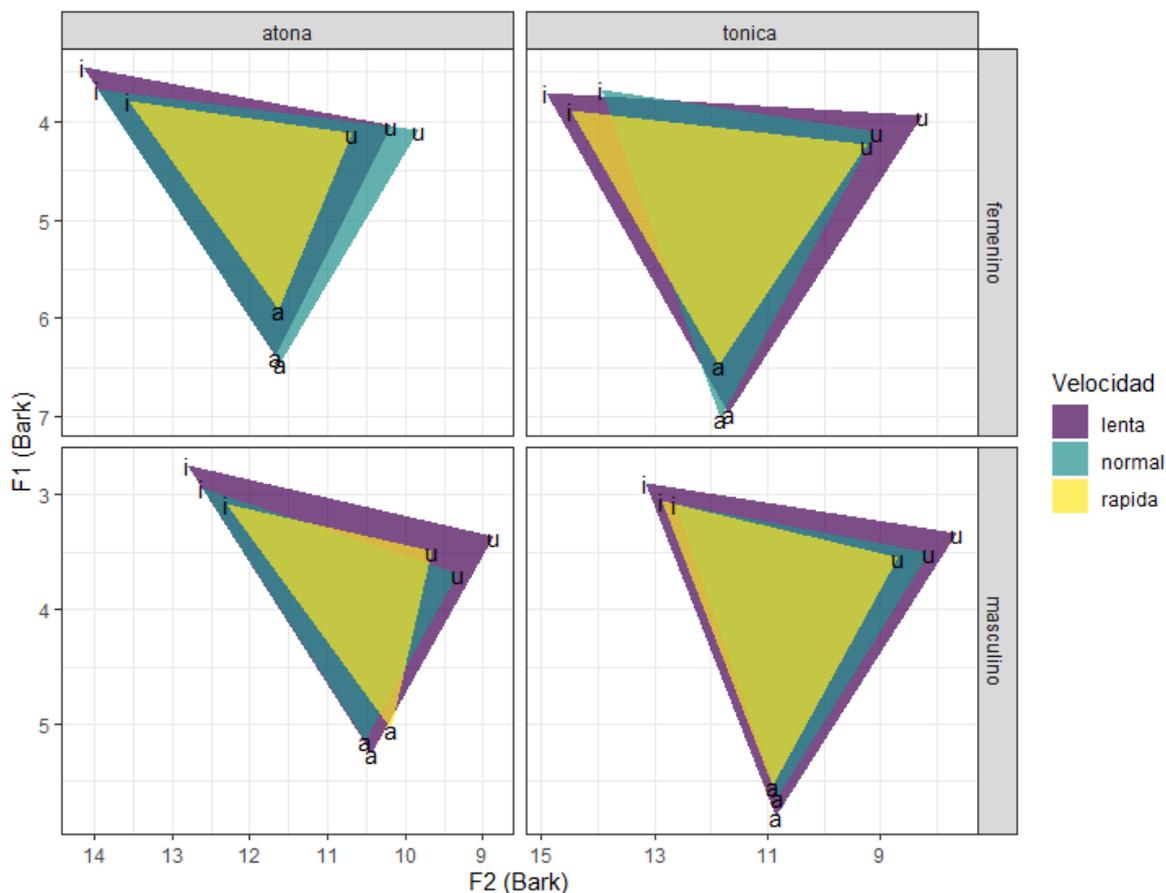


Figura 32. Área del espacio vocálico (AEV) en Bark² en función de la velocidad de habla, la tonicidad y el sexo del hablante.

Tabla 7. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de Área del espacio vocálico (AEV) con un total de 132 observaciones y 22 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	5.47	0.47	11.58	<.001
Tonicidad (tónica)	4.51	0.31	14.36	<.001
Velocidad (normal)	-1.03	0.32	-3.20	.001
Velocidad (rápida)	-2.09	0.35	-5.87	<.001
Sexo (masculino)	-1.72	0.56	-3.05	.005
Tonicidad(tónica):Velocidad(normal)	-0.19	0.44	-0.44	.660
Tonicidad(tónica):Velocidad(rápida)	-1.29	0.44	-2.91	.004
Factores aleatorios		Varianza	Std.dev	
Participante Velocidad lenta	2.063	1.43		
Velocidad normal	1.864	1.36		
Velocidad rápida	1.413	1.18		
Residuos	1.085	1.04		
$R^2_m =$	0.68	$R^2_c =$	0.87	

Las mujeres tienen un AEV con un valor de 1.72 Bark² por encima del AEV de los hombres. Esta diferencia es muy significativa ($\chi^2(1)=7.9$, $p=.008$), pero esto no quiere decir que todas las mujeres tengan un AEV amplia: algunas de ellas se caracterizan por tener un AEV con valores cercanos a los masculinos (por ejemplo, BOG18, BOG20, BOG22), mientras que algunos hombres (por ejemplo, BOG03, BOG05, BOG09, BOG15) tienen un área vocálica tan amplia como la femenina. En este sentido, nuestros resultados demuestran que las diferencias entre sexos son estadísticamente significativas, pero la variable sexo no es categórica, sino que depende de la actuación individual.

La Figura 33 presenta la variación fonética individual del AEV; en la ordenada pueden verse los hablantes en orden ascendente y en la abscisa el AEV expresada en Bark². En primer lugar, y como se mostró anteriormente, los cambios producidos por la velocidad de habla son menores a los que provoca la tonicidad de las vocales: todos los hablantes modifican a su manera el AEV en las diferentes velocidades de habla, pero todos, de manera sistemática, usan un espacio vocálico amplio cuando producen vocales tónicas y un espacio con menor área acústica cuando producen vocales átonas. En segundo lugar, los efectos aleatorios (Figura 34) sugieren que hay un grupo de hablantes (BOG01, BOG07, BOG11, BOG17, BOG18, BOG20) que usa una menor AEV en las tres velocidades de habla. Por esta razón, se ubican en el rango más bajo de valores obtenidos para todos los hablantes y por debajo de los valores predichos por el modelo de efectos mixtos.

La variación fonética individual se ve reflejada en los valores asociados con los factores aleatorios, pues gran parte de la varianza está relacionada con las pendientes para la velocidad de elocución. En particular, la varianza de la velocidad lenta en el valor de los residuos se duplica en este modelo. Finalmente, la varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 68 %, mientras que la varianza explicada por el modelo (R^2_c) es del 87 %.

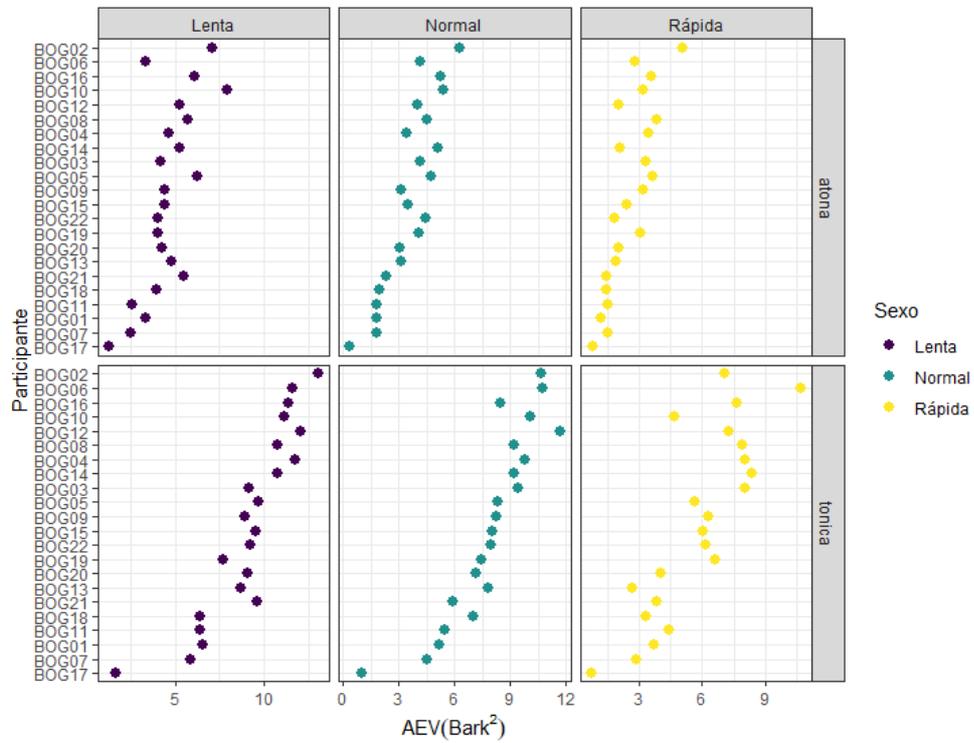


Figura 33. Área del espacio vocálico (AEV) en Bark^2 en función de la velocidad de habla, la tonicidad y el hablante.

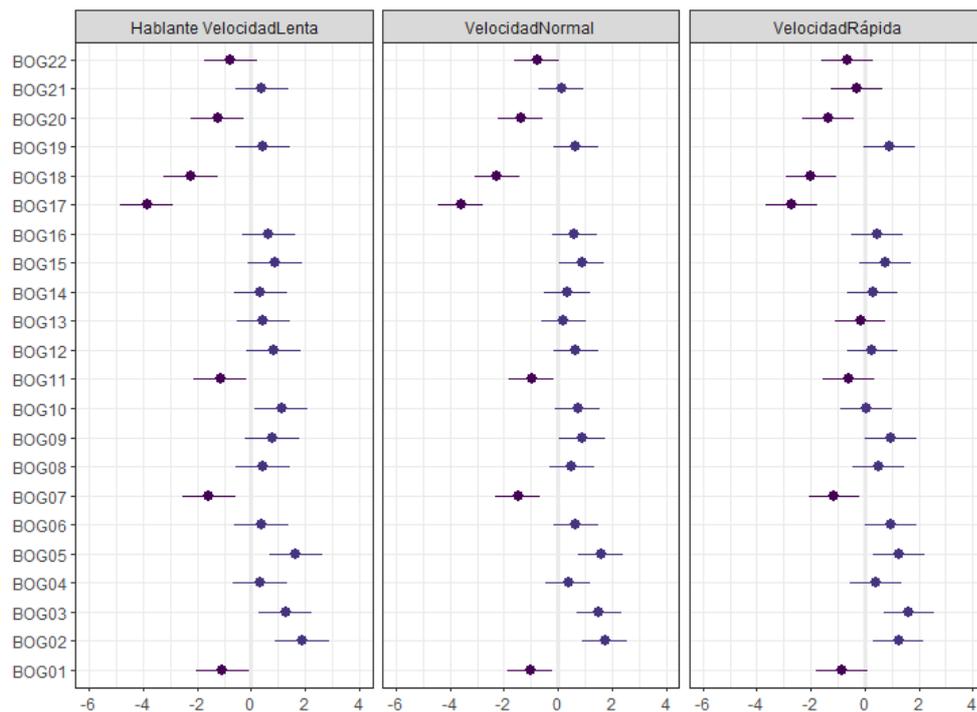


Figura 34. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo del Área del espacio vocálico (AEV).

4.3.6 Dispersión vocálica

Como se dijo en la sección 3.4.3, en esta investigación la dispersión vocálica se calcula por medio de la distancia euclídea entre el centroide de cada hablante (el valor medio de F1 y F2) y las frecuencias de las cinco categorías vocálicas (/i e a o u/) estudiadas. Las distancias se calcularon por categoría vocálica, pero también considerando la tonicidad y, por supuesto, la velocidad de elocución para cada uno de los 22 participantes. Teniendo en cuenta estos factores, se obtuvieron un total de 660 distancias (5 vocales x 22 hablantes x 3 velocidades x 2 condiciones para la tonicidad de la vocal). Con estos resultados se construyó un modelo con la distancia euclídea (expresada en escala Bark) como variable dependiente, y la tonicidad, la velocidad de habla y el sexo como predictores. Se incluyeron interceptos aleatorios para los hablantes y pendientes aleatorias para cada velocidad, e interacciones entre el timbre vocálico y la tonicidad, y entre la tonicidad y la velocidad de elocución.

Los resultados indican que el timbre vocálico tiene un efecto altamente significativo sobre la dispersión vocálica ($\chi^2(4)=1097$, $p < .001$). Como ilustra la Figura 35, las vocales que más se alejan del intercepto (vocal /a/ átona en velocidad lenta) son /i/ y /u/, mientras que las vocales medias /e/ y /o/ tienen pequeñas diferencias con /a/. De hecho, la diferencia no es estadísticamente significativa en el caso de la vocal posterior media /o/. Al igual que ocurre con otros parámetros analizados en este capítulo, la tonicidad produce cambios altamente significativos sobre la dispersión vocálica ($\chi^2(1)=526$, $p < .001$). Como puede verse en la Figura 35, este efecto depende del timbre: las tónicas tienen menor dispersión que las átonas, pero la diferencia es especialmente notoria en la vocal posterior /u/ y las vocales anteriores /i, e/. Por esta razón, la interacción entre tonicidad y tipo de vocal es significativa ($\chi^2(4)=169.93$, $p < .001$) para todas las vocales menos para /o/, que no presenta mayores diferencias entre tónicas y átonas.

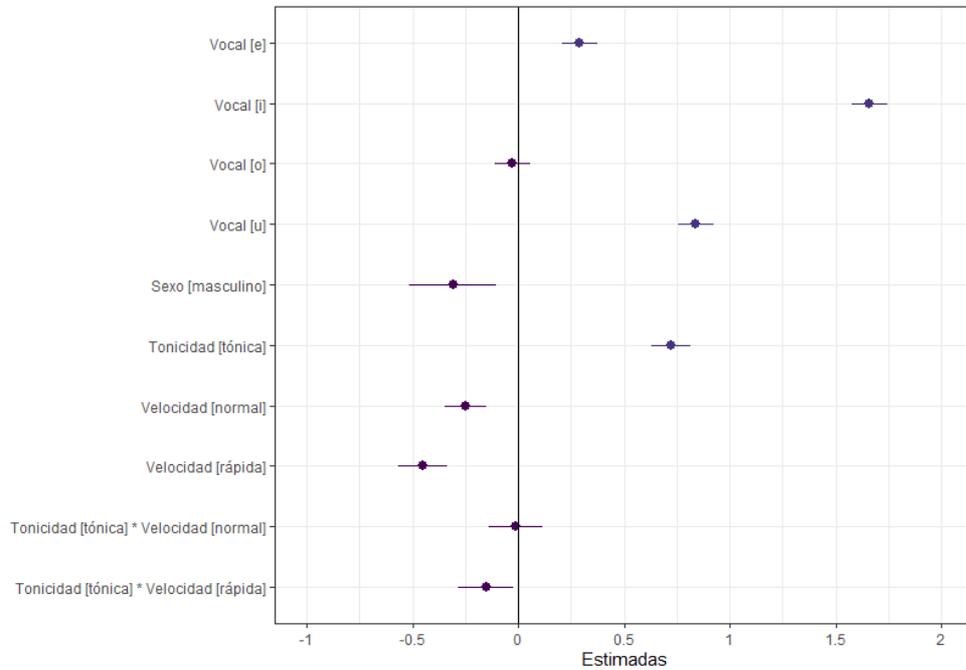


Figura 35. Coeficientes de los factores con relación al intercepto del modelo dispersión vocálica (Distancia euclídea).

Tabla 8. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la dispersión vocálica (distancia euclídea) con 660 observaciones y 22 hablantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	1.34	0.08	15.49	<.001
/e/	0.41	0.05	8.11	<.001
/i/	1.74	0.05	33.83	<.001
/o/	-0.01	0.05	-0.23	.81
/u/	0.51	0.05	9.90	<.001
Tonicidad (tónica)	0.68	0.06	11.18	<.001
Velocidad (normal)	-0.24	0.04	-5.56	<.001
Velocidad (rápida)	-0.44	0.05	-7.86	<.001
Sexo (masculino)	-0.30	0.10	-2.95	<.01
Tonicidad(tónica):Velocidad(normal)	-0.01	0.05	-0.20	.83
Tonicidad(tónica):Velocidad(rápida)	-0.15	0.05	-2.70	<.01
Vocal(/e/): Tonicidad (tónica)	-0.25	0.07	-3.54	<.001
Vocal(/i/): Tonicidad (tónica)	-0.17	0.07	-2.46	.01
Vocal(/u/): Tonicidad (tónica)	0.65	0.07	-8.96	<.001
Factores aleatorios	Varianza	Std.dev		
Participante Velocidad lenta	0.065	0.25		
Velocidad normal	0.066	0.25		
Velocidad rápida	0.066	0.24		
Residuos	0.088	0.29		
$R^2_m = 0.80$	$R^2_c = 0.88$			

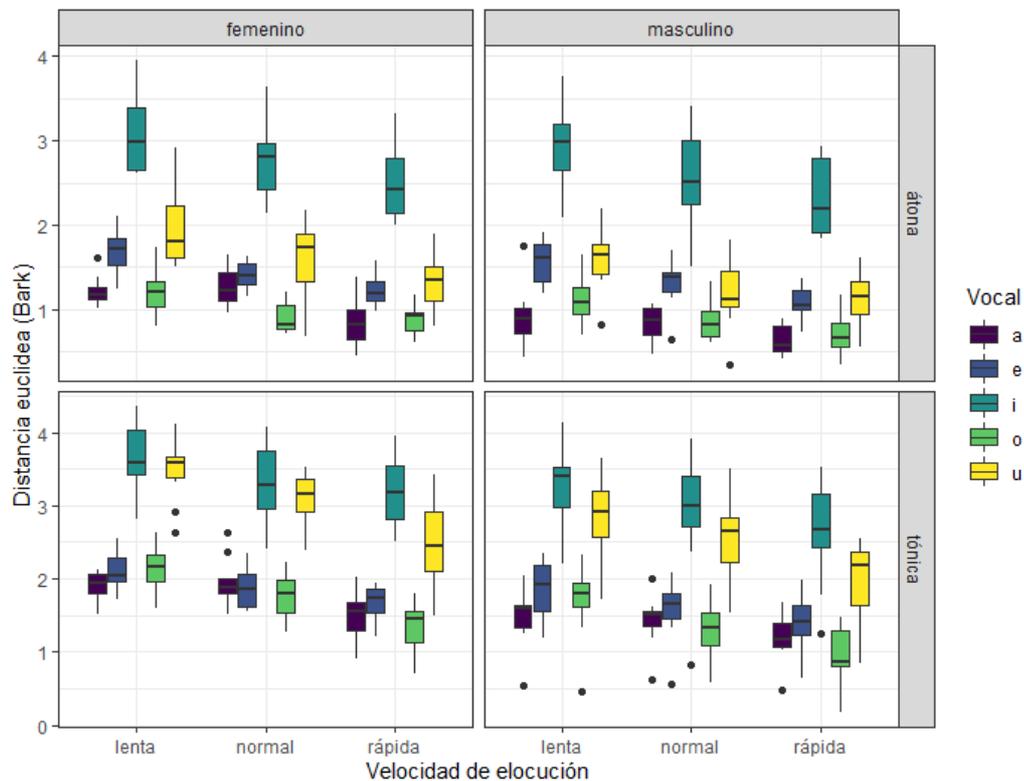


Figura 36. Dispersión vocálica en función de la velocidad de elocución, la tonicidad de la vocal y el timbre vocálico.

La dispersión vocálica disminuye a medida que la velocidad de elocución se hace más rápida ($\chi^2(2)=39.79$, $p < .001$). Las diferencias entre velocidades aumentan cuando se incluye en el modelo una interacción entre la velocidad y la tonicidad ($\chi^2(2)=9.01$, $p = .011$), pero este efecto solamente resulta estadísticamente significativo entre la velocidad lenta y la velocidad rápida. Eso quiere decir que la distancia euclídea entre tónicas y átonas tiene valores cercanos en velocidad normal y velocidad lenta.

Las mujeres tienen mayor dispersión vocálica que los hombres ($\chi^2(1)=7.23$, $p = .007$). En promedio, la dispersión de las vocales femeninas es 0.3 Bark más alta (147 Hz aprox.) que las vocales masculinas, es decir que las vocales producidas por las mujeres están más alejadas del centro acústico. Esto no quiere decir, naturalmente, que todos los participantes se comportan de la misma manera: un grupo de hombres y mujeres tienen interceptos por debajo de los valores predichos (BOG01, BOG04, BOG07, BOG11, BOG17, BOG18, BOG20), mientras que otro grupo de hablantes de ambos sexos presenta valores cercanos o más altos al intercepto de nuestro modelo (vocal /a/ del habla

femenina en velocidad lenta). Como se ve en la Figura 37, las variaciones individuales tienen la misma distribución en las tres velocidades.

La variación fonética individual se ve reflejada en los valores asociados con los factores aleatorios, pues, como recoge la Tabla 8, gran parte de la varianza está relacionada con las pendientes para la velocidad de elocución. En este modelo, la varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 80 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) es del 88 %.

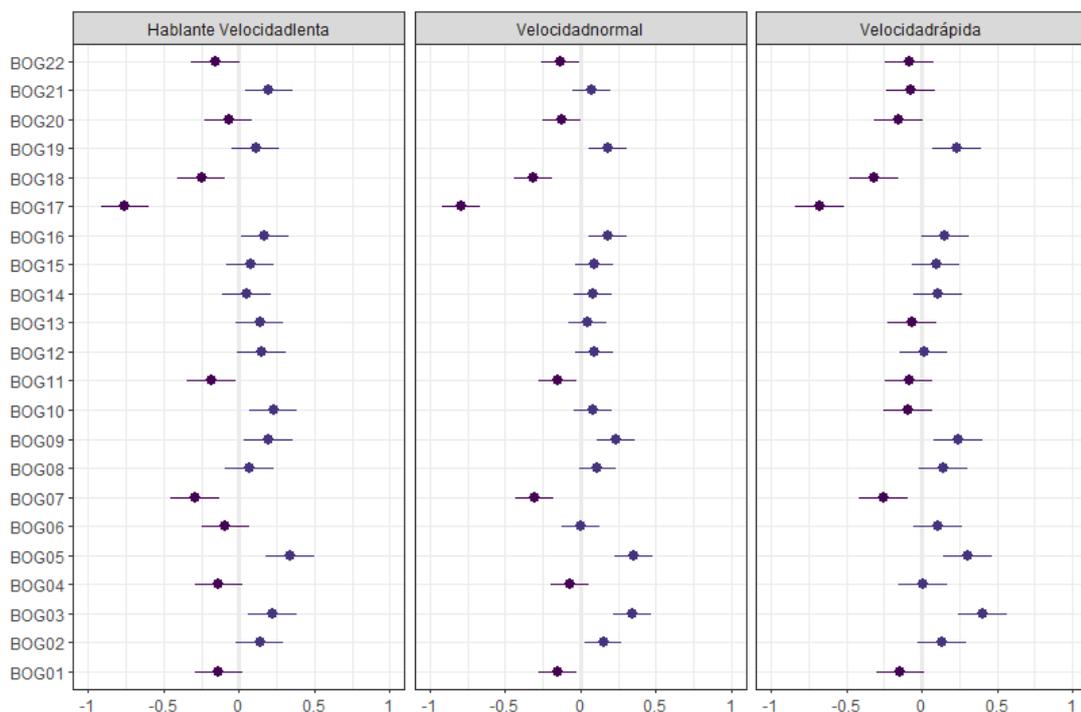


Figura 37. Interceptos aleatorios por hablante y pendientes aleatorias por velocidad de elocución para el modelo de la dispersión vocálica.

4.3.7 Tasa de centralización de los formantes (TCF)

Como se mencionó en la sección 3.4.3, la tasa de centralización de los formantes (TCF) es una medición normalizada del grado de centralización del espacio vocálico (Sapir, Raming y Spielman, 2010). Los valores varían entre 1 y 2, donde un valor de 1 está asociado a una realización hiperarticulada y uno de 2 indica la máxima centralización. El modelo estadístico incluye la TCF como variable dependiente y la tonicidad, la velocidad y el sexo como factores fijos. Como se observa en la Tabla 9 y en la Figura 38, las vocales tónicas son menos centralizadas que las átonas en todos los hablantes

($\chi^2(1)=153.56$, $p < .001$). La velocidad de elocución es un factor altamente significativo ($\chi^2(2)=38.35$, $p < .001$), pero las diferencias son pequeñas entre las diferentes velocidades de habla. Las diferencias entre sexos no son significativas para el grado de centralización ($\chi^2(1)=0.32$, $p = .570$).

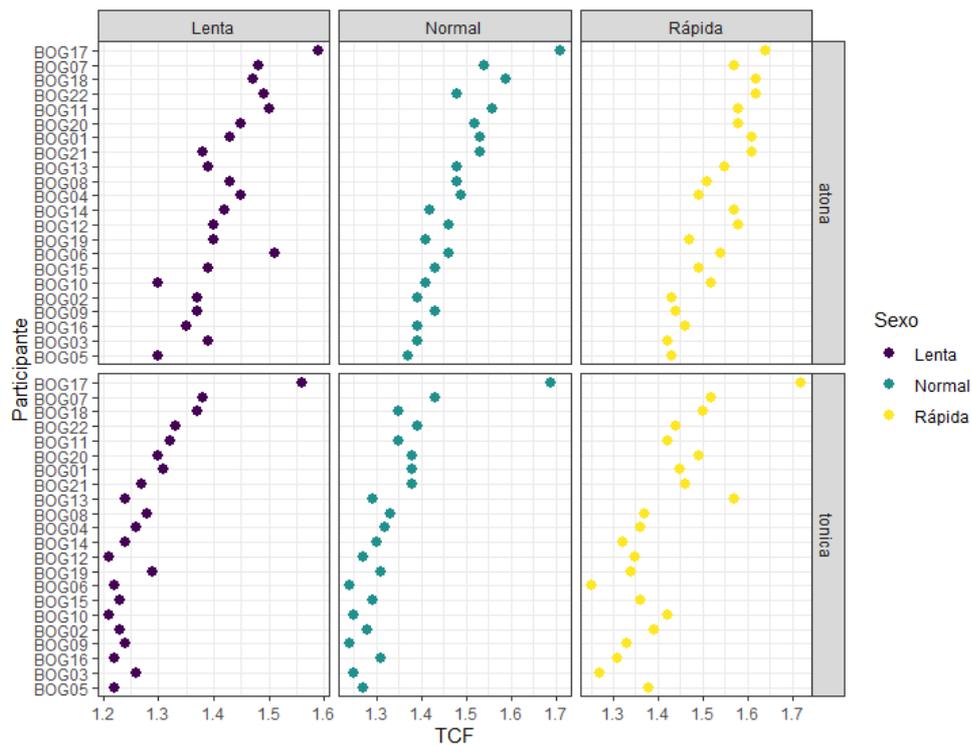


Figura 38. Tasa de Centralización de los Formantes (TCF) en función de la velocidad de habla, la tonicidad y el hablante.

Tabla 9. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la Tasa de Centralización de los Formantes (TCF) con un total de 132 observaciones y 22 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	1.41	0.020	68.28	<.001
Tonicidad (tónica)	-0.13	0.007	-19.36	<.001
Velocidad (normal)	0.05	0.009	5.38	<.001
Velocidad (rápida)	0.12	0.011	10.16	<.001
Sexo (masculino)	0.01	0.028	0.57	.56
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	0.003	0.062	
	Velocidad normal	0.006	0.078	
	Velocidad rápida	0.005	0.072	
Residuos		0.001	0.040	
$R^2_m = 0.51$		$R^2_c = 0.88$		

Los resultados sugieren que la TCF depende de la variación fonética individual al igual que ocurre con las dimensiones del área vocálica, pues, si bien las átonas y la velocidad rápida son, en general, las condiciones que determinan la centralización, las preferencias individuales de los hablantes se ubican a lo largo del continuo de hiperarticulación, lo cual es especialmente notorio en la distribución de las vocales tónicas (Figura 38).

La variación fonética individual se refleja en los valores de los factores aleatorios. En este modelo, gran parte de la varianza está relacionada con las pendientes para la velocidad de elocución: la varianza del habla normal y del habla rápida es cinco veces más alta que los residuos del modelo. En nuestro modelo para la TCF, la varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 51 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) es del 88 %.

4.4 Conclusiones parciales

La primera pregunta abordada en este capítulo es ¿cómo afectan las tres velocidades de elocución en las características temporales y espectrales de las vocales considerando el timbre vocálico, la tonicidad, el entorno consonántico, el sexo del hablante, el tipo de palabra (función vs. contenido) y la posición de la vocal dentro de la palabra y dentro del grupo fónico? A partir de los datos presentados en este capítulo podemos formular las siguientes conclusiones al respecto:

- 1) La velocidad de elocución modifica, como predecíamos, la duración de las vocales del español hablado en Bogotá. En velocidad lenta las vocales presentan una diferencia del 14 % con respecto a la velocidad normal y disminuyen el 19.4 % con respecto a la velocidad rápida. Las vocales tónicas tienen un incremento de la duración (19.5 %) con respecto a las átonas, a excepción de /i/, cuya duración no cambia con la tonicidad de la vocal. En cuanto al timbre, /a/ es la vocal más larga, seguida, de mayor a menor duración, por /o/, /e/, /u/ y /i/ (Anexo C). Otros de los factores que modifican la duración vocálica son el punto de articulación de la consonante precedente, con una mayor duración para las vocales precedidas por una consonante dentoalveolar. La posición de la vocal dentro de la frase también influye sobre los valores de duración, ya que las vocales son más largas en posición final que en posición inicial e intermedia. La

posición de la vocal dentro de la palabra no resultó significativa. En cambio, los hombres bogotanos producen, como se ha reportado en otras lenguas, vocales más breves que las mujeres. Los datos relacionados con el tipo de palabra son contrarios a nuestras expectativas, dado que las vocales tienen menor duración en las palabras con contenido léxico. Estos resultados no son concluyentes debido a que los interceptos de los ítems léxicos tienen una alta varianza y, en particular, porque dos hablantes produjeron vocales con duraciones muy largas en velocidad lenta.

- 2) La velocidad de elocución tiene un efecto estadísticamente significativo sobre los valores de frecuencia de F1, sin embargo, resulta bastante interesante que no hay una diferencia significativa entre el valor de frecuencia de la velocidad lenta y la velocidad rápida. Esto sugiere que los hablantes usaron un grado de apertura similar al habla lenta y normal. En todo caso, los cambios provocados por la velocidad de habla son pequeños en comparación con los efectos del timbre, la tonicidad, la posición de la vocal dentro de la palabra y el sexo del hablante. Así, las vocales medias /e, o/ mantienen, en promedio, a 160 Hz de la vocal abierta /a/. La vocal /i/ presenta un primer formante más bajo y, en ese sentido, su articulación es más cerrada y periférica que la de /u/. Las vocales tónicas tienen un F1 más alto que las átonas, pero esta diferencia es menor en los hombres que en las mujeres. Finalmente, el primer formante es más bajo al final de palabra que en posición inicial e intermedia. Estos resultados se obtuvieron tanto en el modelo con datos en hercios como en los datos normalizados, exceptuando los resultados relacionados con el sexo de los hablantes.
- 3) La velocidad de elocución no tiene ningún efecto sobre el segundo formante. El timbre, como es predecible, produce cambios dependiendo de la articulación de la vocal. Las vocales medias /e, o/ se mantienen a unos 208 Hz de la vocal central abierta /a/. La vocal /i/ tiene el segundo formante más alto y la vocal posterior /u/ tiene una realización que podríamos considerar avanzada y medio cerrada [ɯ]. Las vocales precedidas por una consonante con punto de articulación labial tienen un segundo formante más bajo que las dentoalveolares y velares. Otros factores, como el tipo de palabra y la tonicidad, son significativos dentro del modelo, pero no modifican significativamente los valores de frecuencia de F2.

En los datos normalizados se observan las mismas tendencias, pero, además, encontramos que las vocales tienen un segundo formante más alto al inicio de palabra.

- 4) El sistema vocálico del español hablado en Bogotá se caracteriza, entonces, porque tiene una vocal /u/ avanzada y medio cerrada; el primer formante tiene valores de frecuencia tan cercanos a los de /o/, que se pueden superponer en velocidad rápida. En segundo lugar, la vocal /a/ tiene realizaciones abiertas, medio abiertas y centralizadas [a, a̠, ə] en la mayoría de los hablantes. Las desviaciones de los valores de F1 de /a/ son más altas y ocupan un mayor rango de valores en el habla femenina. En los hombres predomina, a juzgar por la distribución de los datos, una realización centralizada de /a/. La vocal /i/ tiene una realización periférica y varía a lo largo del segundo formante. La vocal cerrada /i/ se mantiene completamente diferenciada de la vocal media /e/ en las tres velocidades.
- 5) El AEV cambia en función de la tonicidad, la velocidad de elocución y el sexo del hablante. El AEV, expresada en Bark², es más amplia en las vocales tónicas y en aquellas producidas con velocidad lenta, mientras es menor para las vocales átonas y el habla rápida. La tendencia general es que las mujeres tengan un sistema vocálico con un AEV más amplia que la usada por los hombres, pero cabe resaltar que esta diferencia no es categórica, de manera que algunos hombres tienen un espacio acústico tan amplio como el femenino y viceversa.
- 6) La velocidad rápida y la velocidad normal tienen menor dispersión que la velocidad lenta. No obstante, la dispersión de las vocales analizadas depende, en gran parte, del timbre y de la tonicidad. Las vocales más cercanas al centro acústico son /a/ y /o/, seguidas de /e/ y /u/. La vocal anterior cerrada /i/ es la vocal más alejada del centroide acústico, lo cual es consistente con los resultados del análisis de las frecuencias formánticas presentados anteriormente. Las tónicas tienen mayor dispersión —menos centralización— que las vocales átonas, lo cual es más sobresaliente en los resultados de las vocales /u/, /i/ y /e/.
- 7) La tasa de centralización de los formantes (TCF), que determina el grado de centralización de las vocales periféricas /i, a, u/, depende principalmente del

hablante pues, si bien las átonas y la velocidad rápida son más centralizadas en todos los hablantes, predominan las características individuales sobre otros factores.

Nuestra segunda pregunta, ¿cómo afecta la duración en los valores de los formantes vocálicos considerando el timbre de la vocal?, tiene como objetivo descubrir si hay una correlación entre los parámetros temporales y los parámetros espectrales. Sobre este punto podemos concluir que:

- 8) La correlación entre la duración y los valores de frecuencia de los formantes solamente es significativa para las vocales /a/ e /i/. En el caso de la vocal abierta /a/, a mayor duración observamos un primer formante más alto —con una articulación más abierta y menos centralizada—. En el caso de /i/, se obtuvieron correlaciones significativas entre la duración y F2 en las tres velocidades, es decir, un incremento de la duración de la vocal está acompañado de una articulación más anteriorizada. Los demás timbres vocálicos arrojan correlaciones débiles y/o no significativas.

Por último, nos preguntábamos cómo afectan las características fonéticas individuales en los resultados obtenidos para cada uno de los parámetros acústicos. Para responder esta pregunta se tuvieron en cuenta los interceptos aleatorios por hablante y las pendientes aleatorias por velocidad. Al respecto, podemos anotar que:

- 9) Al analizar la varianza de los factores aleatorios concluimos que los hablantes se agrupan en, al menos, tres grupos. Un primer grupo de hablantes, la mayoría, presenta coeficientes cercanos a los predichos por el modelo estadístico. El segundo grupo favorece las duraciones breves, valores de frecuencia de F1 bajos, tienen una AEV pequeña y un espacio acústico centralizado, esto es, una baja dispersión y una mayor TCF. El tercer grupo produce vocales más largas, abiertas (F1 alto) con una AEV amplia y una baja TCF. Es importante señalar que, si bien hay diferencias relacionadas con el sexo del hablante, la producción de vocales fonéticamente reducidas es común en ambos sexos.

5. Resultados: La influencia del entorno consonántico sobre los formantes vocálicos

5.1 Introducción

En este capítulo se reportan los resultados del análisis de los valores de frecuencia de los formantes de las vocales del español bogotano en función del entorno consonántico. Los datos corresponden, como en el capítulo anterior, a la lectura de cuatro textos producidos a tres velocidades de elocución por 22 hablantes. Para cuantificar el grado de coarticulación e inferir la velocidad de los movimientos articulatorios, el análisis incluye datos obtenidos en la parte estable de los formantes, el cambio o diferencia absoluta de los formantes entre el inicio y el centro de la vocal, y la velocidad de cambio de F1 y F2. Las preguntas de investigación que se responden en este capítulo son las siguientes:

- 1) ¿Cómo afecta el entorno consonántico en las variaciones espectrales de los dos primeros formantes de las vocales del español bogotano considerando el timbre, la tonicidad, el tipo de palabra, la velocidad de elocución, el sexo del hablante, la posición de la vocal dentro de la palabra y dentro del grupo fónico?
- 2) ¿Cómo afecta el comportamiento individual de los participantes en los resultados obtenidos para cada uno de los parámetros analizados?

En cuanto a la primera pregunta, se espera que el contexto consonántico tenga un efecto sobre el timbre vocálico pues, de acuerdo con los estudios revisados en el segundo capítulo (ver apartado 2.2.1), varias investigaciones muestran que la reducción vocálica, y particularmente en lenguas con reducción fonológica, es una consecuencia de la coarticulación de las consonantes sobre las vocales. No esperamos encontrar modificaciones sistemáticas como observadas en las lenguas germánicas (van Bergem, 1994; Mooshammer y Geng, 2008; Flemming, 2009), sin embargo, de acuerdo con la teoría acústica de la producción del habla, se espera que los valores de frecuencia sean más altos en el entorno de consonantes dentoalveolares y velares que en el entorno de

consonantes labiales (Harrington y Cassidy, 1999, p. 88-90). El primer formante ha recibido menos atención en la bibliografía, no obstante, de acuerdo con el estudio reciente de Romanelli y Menegotto (2018) sobre las vocales del español rioplatense, las vocales pueden tener un F1 más bajo en el contexto de las consonantes dentales y velares.

Con relación a la influencia de las consonantes sobre la trayectoria de los formantes vocálicos, se predice que las vocales que estén asimiladas al contexto consonántico tendrán valores que dependen del punto de articulación, lo cual se refleja en menores cambios espectrales y en menores diferencias entre los valores frecuencia en el inicio y en el centro de la vocal. La viabilidad de esta hipótesis está apoyada por varios estudios que han demostrado que el grado de coarticulación y el grado de precisión articulatoria —como la distinción entre habla hiperarticulada y habla espontánea— se puede determinar acústicamente calculando las variaciones dinámicas de los valores de frecuencia de los formantes (Moon y Lindblom, 1994; van Bergem, 1994; Ferguson y Kewley-Port, 2007; Herrmann, Cunningham y Whiteside, 2014).

También nos interesa saber si otros factores (timbre, tonicidad, el tipo de palabra, la posición de la vocal dentro de la palabra y dentro de la frase y el sexo del hablante) contribuyen al efecto que produce el entorno consonántico sobre los valores de frecuencia de los formantes. En cuanto al timbre, esperamos encontrar efectos significativos sobre los diferentes parámetros, pues es bien sabido que las transiciones formánticas están determinadas por la magnitud del desplazamiento de los articuladores desde el gesto consonántico hasta posición lingual de la vocal (Stevens, 1998, p. 377). En cuanto a la importancia de los factores prosódicos y extralingüísticos, en los capítulos anteriores, hemos presentado argumentos y resultados que justifican su inclusión dentro del análisis.

Por último, se describen las diferencias fonéticas individuales pues, como se mencionó en el capítulo anterior, estas permiten inferir estrategias articulatorias que utilizan los hablantes para contrarrestar los efectos de la duración vocálica. Recordemos que Kuehn y Moll (1976, p. 318) afirman que los hablantes pueden incrementar la velocidad de los movimientos linguales o disminuir la magnitud de los desplazamientos articulatorios cuando usan una velocidad rápida. Afortunadamente, estas modificaciones se pueden

estimar acústicamente por medio de parámetros como la velocidad de cambio de los formantes (Gay, 1978; Moon y Lindblom, 1994).

5.2 Procedimiento de análisis

La mayoría de los estudios que investigan el efecto del contexto consonántico sobre las vocales analizan listas de palabras o palabras dentro de frases marco con el objetivo de controlar la variabilidad a nivel estadístico. En nuestro caso, trabajamos con habla conectada producida en tres velocidades y un número importante de hablantes, así que nos enfrentamos ante una gran variabilidad y los datos pueden ubicarse en un amplio rango. Para contrarrestar estas condiciones y extraer generalizaciones más acertadas, los análisis estadísticos se realizan por timbre, y se elaboran modelos estadísticos con los datos en hercios y normalizados siguiendo el método de Lobanov.

Para determinar el efecto del entorno consonántico sobre el centro de la vocal se utiliza la técnica de agrupamiento de datos *k-means* (Harrington, 2010, p. 174) y la desviación típica en función del punto de articulación y el timbre vocálico (Stevens y House, 1963, p. 120; Recasens, 1985, p. 100; Recasens y Espinosa, 2006, p. 653). Para medir los cambios dinámicos de las transiciones formánticas se construyeron modelos mixtos con las diferencias absolutas ($\Delta F1$ y $\Delta F2$) y la velocidad de cambio de los dos primeros formantes ($rF1$ y $rF2$) como variables de respuesta. Como factores fijos se incluyeron el timbre (/i, e, a, o, u/), la tonicidad (tónica vs. átona), el punto de articulación de la consonante precedente, el tipo de palabra (palabra con contenido léxico vs. palabra funcional), la posición de la vocal dentro de la palabra y dentro del grupo fónico y, finalmente, el sexo del hablante. Los modelos incluyen interceptos aleatorios para los ítems léxicos y para los hablantes, y pendientes aleatorias para la velocidad de elocución.

5.3 Resultados según los parámetros analizados

5.3.1 Frecuencias de formantes en función del contexto consonántico

La Figura 39 ilustra la distribución de los valores de frecuencia de F1 y F2 en función de los fonemas /f, p, b, d, r, l, ʀ, t, s, k, g/, que constituyen las consonantes del contexto consonántico considerado en este trabajo. Los datos están clasificados en dos grupos tomando como referencia el valor medio del segundo formante de cada vocal; en la figura

se distinguen con el número 1 o 2. Este agrupamiento muestra, en primer lugar, que la vocal abierta /a/ tiene un segundo formante más bajo en el contexto de las consonantes labiales /f, p/. La clasificación también deja ver que las consonantes dentoalveolares y velares incrementan los valores de frecuencia de F2 y, en consecuencia, las realizaciones de /a/ son más anteriorizadas en estos contextos.

Las vocales posteriores siguen, en general, el mismo patrón: la vocal posterior /u/ tiene un segundo formante bajo en el entorno de las consonantes labiales, velares /f, p, k/ y en el contexto de la vibrante múltiple /r/; mientras que el segundo formante se centraliza cuando las vocales se ubican en el contexto de consonantes dentoalveolares /t, d, s/. La vocal posterior media /o/ también tiene un F2 con valores de frecuencia bajos en el contexto de las oclusivas labiales y velares /b, p, k/, mientras que las dentoalveolares /t, s, l, r/ incrementan los valores de frecuencia.

Las vocales anteriores /i, e/, de acuerdo con los datos que arroja este método, no siguen las tendencias descritas para los demás timbres: la distribución de los valores de frecuencia del segundo formante de /i/ sugieren que ninguna consonante o punto de articulación ejerce una influencia clara sobre esta vocal, pues la distribución de los formantes no sigue un patrón específico. Lo propio ocurre con la distribución de los valores del segundo formante de la vocal media /e/, si exceptuamos el incremento del F2 en el contexto de las consonantes velares /k, g/. Los demás valores de frecuencia de /e/ se distribuyen entre los 1500 y 2300 Hz sin agruparse claramente según la consonante o el punto de articulación.

Como se mencionó anteriormente, la desviación típica se usa en la investigación fonética como una medida para determinar la resistencia coarticulatoria, es decir, del grado de variabilidad de un fono en función de su contexto fonético (Recasens y Espinosa, 2009, p. 2288). Las desviaciones típicas largas indican una mayor variabilidad y las desviaciones pequeñas mayor grado de estabilidad y mayor resistencia coarticulatoria (Iskarous *et al.*, 2013, p. 1272). La Figura 40 muestra que las desviaciones típicas de los dos primeros formantes siguen, en general, el mismo patrón en las tres velocidades: El primer formante varía en función de la altura lingual, de manera que la variación es más amplia para la vocal /a/, las vocales /e/ y /o/ tienen valores intermedios y las vocales cerradas /u, i/ tienen una menor variación en el contexto de consonantes dentoalveolares y velares. En el contexto de las consonantes labiales, las vocales medias y cerradas tienen desviaciones bajas que varían entre los 51.31 y 57.67 Hz (Tabla 10).

El segundo formante de las vocales /a/ y /o/ en contexto dentoalveolar es menos variable para el F2 de /e/ y /u/ y, contrario a lo que sugiere la bibliografía, la vocal anterior cerrada /i/ es bastante variable en nuestro corpus. En el contexto de las consonantes labiales, la desviación típica es baja para las vocales posteriores y para la vocal abierta /a/, y es alta para los datos de las vocales anteriores: /o/ > /a/ > /u/ > /e/ > /i/. El contexto velar sigue el mismo patrón, salvo que en nuestro corpus no se incluyen casos de la vocal anterior cerrada /i/ en este entorno consonántico.

La desviación típica de F1, calculada por hablante (Figura 41), muestra que las vocales son más variables en los entornos consonánticos dentoalveolar y labial que en el contexto velar. En concordancia con los datos presentados anteriormente, el primer formante de las vocales /a/ y /e/ es el más variable, y es más o menos estable en el caso de las vocales /o, u, e, i/. Los hablantes presentan mayores diferencias en los datos correspondientes a las vocales /a/, /e/ y /o/ en el contexto consonántico dentoalveolar, y los resultados de las vocales cerradas son más estables y menos variables. Por otra parte, las desviaciones del segundo formante (Figura 42) son más altas y variables en el contexto dentoalveolar, especialmente en los valores de las vocales /u/, /e/, /i/. Es importante anotar que cuatro mujeres (BOG10, BOG12, BOG20, BOG22) producen vocales con desviaciones típicas por encima del promedio del primer y segundo formantes. Estos resultados explican, en parte, las desviaciones obtenidas al agrupar a todos los hablantes (Figura 40 y Tabla 10).

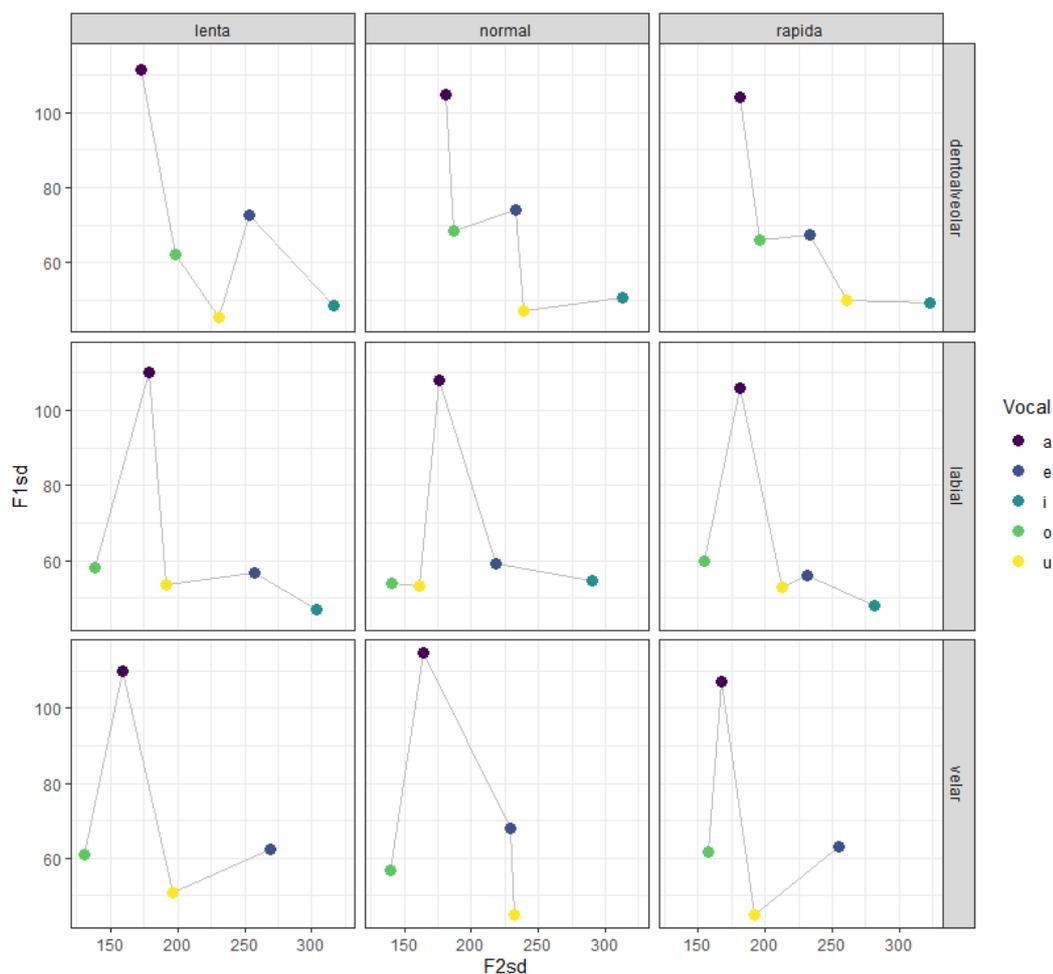


Figura 40. Desviaciones típicas de los valores frecuencia de F1 y F2 (en hercios) en función del punto de articulación y el timbre.

Tabla 10. Desviaciones típicas de los valores frecuencia de F1 y F2 (en hercios) en función del punto de articulación y el timbre.

	F1			F2		
	Labial	Dentoalveolar	Velar	Labial	Dentoalveolar	Velar
/i/	51.31	50.97		296.98	322.80	
/e/	57.67	71.35	64.70	245.18	244.95	252.85
/a/	110.10	108.08	112.89	179.47	178.62	163.56
/o/	57.16	65.77	59.83	151.45	196.43	147.86
/u/	53.37	48.65	47.09	195.89	251.69	215.32

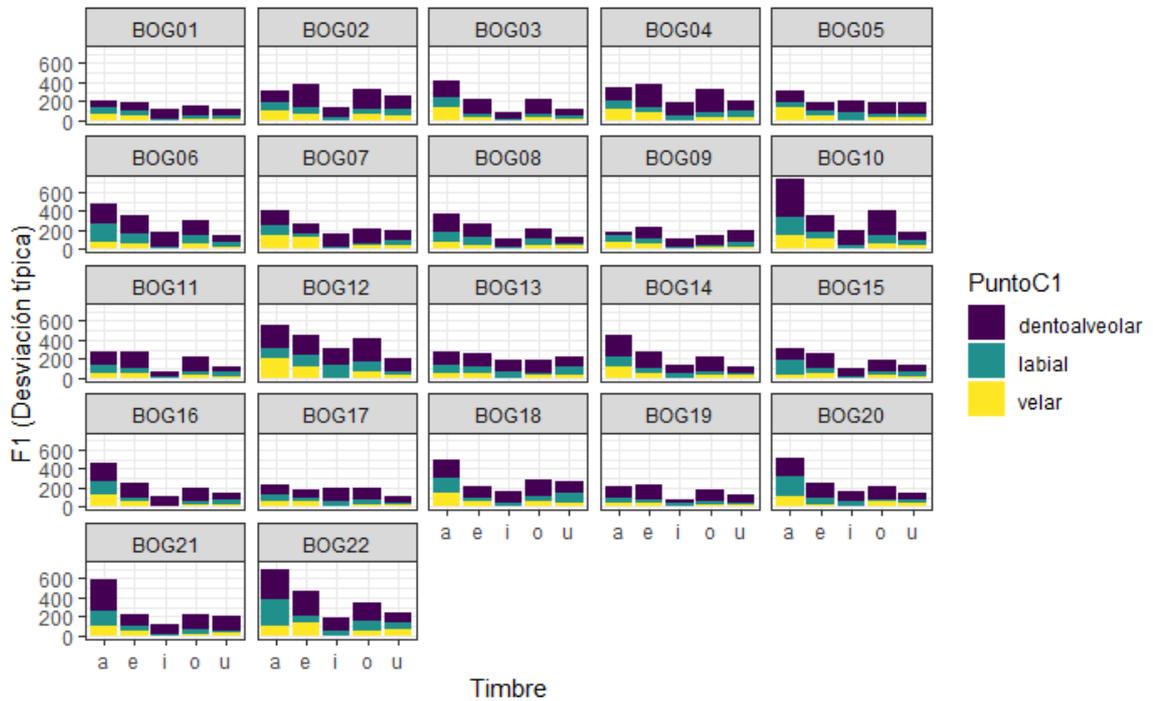


Figura 41. Desviaciones típicas de F1 (en hercios) en función del hablante, el punto de articulación y el timbre vocálico.

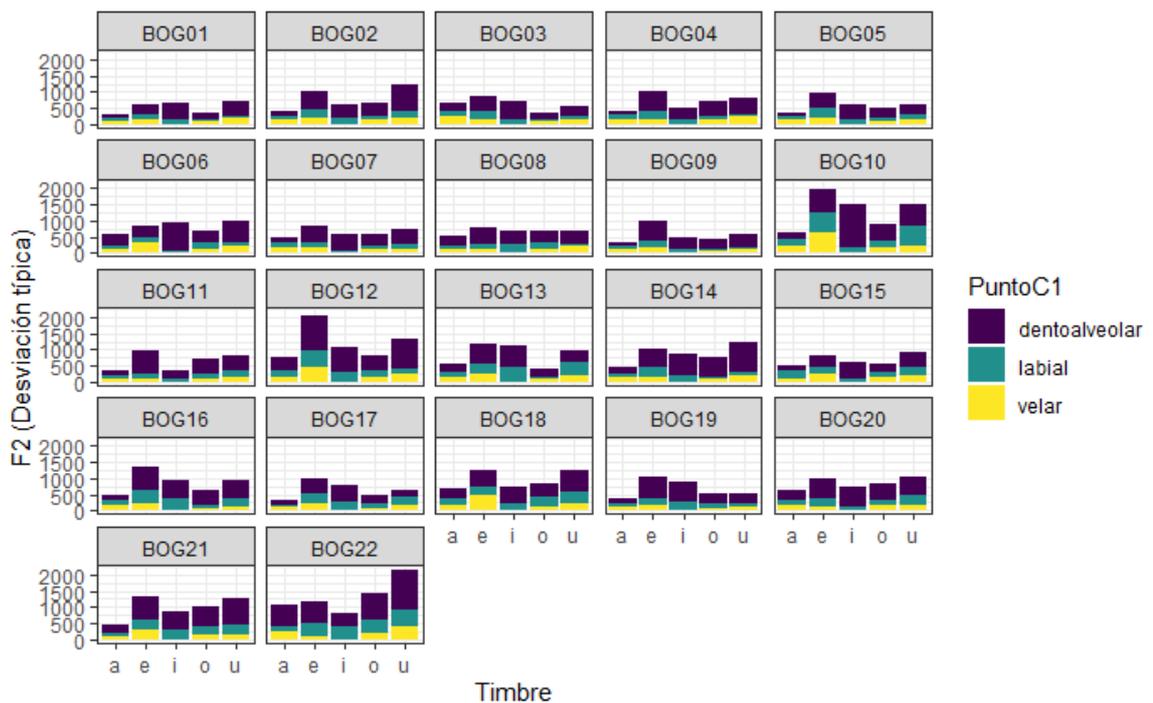


Figura 42. Desviaciones típicas de F2 (en hercios) en función del hablante, el punto de articulación y el timbre vocálico.

5.3.2 Variaciones dinámicas de las frecuencias de los formantes de /i/

Como se mostró anteriormente, el primer formante de /i/ presenta, aparentemente, poca variación cuando se analizan sus valores de frecuencia en función del punto de articulación de las consonantes de su entorno. Para verificar esta observación, en este apartado se presentan los modelos de efectos mixtos para el cambio absoluto $\Delta F1$ y la velocidad de cambio de rF1. Los resultados obtenidos indican que, al analizar los valores en hercios, estos dos parámetros solamente son influenciados por la consonante precedente ($\chi^2(5)=16.16$, $p < .001$) y la velocidad de elocución ($\chi^2(2)=3.41$, $p = .053$). Dentro de las consonantes analizadas, los segmentos que presentan cambios significativos son, de menor a mayor, /p/, /d/ y /s/. Esta diferencia en los valores de frecuencia entre el centro y el inicio de la vocal es mayor en velocidad rápida que en velocidad normal y lenta. La significancia es mayor en los valores normalizados que en los valores en hercios, y este resultado es semejante al obtenido para la consonante precedente ($\chi^2(5)=25.79$, $p < .001$) y para la velocidad de habla ($\chi^2(2)=10.03$, $p < .001$).

En el mismo sentido, la velocidad de cambio del primer formante (rF1) de /i/ está influenciada por la consonante precedente ($\chi^2(5)=12.91$, $p < .001$) y por la velocidad de elocución ($\chi^2(2)=4.37$, $p = .027$); la direccionalidad de estas modificaciones es la misma del cambio o diferencia absoluta de los valores de frecuencia, es decir, que rF1 presenta cambios significativos en /p/ /d/ y /s/, lo cual es más sobresaliente en velocidad rápida que en velocidad normal y lenta. La significancia del modelo es la misma cuando se analizan los datos en hercios que cuando se analizan los datos normalizados.

De acuerdo con los valores obtenidos para los interceptos aleatorios por velocidad de habla y para las pendientes aleatorias por hablante, un grupo de participantes produce la vocal /i/ con menores diferencias o cambios absolutos en las tres velocidades (BOG03, BOG05, BOG07, BOG13, BOG15, BOG19). Estos mismos hablantes disminuyen el valor de rF1 cuando usan una velocidad de habla rápida; el otro grupo de hablantes presenta valores cercanos al intercepto o aumenta la velocidad de cambio en velocidad rápida, lo cual es especialmente notorio en el caso de dos mujeres (BOG06 y BOG12). Los interceptos aleatorios para los ítems léxicos indican que parte de la varianza de $\Delta F1$ se debe a las realizaciones de la palabra *existencia* (entorno /s_s/), pues esta presentó cambios bastante bajos en los valores de frecuencia.

En cuanto al segundo formante, los resultados correspondientes al cambio absoluto ($\Delta F2$) de /i/ se modifican en función del punto de articulación ($\chi^2(5)=34.96$, $p < .001$), y, en menor medida, por la velocidad de elocución ($\chi^2(2)=3.47$, $p = .051$). En el modelo para esta variable, el intercepto corresponde a la vibrante simple /r/ en velocidad lenta y, como se ve en la Figura 43 y en la Tabla 11, las consonantes que producen mayores cambios sobre los valores de frecuencia de /i/ son, primero, las líquidas /r, l/ y, segundo, la oclusiva bilabial sorda /p/. En tanto, las consonantes que causan menores cambios sobre $\Delta F2$ son las obstruyentes dentoalveolares /d/ y /s/ y la fricativa labiodental sorda /f/. Los coeficientes del modelo y los valores obtenidos para los factores aleatorios (Tabla 11) apuntan a que la velocidad lenta introduce mayor variabilidad en los datos y que la diferencia entre velocidad lenta y velocidad rápida es altamente significativa. Al aplicar este modelo a los datos normalizados, se obtiene la misma significatividad para la consonante precedente ($\chi^2(5)=26.62$, $p < .001$) y para la velocidad de elocución ($\chi^2(2)=3.48$, $p = .050$). La varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 26 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) es del 29 %.

Los resultados obtenidos en el modelo de la velocidad de cambio del segundo formante (rF2) de /i/ confirman, en general, los datos reportados anteriormente. En primer lugar, como muestra la Figura 44 y la Tabla 12 la velocidad de cambio es mayor en el contexto de las consonantes /r, l, p/, si bien los coeficientes solamente son altamente significativos para la vibrante simple. En segundo lugar, las consonantes /f, d, t/ presentan una menor rF2, lo cual indica una alta coarticulación de /i/ con el entorno consonántico. Nótese que, si bien la consonante lateral /l/ no es significativa, esta consonante introduce una gran variabilidad en los valores de frecuencia de /i/. El entorno consonántico tiene un efecto altamente significativo sobre rF2 en los datos en hercios ($\chi^2(5)=16.89$, $p < .001$) y en los datos normalizados ($\chi^2(5)=25.08$, $p < .001$), mientras que la velocidad de elocución es marginalmente significativa cuando se analizan los datos en hercios ($\chi^2(2)=19.97$, $p = .990$) y tiene un efecto significativo en los datos normalizados ($\chi^2(2)=3.99$, $p = .031$). El modelo para nuestro corpus tiene una varianza explicada por los factores fijos (R^2_m) del 11 %, mientras que la varianza explicada por el modelo es del 22 %.

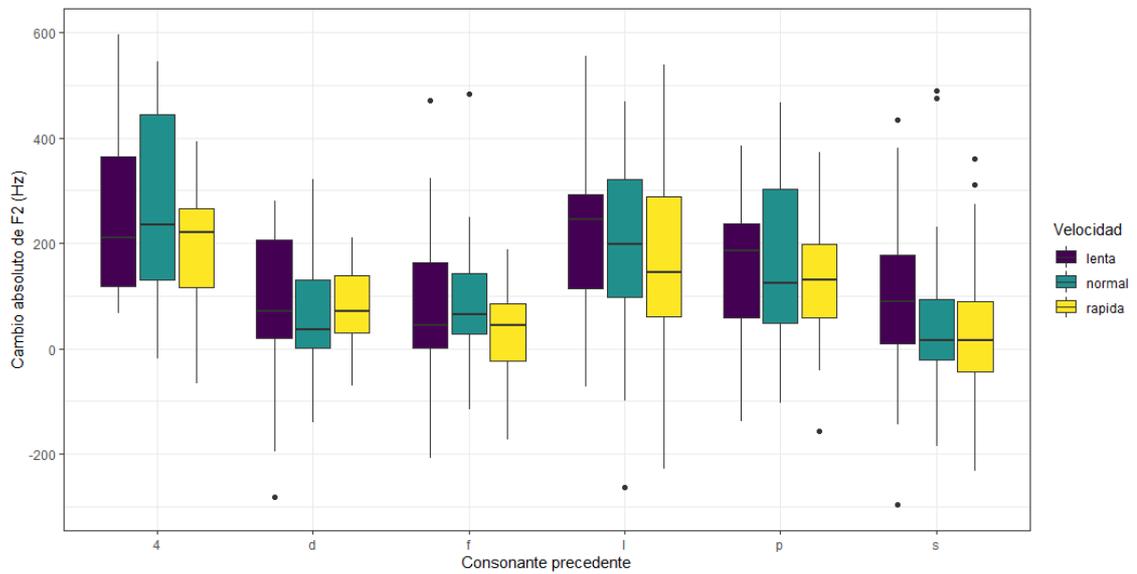


Figura 43. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /i/ en función de la consonante precedente y la velocidad de elocución.

Tabla 11. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de F2 (Hz) de la vocal /i/. El intercepto corresponde a la vibrante simple /t/, el modelo se elaboró sobre un total de 573 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	273.63	26.62	10.28	<.001
/d/	-168.52	26.56	-6.34	<.001
/f/	-187.52	27.10	-6.91	<.001
/l/	-60.75	23.06	-2.63	0.00
/p/	-79.38	26.45	-3.00	0.00
/s/	-191.02	21.90	-8.72	<.001
Velocidad (normal)	-17.50	17.43	-1.00	0.32
Velocidad (rápida)	-42.65	15.93	-2.67	0.00
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	5879	76.68	
	Velocidad normal	1598	39.98	
	Velocidad rápida	393	19.82	
	Residuos	22338	149.46	
$R^2_m = 0.16$		$R^2_c = 0.29$		

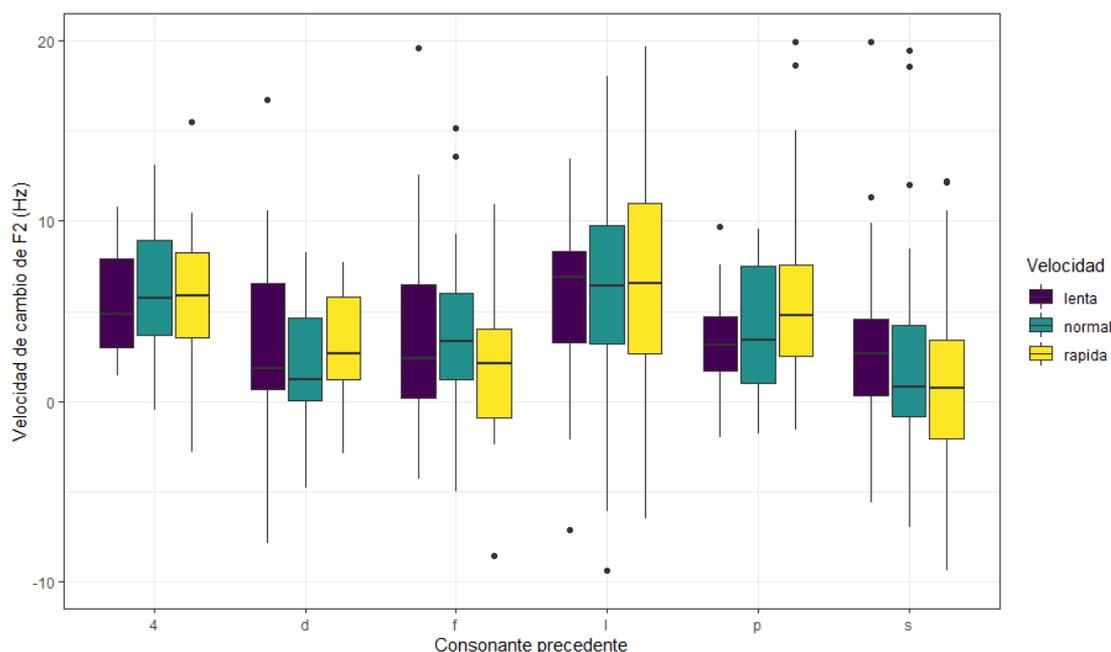


Figura 44. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /i/ en función de la consonante precedente y la velocidad de elocución.

Tabla 12. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /i/. El intercepto corresponde a la vibrante simple /t/, el modelo se elaboró sobre un total de 573 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	6.00	0.76	7.87	<.001
/d/	-3.14	0.85	-3.68	<.001
/f/	-3.18	0.87	-3.65	<.001
/l/	0.16	0.74	0.21	0.82
/p/	-1.54	0.85	-1.81	0.06
/s/	-4.26	0.70	-6.05	<.001
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	2.76	1.66	
	Velocidad normal	1.06	1.02	
	Velocidad rápida	0.13	0.37	
	Residuos	23.10	4.80	
$R^2_m = 0.11$		$R^2_c = 0.22$		

5.3.3 Variaciones dinámicas de las frecuencias de los formantes de /e/

Al inicio de este capítulo se ha señalado que la distribución de los valores del segundo formante de la vocal media /e/ se distribuyen entre los 1500 y 2300 Hz sin agruparse claramente según la consonante o el punto de articulación. El único punto de articulación

que ejerce una influencia clara según la clasificación realizada a partir del método *k-means* es el contexto de las consonantes velares /k, g/, que produce un incremento de los valores de frecuencia de F2. A continuación, se ampliarán estas observaciones con la ayuda de los modelos lineales de efectos mixtos.

La diferencia o cambio absoluto ($\Delta F1$) del primer formante de /e/ está influenciada por la consonante precedente ($\chi^2(9)=49.13, p < .001$) y por el sexo del hablante ($\chi^2(1)=4.59, p = .045$). Las consonantes que ejercen mayor influencia sobre $\Delta F1$ son las oclusivas sonoras /b, d, g/, las obstruyentes sordas /k, s/ y, en último lugar, la consonante lateral /l/. Todas estas consonantes producen cambios positivos en los valores de $\Delta F1$, lo cual significa que hay una mayor abertura lingual durante la producción de /e/ en estos contextos. Los hombres tienen un $\Delta F1$ con valores por debajo de las mujeres. De acuerdo con los coeficientes de los factores aleatorios, estos resultados están asociados a la variación fonética individual, pues un grupo de hablantes, entre hombres y mujeres (BOG03, BOG05, BOG12, BOG13, BOG14, BOG16, BOG18, BOG21), presentan valores de $\Delta F1$ por debajo del intercepto, mientras que un segundo grupo, también con miembros de ambos sexos (BOG02, BOG04, BOG06, BOG11, BOG15), tiene cambios por encima del promedio de todos los hablantes. El modelo con datos normalizados es significativo para la consonante precedente ($\chi^2(9)=55.61, p < .001$) y, como es de esperar, no es significativo para el sexo de los hablantes ($\chi^2(1)=0.01, p = .928$). La varianza explicada por los factores fijos (R^2_m) es del 21 %, mientras que la varianza explicada por el modelo (R^2_c) es del 26 %.

La velocidad de cambio (rF1) de la vocal anterior /e/ está influenciada por la consonante precedente ($\chi^2(9)=28.32, p < .001$) y la velocidad de elocución ($\chi^2(2)=3.60, p < .046$). Las consonantes que provocan mayor rF1 de la vocal /e/ son, como en el parámetro anterior, las oclusivas sonoras /b, d, g/, las obstruyentes sordas /k, s/ y, en menor medida, la consonante lateral. En velocidad normal y rápida la rF1 está por debajo de la velocidad lenta. La variación individual sigue el mismo patrón observado en los resultados de $\Delta F1$, es decir, un grupo de hablantes (BOG03, BOG05, BOG12, BOG13, BOG14, BOG16, BOG18, BOG21) tiene valores de rF1 por debajo del intercepto, mientras que un segundo grupo, conformado por hombres y mujeres, presenta una velocidad de cambio por encima del valor al estimado por el modelo (BOG02, BOG04, BOG06, BOG11, BOG15). Por último, la significancia de los valores de frecuencia en hercios es la misma

que la del modelo con datos normalizados. La varianza explicada por los factores fijos (R^2_m) es del 13 %, mientras que la varianza explicada por el modelo (R^2_c) es del 19 %.

En cuanto al segundo formante de la vocal media /e/, el cambio absoluto ($\Delta F2$) varía en función de la consonante precedente ($\chi^2(9)= 91.81, p < .001$), la velocidad de elocución ($\chi^2(2)= 12.33, p < .001$), la tonicidad ($\chi^2(1)= 27.82, p < .001$) y el sexo del hablante ($\chi^2(1)= 14.28, p = .001$). Las consonantes que se caracterizan por tener un valor mayor de $\Delta F2$ son la vibrante simple /r/ (el intercepto del modelo) y las oclusivas bilabiales /b, p/, las cuales presentan, además, un amplio rango de variación. En cambio, las consonantes con menor $\Delta F2$, es decir, con mayor coarticulación con la vocal media /e/, son las consonantes velares /k, g/ y la sibilante dentoalveolar /s/ (Figura 45). En cuanto a los efectos del tempo, la velocidad rápida y normal tienen valores más bajos de $\Delta F2$ que la velocidad lenta (Tabla 13). Por último, las realizaciones tónicas de /e/ presentan un $\Delta F2$ por encima de las átonas, y las vocales producidas por los hombres tienen menor $\Delta F2$ (mayor coarticulación) que las realizaciones de las mujeres.

En los datos correspondientes a la velocidad de cambio del segundo formante (rF2) de /e/ se observan cambios altamente significativos en función de la consonante precedente ($\chi^2(9)= 96.16, p < .001$) y la tonicidad ($\chi^2(1)= 13.24, p < .001$), y cambios muy significativos para el sexo del hablante ($\chi^2(1)= 10.16, p = .005$). Como recoge la Figura 46 y la Tabla 14, las consonantes asociadas con una mayor rF2 son la vibrante simple /r/ (el intercepto del modelo), y las oclusivas bilabiales /b, p/, en tanto, las oclusivas velares /k, g/ y la sibilante presentan poca velocidad de cambio. Las vocales tónicas tienen mayor rF2 que las átonas (1.3 Hz/ms) y las vocales producidas por los hombres tienen una rF2 con un valor por debajo (-1.5 Hz/ms) del estimado para las mujeres.

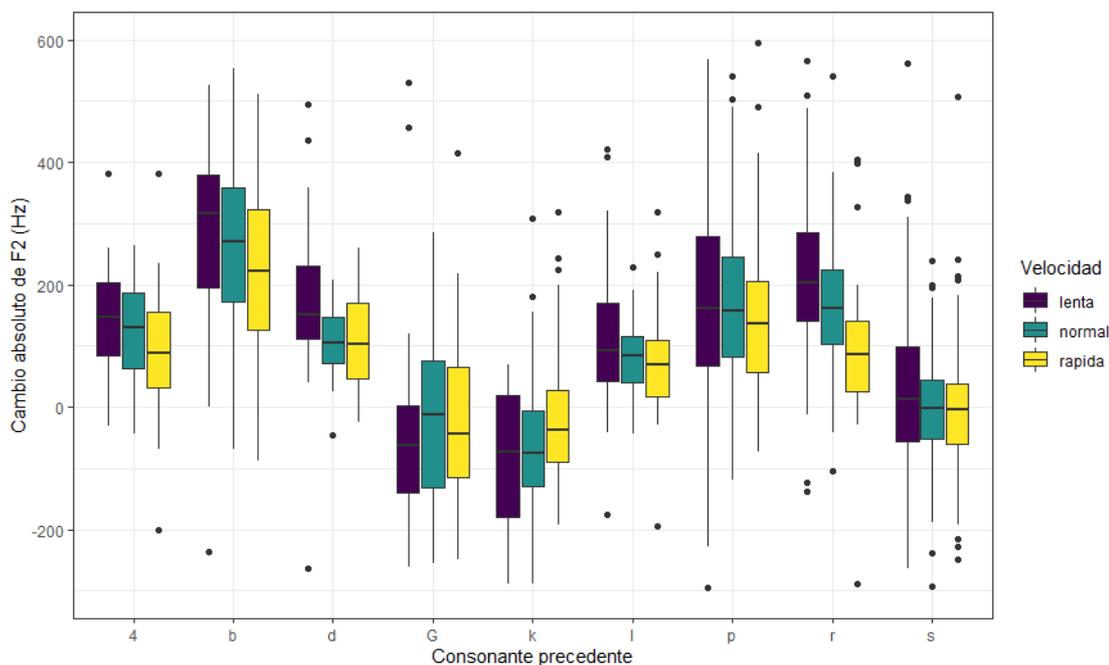


Figura 45. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F_2/\text{Hz}$) de /e/ en función de la consonante precedente y la velocidad de elocución.

Tabla 13. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de F_2 (Hz) de la vocal /e/. El intercepto corresponde a la vibrante simple /t/, el modelo se elaboró sobre un total de 1604 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	165.70	19.66	8.42	<.001
/b/	169.42	20.15	8.40	<.001
/g/	-180.89	25.97	-6.96	<.001
/k/	-175.79	20.35	-8.63	<.001
/p/	57.71	17.16	3.36	<.001
/s/	-135.62	17.94	-7.55	<.001
Velocidad (normal)	-32.82	9.13	-3.59	<.001
Velocidad (rápida)	-45.59	8.91	-5.11	<.001
Tonicidad (tónica)	56.66	10.70	5.29	<.001
Sexo (masculino)	-60.10	13.86	-4.33	<.001
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	1233	35.12	
	Velocidad normal	175	13.25	
	Velocidad rápida	107	10.38	
	Residuos	19867	140.95	
$R^2_m = 0.35$		$R^2_c = 0.38$		

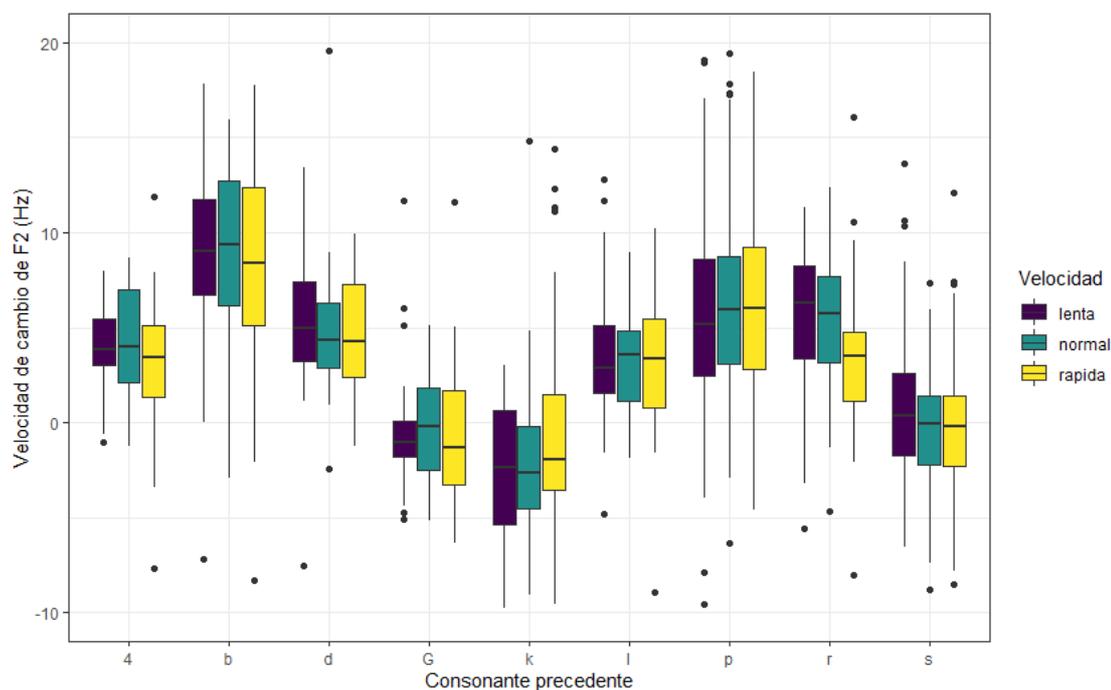


Figura 46. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /e/ en función de la consonante precedente y la velocidad de elocución.

Tabla 14. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /e/. El intercepto corresponde a la vibrante simple /r/, el modelo se elaboró sobre un total de 1603 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	4.16	0.57	7.20	<.001
/b/	5.69	0.69	8.22	<.001
/g/	-5.34	0.89	-5.98	<.001
/k/	-5.59	0.69	-8.00	<.001
/p/	3.01	0.58	5.11	<.001
/s/	-4.32	0.61	-7.02	<.001
Tonicidad (tónica)	1.32	0.36	3.58	<.001
Sexo (masculino)	-1.58	0.31	-4.95	<.001
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	0.72	0.84	
	Velocidad normal	0.58	0.76	
	Velocidad rápida	1.32	1.15	
	Residuos	23.23	4.84	
$R^2_m = 0.35$		$R^2_c = 0.38$		

Los factores fijos analizados en los modelos correspondientes al cambio absoluto y a la velocidad de cambio son altamente significativos cuando se analizan los datos en hercios y los datos normalizados, a excepción del sexo del hablante. En ambos casos, la varianza explicada por los factores fijos (R^2_m) alcanza el 35 % y la varianza explicada por el modelo (R^2_c) alcanza el 38 %.

5.3.4 Variaciones dinámicas de las frecuencias de los formantes de /a/

La diferencia absoluta del primer formante ($\Delta F1$) de la vocal abierta /a/ cambia, primero, por efecto de la velocidad de elocución ($\chi^2(2)= 40.14$, $p < .001$). En velocidad lenta, $\Delta F1$ es mayor que en velocidad normal y rápida en todos los entornos consonánticos. Segundo, el sexo del hablante ($\chi^2(1)= 19.51$, $p < .001$) tiene un efecto altamente significativo, dado que los hombres presentan menor $\Delta F1$ que las mujeres. La consonante precedente tiene una influencia muy significativa ($\chi^2(6)= 6.58$, $p = .002$) cuando la vocal abierta /a/ está en el entorno de la fricativa labiodental /f/ (el intercepto del modelo) y de las oclusivas velares /g, k/. Tercero, las palabras funcionales tienen mayor $\Delta F1$ (+ 46.05 Hz) que las palabras con contenido léxico ($\chi^2(1)= 11.78$, $p = .004$). Finalmente, las vocales tónicas incrementan el valor $\Delta F1$ con respecto a las átonas ($\chi^2(1)= 7.31$, $p = .016$). En el modelo con datos normalizados se observa la misma significancia para el entorno consonántico, la velocidad, el tipo de palabra y la tonicidad. En este caso, la varianza explicada por los factores fijos (R^2_m) es del 39 % y la varianza explicada por el modelo (R^2_c) alcanza el 49 %.

La velocidad del cambio del primer formante ($rF1$) de /a/ se incrementa en las palabras funcionales ($\chi^2(1)= 19.90$, $p < .001$). En cuanto a la consonante precedente, el efecto es muy significativo ($\chi^2(6)= 5.87$, $p = .007$) debido a la influencia de las consonantes labiales /f, p/ y de la oclusiva velar sonora /g/. Por último, la tonicidad tiene un efecto muy significativo sobre la $rF1$ ($\chi^2(1)= 9.72$, $p = .007$). Estos efectos son altamente significativos ($p < .001$) en el modelo con datos normalizados.

La diferencia absoluta del segundo formante ($\Delta F2$) de /a/ varía en función de la consonante precedente ($\chi^2(6)= 40.73$, $p < .001$). Como es de esperar, la vocal abierta /a/ presenta cambios positivos (una transición ascendente) en el contexto de las consonantes labiales y cambios negativos (una transición descendente) en el entorno de las

consonantes velares y dentoalveolares. El valor de $\Delta F2$ es menor en velocidad rápida que en velocidad lenta y normal ($\chi^2(2)= 4.09$, $p = .031$), y aumenta cuando la palabra a la que pertenece la vocal se ubica en posición intermedia del grupo fónico ($\chi^2(2)= 3.42$, $p = .034$). El efecto de la consonante precedente es significativo en el modelo con datos normalizados ($\chi^2(6)= 40.73$, $p = .016$), es muy significativo para la velocidad de elocución ($\chi^2(2)= 6.23$, $p = .007$) y, al igual que en el modelo con datos en hercios, la posición dentro del grupo fónico es significativa ($\chi^2(2)= 4.34$, $p = .013$).

La velocidad de cambio de F2 (rF2) cambia en función del entorno consonántico únicamente, y los valores obtenidos siguen el mismo patrón del $\Delta F2$: en el contexto de las consonantes labiales la rF2 es positiva, mientras que en el contexto velar y dentoalveolar es baja. La significancia del entorno consonántico sobre rF2 es altamente significativa en el modelo en hercios ($\chi^2(6)= 28.32$ $p < .001$) y significativa en el modelo con datos normalizados ($\chi^2(6) = 3.65$ $p = .019$). El mejor modelo es el calculado con datos en hercios, pues la varianza explicada por los factores fijos (R^2_m) alcanza el 46 % y la varianza explicada por el modelo (R^2_c) es del 54 %.

De acuerdo con los factores aleatorios, gran parte de la varianza de $\Delta F2$ y rF2 se debe a los ítems léxicos. Al observar los valores que ofrece el modelo, se puede comprobar que la vocal /a/ aparece en el entorno de obstruyentes sordas (p.ej. *cosas*, *paso*, *pasarme*) y después de una vibrante múltiple, como es el caso de la palabra *guerra*. Esto se refleja en el valor de la varianza y en altos valores para las desviaciones típicas correspondientes a los factores aleatorios. En velocidad lenta y en velocidad normal, los hablantes presentan una gran variación para $\Delta F2$, mientras que para rF2 hay mayor varianza en habla rápida, lo cual indica que los hablantes usan un amplio rango de valores para rF2 (Tabla 15 y Tabla 16).

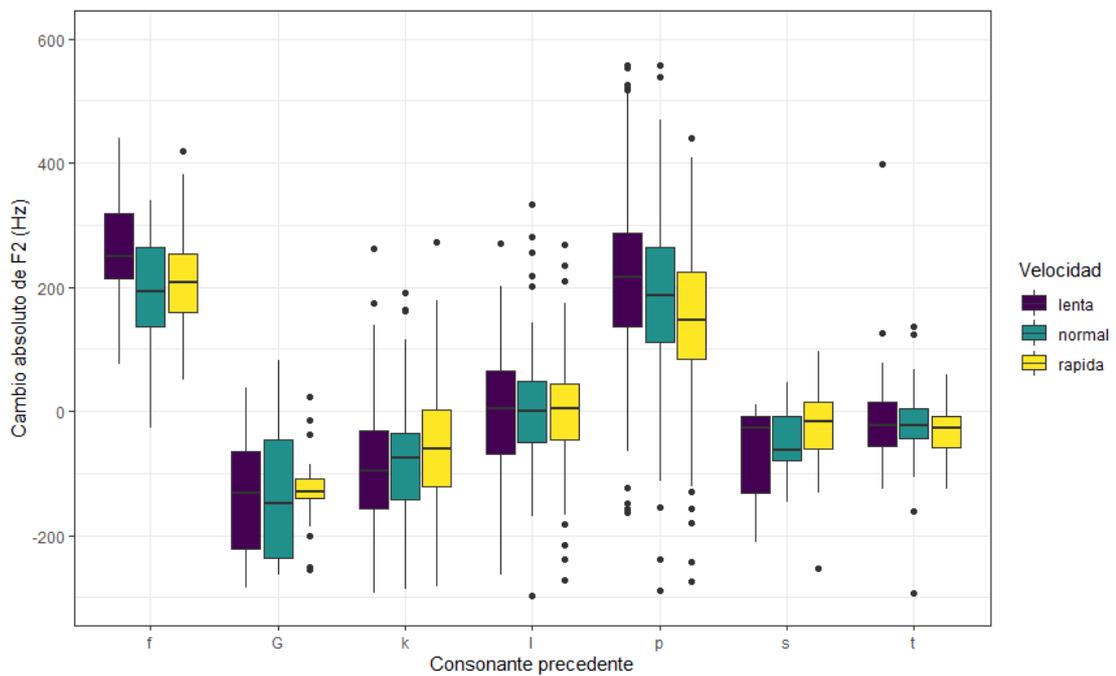


Figura 47. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F_2/\text{Hz}$) de /a/ en función de la consonante precedente y la velocidad de elocución.

Tabla 15. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de ΔF_2 (Hz) de la vocal /a/. El intercepto corresponde a la vibrante simple /f/, el modelo se elaboró sobre un total de 1747 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	194.46	31.59	6.15	<.001
/g/	-364.57	39.78	-9.16	<.001
/k/	-292.00	31.05	-9.40	<.001
/l/	-223.15	34.44	-18.90	<.001
/p/	-44.50	30.43	-1.46	0.15
/s/	-255.42	40.12	-6.36	<.001
/t/	-240.86	34.90	-6.90	<.001
Velocidad (rápida)	-21.16	7.01	-3.01	0.00
Contexto de Frase (intermedia)	34.85	12.39	2.81	0.00
Factores aleatorios		Varianza	Std.dev	
Palabra		614.9	24.79	
Participante	Velocidad lenta	726.2	26.94	
	Velocidad normal	722.2	0.31	
	Velocidad rápida	1.64	1.28	
	Residuos	13120	114.55	
$R^2_m = 0.48$		$R^2_c = 0.54$		

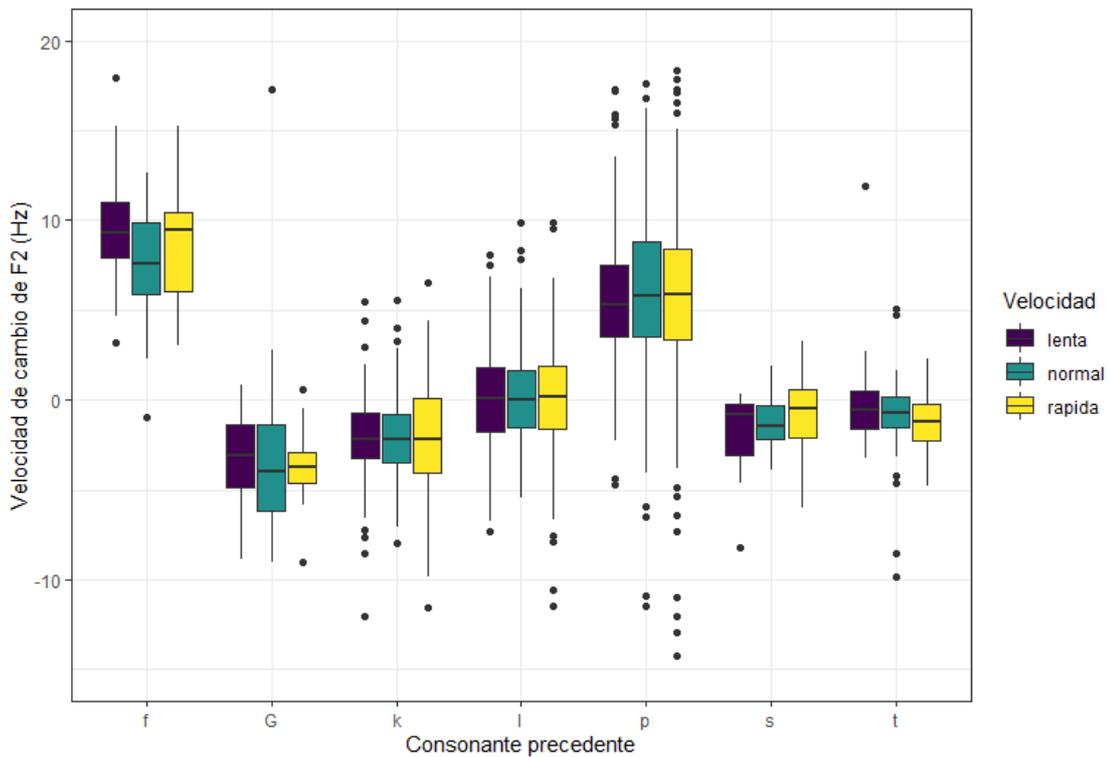


Figura 48. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /a/ en función de la consonante precedente y la velocidad de elocución.

Tabla 16. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /a/. El intercepto corresponde a la vibrante simple /f/, el modelo se elaboró sobre un total de 1747 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	8.97	21.75	8.31	<.001
/g/	-11.81	1.48	-7.96	<.001
/k/	-11.04	1.15	-9.55	<.001
/l/	-9.11	1.28	-7.08	<.001
/p/	-3.61	1.14	-3.16	.000
/s/	-10.56	1.42	-7.42	<.001
/t/	-9.76	1.30	-7.45	<.001
Factores aleatorios		Varianza	Std.dev	
Palabra	0.95	0.97		
Participante				
Velocidad lenta	0.39	0.63		
Velocidad normal	0.04	0.21		
Velocidad rápida	0.57	0.75		
Residuos	12.28	3.50		
$R^2_m = 0.46$		$R^2_c = 0.54$		

5.3.5 Variaciones dinámicas de las frecuencias de los formantes de /o/

La vocal posterior media /o/ presenta pequeños cambios absolutos del primer formante ($\Delta F1$) y en su velocidad de cambio (rF1). En otras palabras, el F1 se mantiene estable y no se modifica por efecto de la consonante precedente, la tonicidad, el sexo del hablante o la velocidad de elocución. Esto quiere decir que ninguno de los factores mencionados resulta estadísticamente significativo dentro del modelo del primer formante, lo cual puede observarse tanto en los modelos con datos en hercios como en el modelo con datos normalizados.

Como se observa en la Tabla 17, el entorno consonántico ofrece un valor de significación altamente significativo para el modelo de $\Delta F2$ de la vocal /o/ ($\chi^2(5)=14.47$, $p < .001$), lo mismo se observa en el caso de la velocidad de elocución ($\chi^2(2)=11.98$, $p < .001$) y del sexo del hablante ($\chi^2(1)=11.01$, $p = .003$). En cuanto al entorno consonántico, el valor de $\Delta F2$ es negativo para las consonantes dentoalveolares, lo cual quiere decir que la transición tiene una trayectoria descendente. El mayor grado de coarticulación se observa en el contexto velar, pues en este entorno hay menor cambio y rango de variación (Figura 49). Los valores de $\Delta F2$ son mayores en velocidad normal que en velocidad lenta y rápida y, adicionalmente, son mayores para los hombres que para las mujeres.

En cuanto a la velocidad de cambio de la frecuencia del segundo formante (rF2), se obtuvo nuevamente un efecto significativo para el entorno consonántico ($\chi^2(5)=7.81$, $p < .001$), la velocidad de elocución ($\chi^2(2)=8.21$, $p = .002$) y el sexo ($\chi^2(1)=7.67$, $p = .012$). Respecto al entorno consonántico, hay poca velocidad de cambio en el contexto velar, lo cual confirma que en este contexto hay un alto grado de coarticulación con la vocal posterior media /o/. Las consonantes bilabiales presentan cambios positivos /b, p/, pero el rango de variación es tan alto que los coeficientes no son estadísticamente significativos o, en otras palabras, no se puede considerar que las estimadas sean representativas de los datos (Tabla 18). Las consonantes dentoalveolares arrojan valores negativos para rF2 y, a pesar del amplio rango de valores, las diferencias son significativas. Al igual que se observa en el modelo para $\Delta F2$, la velocidad de cambio del segundo formante de /o/ es mayor en velocidad normal que en velocidad lenta y rápida y, además, los hombres presentan una rF2 más alta que las mujeres.

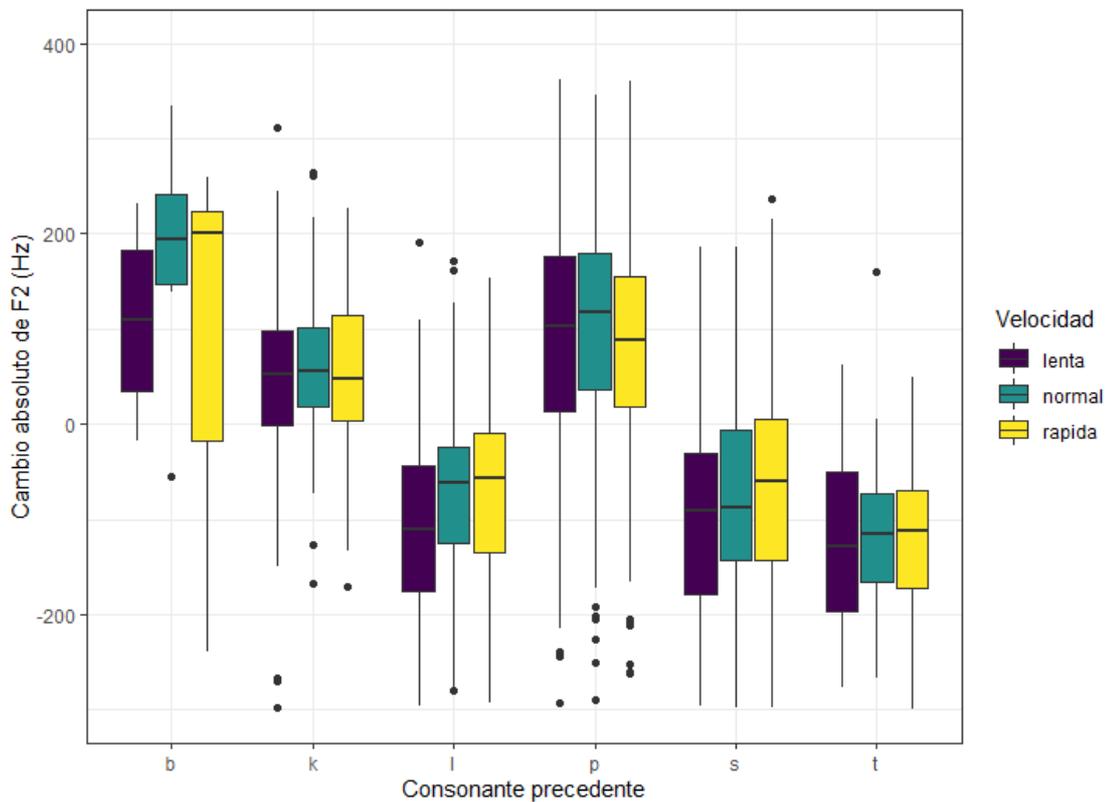


Figura 49. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /o/ en función de la consonante precedente y la velocidad de elocución.

Tabla 17. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de $\Delta F2$ (Hz) de la vocal /o/. El intercepto corresponde a la vibrante simple /b/, el modelo se elaboró sobre un total de 1747 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	14.72	55.55	0.26	0.79
/k/	-27.43	59.64	-0.46	0.64
/l/	-160.03	64.16	-2.49	0.02
/p/	-13.19	58.59	-0.22	0.82
/s/	-182.91	61.28	-2.98	0.00
/t/	-211.34	59.93	-3.52	.001
Velocidad (normal)	53.24	10.78	4.93	<.001
Velocidad (rápida)	39.05	9.19	4.24	<.001
Sexo (masculino)	50.30	13.39	3.75	.001
Factores aleatorios		Varianza	Std.dev	
Palabra		2207	46.98	
Participante				
Velocidad lenta		1728	41.57	
Velocidad normal		527	22.96	
Velocidad rápida		1068	32.67	
Residuos		19456	139.49	
$R^2_m = 0.21$		$R^2_c = 0.36$		

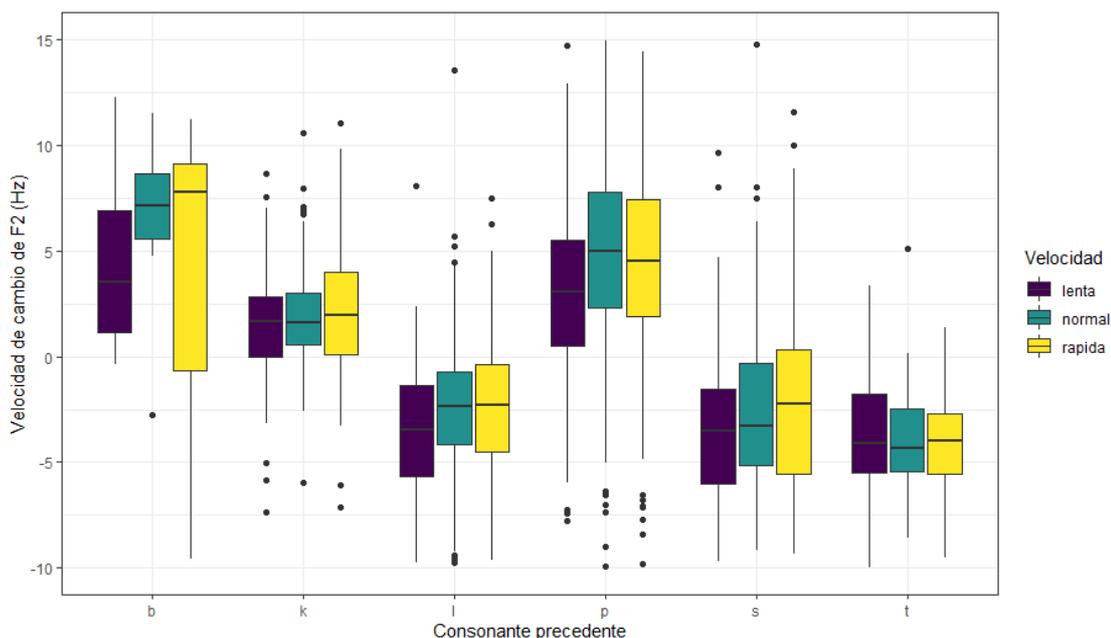


Figura 50. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /o/ en función de la consonante precedente y la velocidad de elocución.

Tabla 18. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /o/. El intercepto corresponde a la vibrante simple /b/, el modelo se elaboró sobre un total de 1877 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	1.05	1.98	0.53	0.59
/k/	-1.18	2.13	-0.55	0.58
/l/	-5.46	2.29	-2.38	0.02
/p/	-0.04	2.09	-0.02	0.98
/s/	-5.63	2.19	-2.56	0.01
/t/	-7.34	2.14	-3.41	0.00
Velocidad (normal)	1.27	0.30	4.24	0.00
Velocidad (rápida)	0.60	0.30	1.99	0.04
Sexo (masculino)	1.59	0.50	3.13	0.00
Factores aleatorios		Varianza	Std.dev	
Palabra		2.78	1.66	
Participante Velocidad lenta		1.06	1.03	
Velocidad normal		0.11	0.33	
Velocidad rápida		0.12	0.36	
Residuos		26.39	5.13	
$R^2_m = 0.19$		$R^2_c = 0.33$		

De acuerdo con los factores aleatorios (Tabla 17 y Tabla 18), la varianza de los parámetros analizados está influenciada principalmente por los ítems léxicos y por las variaciones individuales. Al analizar los factores de manera individual, se puede observar que parte de la variación está asociada con palabras seguidas de una consonante lateral (p.ej., *sol*, *policial*, *árboles*, *sol*). La variación individual se debe a que en velocidad lenta y rápida un grupo de hablantes tienen valores por encima del intercepto en los dos parámetros analizados.

5.3.6 Variaciones dinámicas de las frecuencias de los formantes de /u/

La vocal posterior cerrada /u/, al igual que la vocal posterior media /o/, presenta cambios absolutos del primer formante ($\Delta F1$) que no son significativos para ninguno de los factores fijos analizados. En la misma línea, la velocidad de cambio ($rF1$) no varía en función de la consonante precedente, la tonicidad o el sexo del hablante, como sí ocurre con otros timbres vocálicos analizados.

En cuanto al segundo formante de /u/, el valor de $\Delta F2$ cambia en función de la consonante precedente ($\chi^2(6)=8.31$, $p=.046$). El efecto es significativo y está relacionado con las diferencias entre el intercepto del modelo, la oclusiva dentoalveolar sonora /d/, la oclusiva velar sorda /k/ y la oclusiva dentoalveolar sorda /t/. Como se ve en la Tabla 19, los coeficientes para las consonantes labiales no son significativos debido a que el error estándar iguala o supera las estimadas. Esto también se refleja en los amplios rangos de variación que representan por medio de los diagramas de caja (Figura 51).

Al analizar los valores de $rF2$ de /u/ se obtienen resultados semejantes a los descritos anteriormente (Tabla 20), si bien la significancia es altamente significativa para la consonante precedente ($\chi^2(6)=17.64$, $p < .001$). Como se ve en la Figura 52, la velocidad de cambio se acerca a cero (hay mayor coarticulación) en el contexto de las consonantes velares y dentoalveolares, mientras que en el contexto labial los datos son variables y no significativos. Finalmente, en el modelo de /u/ los interceptos aleatorios para los ítems léxicos no son significativas. En este caso, la varianza del modelo se explica porque la mayoría de los hablantes tiene un amplio rango de valores en velocidad normal y rápida (Tabla 20).

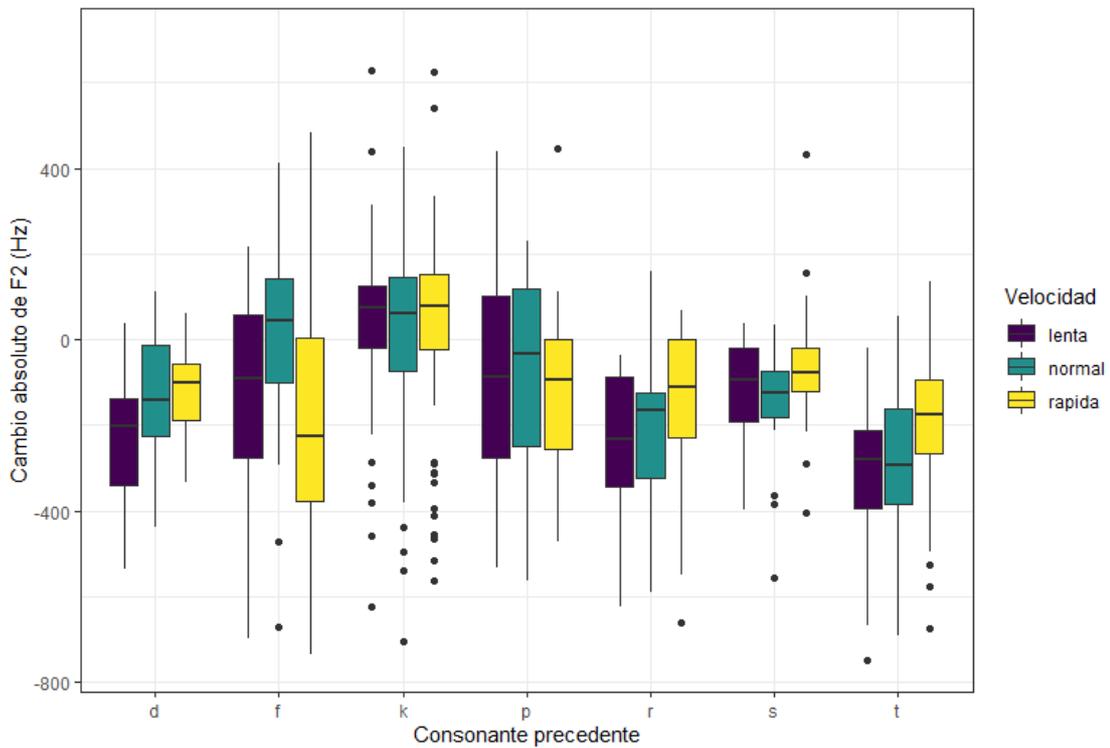


Figura 51. Cambios absolutos de los valores de frecuencia del segundo formante ($\Delta F2/Hz$) de /u/ en función de la consonante precedente y la velocidad de elocución.

Tabla 19. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de los cambios absolutos de $\Delta F2$ (Hz) de la vocal /u/. El intercepto corresponde a la vibrante simple /d/, el modelo se elaboró sobre un total de 722 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	-204.08	35.54	-5.74	<.001
/f/	46.21	50.98	0.90	0.36
/k/	220.35	39.89	5.52	<.001
/p/	47.50	49.69	0.95	0.33
/p/	27.28	51.13	0.53	0.59
/s/	60.59	43.56	1.39	0.16
/t/	-75.25	39.73	-1.89	0.05
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	3435	58.61	
	Velocidad normal	5582	74.71	
	Velocidad rápida	2017	44.91	
	Residuos	61629	248.25	
$R^2_m = 0.15$		$R^2_c = 0.23$		

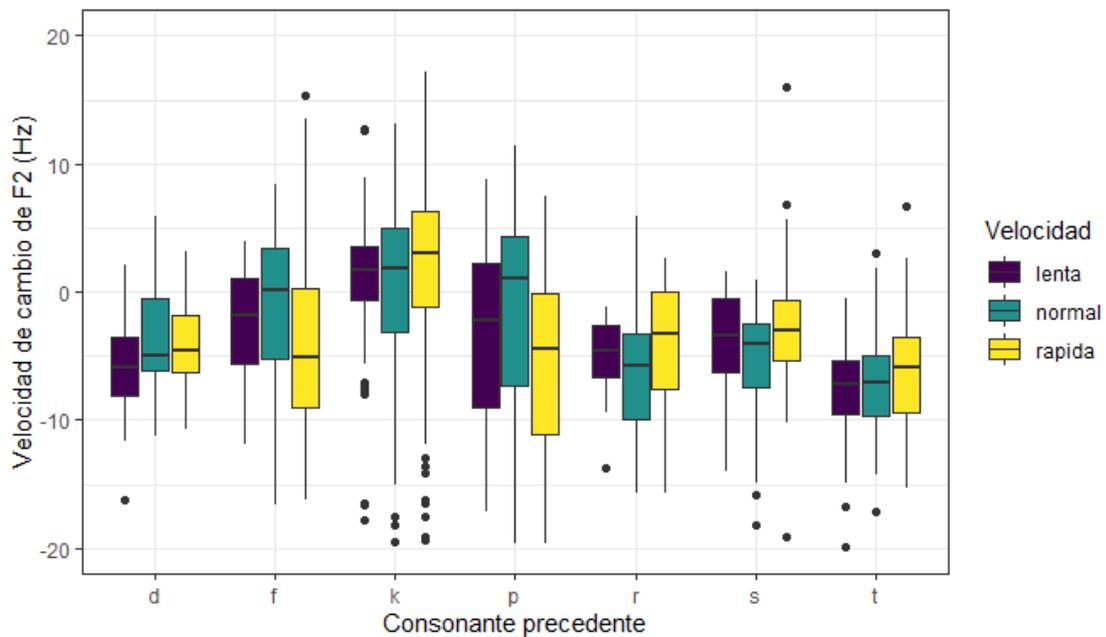


Figura 52. Velocidad de cambio de los valores de frecuencia de F2 (Hz/ms) de /u/ en función de la consonante precedente y la velocidad de elocución.

Tabla 20. Coeficientes de los efectos fijos y de los efectos aleatorios para un modelo de la velocidad de cambio de F2 (Hz/ms) de la vocal /u/. El intercepto corresponde a la vibrante simple /d/, el modelo se elaboró sobre un total de 722 observaciones y 22 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	-4.70	1.01	-4.62	<.001
/f/	0.97	1.36	0.71	0.47
/k/	5.05	1.07	4.68	<.001
/p/	-1.39	1.34	-1.03	0.30
/r/	-0.90	1.36	-0.66	0.50
/s/	0.24	1.37	0.18	<.001
/t/	-2.46	1.17	-2.08	0.03
Factores aleatorios		Varianza	Std.dev	
Participante	Velocidad lenta	0.39	0.63	
	Velocidad normal	7.02	2.65	
	Velocidad rápida	6.43	2.53	
	Residuos	61.21	7.82	
$R^2_m = 0.12$		$R^2_c = 0.18$		

5.4 Conclusiones parciales

Dado que buena parte de los estudios sobre la reducción vocálica sostienen que la reducción fonética de las vocales es causada, en parte, por efecto de la coarticulación, en este capítulo nos preguntamos cómo afecta el entorno consonántico en los valores de frecuencia de los formantes vocálicos. A partir del análisis de los valores de frecuencia de F1 y F2 en el centro de la vocal y el análisis de la desviación típica, podemos concluir que este efecto depende, en primer lugar, de la altura lingual y del timbre:

- 1) De acuerdo con la desviación típica, la variabilidad del F1 depende de la altura lingual: los valores de la vocal abierta /a/ son altos, las vocales /e, o/ tienen valores intermedios, y las vocales cerradas /i, u/ tienen menos variabilidad. En tanto, la desviación del F2 es alta para las vocales anteriores /i, e/ y es más baja para la vocal abierta /a/ y la vocal posterior media /o/. La vocal cerrada /u/ tiene una desviación baja en los contextos labial y velar, pero una desviación típica más alta en el contexto dentoalveolar.
- 2) El segundo formante de la vocal abierta /a/ disminuye en el contexto de las consonantes labiales y es más alto en el entorno de las velares y dentoalveolares, lo cual sugiere una pronunciación anteriorizada en este contexto. Este efecto también se observa en el caso de la vocal anterior media /e/. Las vocales posteriores /o, u/ tienen un F2 bajo en el entorno de consonantes labiales y velares, mientras que los valores de F2 se incrementan en el contexto de las dentoalveolares y, por tanto, en este contexto su articulación es más centralizada. La clasificación de los valores de frecuencia de /i/ derivados del algoritmo *k-means*, no muestra un patrón claro asociado al lugar de articulación.

En cuanto a las variaciones dinámicas de los formantes, que arrojan información sobre el grado de coarticulación y la trayectoria de los formantes de las vocales, nuestro análisis arrojó los siguientes resultados:

- 3) El cambio absoluto y la velocidad de cambio del primer formante ($\Delta F1$ y $rF1$) de las vocales /a, e, i/ aumentan en el contexto de las consonantes dentoalveolares y velares. Esto significa que la trayectoria de F1 es ascendente como supone la teoría del locus (Johnson, 2003, p. 178). En el caso de las vocales posteriores /o, u/, los cambios no

son estadísticamente significativos. Como se verá más adelante, la mayoría de cambios espectrales del primer formante están asociados a la velocidad de elocución y a otros factores como el sexo del hablante.

- 4) La diferencia absoluta y la velocidad de cambio del segundo formante de /i/ se incrementan en el contexto de las líquidas /l, r/ y de la oclusiva bilabial /p/. En particular, la consonante lateral /l/ introduce una gran variabilidad sobre esta vocal. Los valores de $\Delta F2$ y $rF2$ se acercan a cero en el entorno de /f/ y de las dentoalveolares /d/ y /s/, lo cual indica una mayor coarticulación de estas consonantes sobre /i/. La vocal anterior media /e/ presenta cambios semejantes: las trayectorias de los formantes tienen cambios significativos en el entorno de /r/ y de las oclusivas bilabiales /b, p/, mientras que hay mayor coarticulación en el contexto de las oclusivas velares /k, g/ y de la sibilante /s/.
- 5) El F2 de la vocal abierta /a/ tiene una transición ascendente en el contexto de las consonantes labiales y una transición descendente en el contexto de las velares y dentoalveolares. De acuerdo con los modelos estadísticos para $\Delta F2$ y $rF2$, la coarticulación de /a/ es mayor en los contextos velar /g, k/ y dentoalveolar /s, t, l/.
- 6) Las vocales posteriores /o, u/ son acústicamente variables en los diferentes entornos consonánticos. De acuerdo con los interceptos aleatorios del modelo mixto, esta variabilidad está asociada con las palabras de contenido léxico en las que /o/ aparece ante una consonante lateral (p.ej., *sol*, *policial*, *árboles*, *sol*). Finalmente, el F2 de la vocal posterior cerrada /u/ tiene un cambio espectral y velocidades de articulación cercanas a cero (mayor coarticulación) en el entorno de las velares y dentoalveolares, y especialmente en el contexto de /s/.
- 7) En ese sentido, podemos concluir que el contexto dentoalveolar ejerce mayor influencia sobre las vocales bogotanas. Entre las consonantes dentoalveolares, la sibilante /s/ causa un mayor grado de coarticulación sobre las cinco vocales.

Sobre la influencia de factores como la velocidad de habla, la tonicidad, el tipo de palabra, el sexo del hablante, la posición de la vocal dentro de la palabra y dentro del grupo fónico, concluimos que:

- 8) Los hablantes producen la vocal /i/ con el F1 con mayor cambio y velocidad de cambio espectral en velocidad rápida, mientras que cuando producen la vocal abierta /a/ y la anterior media /e/ utilizan la estrategia contraria, es decir, las

variaciones dinámicas de F1 son mayores en velocidad lenta que en velocidad normal y rápida. La diferencia absoluta de F2 ($\Delta F2$) es menor en velocidad rápida y normal que en habla lenta, lo cual confirma que hay mayor coarticulación de las consonantes sobre las vocales cuando disminuye el tiempo para realizar los movimientos articulatorios.

- 9) La tonicidad modifica la trayectoria de los formantes de las vocales /e/ y /a/: las realizaciones tónicas de la vocal media /e/ incrementan los valores de $\Delta F2$ y rF2, mientras que en el caso de la vocal abierta /a/, las tónicas incrementan los valores de $\Delta F1$ y rF1, es decir, modificaron el grado de abertura vocálica. El sexo del hablante también produce cambios significativos, pues las realizaciones masculinas de las vocales /a/ y /e/ tienen mayor coarticulación que las vocales producidas por las mujeres. Adicionalmente, cuando la vocal /a/ forma parte de una palabra funcional, incrementan los cambios espectrales del primer formante. Por último, en posición final hay mayor coarticulación (menor $\Delta F2$) que en posición intermedia del grupo fónico.

Finalmente, respecto a la influencia del comportamiento individual de los participantes sobre los resultados obtenidos podemos decir que:

- 10) Las desviaciones típicas de los valores de frecuencia de los formantes muestran que, en lo que respecta a F1, hay mayor variación interlocutor en la producción de las vocales /a/, /e/ y /o/, y esta variabilidad es mayor en el contexto dentoalveolar y labial que en el contexto velar (Figura 41 y Figura 42). En cuanto al F2, hay mayor variación entre los hablantes en la producción de /e/ y las vocales cerradas /i, u/.
- 11) Las diferencias individuales en las variaciones dinámicas de los formantes se manifiestan de manera clara en el primer formante ($\Delta F1$ y rF1) de las vocales anteriores /i, e/, pues algunos hablantes producen estas vocales con valores por debajo del intercepto del modelo, mientras que un segundo grupo tiene mayor variabilidad en el grado de abertura de las vocales.

6. Resultados: Duración, formantes, espacio acústico y centralización de las vocales en habla leída y en habla espontánea

6.1 Introducción

El estilo de habla es uno de los factores más estudiados con relación a la reducción fonética de las vocales. En la sección 2.2.6 vimos que, en un gran número de lenguas, los valores de frecuencia de F1 y F2 se acercan al centro del espacio acústico y el área vocálica se estrecha en habla espontánea, mientras que, en habla leída y en las pronunciaciones enfáticas, las vocales tienen una mayor duración y las frecuencias de los formantes están más diferenciadas acústicamente. De acuerdo con la teoría de la hiper e hipoarticulación (Lindblom, 1990), la reducción fonética de las vocales se manifiesta de esta manera porque la articulación está prospectivamente organizada y orientada hacia el oyente. Esto significa que los hablantes acudimos a una pronunciación hiperarticulada o que favorece la hipoarticulación dependiendo de las demandas del oyente y de la situación comunicativa. Por esta razón, la lectura en voz alta, que implica una mayor distintividad de la información, requiere la hiperarticulación de las vocales y, en general, favorece las realizaciones fonéticas canónicas consideradas como normativas en una lengua.

Este capítulo se estructura alrededor de dos preguntas de investigación. Primero, nos interesa saber cómo afecta el estilo de habla (habla leída vs. habla espontánea) en los valores de frecuencia de los formantes vocálicos y en la duración considerando el timbre, la tonicidad, el contexto consonántico, el tipo de palabra, la posición de la vocal dentro del grupo fónico y el sexo del hablante. Gracias a los resultados reportados en el Capítulo 4, sabemos que estas variables modifican significativamente los valores de la duración y de la frecuencia de F1 y F2. Con esta pregunta, buscamos determinar si el estilo tiene un efecto similar a la velocidad de elocución o interactúa con otros factores y, en ese

sentido, incrementa o contrarresta los efectos de otras variables. De acuerdo con el *principio de atención* de Labov (1972, p. 112) y la teoría de la hiper e hipoarticulación (Lindblom, 1990), los estilos de habla se organizan en una escala determinada por el grado de atención que presta el locutor a su pronunciación. Así que esperamos encontrar una mayor variabilidad acústica en nuestro corpus de conversaciones que en los datos del corpus de textos leídos. Sin embargo, también es posible que la línea divisoria entre habla leída y habla espontánea no sea tan fácil de demarcar con los parámetros analizados. Así lo sugieren los resultados preliminares de esta investigación sobre las características acústicas de las vocales bogotanas en habla leída y habla narrada (Correa, 2017). De este estudio se desprende, por ejemplo, que en habla leída las vocales tienen una realización más abierta, pero la duración, el segundo formante y la dispersión vocálica no se diferencian de los valores obtenidos de los datos extraídos de narraciones producidas por los mismos hablantes.

La primera pregunta también busca caracterizar la organización del espacio acústico y la centralización, entendida como la distancia entre la frecuencia de un timbre vocálico y el centro de los valores acústicos del hablante, entendido como el valor medio de F1 y F2 de todas las vocales. Gracias a esta comparación podremos saber cómo organizan los hablantes su espacio acústico en las diferentes velocidades de lectura (normal, lenta y rápida) y en habla espontánea. Vimos que una posibilidad, como predicen el principio de atención de Labov (1972) y los planteamientos de Lindblom (1990), es que el espacio acústico y la centralización se modifiquen en función de las diferentes tareas utilizadas en este trabajo para la recolección de datos. Como demuestran los resultados presentados en el Capítulo 4, otro resultado posible es que el AEV y la TCF también dependan de la tonicidad, la variación fonética individual y el sexo del hablante. Por esta razón, la segunda pregunta abordada en este capítulo pretende responder a la cuestión de cómo afecta la variación individual en los resultados obtenidos para los parámetros acústicos en los dos estilos de habla.

6.2 Procedimiento de análisis

Para responder a la primera pregunta, que indaga la influencia del estilo de habla sobre la duración y los valores de frecuencia de F1 y F2, comparamos el habla leída a una velocidad normal con el habla espontánea producida por el mismo hablante en una

conversación con un amigo (Sección 6.3). Ambos estilos pueden categorizarse como *habla conectada*, pero hay una diferencia importante entre ambas tareas: en la primera, los hablantes prestaron máxima atención al contenido lingüístico de los textos y a la velocidad de habla, mientras que la segunda tarea se trató de una conversación entre dos personas que tenían una estrecha relación de amistad. Usamos los datos de 18 hablantes de los 22 hablantes; no se incluyeron dos conversaciones puesto que, en un caso, hubo fallos de grabación en la tarea de habla espontánea que comprometieron la calidad de la señal (BOG01 y BOG03). En otro caso, la longitud de las grabaciones no permitió obtener el suficiente número de vocales y contextos diferentes de una misma vocal para comparar los timbres vocálicos en los dos estilos de habla (BOG20 y BOG22). También se compararon las áreas del espacio vocálico (AEV), la dispersión acústica, el índice δ , y la tasa de centralización de los formantes (TCF), calculados con los datos obtenidos en habla lenta, normal, rápida y en habla espontánea.

Para analizar los datos se construyeron modelos de regresión de efectos mixtos con los dos primeros formantes, la duración vocálica, el área del espacio vocálico (AEV), la dispersión acústica y la tasa de centralización de los formantes (TCF) como variables respuesta. Como factores fijos incluimos el timbre (/i, e, a, o, u/), la tonicidad, el punto de articulación de la palabra precedente, el tipo de palabra, el sexo del hablante y la posición dentro del grupo fónico. Los modelos incluyeron, además del intercepto del modelo, interceptos aleatorios para los ítems léxicos y para los hablantes.

6.3 Análisis del habla leída en velocidad normal vs. habla espontánea

6.3.1 Duración

En el modelo de la duración de las vocales producidas en habla leída y habla espontánea (Tabla 21), el intercepto tiene una duración de 80.25 ms y corresponde a la duración de una vocal abierta /a/, átona, producida en habla espontánea por hablantes femeninas, en el entorno consonántico dentoalveolar y en posición final del grupo fónico. De acuerdo con los coeficientes de la regresión, el timbre tiene un efecto altamente significativo sobre la duración ($\chi^2(4)=125.34$, $p < .001$). La duración de las vocales /e, i, u/ está 11 ms por debajo de la duración de /a/, mientras que la vocal posterior media /o/ está más

cercana al intercepto, con una diferencia de -9.17 ms. En el Anexo F, que presenta las mediciones discriminadas por tonicidad y sexo, puede observarse que la duración está relacionada con el grado de abertura: la vocal anterior cerrada /i/ es breve, la vocal abierta /a/ es larga, y las vocales /o, e, u/ tienen duraciones intermedias y con valores semejantes. La tonicidad ($\chi^2(1)=275.35$, $p < .001$) y la posición de la vocal dentro del grupo fónico ($\chi^2(2)=113.85$, $p < .001$) también producen cambios altamente significativos. Como se observa en la Figura 53, las vocales tónicas son 16.34 ms más largas que las átonas y tienen mayor duración en posición final e inicial que en posición intermedia del grupo fónico. El punto de articulación de la consonante precedente tiene un efecto altamente significativo ($\chi^2(2)=16.38$, $p < .001$), ya que las vocales en entorno labial y velar son más breves que en el contexto dentoalveolar. Finalmente, en este modelo los hombres tienen vocales más breves (-6.75 ms) que las mujeres ($\chi^2(1)=9.34$, $p < .002$).

Es importante resaltar dos aspectos. Primero, la diferencia entre habla leída y habla espontánea no resultó estadísticamente significativa. Podemos decir que la duración está influida más bien por factores segmentales y prosódicos como el timbre vocálico, la tonicidad, el punto de articulación de la consonante precedente, la ubicación de la vocal dentro de la frase y, factores extralingüísticos como el sexo del hablante. Pero, contrario a lo que suponen el principio de atención y la teoría de la hiper e hipoarticulación, la duración no se modificó de manera significativa debido al estilo de habla. Segundo, los coeficientes concuerdan con el modelo de la duración en habla leída a tres velocidades (Tabla 3), lo cual confirma los resultados reportados en el Capítulo 4 pero, más importante, indican que la duración no depende de la distinción entre habla leída y habla espontánea.

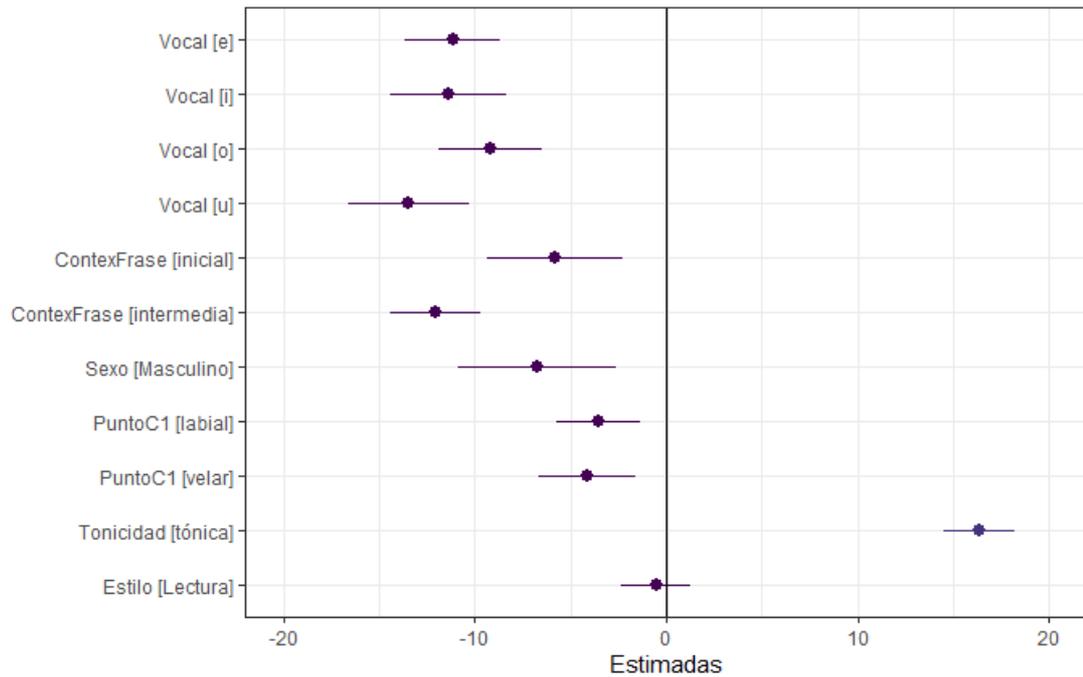


Figura 53. Coeficientes de los factores fijos del modelo de la duración vocálica (ms) en habla leída y habla espontánea.

Tabla 21. Coeficientes de los efectos fijos y de los efectos aleatorios fijos para el modelo de la duración vocálica en habla leída y en habla espontánea con un total de 4223 observaciones y 18 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	80.25	2.12	37.73	<.001
/e/	-11.16	1.25	-8.97	<.001
/i/	-11.36	1.55	-7.30	<.001
/o/	-9.17	1.37	-6.66	<.001
/u/	-13.45	1.60	-8.03	<.001
Tonicidad (tónica)	16.34	0.95	17.09	<.001
Contexto Frase (inicial)	-5.84	1.17	-3.25	.001
Contexto Frase (intermedia)	-12.06	1.18	-10.13	<.001
Sexo (Hombres)	-6.75	2.10	-3.20	.004
PuntoC1(labial)	-3.54	1.13	-3.13	.001
PuntoC1(velar)	-4.13	1.28	-3.21	.001
Estilo(lectura)	-0.5	0.90	-0.60	0.54
Factores aleatorios	Varianza	Std.dev		
Palabra	104.45	10.22		
Participante	18.74	4.32		
Residuos	330.18	18.17		
R ² Marginal = 0.21		R ² Condicional = 0.42		

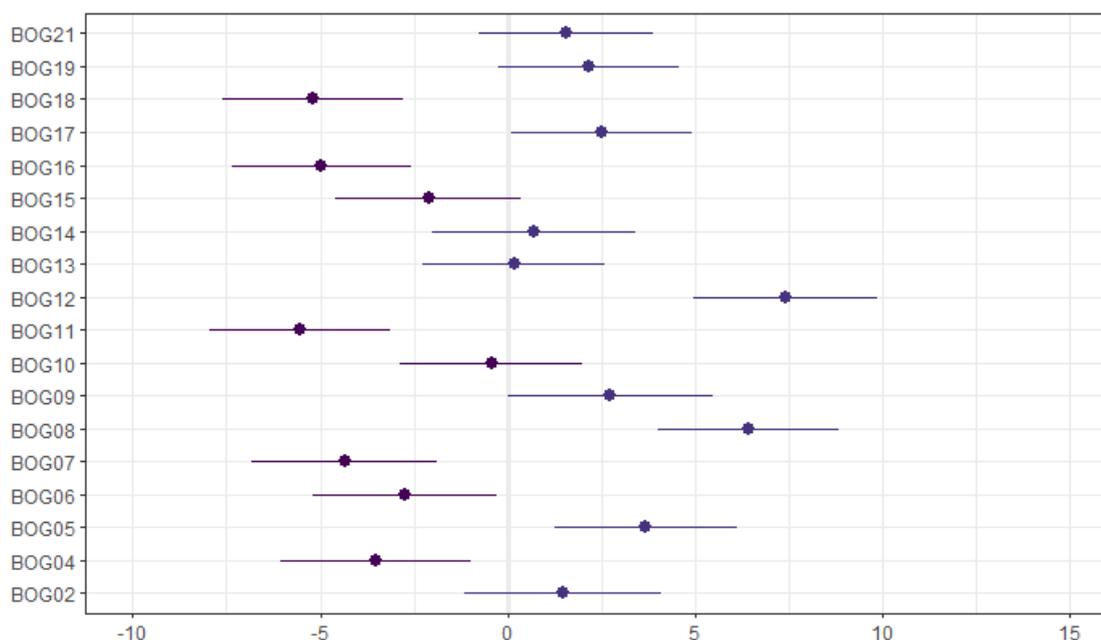


Figura 54. Interceptos aleatorios por hablante para el modelo de la duración en habla leída y habla espontánea.

De acuerdo con los interceptos aleatorios, es claro que los hablantes realizan las vocales en habla leída y en habla espontánea con duraciones muy cercanas y, por tanto, las diferencias no son estadísticamente significativas. Sin embargo, la variación individual refleja algunas tendencias interesantes. Como ilustra la Figura 54, los interceptos aleatorios de algunos participantes tienen duraciones por debajo de los valores estimados (BOG04, BOG06, BOG07, BOG11, BOG16, BOG18), mientras que otros tienen duraciones por encima de los valores de referencia (BOG05, BOG08, BOG12). A juzgar por los valores medios y las desviaciones típicas de cada hablante, es claro que el modelo ubica por encima del intercepto a aquellos participantes que producen una vocal con valores por encima de 100 ms (Anexo F). Un hablante (BOG05) produce las vocales en habla espontánea con una duración media y una desviación más altas que las vocales en habla leída. Otro (BOG09), presenta realizaciones fonéticas de /i/ con una duración particularmente larga en habla espontánea.

6.3.2 Valores de frecuencia de F1 y F2

La Tabla 22 presenta los coeficientes de los factores fijos y de los efectos aleatorios del modelo del primer formante para el habla leída en velocidad normal y el habla

espontánea. El intercepto, como en el caso anterior, corresponde a la vocal abierta /a/, átona, en contexto dentoalveolar, al final de un grupo fónico, y producida por las mujeres en habla espontánea. La Figura 55 muestra los efectos de los factores fijos con relación al intercepto, de tal manera que los factores más alejados son aquellos que ejercen mayor influencia sobre los valores de frecuencia de F1, mientras que los más próximos al intercepto producen menos cambios. Como puede verse, el timbre vocálico es el factor que produce mayores cambios en los valores de frecuencia de F1 ($\chi^2(4)=2955$, $p < .001$). Las vocales medias /e, o/ tienen una diferencia de -127 y 138 Hz respecto de la vocal abierta /a/; las vocales cerradas tienen mayor separación: la vocal anterior /i/ es la vocal más cerrada con una diferencia de -245 Hz respecto de /a/, seguida de /u/, con una diferencia de -209 Hz. Estos valores de frecuencia del F1 son menores a los reportados en el Capítulo 4, pero, en líneas generales, concuerdan con los resultados obtenidos para el habla leída a tres velocidades (Tabla 4, sección 4.3.1).

Ahora bien, en el modelo estadístico para el F1, la tonicidad de la vocal produce cambios significativos ($\chi^2(1)=87.62$, $p < .001$), lo cual se refleja en un incremento de 70.73 Hz en las vocales tónicas respecto de las átonas. Además, hay una interacción altamente significativa entre el timbre vocálico y la tonicidad ($\chi^2(4)=79.71$, $p < .001$): el F1 se incrementa en las realizaciones tónicas de /a/ alrededor de 70 Hz, mientras que disminuye en las vocales medias /o, e/ y, especialmente, en las vocales cerradas /i, u/. Esta interacción también indica que la diferencia entre tónicas y átonas modifica principalmente el F1 de la vocal abierta /a/ y de las vocales medias, sin embargo, el primer formante de las vocales cerradas /i, u/ no cambia significativamente por efecto de la tonicidad (Anexo F).

El punto de articulación de la consonante precedente es una variable altamente significativa para los valores de frecuencia del F1 ($\chi^2(2)=47.43$, $p < .001$). En concreto, las vocales en contexto labial y velar tienen un F1 más alto que el entorno dentoalveolar, si bien la diferencia es mayor en el entorno labial (20.10 Hz) que en el velar (12.39 Hz).

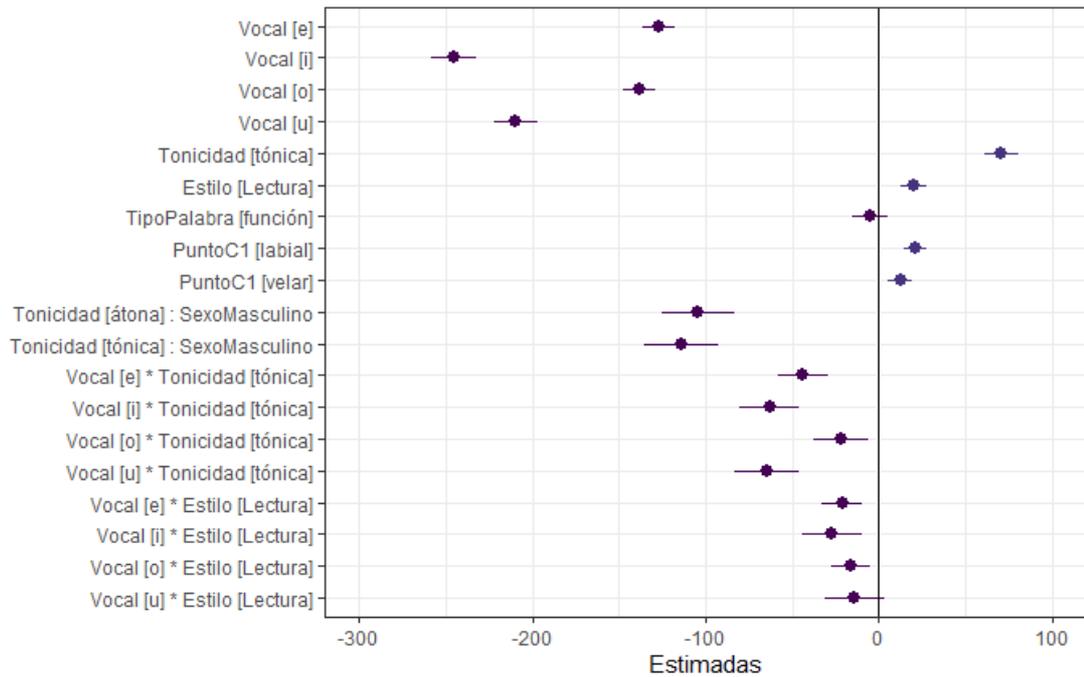


Figura 55. Diagrama de puntos con los coeficientes de los factores fijos en el modelo de los valores de frecuencia de F1 (Hz) de las vocales producidas en habla leída y en habla espontánea.

Tabla 22. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F1 (Hz) con un total de 4071 observaciones y 18 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	635.93	8.53	74.51	<.001
/e/	-127.09	4.69	-27.09	<.001
/o/	-138.40	4.79	-28.89	<.001
/i/	-245.85	6.66	-36.90	<.001
/u/	-209.68	6.44	-32.55	<.001
Tonicidad (tónica)	70.73	5.10	-13.85	<.001
Estilo(lectura)	20.10	4.01	3.66	<.001
Punto art. C1 (labial)	20.99	3.12	6.71	<.001
Punto art. C1 (velar)	12.39	3.60	3.43	<.001
Sexo (masculino)	-104.54	10.73	-9.73	<.001
Tonicidad(tónica): Sexo (masculino)	-9.70	3.63	-2.66	.007
Vocal (/e/): Tonicidad (tónica)	-43.80	7.52	-5.82	<.001
Vocal (/i/): Tonicidad (tónica)	-63.12	8.68	-7.26	<.001
Vocal (/o/): Tonicidad (tónica)	-21.48	8.05	-2.66	.007
Vocal (/u/): Tonicidad (tónica)	-64.84	9.43	-6.87	<.001
Vocal (/e/): Estilo (lectura)	-21.32	5.99	-3.55	<.001
Vocal (/i/): Estilo (lectura)	-27.15	8.72	-3.11	.001
Vocal (/o/): Estilo (lectura)	-16.40	5.62	-2.91	.003
Factores aleatorios	Varianza	Std.dev		
Palabra	577.2	24.03		
Participante	538.2	23.20		
Residuos	2720	52.16		
$R^2_m = 0.76$	$R^2_c = 0.83$			

Los valores de frecuencia de F1 también presentan diferencias muy significativas asociadas al sexo del hablante ($\chi^2(1)=40.19$, $p < .001$). Los hombres tienen un primer formante -104.54 Hz por debajo del valor de frecuencia de las mujeres, lo cual equivale a una diferencia entre sexos del 15.9 %. Hay una interacción entre la tonicidad y el sexo, según la cual el primer formante de las vocales tónicas producidas por los hombres es menor que el producido por las mujeres.

El estilo de habla no produce por sí solo un efecto significativo sobre el primer formante ($\chi^2(1)=2.02$, $p = .156$). Sin embargo, al incluir una interacción entre el estilo y el timbre, el efecto resulta altamente significativo ($\chi^2(4)=18.48$, $p < .001$). La vocal /a/ (el intercepto) es más abierta en habla leída, mientras que las vocales /i, e, o/ son más cerradas en la lectura que en el habla espontánea. Por esta razón, las pendientes de la interacción son negativas y el efecto de la variable estilo es positiva para la vocal abierta /a/ (Tabla 22). En todo caso, los cambios asociados al estilo de habla son bajos si los comparamos con los efectos de factores como el timbre y el sexo (Figura 55). Llama la atención que, de acuerdo con los datos analizados, variables como la posición de la vocal dentro del grupo fónico o el tipo de palabra no tienen efectos significativos sobre el valor de frecuencia de F1.

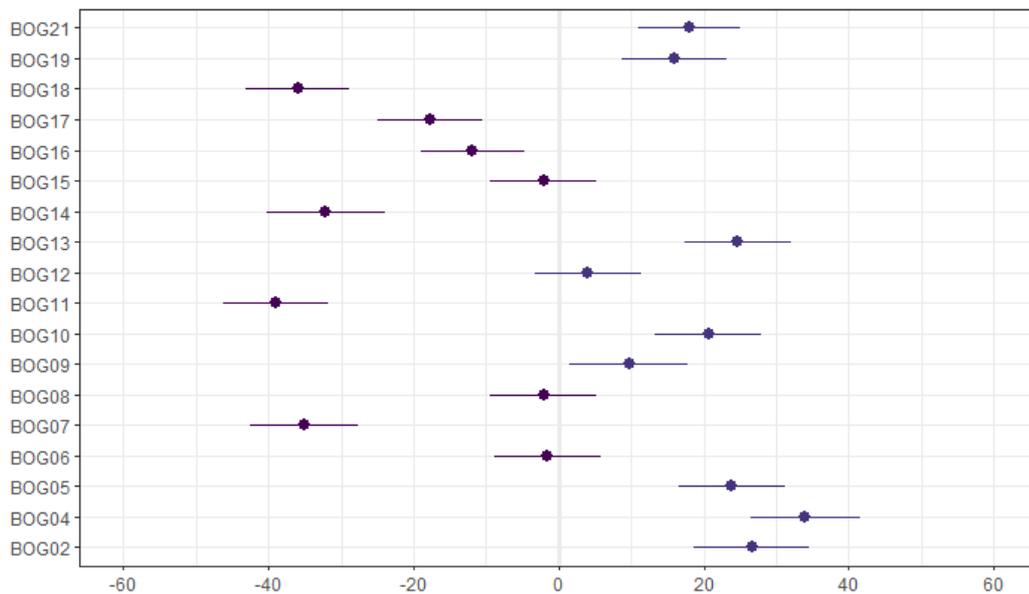


Figura 56. Interceptos aleatorios por hablante para el modelo de los valores de frecuencia de F1.

Ahora bien, nuestro modelo incluyó interceptos aleatorios por participante y por ítem léxico con el objetivo de analizar la variación fonética individual. Como se ve en la Figura 56, algunos hablantes presentaron coeficientes menores a los estimados (BOG07, BOG11, BOG14, BOG18), otros se agruparon alrededor del intercepto, y un tercer grupo tiene vocales más abiertas que el promedio (BOG02, BOG04, BOG13, BOG19, BOG21). Es interesante anotar aquí que la variación sugerida por los interceptos es la misma que la reportada para el modelo de habla leída a tres velocidades, lo cual significa que los valores de frecuencia del F1 característicos de cada locutor se mantienen en los diferentes estilos de habla. De acuerdo con la información de la Tabla 22, los interceptos aleatorios para el ítem léxico y los hablantes explican parte de la varianza observada en el modelo. La varianza explicada por los factores fijos o varianza marginal (R^2_m) fue del 76 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) fue del 83 %.

Construimos un modelo con los valores normalizados del primer formante y encontramos los mismos factores y cambios descritos para los datos en hercios. La única excepción es la interacción entre el timbre y la tonicidad, que resulta significativa únicamente para la vocal abierta /a/ y las vocales medias /e, o/. Los estadísticos del modelo con datos normalizados son bastante cercanos a los reportados en el modelo elaborado con datos en hercios (Tabla 4, secciones 4.3.2 y 4.3.3). En este caso, la varianza explicada por los factores fijos o varianza marginal (R^2_m) fue del 76 %, mientras que la varianza explicada por el modelo o varianza condicional (R^2_c) fue del 81 %.

En cuanto al modelo para el segundo formante (Figura 57 y Tabla 23), los resultados muestran que los valores de frecuencia de F2 están influidos por el timbre vocálico ($\chi^2(4)=2879$, $p < .001$), el sexo ($\chi^2(1)=36.68$, $p < .001$), el punto de articulación de la consonante precedente ($\chi^2(2)=186.62$, $p < .001$), por el estilo de habla ($\chi^2(1)=21.43$, $p < .001$) y la tonicidad ($\chi^2(1)=12.38$, $p < .001$). Encontramos interacciones altamente significativas entre el timbre y el estilo habla ($\chi^2(4)=48.43$, $p < .001$), y entre el timbre y la tonicidad ($\chi^2(4)=155.63$, $p < .001$). Los coeficientes indican, de una parte, que las vocales anteriores /i, e/ tienen un valor de F2 más alto en habla leída que en habla espontánea y, de otra, que las realizaciones átonas de las vocales anteriores son retraídas [i̠, e̠] y las realizaciones átonas de las vocales posteriores son avanzadas [o̠, u̠], pero, especialmente, la interacción hace ver que este efecto produce cambios importantes

sobre el timbre de la vocal posterior /u/. Otros factores, como el tipo de palabra y la posición de la vocal dentro del grupo fónico, no resultaron estadísticamente significativos.

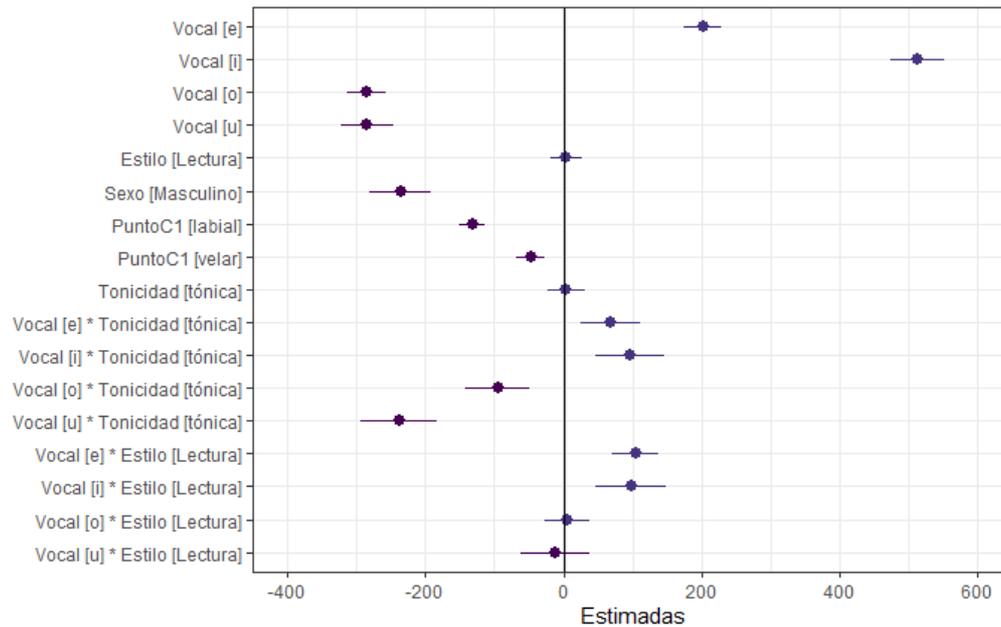


Figura 57. Coeficientes de los factores fijos del modelo de los valores de frecuencia de F2 (Hz) de las vocales producidas en habla leída y en habla espontánea.

Tabla 23. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de F2 (Hz) con un total de 4071 observaciones y 18 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	1701.77	19.60	86.79	<.001
/e/	202.23	13.88	14.56	<.001
/o/	-285.57	14.23	-20.06	<.001
/i/	513.12	19.67	26.08	<.001
/u/	-284.15	19.06	-14.90	<.001
Punto art. C1 (labial)	-132.08	9.33	-14.14	<.001
Punto art. C1 (velar)	-47.62	10.66	-4.46	<.000
Sexo (masculino)	-236.27	23.03	-10.26	<.001
Vocal (/e/): Tonicidad (tónica)	68.62	22.30	3.07	.002
Vocal (/i/): Tonicidad (tónica)	96.83	25.76	3.75	.000
Vocal (/o/): Tonicidad (tónica)	-95.25	24.07	-3.95	<.001
Vocal (/u/): Tonicidad (tónica)	-238.39	28.41	-8.38	<.001
Vocal (/e/): Estilo (lectura)	104.13	17.61	5.91	<.000
Vocal (/i/): Estilo (lectura)	97.81	25.96	3.76	.000
Factores aleatorios		Varianza	Std.dev	
Palabra	6457	80.35		
Participante	2368	48.66		
Residuos	22218	149.06		
$R^2_m = 0.78$		$R^2_c = 0.84$		

La Figura 57 y la Tabla 23 muestran que el intercepto corresponde a la vocal central abierta /a/ y, por tanto, las vocales posteriores tienen un F2 con valores por debajo de la estimada (/o/ = -285.57 Hz, u = -284.15 Hz), mientras que las vocales anteriores se ubican por encima de esta (/e/ = 202.23 Hz, /i/ = 513.12 Hz). Como ya se ha observado para los valores del F1, el F2 tiene un valor más bajo en habla masculina (-236.27 Hz) y esta diferencia entre hombres y mujeres es altamente significativa ($\chi^2(1)=36.68$, $p < .001$). El punto de articulación de la consonante precedente también tiene un efecto altamente significativo sobre el F2 de la vocal ($\chi^2(2)=186.62$, $p < .001$). En concreto, las vocales en contexto labial y velar tienen un F2 más bajo que el entorno de consonantes dentoalveolares, si bien la diferencia es mayor en el entorno labial (-132.08 Hz) que en el velar (-47.62 Hz).

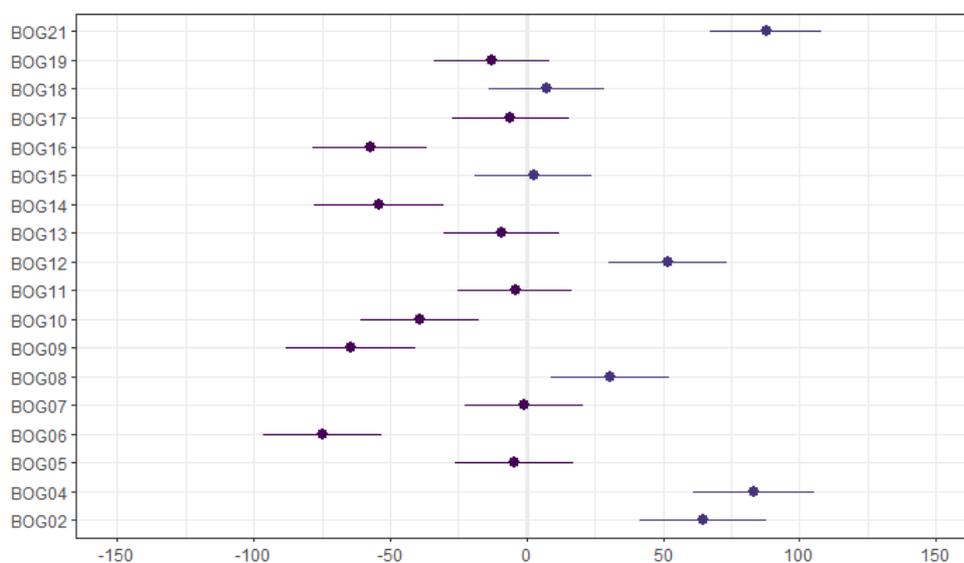


Figura 58. Interceptos aleatorios por hablante para el modelo de los valores de F2.

Los interceptos aleatorios (Figura 58) reflejan que, como hemos visto en los resultados reportados a lo largo de esta tesis, algunos hablantes tienen interceptos por encima o por debajo del valor estimado. En el caso de los valores de frecuencia de F2, algunos hombres presentan valores por encima del intercepto (BOG21), y algunas mujeres tienen interceptos con valores por debajo de las estimadas (BOG06, BOG09, BOG14, BOG16). Esta variación fonética individual manifiesta las mismas tendencias en habla leída a tres velocidades (Tabla 5, sección 4.3.2). La varianza explicada por los factores fijos o varianza marginal (R^2_m) es del 78 %, mientras que la varianza explicada por el modelo

o varianza condicional (R^2_c) es del 84 %. Se construyó un modelo con los valores normalizados del segundo formante y se encontraron los mismos factores y cambios descritos para los datos en hercios.

En este punto de la exposición, los diagramas de la distribución y las desviaciones de las frecuencias de F1 y F2 resultan útiles para caracterizar los resultados de los modelos estadísticos presentados con anterioridad. La Figura 59 es un diagrama de los valores de frecuencia de los dos primeros formantes en hercios (Hz) que ilustra claramente la magnitud de las diferencias acústicas entre hombres y mujeres. En esta figura también podemos ver que, como sugieren los modelos estadísticos, los formantes en habla espontánea son más variables y tienden más a la centralización que los valores de los formantes del habla leída. Esta tendencia es más sobresaliente en la vocal anterior media /e/, pero, como se puede observar, afecta en general a todas las vocales bogotanas.

La Figura 60 ilustra los datos normalizados con el método de Lobanov (1971). En la izquierda se comparan las vocales tónicas con las átonas, y en la derecha se compara nuevamente el habla leída con el habla espontánea. La tonicidad es un factor que modifica las frecuencias de los formantes, pero de una manera no uniforme, esto quiere decir que el efecto depende del timbre. Nótese, por ejemplo, que la vocal posterior cerrada /u/ y la vocal abierta /a/ se desplazan al centro del espacio acústico, pero, en un caso la centralización afecta al grado de abertura y, en el otro, al grado de anterioridad/posterioridad de la vocal.

Las variaciones relacionadas con el timbre siguen, en general, los mismos patrones reportados en el análisis de los formantes producidos a tres velocidades de habla. El primer formante de las vocales posteriores /o, u/ se solapa en los dos estilos habla, pero especialmente cuando la vocal es átona. Las desviaciones típicas del segundo formante de estas vocales son altas, lo cual, como se explicó en el capítulo anterior, indica una fuerte influencia de las consonantes dentoalveolares. La vocal abierta /a/ tiene realizaciones abiertas, medio abiertas y centralizadas en ambos estilos, lo cual es más sobresaliente en los formantes producidos por los hombres que en los producidos por las mujeres. La vocal anterior media /e/ también es bastante variable y, como se dijo anteriormente, es el timbre que presenta mayores cambios en habla espontánea.

Finalmente, la vocal cerrada /i/, si bien tiene una realización que varía a lo largo del segundo formante, es la vocal más estable en las diferentes condiciones estudiadas.

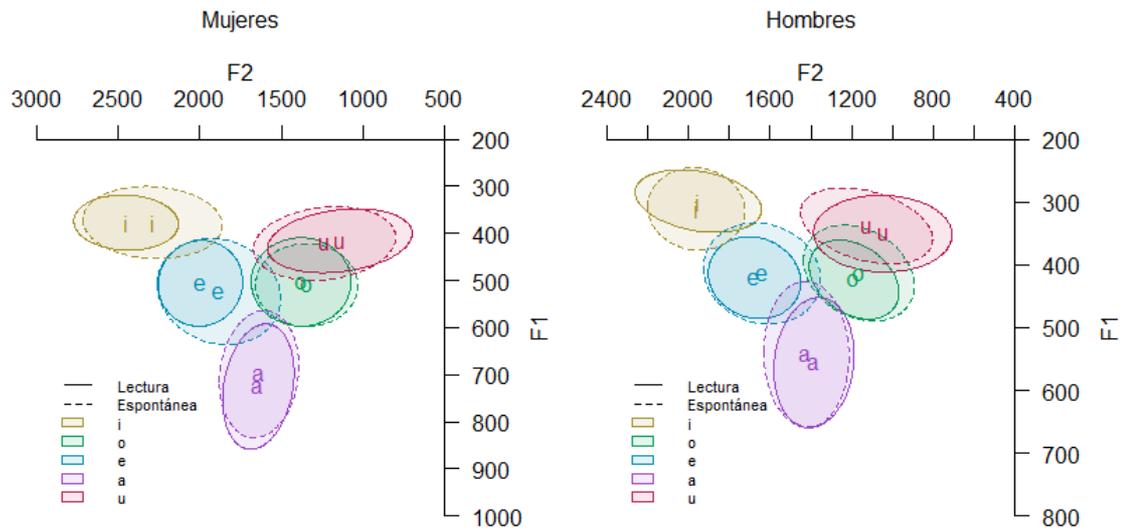


Figura 59. Valores medios y las desviaciones típicas de los valores de frecuencia de F1 y F2 (Hz) en función del sexo del hablante y el estilo de habla.

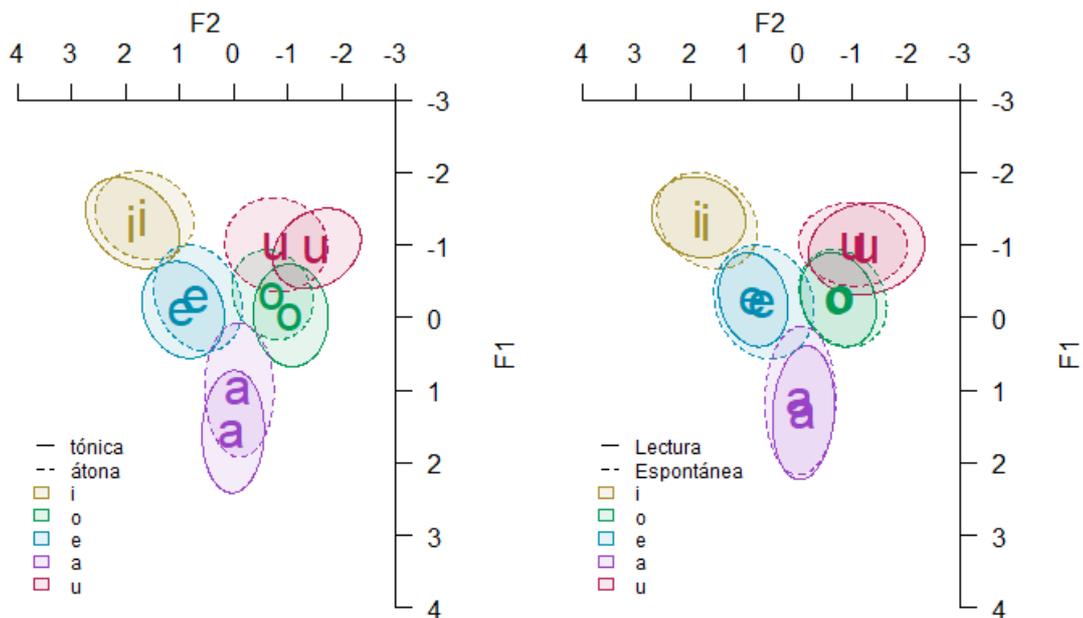


Figura 60. Diagrama con los valores medios y las desviaciones típicas de los valores de frecuencia de F1 y F2 (Lobanov) de las vocales tónicas y átonas producidas en habla leída y habla en espontánea.

6.4 Área vocálica, centralización y dispersión de las vocales

6.4.1 Área del espacio vocálico (AEV)

En el modelo estadístico del área del espacio vocálico (AEV) se incluyen como factores fijos el estilo de habla, la tonicidad de la vocal y el sexo del hablante. En este análisis, hay cuatro niveles en la variable estilo (habla leída a una velocidad normal, lenta y rápida, y habla espontánea), pues, recordemos, el objetivo de este diseño estadístico es saber cómo organizan los hablantes su espacio acústico en una escala desde un habla más hiperarticulada hasta un habla más hipoarticulada, y observar si el habla a velocidad rápida posee el mismo grado de centralización que el habla espontánea. El modelo también incorpora el hablante como intercepto aleatorio y una interacción entre la tonicidad y el estilo. En la Tabla 24 se resumen los coeficientes y el nivel de significación del modelo del AEV. Como puede verse, la tonicidad produce cambios altamente significativos sobre el AEV ($\chi^2(1)=173.78$, $p < .001$), ya que el AEV de las vocales tónicas se incrementa 2.45 Bark² respecto del AEV de las vocales átonas. El estilo de habla también produce cambios muy significativos ($\chi^2(3)=90.23$, $p < .001$), que se reflejan en un área más amplia para la velocidad lenta, un área intermedia en la velocidad normal y en el habla espontánea, y la menor área acústica en la velocidad rápida. Hay una interacción altamente significativa entre el estilo y la tonicidad ($\chi^2(3)=9.36$, $p < .001$) de acuerdo con la cual la diferencia entre el AEV de las átonas y las tónicas es mayor en velocidad lenta y velocidad normal que en habla espontánea. El AEV de los hablantes masculinos es menor que el de los femeninos y, de acuerdo con la prueba de significación, esta diferencia es altamente significativa ($\chi^2(1)=10.04$, $p < .001$).

Como se observa en la Figura 61, el AEV varía en función de la tonicidad, el estilo de habla y el sexo del participante. En concordancia con los coeficientes del modelo, podemos ver que, en las diferentes condiciones, el habla lenta tiene la mayor AEV y el habla rápida tiene la menor AEV. En cambio, el habla espontánea no se diferencia del habla leída en velocidad normal. Esto se refleja en una superposición de las áreas vocálicas y en la poca diferenciación del habla espontánea de las demás condiciones de elocución.

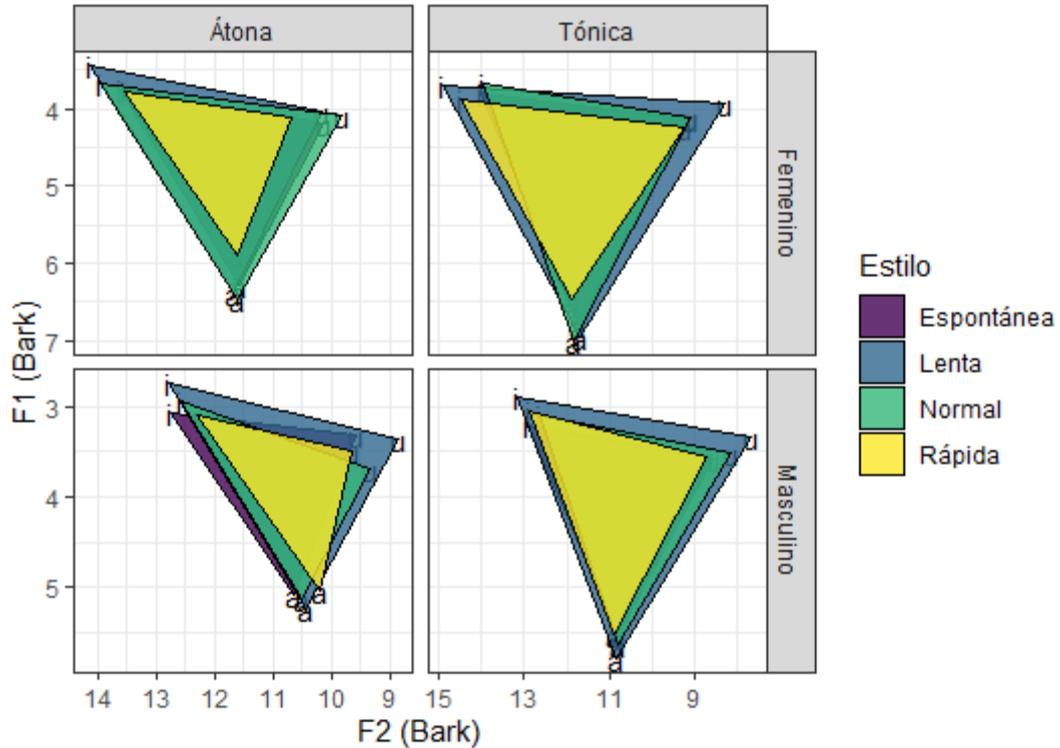


Figura 61. Área del espacio vocálico (Bark) en función de la velocidad y el estilo de habla de 18 hablantes del español bogotano.

Tabla 24. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de Área del espacio vocálico (AEV) con un total de 144 observaciones y 18 participantes.

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	4.56	0.48	9.44	<.001
Tonicidad (tónica)	2.45	0.37	6.54	<.001
Espontánea Vs. Lenta	1.25	0.37	3.34	.001
Espontánea Vs. Normal	0.12	0.37	-5.87	.73
Espontánea Vs. Rápida	-0.88	0.37	0.33	.019
Sexo (masculino)	-2.15	0.58	-3.66	.001
Tonicidad(tónica):Velocidad(lenta)	2.06	0.53	3.88	.000
Tonicidad(tónica):Velocidad(normal)	1.92	0.53	3.61	.000
Factores aleatorios	Varianza	Std.dev		
Participante	1.39	1.17		
Residuos	1.08	1.04		
$R^2_m = 0.68 \quad R^2_c = 0.85$				

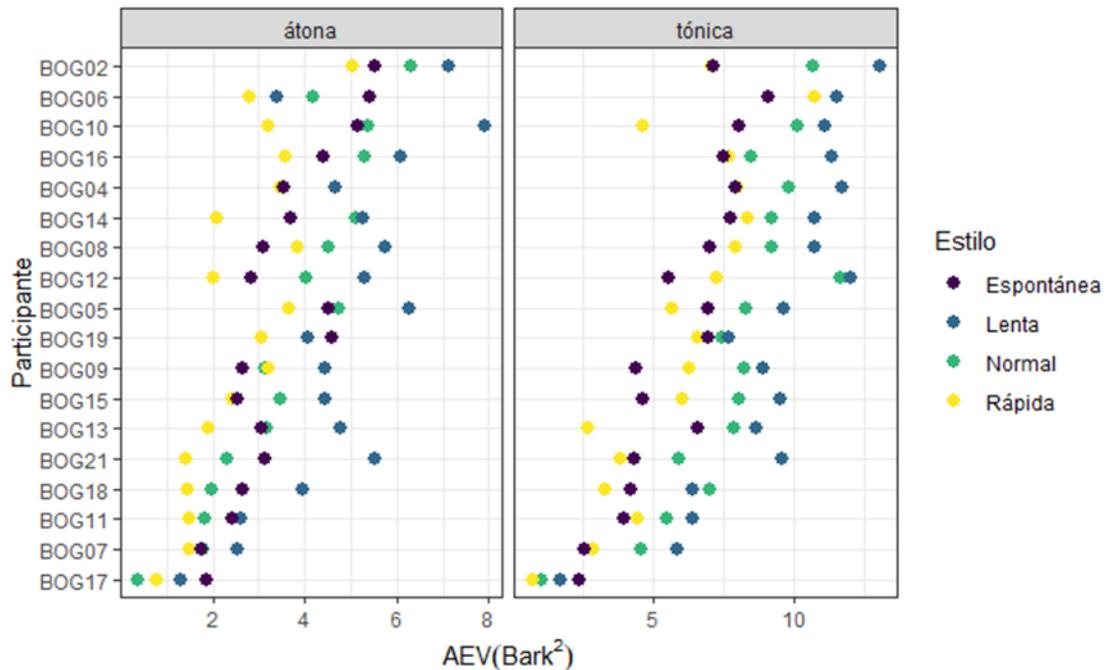


Figura 62. Área del espacio vocálico (AEV) en Bark² por hablante y en función de la velocidad y el estilo de habla.

La Figura 62 presenta la variación fonética individual del AEV en habla leída en tres velocidades y en habla espontánea; la ordenada muestra los hablantes de menor a mayor AEV y la abscisa el área acústica expresada en Bark². Como predice el modelo, las vocales tónicas tienen un espacio acústico más amplio que las átonas en las producciones de todos los hablantes. Como se recoge en la Figura 62, la variación individual es amplia y no resulta fácil de generalizar. En todo caso, podemos decir, en primer lugar, que el AEV más amplia corresponde a las tónicas producidas en velocidad lenta, tiene un valor intermedio en velocidad normal, y la menor AEV se observa en las átonas producidas en velocidad rápida. Por su parte, el AEV del habla espontánea depende del hablante y de la tonicidad. Las tónicas en habla espontánea tienen menor área que la lectura en velocidad rápida y normal, si bien algunos participantes producen las vocales en habla espontánea con mayor AEV que en velocidad rápida. En tanto, las átonas producidas en habla espontánea tienen mayor AEV que la velocidad rápida y menor área que la lectura en velocidad normal. Llama la atención que algunos hablantes (BOG07, BOG17, BOG11 y BOG18) tienen un AEV bastante reducida en todos los estilos habla, en las realizaciones tónicas y en las átonas.

6.4.2 Tasa de centralización de los formantes (TCF)

Recordemos que la Tasa de centralización de los formantes (TCF) es una medición que normaliza los valores de frecuencia de los formantes y estima el grado de centralización del espacio vocálico (Sapir, Raming y Spielman, 2010). Los valores varían entre 1 y 2, donde un valor de 1 está asociado a una realización hiperarticulada y uno de 2 indica la máxima centralización. El modelo mixto incluye la TCF como variable dependiente y la tonicidad, el estilo de habla y el sexo como factores fijos. De acuerdo con los coeficientes del modelo (Tabla 25 y Figura 63), las vocales tónicas son menos centralizadas que las átonas ($\chi^2(1)=136.95$, $p < .001$), el estilo de habla tiene un efecto altamente significativo ($\chi^2(3)=89.34$, $p < .001$), pero el sexo no tiene un efecto sobre el grado de centralización de las vocales ($\chi^2(1)=0.32$, $p = .570$).

En la Figura 63 puede verse que las vocales se organizan, de menor a mayor centralización, así: *habla lenta* > *habla normal* > *habla espontánea* > *habla rápida*. Los resultados muestran que los valores de la TCF también dependen de la variación fonética individual. Por ejemplo, unos hablantes tienen un grado de centralización para cada estilo, mientras que otros, como predice el modelo, tienen los mismos valores para el habla espontánea y el habla rápida, pero otros tienen una tasa de centralización similar en el habla espontánea y en la velocidad normal. Es importante reiterar que la variación individual determina en parte la tasa de centralización de los formantes, pero no está asociada al sexo del hablante.

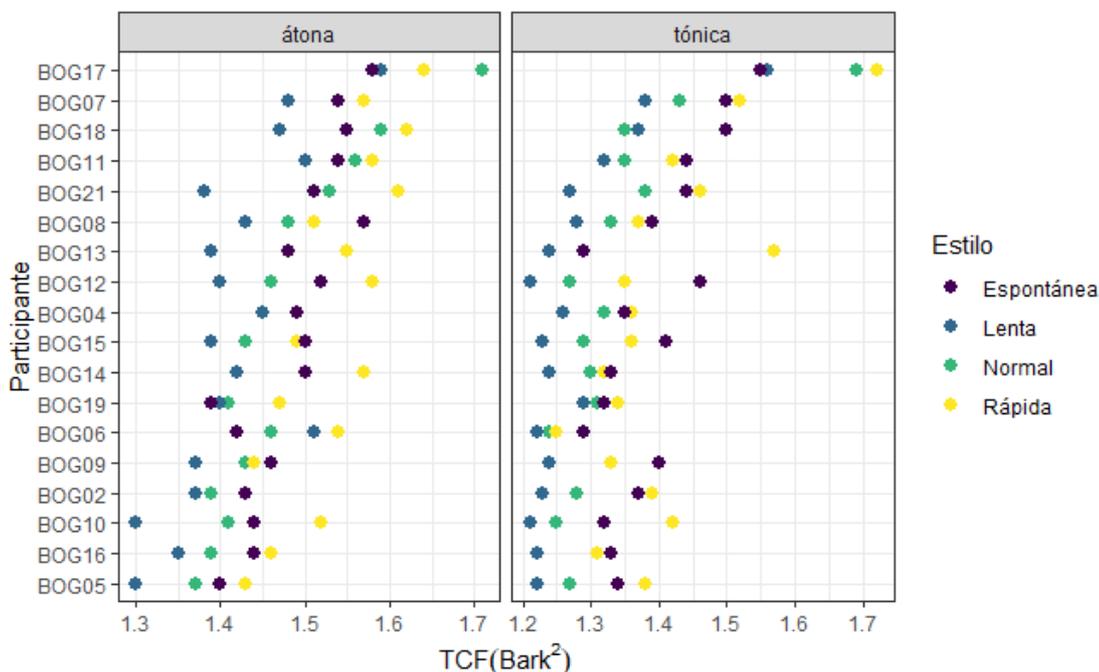


Figura 63. Diagrama de puntos con el Área del espacio vocálico (AEV) en Bark² por hablante y en función de la velocidad y estilo de habla.

Tabla 25. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la Tasa de Centralización de los Formantes (TCF) con un total de 144 observaciones y 18 participantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto	1.48	0.023	62.66	<.001
Tonicidad (tónica)	-0.12	0.007	-15.73	<.001
Espontánea Vs. Lenta	-0.09	0.011	-8.14	<.001
Espontánea Vs. Normal	-0.03	0.011	-3.30	.001
Sexo (masculino)	0.03	0.031	62.66	0.22
Factores aleatorios	Varianza	Std.dev		
Participante	0.004	0.064		
Residuos	0.002	0.047		
$R^2_m = 0.49$	$R^2_c = 0.82$			

6.4.3 Correlación entre el espacio vocálico y la tasa de centralización

Es importante anotar que los hablantes que presentan un AEV amplia tienen, en general, una menor tasa de centralización y, viceversa, los hablantes con mayor TCF tienen una menor AEV. Realizamos una correlación entre la AEV y TCF para saber si, como predice la teoría de Lindblom (1990), el espacio acústico y la centralización se organizan en un continuo entre la hiper e hipoarticulación.

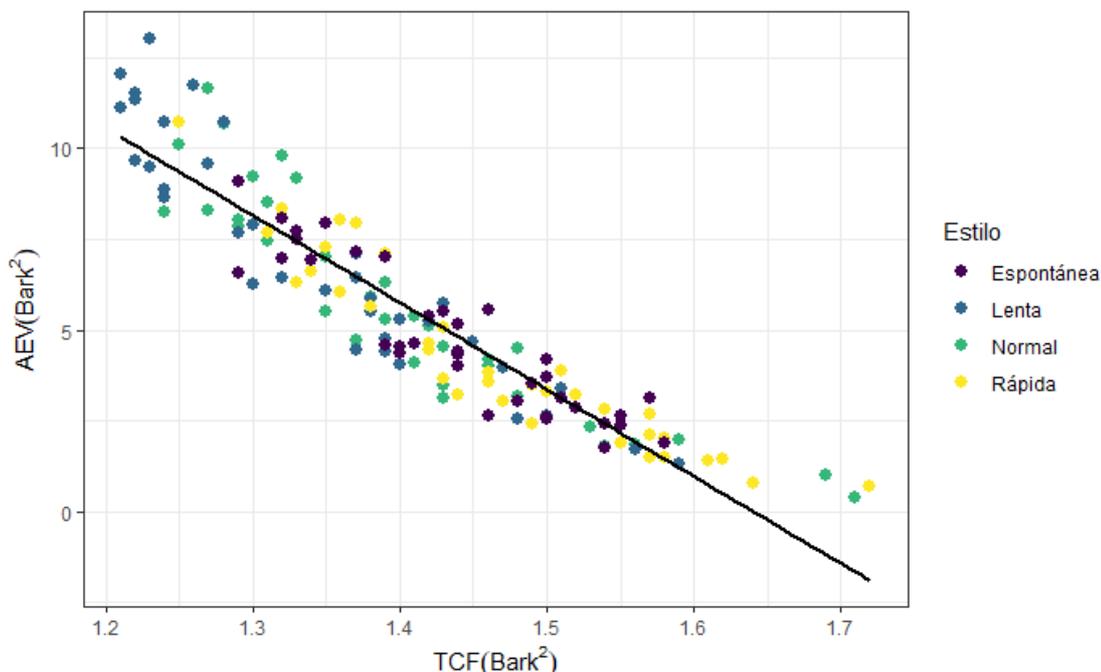


Figura 64. Correlación entre el área del espacio vocálico y tasa de centralización de los formantes.

La correlación entre el AEV y la TCF es negativa, es decir, cuando el área del espacio vocálico aumenta, disminuye la centralización de las vocales y cuando aumenta la centralización, disminuye el AEV. La correlación entre las variables es fuerte y estadísticamente significativa ($r = -0.932$, $p < .001$). La Figura 64 presenta los datos correspondientes a las vocales producidas por 18 hablantes a tres velocidades de habla y en habla espontánea. Nótese que, si bien las diferencias entre estilos no son categóricas, podemos ver claramente que los participantes en nuestro estudio organizaron el espacio acústico de las vocales en un continuum entre realizaciones hiperarticuladas (habla lenta) y realizaciones hipoarticuladas, que estuvieron asociadas a la velocidad rápida, a la velocidad normal y al habla espontánea según las características individuales del hablante.

6.4.4 Dispersión vocálica

La dispersión vocálica se calculó por medio de la distancia euclídea en función del timbre vocálico, la tonicidad, el sexo y el estilo. Obtuvimos 720 valores de las vocales tónicas y átonas producidas por 18 hablantes en habla leída a tres velocidades y en habla espontánea. Con estos resultados se construyó un modelo con la distancia euclídea (expresada en escala Bark) como variable dependiente, y la tonicidad, el estilo y el sexo

como predictores. Se incluyeron interceptos aleatorios para los hablantes y los ítems léxicos, una interacción entre el timbre vocálico y la tonicidad, y otra, entre la tonicidad y el estilo de habla.

En el modelo de la dispersión vocálica (Tabla 26), el intercepto corresponde a la distancia de la vocal abierta /a/, realizada como una átona, y producida en habla espontánea por el grupo de mujeres. Los coeficientes de los factores fijos indican que el timbre tiene un efecto altamente significativo sobre la dispersión vocálica ($\chi^2(4)=1072$, $p < .001$). En este sentido, las vocales cerradas (/i/ y /u/) están más alejadas del intercepto y las vocales medias /e/ y /o/ están bastante próximas. En otras palabras, las vocales medias tienen realizaciones tan centralizadas como las de la vocal abierta /a/. De hecho, como se ve en la Figura 65, la distancia entre estas vocales es bastante próxima en las diferentes condiciones analizadas.

La tonicidad produce cambios altamente significativos sobre la dispersión vocálica ($\chi^2(1)=435$, $p < .001$). Este efecto depende del timbre, ya que, si bien las tónicas tienen menor dispersión que las átonas, la diferencia es especialmente notoria en la vocal posterior /u/ y en la vocal media /e/ ($\chi^2(4)=129.67$, $p < .001$). La dispersión vocálica se organizó en función del estilo, con una mayor distancia para la velocidad lenta, una distancia intermedia para la velocidad normal y el habla espontánea, y la menor distancia para la velocidad rápida ($\chi^2(3)=222$, $p < .001$). Hay una interacción entre la tonicidad y el estilo que produce un incremento de la distancia euclídea de las tónicas en habla leída a velocidad normal y lenta.

De acuerdo con nuestros resultados, las mujeres tienen un valor más alto para la distancia euclídea que los hombres ($\chi^2(1)=7.45$, $p = .006$). En línea con los resultados reportados en la sección 4.3.6, la dispersión de las vocales femeninas es 0.3 Bark más alta (147 Hz aprox.) que las vocales masculinas. Esto no quiere decir, insistimos, que todos los participantes se comportan de la misma manera: un grupo de hombres y mujeres presenta interceptos por debajo de los valores predichos (BOG04, BOG07, BOG17, BOG18), mientras que otro grupo de hablantes de ambos sexos tiene valores cercanos al intercepto del modelo (vocal /a/ del habla femenina en velocidad lenta).

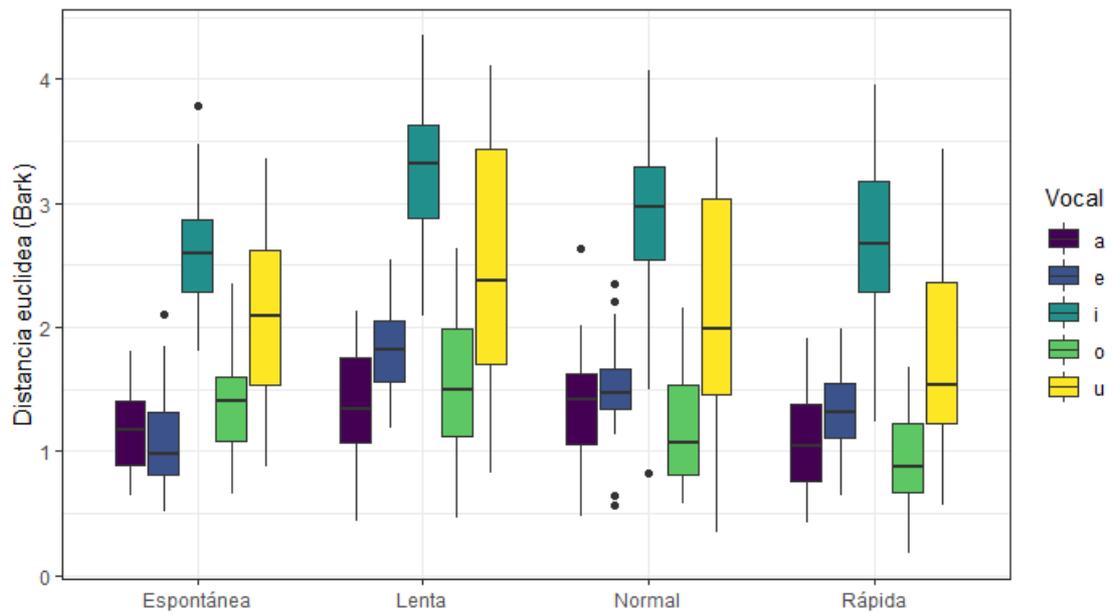


Figura 65. Dispersión vocálica en función de la velocidad de elocución, el estilo y el timbre.

Tabla 26. Coeficientes de los efectos fijos y los efectos aleatorios incluidos en un modelo de la dispersión vocálica (distancia euclídea) con 720 observaciones y 18 hablantes

Predictores	Estimada	Error estándar	Valor de t	Valor de p
Intercepto /a/	1.07	0.09	11.79	<.001
/e/	0.28	0.05	4.99	<.001
/i/	1.70	0.05	30.34	<.001
/o/	0.00	0.05	10.08	.88
/u/	0.56	0.05	9.90	<.001
Tonicidad (tónica)	0.41	0.07	5.82	<.001
Espontánea vs. Lenta	0.31	0.05	6.25	<.001
Espontánea vs. Normal	0.04	0.05	-2.87	0.37
Espontánea vs. Rápida	-0.14	0.05	-7.86	.004
Sexo (masculino)	-0.33	0.10	-3.03	.007
Tonicidad(tónica):Velocidad(lenta)	0.23	0.07	3.28	.001
Tonicidad(tónica):Velocidad(normal)	0.22	0.07	3.15	.001
Vocal(/u/): Tonicidad (tónica)	0.65	0.07	8.20	<.001
Factores aleatorios		Varianza	Std.dev	
Participante		0.05	0.22	
Residuos		0.11	0.33	
$R^2_m = 0.78$		$R^2_c = 0.84$		

6.4.5 Índice ∂

El índice ∂ permite establecer, para cada timbre vocálico, la diferencia entre las distancias euclídeas de dos estilos de habla (Poch-Olivé, Dhainaut y Harmegnies, 2003, p. 523). Los valores cercanos a cero indican que no hay diferencia entre los dos estilos; mientras un incremento en el índice ∂ sugiere que la dispersión de las vocales es diferente en las condiciones comparadas. Para analizar nuestros datos, calculamos la diferencia entre la distancia euclídea del habla lenta (la pronunciación más hiperarticulada de nuestro corpus) y la distancia euclídea de las vocales pronunciadas en habla leída a velocidad rápida, normal y en habla espontánea. También comparamos la velocidad normal con el habla espontánea, y la velocidad normal con el habla rápida. Finalmente, se calculó el índice ∂ para saber si hay distinción entre las condiciones más hipoarticuladas, a saber, el habla espontánea y la velocidad rápida. En resumen, se compararon las siguientes condiciones:

- ∂_1 : Habla lenta–habla rápida
- ∂_2 : Habla lenta–habla espontánea
- ∂_3 : Habla lenta–habla normal
- ∂_4 : Habla normal–habla espontánea
- ∂_5 : Habla normal–habla rápida
- ∂_6 : Habla espontánea–habla rápida

La Figura 66 y la Tabla 27 representan el índice ∂ y sus valores en las seis situaciones en función del timbre vocálico. Las diferencias son difíciles de generalizar, pero es claro que los estilos son más diferentes cuando se compara el habla lenta con las demás condiciones de elocución (índices ∂_1 , ∂_2 y ∂_3) y, en segundo lugar, cuando se compara el habla normal con el habla espontánea o rápida (∂_4 y ∂_5). Las menores diferencias se observan entre el habla espontánea y el habla rápida (∂_6), pero, en todo caso, no se trata de estilos equivalentes. También se aprecia un efecto del timbre: las vocales anteriores /i, e/ presentan diferencias en todas las comparaciones, a excepción de ∂_6 . De la misma manera, los índices de la vocal posterior cerrada /u/ se diferencian en todas las

condiciones, menos entre el habla normal y el habla espontánea (∂_4). La vocal abierta /a/ tiene la menor influencia de la condición de elocución, aunque se aprecian diferencias entre el habla lenta y rápida (∂_1) y entre el habla normal y rápida (∂_5). La vocal posterior media /o/ es bastante variable, pero las diferencias sobresalen cuando se compara cualquier condición con el habla rápida.

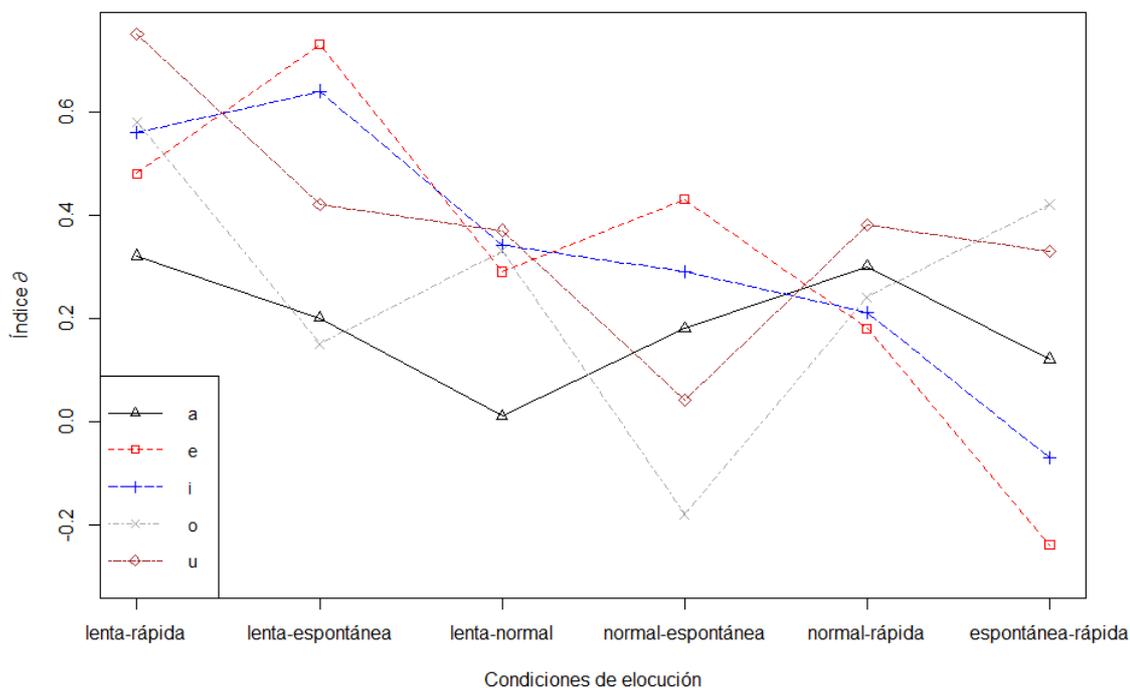


Figura 66. Índice ∂ para cuatro condiciones de elocución en función del timbre.

Tabla 27. Índice ∂ para las vocales /i, e, a, o, u/ según la tonicidad en cuatro condiciones de elocución

Vocal	D1	D2	D3	D4	D5	D6
/i/	0.56	0.64	0.34	0.29	0.21	-0.07
/e/	0.48	0.73	0.29	0.43	0.18	-0.24
/a/	0.32	0.20	0.01	0.18	0.30	0.12
/o/	0.58	0.15	0.33	-0.18	0.24	0.42
/u/	0.75	0.42	0.37	0.04	0.38	0.33

En resumen, el índice ∂ permite ver que las diferencias acústicas entre estilos son grandes cuando se compara una pronunciación hiperarticulada (habla lenta) con una pronunciación hipoarticulada (habla rápida o habla espontánea), y son pequeñas cuando

se compara el habla leída a velocidad normal con el habla espontánea, y cuando se compara el habla normal y el habla espontánea con el habla rápida.

6.5 Conclusiones parciales

La primera pregunta de investigación abordada en este capítulo fue ¿cómo afectan los dos estilos de habla (habla leída vs. habla espontánea) en las características temporales y espectrales de las vocales considerando el timbre vocálico, la tonicidad, el contexto consonántico, el tipo de palabra, la posición de la vocal dentro del grupo fónico y el sexo del hablante? A partir del análisis cuantitativo de las vocales leídas a velocidad normal y de las vocales producidas en habla espontánea, podemos concluir que:

- 1) La duración del habla normal no presenta diferencias significativas con el habla espontánea. La duración está influida, en cambio, por factores segmentales y prosódicos como el timbre, la tonicidad, el entorno consonántico y la posición de la vocal dentro del grupo fónico. Con relación al timbre, las vocales del español hablado en Bogotá se organizan, de mayor a menor duración, así: /a/ > /o/ > /e/ > /u/ > /i/ (Anexo F). Las vocales tónicas son, de acuerdo con nuestro modelo, 16.34 ms más largas que las átonas y tienen mayor duración en posición final e inicial que en posición intermedia del grupo fónico. El punto de articulación de la consonante precedente tiene un efecto altamente significativo, ya que las vocales en un entorno labial o velar son más breves que en el contexto dentoalveolar. Finalmente, en nuestros datos, los hombres tienen vocales más breves que las mujeres.
- 2) Los valores de frecuencia del primer formante están determinados por el timbre, el sexo del locutor, la tonicidad, el estilo y el punto de articulación de la consonante precedente. Variables como la posición de la vocal dentro del grupo fónico y el tipo de palabra no tienen efectos significativos sobre el primer formante de las vocales de nuestro corpus. En cuanto a los cambios producidos por cada uno de estos factores, debemos decir que el timbre produce mayores cambios sobre los valores de frecuencia de F1 debido a su relación con la abertura vocálica. Después podemos mencionar el sexo, pues los hombres tienen un primer formante por debajo del valor de frecuencia de las mujeres, el cual

equivale a una diferencia del 15.9 %. Además, hay una interacción entre la tonicidad y el sexo del participante de acuerdo con la cual el F1 de los hombres es más bajo que el de las mujeres en las vocales tónicas. En cuanto a la tonicidad, vimos que las realizaciones tónicas de las vocales tuvieron un F1 más alto (70.73 Hz) que las átonas. La diferencia entre tónicas y átonas modifica principalmente el F1 de la vocal abierta /a/ y las vocales medias /e, o/. En cuanto al estilo de habla, solo encontramos un efecto homogéneo sobre el F1 cuando este factor interactúa con el timbre: en habla leída la vocal central /a/ es más abierta que en habla espontánea, mientras que las vocales /i, e, o/, por el contrario, son más cerradas en la lectura. Finalmente, el punto de articulación de la consonante precedente también tiene un efecto altamente significativo sobre el valor de frecuencia F1. En concreto, las vocales en contexto labial y velar tienen un F1 más alto que en el entorno dentoalveolar, si bien la diferencia es mayor en el entorno labial (20.10 Hz) que en el velar (12.39 Hz).

- 3) En cuanto al segundo formante, los resultados muestran que los valores de frecuencia de F2 están influidos por el timbre vocálico, el sexo, el punto de articulación de la consonante precedente, por el estilo de habla y la tonicidad. El timbre vocálico interactúa con el estilo y la tonicidad: de una parte, las vocales anteriores /i, e/ tienen un valor de F2 más alto en habla leída que en habla espontánea y, de otra, que las realizaciones átonas de las vocales anteriores son retraídas [i̠, e̠] y las realizaciones átonas de las vocales posteriores son avanzadas [o̟, u̟], pero, especialmente, este efecto produce cambios importantes sobre el timbre de la vocal posterior /u/. En cuanto al sexo del locutor, encontramos que, en nuestro corpus, los hombres tienen un F2 más bajo que las mujeres (-236.27 Hz). Las vocales en contexto labial y velar tienen un F2 más bajo que en el entorno dentoalveolar, si bien la diferencia es mayor en las labiales (-132.08 Hz) que en las velares (-47.62 Hz). Otros factores, como el tipo de palabra y la posición de la vocal dentro del grupo fónico no modificaron de manera significativa los valores de F2.

En cuanto a la organización del espacio vocálico, la dispersión y la centralización en las diferentes condiciones de elocución, se concluye:

- 4) El AEV se organiza en función de la tonicidad, el estilo de habla y el sexo del participante. Así, el AEV de las vocales tónicas se incrementa respecto del AEV de las vocales átonas, el habla lenta tiene la mayor AEV de todas las condiciones de elocución analizadas y el habla rápida tiene la menor área acústica. En cambio, el habla espontánea no se diferencia del habla leída en velocidad normal. Esto se refleja en una superposición de las áreas vocálicas y en la poca diferenciación del habla espontánea de las demás condiciones de elocución. El AEV masculina es menor que la femenina, aunque varios participantes tuvieron áreas semejantes a las del sexo opuesto. El índice ∂ , que establece la diferencia entre la distancia euclídea de dos estilos de habla, permite concluir que las diferencias acústicas entre estilos son más grandes cuando se compara una pronunciación hiperarticulada (habla lenta) con una hipoarticulada (habla rápida o habla espontánea), y son pequeñas cuando se compara el habla leída a velocidad normal con el habla espontánea, y cuando se compara el habla normal y el habla espontánea con el habla rápida.
- 5) En cuanto a la centralización del espacio vocálico y de cada timbre, determinadas a partir de la TCF y la dispersión, podemos decir que el espacio vocálico cambia en función de la tonicidad y la condición de elocución: las vocales tónicas son menos centralizadas que las átonas, y el espacio acústico se organiza, de mayor a menor tasa de centralización, así: *habla lenta* > *habla normal* > *habla espontánea* > *habla rápida*. De acuerdo con los datos sobre la dispersión, las vocales cerradas (/i/ y /u/) son las vocales menos centralizadas, mientras que las vocales más cercanas al centro acústico son la vocal /a/ y las vocales medias /e/ y /o/. En otras palabras, las vocales medias tienen realizaciones tan centralizadas como las de la vocal abierta /a/. La dispersión vocálica varía de la misma manera que la TCF, pues se encontró una mayor distancia para la velocidad lenta, una distancia intermedia para la velocidad normal y el habla espontánea, y la menor distancia para la velocidad rápida. Finalmente, de acuerdo con nuestros resultados las mujeres tienen un valor más alto para la distancia euclídea (menor centralización) que los hombres. De acuerdo con el índice ∂ , la vocal abierta /a/ es menos sensible al estilo de habla que las vocales anteriores /e, i / y la vocal posterior cerrada /u/.

- 6) Existe una relación inversa entre el AEV y el TCF. Los hablantes que presentan un AEV amplia tienen, en general, una menor tasa de centralización y, viceversa, los hablantes con mayor TCF tienen una menor AEV. Si bien las diferencias entre estilos no son categóricas, los participantes en nuestro estudio organizaron el espacio acústico de las vocales en un continuum entre realizaciones hiperarticuladas (habla lenta) y realizaciones hipoarticuladas, que estuvieron asociadas a la velocidad rápida, a la velocidad normal y al habla espontánea.

Finalmente, respondiendo a la segunda pregunta de investigación, ¿cómo afecta la variación individual en los resultados obtenidos para los parámetros acústicos en los dos estilos de habla?, debemos señalar que, de acuerdo con los resultados de nuestro análisis, podemos afirmar que:

- 7) Los hablantes realizan fonéticamente las vocales en habla leída y en habla espontánea con duraciones muy cercanas y, por tanto, las diferencias no son estadísticamente significativas. En general, la desviación típica de la duración es mayor en habla espontánea que en habla leída y en las vocales átonas que en las tónicas (Anexo F). Sin embargo, se pueden observar variaciones individuales que reflejan ciertas tendencias ya mencionadas en el apartado 6.3.1.
- 8) En cuanto al primer formante, que tiene relación con el timbre y la abertura vocálica, podemos decir que un grupo de hablantes tienen coeficientes menores al intercepto (BOG07, BOG11, BOG14, BOG18), es decir, producen vocales más centralizadas; y otro grupo tiene vocales más abiertas que el promedio (BOG02, BOG04, BOG13, BOG19, BOG21). En los valores del F2 solo un hombre presenta datos por encima del intercepto. Es interesante anotar aquí que la variación sugerida por los interceptos es la misma que la reportada para el modelo de habla leída a tres velocidades, lo cual indica que los valores de frecuencia de F1 y F2 característicos de cada locutor, se mantienen en los diferentes estilos de habla.
- 9) A partir de los datos del AEV se desprende que la variación individual es amplia y no resulta fácil de generalizar. En todo caso, el AEV más amplia corresponde a las tónicas producidas en velocidad lenta, tiene un valor intermedio en velocidad normal, y la menor AEV se observa en las átonas producidas en

velocidad rápida. No obstante, cabe señalar que algunos hablantes tienen un AEV bastante reducida en todos los estilos habla, en las realizaciones tónicas y en las átonas (BOG07, BOG17, BOG11 y BOG18).

- 10) La tasa de centralización está determinada, en parte, por la variación individual. Por ejemplo, mientras unos hablantes tienen un grado de centralización para cada estilo, otros tienen los mismos valores para el habla espontánea y el habla rápida, un tercer grupo tiene una tasa de centralización similar en el habla espontánea y en la velocidad normal. Es importante reiterar que la variación individual determina en parte la tasa de centralización de los formantes, pero no está asociada únicamente al sexo del hablante.
- 11) Las mujeres tienen una distancia euclídea más alta (menos centralización) que los hombres, lo cual no quiere decir que todos los participantes se comportan de la misma manera. Por ejemplo, un grupo de hombres y mujeres presentaron interceptos por debajo de los valores predichos por los modelos mixtos (BOG04, BOG07, BOG17, BOG18).

7. Discusión y conclusiones

En este capítulo se presenta, en primer lugar, la discusión de los resultados reportados en los capítulos anteriores teniendo en cuenta los factores que, según la bibliografía revisada en el Capítulo 2, pueden influir en la reducción fonética de las vocales. En segundo lugar, se presentan las conclusiones generales que responden a las cinco preguntas de investigación formuladas en la sección 2.4.

7.1 Discusión de los resultados

7.1.1 Características temporales y espectrales de las vocales bogotanas

La duración de las vocales tónicas del español hablado en Bogotá está relacionada con la abertura linguomandibular; la vocal abierta /a/ es larga, la vocal anterior cerrada /i/ es breve y las vocales /o, e, u/ tienen una duración intermedia. De acuerdo con los datos estadísticos, estas vocales se organizan, de mayor a menor duración, así: /a/ > /o/ > /e/ > /u/ > /i/. Este patrón concuerda, en líneas generales, con los datos reportados en otros estudios sobre el español (Navarro Tomás, 1916, p. 403; Marín Gálvez, 1994-1995, p. 224; Correa, 2017) y, según Lehiste (1976, p. 227), también se ha documentado en lenguas como el danés, inglés, sueco, thai y sami. La duración intrínseca de las vocales tónicas del español bogotano se ordena de la misma manera en habla leída y en habla espontánea, a excepción de las vocales átonas del habla leída que tienen duraciones variables.

Las vocales del español hablado en Bogotá presentan variaciones importantes en los valores de frecuencia de los dos primeros formantes. La vocal abierta /a/ tiene el primer formante con la mayor variación y, por tanto, presenta realizaciones abiertas, medio abiertas y centralizadas [a, a̠, ə]. Para el caso de las vocales medias /e, o/, el primer formante es estable y la reducción se manifiesta en un desplazamiento del F2 hacia los 1500 Hz. La vocal posterior /u/ tiene valores de F1 cercanos a los de la vocal posterior

media /o/ y el segundo formante refleja realizaciones adelantadas en todos los hablantes, de manera que resulta más apropiado describirla como una vocal posterior medio-cerrada y avanzada [ɤ]. La vocal /i/ tiene una realización periférica y se mantiene diferenciada de los demás timbres vocálicos en todas las condiciones. De acuerdo con los datos de la dispersión vocálica, calculada por medio de la distancia euclídea, las vocales más cercanas al centro acústico son /a/ y /o/, seguidas de /e/ y /u/. La vocal anterior cerrada /i/ es la vocal más alejada del centroide acústico, lo cual es consistente con los resultados del análisis de las frecuencias formánticas. A continuación, se discute la importancia y cómo interactúan los diferentes factores lingüísticos analizados, orgánicos y estilísticos, sobre estas variaciones temporales y espectrales de las vocales bogotanas.

7.1.2 Efectos de la velocidad de elocución sobre la duración y los valores de frecuencia de F1 y F2

La velocidad de elocución modifica, como predecíamos, la duración vocálica: en velocidad lenta las vocales presentan una diferencia del 14 % con respecto a la velocidad normal y disminuyen el 19.4 % con relación a la velocidad rápida. Los valores de frecuencia del primer formante son ligeramente más altos en velocidad normal que en velocidad lenta, pero no hay diferencias significativas entre los valores de frecuencia de F1 en la velocidad lenta y en la velocidad rápida. Esto muestra que, si bien los hablantes modifican la duración cuando pasan de una velocidad lenta a una normal, y de una velocidad lenta a una rápida, el grado de abertura es semejante en las tres velocidades. El segundo formante no cambia tampoco cuando se comparan las tres condiciones de elocución.

Estos resultados sugieren que es posible incrementar los valores de F1 en velocidad rápida y contrarrestar los efectos de la duración sobre los del F2. Esto ocurre, según diferentes autores, debido a que los hablantes aumentan la velocidad de los movimientos articulatorios para evitar la reducción fonética de las vocales (Engstrand, 1988; Kuehn y Moll, 1976; Moon y Lindblom, 1994). Por ejemplo, van Son y Pols (1990) consideran que este resultado puede estar relacionado con la experiencia del hablante, pues un locutor profesional o un lector experimentado podrá adaptar más fácilmente su articulación al habla rápida. Todas las personas que aportaron información para esta investigación tenían estudios universitarios, lo cual hace viable esta explicación. Tsao,

Weismer e Iqbal, (2006) plantean que la velocidad de habla no produce variaciones significativas sobre la duración y sobre los dos primeros formantes porque los fonetistas acostumbramos a obtener los datos acústicos en el centro de las vocales. Como veremos más adelante, nuestros resultados hacen viable esta hipótesis, pues la diferencia absoluta de F2 sí se modifica por efecto de la velocidad de habla.

De acuerdo con nuestros datos, el resultado más destacado de los resultados es que el efecto de la velocidad de habla sobre la duración y los valores de frecuencia de F1 y F2 varía en función del hablante. Considerando los interceptos y las pendientes aleatorias, la intervencibilidad fue mayor en habla lenta: unos hablantes produjeron las vocales con duraciones largas y otros con duraciones más breves que el promedio. Esto mismo ocurre con los valores del primer y del segundo formante: un grupo de hablantes tiene un primer formante con valores de frecuencia bajos (menor abertura lingual), un grupo tiene valores cercanos al intercepto del modelo estadístico, y un tercer grupo de hablantes presenta vocales con duraciones más largas, abiertas y más anteriorizadas. La experiencia lectora puede ayudar a evitar los efectos de la velocidad de habla sobre la duración y las frecuencias de F1 y F2, pero, en definitiva, los cambios dependen de las características del tracto vocal de cada hablante.

7.1.3 La correlación entre la duración y los formantes vocálicos

Para entender mejor la relación entre la duración y los formantes vocálicos, realizamos un análisis estadístico para determinar si hay una correlación entre la duración y los valores de frecuencia de F1 y F2. Se encontraron correlaciones positivas, moderadas y altamente significativas entre la duración y los valores del primer formante de /a/, y entre la duración y el segundo formante de /i/ en las tres velocidades de habla. En el caso de la vocal abierta /a/, a mayor duración observamos un valor más alto del primer formante —con una articulación más abierta y menos centralizada—. En el caso de /i/, un incremento de la duración está acompañado de una articulación más anteriorizada. Estos resultados concuerdan con los resultados reportados por Nadeu para el español peninsular (2014, p. 13-15); la única diferencia es que la autora reporta una correlación fuerte entre /e/ y el segundo formante que, en nuestros datos, apenas alcanza una correlación moderada en habla lenta ($r = 0.35$).

Como vemos, la variación acústica de la vocal abierta /a/ se explica como un efecto de la duración, pues varios estudios han mostrado que cuando la longitud decrece, resulta más difícil producir vocales abiertas y, por tanto, las realizaciones son más cerradas (Linblom, 1963, p. 1778; Flemming, 2005, p. 20). De hecho, en opinión de Barnes (2006), desde un punto de vista diacrónico, la reducción vocálica surge de la neutralización de las oposiciones fonológicas basadas en la altura lingual. En el español bogotano la correlación entre duración y la altura lingual de /a/ ocurre, como vimos, en las tres velocidades de habla y se refleja fonéticamente en la aparición de realizaciones abiertas, medio abiertas y centralizadas [a, a̠, ə]. El segundo formante de /i/ cambia en la misma dirección, pues las vocales breves son retraídas [i̠], lo cual resulta más claro en los datos de la velocidad lenta.

7.1.4 La influencia de la tonicidad sobre la duración y los valores de frecuencia de F1 y F2

En velocidad lenta, la duración de las vocales tónicas se incrementa el 19.5 % (17.55 ms) con respecto a las átonas. Este porcentaje se reduce al 13 % (11.55 ms) en velocidad normal, y en velocidad rápida es tan solo del 2 % (2.49 ms). En el modelo en que comparamos la lectura en velocidad normal con habla espontánea hay un incremento del 20 % (16.34 ms) en la duración de las tónicas. Esta diferencia es mayor a la reportada por Nadeu (2013, p.151), quien encontró, en su análisis del español de Madrid, que las átonas son 9.45 % más breves que las tónicas. Como han demostrado diferentes estudios (Llisterri *et al.*, 2003; Ortega-Llebaria y Prieto, 2010; Gendrot, Adda-Decker y Santiago, 2019), los resultados presentados confirman que la duración es uno de los principales correlatos acústicos del acento léxico en el español.

De acuerdo con los resultados obtenidos a partir del análisis de las vocales producidas a tres velocidades (Capítulo 4), las vocales tónicas tienen un F1 más alto que las átonas (28 Hz). En el modelo en que comparamos habla leída con habla espontánea la diferencia fue mayor (70.73 Hz). En este mismo modelo se incluyó una interacción entre la tonicidad y el timbre: los resultados indican que la diferencia entre tónicas y átonas modifica principalmente el primer formante de la vocal abierta /a/ y de las vocales medias /e, o/. El segundo formante es menor en las vocales átonas. Al incluir una interacción entre el timbre y el valor de frecuencia de F2, se puede observar que las

vocales anteriores /e, i/ tienen realizaciones tónicas más adelantadas que las átonas, y las vocales posteriores /o, u/ tienen vocales tónicas más retrasadas que las átonas. La dispersión vocálica, calculada por medio de la distancia euclídea, corrobora las afirmaciones anteriores. Las tónicas tienen mayor dispersión —menos centralización— que las vocales átonas y, muestran, además, que este efecto es particularmente sobresaliente en los resultados correspondientes a la vocal posterior cerrada /u/.

Lo anterior no debe interpretarse como una evidencia a favor de la hipótesis de que el timbre es un correlato acústico del acento léxico en el español bogotano. Se trata de tendencias asistemáticas y no convencionalizadas que dependen del hablante y el dialecto. Por ejemplo, Romanelli y Menegotto (2018, p. 173) reportan realizaciones abiertas para la /o/ tónica, pero, contrario a nuestros resultados, no registran ningún efecto sobre el segundo formante. Nadeu (2013, p. 68) reporta realizaciones más o menos estables para /e/ y / u/, mientras que en nuestros datos hay cambios significativos en el segundo formante de estas vocales. En ese sentido, concordamos con esta última autora en que los efectos del acento sobre el timbre vocálico no son homogéneos entre los hablantes de español (Nadeu, 2013, p. 153).

7.1.5 Factores que inciden sobre el área y la centralización del espacio acústico

El área del espacio vocálico (AEV) cambia en función de la tonicidad, la velocidad de elocución y, como se discutirá más adelante, también varía con el sexo del hablante. El AEV es más amplia en las vocales tónicas y en aquellas producidas con velocidad lenta, mientras que es menor para las vocales átonas y el habla rápida. Podemos afirmar que esta situación de habla (vocales átonas en velocidad rápida) es la que presenta la máxima centralización. Es importante anotar que los efectos de la tonicidad son mayores que los de la velocidad de habla pues, mientras la diferencia entre átonas y tónicas es de 4.51 Bark², el AEV de la velocidad rápida solamente se contrae -2.09 Bark² con respecto a la velocidad lenta y -1.06 Bark² con respecto a la velocidad normal. Este resultado ha sido reportado en español y en catalán (Nadeu, 2014, p. 18), pero también en trabajos sobre otras lenguas como el francés (Audibert *et al.*, 2015), el inglés (Fourakis, 1991, 1829) y fue encontrado también en vocales con duraciones breves en un estudio comparativo sobre el italiano, árabe, inglés, español, chino mandarín, portugués y el

alemán (Gendrot y Adda-Decker, 2007, p. 1419). En ese sentido, podemos afirmar, al menos como hipótesis de trabajo, que la compresión del AEV es un efecto global de la velocidad de elocución y la tonicidad sobre el espacio acústico de los sistemas vocálicos.

La tasa de centralización de los formantes (TCF), que determina el grado de centralización del espacio vocálico formado por las vocales periféricas /i, a, u/, depende principalmente del hablante pues, si bien las átonas y la velocidad rápida son más centralizadas en todos los hablantes, en nuestro corpus predominan las características individuales sobre otros factores. Este resultado es inesperado en cuanto que esta medición se usa para normalizar las frecuencias de los formantes, minimizar la variación interlocutor y maximizar la detección de la centralización (Sapir, Raming y Spielman, 2010, p. 116). Nuestros resultados no apoyan esta idea. Una posible explicación es que la TCF ha sido usada principalmente para analizar habla patológica y para evaluar los efectos de los tratamientos para disartria: consideramos que esta medición puede tener un mayor ámbito de aplicación al considerado en la bibliografía; esta afirmación está confirmada por el estudio reciente de Caverlé y Vogel (2020, p. 1442), quienes mostraron que la TCF también resulta útil para detectar cambios en la producción de las vocales a lo largo del tiempo (como las diferencias intersesión o los efectos de la fatiga), aunque los resultados parecen ser bastante sensibles a las vocales incluidas en el análisis como, por ejemplo, cuando la lengua tiene más de una vocal abierta o se incluye una vocal posterior diferente de [u]. En todo caso, es claro que hace falta investigar más la utilidad de esta medida dentro de los estudios fonéticos y sociofonéticos y con habla no patológica.

7.1.6 La influencia del entorno consonántico

Para determinar los efectos del entorno consonántico sobre los formantes vocálicos analizamos inicialmente los valores de frecuencia de F2 en el centro de la vocal y la desviación típica o estándar. En cuanto a los valores de frecuencia de F2, los resultados arrojados por el algoritmo *k-means* indican que el segundo formante de /a/ y /e/ disminuyen en el contexto de las consonantes labiales y es más alto en el entorno de las velares y dentoalveolares. Como es de esperar, las vocales posteriores /o, u/ tienen un F2 bajo en el entorno de consonantes labiales y velares, mientras que los valores de

frecuencia de F2 se incrementan en el contexto de las dentoalveolares y, por tanto, en este contexto su articulación es más avanzada [ɔ, ʊ].

La desviación típica muestra que, en concordancia con otros estudios (Recasens, 1985; p. 111; Recasens y Espinosa, 2006, p. 662), la variabilidad del primer formante depende de la altura lingual: la vocal abierta /a/ tiene una desviación alta, las vocales /e, o/ tienen valores intermedios y las vocales cerradas /i, u/ los valores más bajos. En cuanto al segundo formante, encontramos mayor variabilidad en las vocales anteriores /i, e/ y menor desviación en los datos correspondientes a la vocal abierta /a/ y a la vocal posterior media /o/. La vocal cerrada /u/ tiene una desviación baja en los contextos labial y velar, pero en el contexto dentoalveolar la desviación típica es alta. Los resultados correspondientes a las vocales anteriores no concuerdan con otros estudios, pues se considera que estas vocales presentan poca variabilidad y, por tanto, son resistentes a la coarticulación (Recasens y Espinosa, 2006, p. 662; Mooshammer y Geng, 2008, p. 119).

Para explicar los resultados mencionados anteriormente y hacer inferencias sobre la trayectoria de los formantes de las vocales, medimos la diferencia o cambio absoluto y la velocidad de cambio de F1 y F2. En el caso del primer formante, los valores del cambio absoluto y de la velocidad de cambio ($\Delta F1$ y $rF1$) de las vocales /a, e, i/ se incrementan en el contexto de las consonantes dentoalveolares y velares. Esto significa que la trayectoria de F1 es cóncava ascendente \cap como supone la teoría del locus (Delattre, 1969b, p. 3 ; Johnson, 2003, p. 178). Estas trayectorias cambian en función de la velocidad de elocución: el primer formante de /i/ tiene mayor diferencia absoluta y mayor velocidad de cambio en velocidad rápida que en velocidad lenta, mientras que en /a/ y /e/ se observa lo contrario, es decir, las variaciones dinámicas de F1 son mayores en velocidad lenta que en velocidad normal y rápida. Los cambios espectrales del primer formante de las vocales posteriores /o, u/ no presentan modificaciones estadísticamente significativas para ninguno de los factores analizados.

En cuanto a los cambios de F2, se encontró que los valores de la diferencia absoluta y de la velocidad de cambio de /i/ se incrementan en el contexto de las líquidas /l, r/ y de la oclusiva bilabial /p/. Es importante destacar que la consonante lateral /l/ introduce una gran variabilidad sobre esta vocal, lo cual explica la alta desviación estándar. En la misma línea, los valores de $\Delta F2$ y $rF2$ se acercan a cero en el entorno de /f/ y de las

dentoalveolares /d/ y /s/, lo cual significa que estas consonantes ejercen una gran influencia sobre la articulación de /i/. La vocal anterior media /e/ presenta cambios semejantes: las trayectorias de los formantes tienen cambios significativos en el entorno de /r/ y de las oclusivas bilabiales /b, p/, mientras que hay mayor coarticulación en el contexto de las oclusivas velares /k, g/ y de la sibilante /s/. En inglés de Escocia se ha encontrado que las laterales y róticas ejercen un gran efecto sobre la articulación de /i/ debido, según Zharkova (2007), a la distancia entre los contextos consonánticos analizados. En el caso de nuestro corpus, una gran cantidad de apariciones de las vocales anteriores se dan en los contextos /l_s, l_t, s_s, d_l/; esto sugiere que se trata más bien de un caso de coarticulación del segundo formante con el contexto dentoalveolar. La variabilidad de /i/ en el contexto de las labiales parece ser un caso de coarticulación con C₂, pues las vocales analizadas aparecen en los contextos /f_s/ y /p_d/.

Como predice la teoría del locus (Delattre, 1969b, p. 3; Johnson, 2003, p. 178), el segundo formante de /a/ tiene una transición ascendente en el contexto de las consonantes labiales y una transición descendente en el contexto de las velares y dentoalveolares. De acuerdo con los modelos estadísticos, la coarticulación de /a/ es mayor en los contextos velar /g, k/ y dentoalveolar /s, t, l/. Las vocales posteriores /o, u/ son acústica y estadísticamente variables en los diferentes entornos consonánticos analizados. De acuerdo con los interceptos aleatorios, la variabilidad de /o/ está asociada con las palabras de contenido léxico en las que esta vocal aparece ante una consonante lateral (p.ej., *sol*, *policial*, *árboles*, *sol*). Finalmente, el F2 de la vocal posterior cerrada /u/ tiene un cambio espectral y velocidades de articulación cercanas a cero (mayor coarticulación) en el entorno de las velares y dentoalveolares, y especialmente en el contexto de /s/.

Como puede verse, el contexto dentoalveolar ejerce gran influencia sobre las cinco vocales bogotanas. Entre las consonantes dentoalveolares, la sibilante /s/ tiene mayor efecto sobre los valores de frecuencia y sobre los cambios dinámicos de los dos primeros formantes. Chládkova, Escudero y Boersma, (2011, p. 426) encontraron que en el español de Madrid y, en menor medida en la variedad de Lima, las vocales posteriores /o, u/ y la vocal anterior cerrada /i/ se centralizan después de las alveolares y,

especialmente, después de la sibilante /s/². Este resultado también es compatible con el proceso de ensordecimiento de las vocales que caracteriza el vocalismo de las zonas altas de América (sección 2.2.8).

Los resultados indican que, en el caso de las vocales /i, e, a/, la diferencia absoluta de F2 es menor en velocidad rápida y normal que en habla lenta, lo cual demuestra que hay mayor coarticulación de las consonantes sobre las vocales cuando disminuye el tiempo para realizar los movimientos articulatorios. Es importante tener en cuenta que, en el modelo en que analizamos los cambios producidos por la velocidad de habla sobre los valores de frecuencia de F2 obtenidos en el centro de la vocal, no obtuvimos un efecto estadísticamente significativo. Es significativa que la velocidad de habla modifica la trayectoria del F2, pero no su parte estable o *target*. La tonicidad modifica la trayectoria de los formantes de las vocales /e/ y /a/: las realizaciones tónicas de la vocal media /e/ tienen valores más altos de $\Delta F2$ y rF2 que las átonas; mientras que las realizaciones tónicas de la vocal abierta /a/ tienen valores de $\Delta F1$ y rF1 más altos que en sus contrapartes átonas.

7.1.7 Factores léxicos

Para analizar la influencia de los factores léxicos diferenciamos las vocales según el tipo de palabra. La predicción, de acuerdo con lo discutido en la sección 2.2.5, es que las vocales de las palabras funcionales se pronuncien con menor duración y mayor grado de centralización. Los resultados son contrarios a esta hipótesis debido a que las vocales analizadas tienen menor duración en las palabras con contenido léxico. Los valores de frecuencia del primer formante no mostraron cambios estadísticamente significativos; en el caso de /a/, los cambios espectrales del primer formante aumentaron —hubo menor coarticulación— en las palabras funcionales. En el modelo estadístico elaborado para las vocales producidas a tres velocidades, el F2 fue menor en las palabras funcionales, pero en el modelo estadístico elaborado para comparar el habla leída con el habla espontánea no se presentó ningún efecto.

² Se debe tener en cuenta que en el español peninsular la sibilante /s/ es apico-alveolar, mientras que en el español colombiano y peruano la articulación es lamino-alveolar.

Estos resultados sugieren que la distinción entre palabras funcionales y palabras con contenido no es suficiente para analizar la influencia de las propiedades léxicas sobre la variación acústica de las vocales. Otra posible explicación es que la varianza asociada a los ítems léxicos se trata como un efecto aleatorio en los modelos mixtos, de manera que los efectos del tipo de palabra pudieron ser atenuados por el tratamiento estadístico utilizado. En estudios posteriores podrían implementarse otros métodos probabilísticos como la *información mutua* o IM (Pluymaekers, Ernestus, y Baayen, 2005b; Bürki, Ernestus, Gendrot, Fougeron, y Frauenfelder, 2011) que permitan analizar los efectos de léxicos como una variable cuantitativa y no como un factor fijo o un efecto aleatorio.

7.1.8 Posición de la vocal dentro de la palabra y el grupo fónico

La posición de la vocal dentro de la palabra no afecta a la duración. Este resultado es diferente a los datos presentados por Gendrot, Adda-Decker y Santiago (2019, p. 88), quienes reportan que la duración vocálica se modifica por efecto de la posición de la vocal dentro de la palabra y con la ausencia/presencia de pausa (Gendrot, Adda-Decker y Santiago, 2019, p. 88). En cuanto al primer formante, en habla lenta es 50 Hz más bajo (menor abertura linguomandibular) al final de palabra, pero esta diferencia no se observa en el modelo en que se compara el habla leída con el habla espontánea. En cambio, el segundo formante no presenta cambios estadísticamente significativos asociados a la posición de la vocal dentro de la palabra.

La posición de la vocal dentro de la frase tuvo una influencia sobre la duración y los valores de frecuencia de F2: las vocales son más largas en posición final, la duración disminuye el 5.9 % en posición inicial y son más breves en posición intermedia (-10.68 %). De acuerdo con los datos de los cambios espectrales, en posición final del grupo fónico hay mayor coarticulación (menor $\Delta F2$) que en posición intermedia e inicial. Sin embargo, en el modelo en que comparamos habla leída con habla espontánea, ninguno de los formantes presenta cambios significativos. Esto sugiere que, si bien la duración es sensible al contexto de frase, no es posible afirmar lo mismo en cuanto a las características espectrales de las vocales. En la investigación futura se debe evaluar si fenómenos entonativos con funciones pragmáticas definidas como, por ejemplo, la focalización, modifican los valores de frecuencia de los formantes. Pero, con los

resultados obtenidos en este estudio, no es posible afirmar que las posiciones prosódicamente relevantes modifican el timbre de las vocales en el español de Bogotá.

7.1.9 Diferencias entre hombres y mujeres

Los hombres bogotanos producen vocales más breves y formantes con valores de frecuencia más bajos que las mujeres: la duración es 5.85 ms más breve en el modelo elaborado para el habla leída a tres velocidades, y 6.75 ms más breve en el modelo que compara la lectura a velocidad normal con el habla espontánea. El primer formante de las vocales producidas por los hombres tiene una diferencia del 15 % con respecto al habla femenina, mientras que para el segundo formante la diferencia es del 13 %. Los resultados de las frecuencias de los formantes en función del sexo son iguales en los modelos presentados en el Capítulo 4 y en el Capítulo 6. A partir del análisis del efecto de las consonantes sobre las vocales, podemos afirmar que las realizaciones masculinas de las vocales /a/ y /e/ tienen un mayor grado de coarticulación que las vocales producidas por las mujeres.

Los datos de la dispersión vocálica muestran que las mujeres tienen vocales menos centralizadas que los hombres. Esto no quiere decir que todos los participantes se comportan de la misma manera, pues algunas mujeres producen vocales con una dispersión por debajo de los valores predichos por el modelo estadístico. De igual manera, la tendencia general es que las mujeres tengan un sistema vocálico con un área del espacio vocálico (AEV) más amplia que la usada por los hombres, pero, insistimos, esta diferencia no es categórica, pues algunos hombres tienen un espacio acústico tan amplio como el de las mujeres, y viceversa.

Estos resultados concuerdan con los estudios discutidos en la sección 2.2.7 en cuanto confirman que los hombres bogotanos, como los hablantes de otras lenguas, producen vocales centralizadas y breves, mientras que las mujeres producen vocales largas, periféricas y con un espacio acústico más amplio (Byrd, 1994; Simpson y Ericsson, 2007; Simpson, 2009; Weirich y Simpson, 2013). Sin embargo, a diferencia de estos estudios, nuestros resultados sugieren que estas diferencias tienen excepciones: algunas mujeres producen vocales centralizadas y algunos hombres, vocales periféricas. A nuestro juicio, la hiper e hipoarticulación es una elección fonética individual y no implica la pertenencia de un hablante a uno u otro sexo.

7.1.10 El estilo de habla vs. velocidad de elocución

En el capítulo 6 se comparó el habla leída a velocidad normal con el habla espontánea obtenida por medio de una conversación entre amigos. Esta comparación busca determinar el efecto del estilo de habla, de manera que el habla leída a una velocidad normal equivale a lo que habitualmente se denomina *habla de laboratorio* frente a la interacción entre dos amigos, que hemos categorizado como *habla espontánea*. La duración de las vocales no se modificó por efecto del estilo, lo cual concuerda con nuestro estudio piloto (Correa, 2017, p. 75), que compara habla leída con narraciones producidas por los mismos hablantes. El primer formante se modifica al incluir una interacción entre el estilo de habla con el timbre: en habla leída, la vocal central /a/ es más abierta que en habla espontánea, mientras que las vocales /i, e, o/, por el contrario, son más cerradas en la lectura. El segundo formante también varía en función del estilo y la tonicidad de la vocal: las vocales anteriores /i, e/ tienen un F2 más alto en habla leída y las realizaciones átonas son más centralizadas que las tónicas, lo cual afecta especialmente a la vocal posterior cerrada /u/.

En la segunda parte del Capítulo 6 comparamos el área del espacio vocálico (AEV), la dispersión de las vocales, la tasa de centralización de los formantes (TCF) y el índice δ entre la distancia euclídea del habla lenta (la pronunciación más hiperarticulada de nuestro corpus) y la distancia euclídea de las vocales pronunciadas en habla leída a velocidad rápida, normal y en habla espontánea. También comparamos la velocidad normal con el habla espontánea, y la velocidad normal con el habla rápida. Encontramos, en primer lugar, que el AEV del habla espontánea no se diferencia del habla leída en velocidad normal. Esto se refleja en una superposición de las áreas vocálicas y en la poca diferenciación del habla espontánea de las demás condiciones de elocución. El índice δ , que permite calcular la diferencia entre la distancia euclídea de dos estilos de habla, confirma estos resultados: las diferencias acústicas entre estilos son grandes solamente cuando se compara una pronunciación hiperarticulada (habla lenta) con una pronunciación hipoarticulada (habla leída a velocidad rápida o habla espontánea), y son pequeñas, contrario a las predicciones, cuando se compara el habla leída a velocidad normal con el habla espontánea, y cuando se compara el habla normal y el habla espontánea con el habla leída a velocidad rápida.

En cuanto a la centralización del espacio vocálico y de cada timbre, determinadas a partir de la TCF y la dispersión vocálica, podemos decir que el espacio vocálico se organiza, de mayor a menor tasa de centralización, así: *habla lenta* > *habla normal* > *habla espontánea* > *habla rápida*. Los hablantes que presentan un AEV amplia tienen, en general, una menor tasa de centralización y, viceversa, los hablantes con mayor TCF tienen una menor AEV. Si bien las diferencias entre estilos no son categóricas, los participantes en este estudio organizaron el espacio acústico de las vocales en un continuum entre realizaciones hiperarticuladas (*habla lenta*) y realizaciones hipoarticuladas.

De acuerdo con los datos sobre la dispersión, las vocales cerradas (/i/ y /u/) tienen mayor distancia del centroide acústico, mientras que la vocal /a/ y las vocales medias /e/ y /o/ son las más cercanas. En otras palabras, las vocales medias tienen realizaciones tan centralizadas como las de la vocal abierta /a/. La dispersión vocálica varía de la misma manera que la TCF, pues se encontró una mayor distancia para la velocidad lenta, una distancia intermedia para la velocidad normal y el habla espontánea, y una menor distancia para la velocidad rápida. Finalmente, de acuerdo con el índice δ , la vocal abierta /a/ es menos sensible al estilo de habla que las vocales anteriores /e, i/ y la vocal posterior cerrada /u/.

La similitud entre el habla leída a velocidad normal con el habla espontánea debe interpretarse como un llamado a la prudencia cuando usamos términos como *habla de laboratorio*, *habla casual* e, incluso, el concepto mismo de *habla espontánea* (Simpson, 2013, p.155). Resulta más apropiada la distinción clásica entre palabras aisladas (*citation form*) y habla conectada (Jones, 1944), pues no hace referencia a una situación de elocución específica. En ese sentido, nuestro estudio es un análisis de habla conectada obtenida por medio de dos tareas de recolección de datos: la lectura y la conversación. En cambio, los estudios discutidos en la sección 2.2.6 comparan palabras aisladas con habla conectada obtenida por medio de tareas tan diversas como los diálogos de radio y televisión, entrevistas, tareas de mapas, monólogos, conversaciones, etcétera (sobre esta discusión véase también Llisterri, 1992). Así mismo, el grado de atención no parece ser el principio que organiza los estilos de habla como sugiere Labov (1972, p.112), pues nuestros hablantes leyeron con máxima atención y conversaron con un amigo y, en todas las situaciones, predominaron los factores lingüísticos y las estrategias fonéticas

individuales sobre las diferentes situaciones comunicativas resultantes de utilizar diferentes tareas en la recolección de los datos.

7.1.11 La variación individual

Por último, nos preguntábamos cómo afectan las características fonéticas individuales en los resultados obtenidos para cada uno de los parámetros acústicos. En la lectura a tres velocidades, los interceptos aleatorios por hablante y las pendientes aleatorias por velocidad muestran que los hablantes se agrupan en tres grupos. El primero presenta coeficientes cercanos a los predichos por el modelo estadístico. El segundo grupo favorece las duraciones breves, un primer formante con valores de frecuencia bajos, tienen una AEV pequeña y un espacio acústico centralizado. El tercer grupo produce vocales largas, abiertas (F1 más alto), con una AEV amplia y una TCF baja. Si bien hay diferencias relacionadas con el sexo del hablante, la producción de vocales fonéticamente reducidas es común en los datos correspondientes a ambos sexos. Adicionalmente, la variación sugerida por los interceptos es la misma en los modelos estadísticos elaborados para el habla leída a tres velocidades que para el modelo que compara habla leída con habla espontánea.

Al analizar los efectos del entorno consonántico sobre las características acústicas de las vocales bogotanas, encontramos que las desviaciones típicas de los formantes indican que, en lo que respecta a F1, hay mayor variación interlocutor en la producción de las vocales /a/, /e/ y /o/, y esta variabilidad es mayor en el contexto dentoalveolar y labial que en el contexto de velar. Las diferencias individuales en las variaciones dinámicas del primer formante ($\Delta F1$ y $rF1$) se manifiestan principalmente en las vocales anteriores /i, e/, pues algunos hablantes producen estas vocales con valores por debajo del intercepto del modelo, mientras que un segundo grupo tiene mayor variabilidad en el grado de abertura de las vocales. En cuanto al F2, las diferencias entre los hablantes se destacan en la producción de /e/ y en los datos correspondientes a las vocales cerradas /i, u/.

Los datos del AEV hacen ver que la variación individual es amplia y no resulta fácil de generalizar. En todo caso, el AEV más amplia corresponde a las tónicas producidas en velocidad lenta, tiene un valor intermedio en velocidad normal, y la menor AEV se observa en las átonas producidas en velocidad rápida. Llama la atención, de un lado, que

algunos hablantes tienen un AEV bastante reducida en todos los estilos de habla, en las realizaciones tónicas y en las átonas. De otro lado, podemos observar que, mientras unos hablantes tienen un grado de centralización diferente para cada estilo, otros tienen los mismos valores para el habla espontánea y el habla rápida, y un tercer grupo tiene una tasa de centralización similar en el habla espontánea y en la velocidad normal.

A la vista de nuestros resultados, podría afirmarse que la reducción fonética de las vocales bogotanas y el efecto específico de los factores lingüísticos, orgánicos y expresivos depende de la variación fonética del locutor. Como afirman Stevens y Harrington (2014, p.8), el cambio y la variación fonética es el resultado de la interacción entre individuos que usan diferentes estrategias articulatorias y responden de manera diferente a las condiciones sociolingüísticas.

7.2 Conclusiones

El español hablado en Bogotá es una de las variedades más prestigiosas de Hispanoamérica. Esta reputación se debe a la importancia de la filología y la corrección lingüística en la identidad cultural colombiana y, de acuerdo con las valoraciones de los hablantes, a una supuesta tendencia conservadora en la pronunciación (Patiño Roselli, 2004; Espejo, 2012; Schwegler y Correa, 2018; Bernal, Díaz y Munévar, 2018). No obstante, de acuerdo con las caracterizaciones fonéticas publicadas (Flórez, 1973; Montes, *et al.*, 1998), el español hablado en Bogotá es una variedad en la que se encuentran numerosos ejemplos de procesos de debilitamiento y simplificación fonética de las unidades lingüísticas. Esta tendencia, denominada reducción fonética, modifica las características temporales y espectrales de las vocales haciéndolas más breves y centralizadas que sus realizaciones canónicas. De acuerdo con la revisión y el análisis de la bibliografía (Capítulo 2), este tipo de variación fonética puede deberse a la influencia de factores lingüísticos (timbre, contexto consonántico, acento, las propiedades léxicas y el estilo de habla), orgánicos (sexo y variación fonética individual) y expresivos (velocidad de elocución) sobre la duración, los valores de frecuencia de los dos primeros formantes y el área del espacio acústico de las vocales. Dado que la mayoría de los estudios sobre la reducción vocálica analizan un conjunto limitado de variables, el principal objetivo de esta tesis es ofrecer una visión de conjunto que permita estimar qué importancia tiene cada uno de los factores mencionados sobre la variabilidad vocálica

del español bogotano. Gran parte de los estudios revisados en el Capítulo 2 se centran en lenguas germánicas, así que otro objetivo es aportar datos que amplíen nuestros conocimientos sobre el vocalismo del español de América.

Respondiendo a la primera pregunta que nos planteábamos en el apartado 2.4 sobre cómo afecta la velocidad de elocución en los valores temporales y espectrales de las vocales, considerando otros factores como el timbre, el sexo, el entorno consonántico, la tonicidad, el tipo de palabra, la posición de la vocal dentro de la palabra y dentro del grupo fónico, podemos señalar que la duración disminuye cuando se incrementa la velocidad de habla, pero los valores de frecuencia de F1 son estables en las tres velocidades de habla y los valores de frecuencia de F2 no presentan diferencias estadísticamente significativas. De acuerdo con la bibliografía, el incremento de la velocidad de elocución provoca una disminución de la duración e, inversamente, la velocidad lenta incrementa la longitud de las vocales. Con relación a los formantes, las investigaciones sugieren dos hipótesis. Primero, se afirma que si la duración vocálica es breve, como podría suceder en una velocidad de habla rápida, los formantes vocálicos no alcanzarán sus valores canónicos y terminarán por acercarse a los valores de las transiciones consonánticas (Lindblom, 1963; Lindblom, Sussman y Agwuele, 2009). Segundo, se sostiene que la velocidad de elocución modifica la duración, pero no produce cambios significativos en los valores de frecuencia de los formantes porque los hablantes pueden controlar parámetros como la velocidad articulatoria o el grado de abertura lingual (Engstrand, 1988; Van Son y Pols, 1990, 1992; Fourakis, 1991; Brzezicha y Kul, 2014; Kul, 2015). No obstante, nuestros resultados sobre el habla leída apoyan parcialmente esta última hipótesis, pues si bien se confirma que la duración disminuye cuando la velocidad de habla aumenta, los valores de frecuencia de F1 permanecen constantes en las tres velocidades, y los de F2, aunque se modifican, esas variaciones no presentan diferencias estadísticamente significativas. Es importante anotar que, si bien la velocidad de habla no afecta los valores de frecuencia de los dos primeros formantes, sí modifica el área del espacio acústico formado por las vocales periféricas /i, a u/, puesto que el AEV es más amplia en velocidad lenta que en velocidad normal y rápida.

En cuanto a los demás factores que influyen sobre las variaciones temporales y espectrales de las vocales bogotanas, encontramos cambios significativos asociados con

el timbre, la tonicidad y el sexo. Con relación al timbre, la vocal abierta /a/ es larga, la vocal anterior cerrada /i/ breve, y las vocales /e, o, u/ tienen valores intermedios. La vocal abierta presenta un primer formante variable, de manera que sus realizaciones fonéticas son abiertas, medio abiertas y centralizadas [a, a̠, ə]. Las vocales medias /e, o/ tienen un primer formante estable y los cambios espectrales se manifiestan principalmente en el segundo formante. La vocal posterior cerrada /u/ es altamente variable y, según se puede inferir de los valores de frecuencia de los formantes, es una vocal que se realiza fonéticamente como una posterior [u] y como una posterior avanzada [ɯ]. La tonicidad produce cambios significativos en todos los parámetros analizados: las tónicas tienen mayor duración y una realización más abierta y periférica, mientras que las átonas son breves, más cerradas y centralizadas. Por lo tanto, en ausencia de tonicidad, las realizaciones de la vocal abierta son medio abiertas y centralizadas [a̠, ə], las anteriores son retraídas [i̠, e̠] y las realizaciones átonas de las posteriores son avanzadas [ɔ, ɯ]. En el mismo sentido, el AEV es amplia en las vocales tónicas y se contrae en las átonas. Como vimos anteriormente, este efecto de la tonicidad sobre el AEV es mayor al que ejerce la velocidad de habla. Con respecto al sexo, las mujeres producen vocales largas, periféricas y con un espacio acústico amplio, mientras que los hombres producen vocales breves, centralizadas y con menor AEV. Esta es una generalización derivada del análisis estadístico de todos los hablantes, y no debe interpretarse como una diferencia categórica entre sexos.

Los factores con menor influencia sobre las características temporales y espectrales son el tipo de palabra, la posición de la vocal dentro de la palabra y dentro del grupo fónico. En cuanto al tipo de palabra, la duración es mayor en las palabras funcionales, el primer formante no presenta cambios significativos, y los valores de frecuencia de F2 disminuyen ligeramente en las palabras funcionales. La posición de la vocal dentro de la palabra no produce cambios temporales ni modifica el segundo formante, pero los valores de frecuencia de F1 disminuyen al final de palabra en habla leída. La duración, como es de esperar, es mayor en posición final del grupo fónico, tiene una duración intermedia en posición inicial y es más breve en posición intermedia del grupo fónico. Sin embargo, en habla leída ninguno de los formantes presentó cambios estadísticamente significativos asociados a este factor.

La segunda pregunta de esta investigación se propone determinar cómo afecta la duración en los valores de frecuencia de los formantes considerando el timbre y la velocidad de habla. En las tres velocidades encontramos correlaciones positivas, moderadas y altamente significativas para las vocales /a, i/. Esto significa que, cuando se incrementa la duración, la vocal /a/ es más abierta y la articulación de /i/ es más anteriorizada o periférica. La vocal /e/ presenta una correlación moderada, pero solamente resulta significativa en habla lenta.

En el Capítulo 5 se aborda la tercera pregunta de investigación, cuyo objetivo es precisar cómo afecta el contexto consonántico en los valores espectrales de las vocales considerando el timbre, la tonicidad, la velocidad de habla, el sexo, la posición de la vocal dentro de la palabra y dentro del grupo fónico. Para cuantificar el grado de coarticulación y hacer inferencias sobre la velocidad de los movimientos articulatorios, el análisis incluye datos obtenidos en la parte estable de los formantes, mediciones de la diferencia absoluta de los formantes entre el inicio y el centro de la vocal, y la velocidad de cambio de F1 y F2. De acuerdo con nuestros resultados, la variabilidad del primer formante depende de la altura lingual, de manera que es más alta en las vocales /a, e, o/ que en las vocales cerradas. La variabilidad del segundo formante es más alta para las vocales anteriores /i, e/ y baja para las vocales /a, o/. La vocal posterior cerrada /u/ tiene una desviación baja en el contexto velar y una desviación alta en el contexto dentoalveolar. En cuanto a los valores de frecuencia de F2 en el centro de la vocal, nuestro análisis sugiere que, en el caso de /a/ y /e/, el segundo formante disminuye en el contexto de las consonantes labiales e incrementa en los contextos dentoalveolar y velar. Los valores de frecuencia de F2 de las vocales posteriores /o, u/ son bajos en los contextos labial y velar, y aumentan en el contexto dentoalveolar. Los valores de frecuencia de /i/ en el centro de la vocal son los más estables de las cinco vocales bogotanas.

La trayectoria y los cambios dinámicos de los formantes también cambian en función de las consonantes adyacentes, de manera que las vocales asimiladas al contexto presentan menor cambio espectral y menores diferencias entre los valores frecuencia obtenidos al inicio y en el centro de la vocal. La viabilidad de esta hipótesis está respaldada por varios estudios que han demostrado que el grado de coarticulación y el grado de precisión articulatoria —como la distinción entre habla hiperarticulada y habla espontánea— se

puede determinar acústicamente calculando las variaciones dinámicas de los valores de frecuencia de los formantes (Moon y Lindblom, 1994; van Bergem, 1994; Ferguson y Kewley-Port, 2007; Herrmann, Cunningham y Whiteside, 2014). En los datos del español hablado en Bogotá, la diferencia absoluta y la velocidad de cambio del segundo formante ($\Delta F2$ y $rF2$) indican que las vocales se coarticulaban principalmente en los entornos consonánticos velar y dentoalveolar. Además, cabe señalar que, en este último contexto, la consonante que más influyó sobre los valores dinámicos de F2 fue la sibilante alveolar /s/.

Respecto a la contribución de otros factores sobre el efecto que produce el entorno consonántico en los valores de frecuencia de los formantes, encontramos que la velocidad de cambio modifica la trayectoria de F1 y F2 de las vocales /i, e, a/: cuando los hablantes producen la vocal cerrada /i/ a velocidad rápida incrementan la velocidad de cambio y aumenta la diferencia absoluta de F1, mientras que cuando producen las vocales /a, e/ a esta velocidad disminuye la magnitud de los cambios espectrales. Este resultado es consistente con los datos relacionados con la desviación típica que, como vimos, se incrementa con la abertura lingual. Los resultados del segundo formante muestran que la diferencia absoluta ($F2\Delta$) es menor en velocidad rápida y normal que en velocidad lenta. Esto significa que hay mayor coarticulación de las vocales cuando disminuye el tiempo para producir una vocal. Debido a que en la velocidad de habla no modifica los valores de F2 en el centro de la vocal, este resultado sugiere que la velocidad de elocución tiene un impacto sobre la trayectoria del segundo formante, pero no necesariamente sobre el valor de frecuencia en el punto estable o *target*.

La tonicidad también modifica los valores dinámicos de los formantes de las vocales /a/ y /e/ pero no inciden sobre el mismo formante: las realizaciones tónicas de /e/ tienen una mayor diferencia y velocidad de cambio espectral del segundo formante que las átonas, mientras que, en el caso de /a/, las tónicas tienen mayor diferencia y velocidad de cambio espectral en el primer formante. Estas vocales tuvieron una mayor coarticulación en el habla masculina. Cuando la vocal /a/ pertenece a una palabra funcional, incrementan los valores de $F1\Delta$, lo cual no concuerda con la hipótesis que sostiene que las palabras funcionales son más centralizadas que las palabras de contenido léxico. Por último, en posición final del grupo fónico hay mayor coarticulación (menor $\Delta F2$) que en posición intermedia.

La cuarta pregunta de esta investigación estaba relacionada con la influencia de los estilos de habla (habla leída vs. habla espontánea) sobre los valores de frecuencia de los formantes vocálicos y los de la duración considerando el timbre, la tonicidad, el contexto consonántico, el tipo de palabra, la posición de la vocal dentro del grupo fónico y el sexo del hablante. El objetivo de esta pregunta es determinar si el estilo tiene un efecto similar a la velocidad de elocución o interactúa con otros factores y, en ese sentido, incrementa o contrarresta los efectos de otras variables. De acuerdo con el principio de atención de Labov (1972, p. 112) y la teoría de la hiper e hipoarticulación (Lindblom, 1990), según los cuales los estilos de habla se organizan en una escala determinada por el grado de atención que presta el locutor a su pronunciación, se predice que hay mayor variabilidad acústica en el corpus de conversaciones que en los datos correspondientes al habla leída. Nuestros resultados muestran que los valores del primer formante se modifican, así la vocal /a/ es más abierta y las vocales /i, e, o/ son más cerradas en habla leída a velocidad normal. El estilo modifica el segundo formante de las vocales anteriores, pues en habla leída tienen una articulación más anteriorizada. No obstante, si consideramos el AEV, no se han hallado diferencias entre el habla leída a velocidad normal y el habla espontánea. Un resultado similar se obtiene con el índice δ , pues las diferencias entre estilos son grandes solamente cuando se compara una pronunciación hiperarticulada, como el habla lenta, con una pronunciación hipoarticulada, como el habla espontánea o el habla leída a velocidad rápida, pero no si se comparan con el habla leída a velocidad normal. Así mismo, los valores del índice δ sugieren que la vocal /a/ es menos sensible a la influencia del estilo de habla que las vocales anteriores /e, i/ y la vocal posterior cerrada /u/. La variación de la vocal /o/ tiene un comportamiento variable que no se puede explicar teniendo en cuenta únicamente el estilo de habla.

En cuanto a la centralización de las vocales, la TCF y la dispersión indican que el espacio acústico se organiza, de mayor a menor tasa de centralización, así: habla lenta > habla normal > habla espontánea > habla rápida. Esto quiere decir que los resultados indican que el grado de centralización en habla espontánea no es comparable con una lectura rápida. En los modelos estadísticos en que comparamos las variaciones temporales y espectrales del habla leída a velocidad normal con el habla espontánea, obtuvimos, para cada uno de los factores analizados, coeficientes semejantes a los obtenidos en el análisis del habla leída a tres velocidades, a excepción de la posición de la vocal dentro de la

palabra y dentro del grupo fónico, que no mostró ningún efecto estadísticamente significativo sobre los valores de frecuencia de F1 y F2.

Finalmente, nos preguntamos cómo afectan las características fonéticas individuales en los resultados obtenidos para cada uno de los parámetros. Al respecto, podemos concluir que la variación individual fue sobresaliente en cada uno de los modelos estadísticos implementados y que los efectos de los factores analizados no son homogéneos ni están convencionalizados entre los hablantes. Sin embargo, podemos decir que los hablantes se dividieron en, al menos, tres grupos. El primero presenta variaciones acústicas cercanas al intercepto de cada uno de los modelos. El segundo grupo favorece las vocales breves, un primer formante con valores de frecuencia bajos, un espacio acústico con un área pequeña y con vocales centralizadas. Mientras que el tercero produce vocales largas, abiertas, con un AEV amplia y poco centralizada. Es importante destacar que algunos hablantes pertenecientes al segundo grupo producen vocales reducidas en todas las condiciones de elocución analizadas. Podemos concluir, por lo tanto, que la hiperarticulación e hipoarticulación de las vocales es una elección fonética individual.

En conclusión, el análisis de las características temporales y espectrales de las vocales del español hablado en Bogotá muestra que la reducción fonética de las vocales es un fenómeno de variación sincrónica determinado por factores segmentales (entorno consonántico y timbre), prosódicos (tonicidad) y por factores como el sexo del hablante, la velocidad de elocución, por el estilo de habla y la actuación fonética individual. Esta tesis ofrece una caracterización fonética detallada que busca mejorar nuestros conocimientos sobre el español de Colombia y sobre el vocalismo del español de América. Dado que los parámetros acústicos y las métricas implementadas son sensibles a las características fonéticas individuales, los datos reportados y la metodología son de gran utilidad para futuros estudios y análisis de identificación del locutor. Finalmente, los resultados permiten afirmar que es necesario revisar de manera crítica los principios y la terminología que los fonetistas y lingüistas usamos para explicar la variación condicionada por los estilos de habla.

Bibliografía

- Adank, P., Smits, R. y Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116(5), 3099-3107.
- Aguilar, L. y Machuca, M. (1994). Problemas de definición y categorización de los estilos de habla. En J. Llisterri y D. Poch (eds.), *Nuevos Horizontes de la Lingüística Aplicada. Actas del XII Congreso Nacional de la Asociación Española de Lingüística Aplicada*, (pp. 17-23). Barcelona.
- Aguilar, L., Machuca, M. J. y Martínez, G. (1991). Analysis of the Spanish sequence "de" in content words and in function words in continuous speech. *ESCA Workshop on Phonetics and Phonology of Speaking Styles*, (pp. 7-1/7-4). Barcelona.
- Albalá, M. J., Battaner, E., Carranza, M., Gil, J., Llisterri, J., Machuca, M. y Ríos, A. (2008). VILE: Nuevos datos acústicos sobre las vocales del español. *Journal of Theoretical and Experimental Linguistics*, 1, 1-14.
- Álvarez González, J. A. (1989). Influencias de los sonidos contiguos en el timbre de las vocales (estudio acústico). *Revista de la Sociedad Española de Lingüística*, 11, 427-445.
- Amir, N. y Amir, O. (2007). Novel measures for vowel reduction. En J. Trouvain y W. Barry (eds.), *16th International Congress of Phonetic Sciences*, (pp. 849-852). Saarbrücken.
- Audibert, N., Fougeron, C., Gendrot, C. y Adda-Decker, M. (2015). Duration vs. style-dependent vowel variation: a multiparametric investigation. En T. S. ICPhS2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, (pp. 0753.1-5).
- Aylett, M. y Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: a functional explanation for the relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31-56.

- Aylett, M. y Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic nuclei. *The Journal of the Acoustical Society of America*, 119(5), 3048-3058.
- Azcúnaga, R. (2010). Fonética del español salvadoreño. En M. Á. Quesada Pachecho (ed.), *El español hablado en América Central: nivel fonético*. (pp. 83-113). Iberoamericana y Vervuert.
- Baayen, H. (2008). *Analyzing Linguistic data. A practical introduction to statistics using R*. Cambridge University Press.
- Baayen, H. (2011). Mixed-Effects Models. En Cohn, A. C., C. Fougeron y M. K. Huffman, *The Oxford Handbook of Laboratory Phonology* (pp. 668-677). Oxford University Press.
- Baayen, H. (2013). Multivariate Statistics. En R. Podesva y D. Sharma (Edits.), *Research Methods in Linguistics* (pp. 337-372). Cambridge University Press.
- Barnes, J. (2006). *Strength and weakness at the interface. Positional neutralization in phonetics and phonology*. Mouton de Gruyter.
- Barry, W. (1998). Time as a factor in the acoustic variation of schwa. En R. Mannell y J. Robert-Ribes (ed.), *The 5th International Conference on spoken Language Processing*. Australian Speech Science and Technology Association, Incorporated (ASSTA).
- Barry, W. y Andreeva, B. (2001). Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association*, 31(1), 51-66.
- Bates, D. M., Bolker, B. y Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Bell, A., Breiner, J., Gregory, M. G. y Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60, 92-111.
- Bernal, J. (2016). *Actitudes lingüísticas de los bogotanos hacia los dialectos del español hablado en Colombia y las variedades nacionales de esta lengua en los demás países de Hispanoamérica* [tesis doctoral, Universidad de Bergen]. Bergen, Noruega.
- Bernal, J., Díaz, C. y Munévar, A. (2018). Linguistic Attitudes of the Spanish Speakers From Latin American Countries and Spain Respect to the Spanish Spoken in

- Colombia. *LILA'18. III. International Linguistics and Language Studies*, (pp. 109-217). Estambul .
- Bladlow, A. (1995). A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America*, 97(3), 1916-1924.
- Blomgren, M., Robb, M. y Chen, Y. (1998). A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research*, 41, 1042-1051.
- Boersma, P. y Weenink, D. (2021). *Praat: doing phonetics by computer [Computer program]. Version 6.1.38*. <http://www.praat.org/>
- Bosque, I. (1990). *Las categorías gramaticales. Relaciones y diferencias*. Editorial Síntesis.
- Brugman, H. y Russel, A. (2004). Annotating Multimedia/ Multi-modal resources with ELAN. *Proceedings of LREC 2004. Fourth International Conference on Language Resources and Evaluation*.
- Bruyninckx, M., Harmegnies, B., Llisterri, J. y Poch, D. (1994). Language-Induced voice quality variability in bilinguals. *Journal of Phonetics*, 1(22), 19-31.
- Brzezicha, B. y Kul, M. (2014). Effects os speech rate, phonetic background and gender on vowel reduction in the speech of non native speakers of english. *Poznań Studies in Contemporary Linguistics*, 50(4), 397-417.
- Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C. y Frauenfelder, U. H. (2011). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *The Journal of the Acoustical Society of America*, 130(6), 3980-3991.
- Bybee, J. (2001). *Phonology and language use*. Cambridge University Press.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language, Variation and Change*, 14, 261-290.
- Bybee, J., File-Muriel, R. y Napoleão de Souza, R. (2016). Special reduction: a usage-based approach. *Language and Cognition*, 8(3), 421-446.
- Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, 15(1-2), 39-54.
- Calamai, S. (2003). The Pisan vowel system of read ans semispontaneous speech. *Quaderni del Laboratorio di Linguistica*, 4, 72-98.

- Canfield, L. (1962). Observaciones sobre la pronunciación del castellano en Colombia. *Hispania*, 45, 247-248.
- Cangemi, F. y Niebuhr, O. (2018). Rethinking reduction and canonical forms. En F. Cangemi, M. Clayards, O. Niebuhr y B. y. Schuppler, *Rethinking Reduction. Interdisciplinary Perspectives on Conditions, Mechanisms, and Domains for Phonetic Variation* (pp. 277-302). De Gruyter Mouton.
- Cangemi, F., Clayards, M., Niebuhr, O., Schuppler, B. y Zellers, M. (eds.). (2018). *Rethinking reduction. Interdisciplinary perspectives on conditions, mechanism, and domains for phonetic variation*. Mouton de Gruyter.
- Cantero, F. (2002). *Teoría y análisis de la entonación*. Edicions Universitat de Barcelona.
- Carbonell, J. y Llisterri, J. (1992). Catalan. *Journal of the International Phonetic Association*, 22(1-2), 53-56.
- Catford, J. C. (1977). *Fundamental problems in Phonetics*. Indiana University Press.
- Catford, J. C. (1990). A proposal concerning central vowels. *Journal of the International Phonetic Association*, 20(2), 26-28.
- Caverlé, M. y Vogel, A. (2020). Stability, reability, and sensivity of acoustic measures of vowel space: a comparison of vowel space area, formant centralization ratio, and vowel articulation index. *The Journal of the Acoustical Society of America*, 148(3), 1436-1444.
- Chiba, T. y Kajiyama, M. (1958). *The vowel. Its nature ans structure*. Phonetic Society of Japan.
- Chiquito, A. y Quesada Pacheco, M. (2014). *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*. Universitetet i Bergen.
- Chládková, K. y Escudero, P. (2012). Comparing vowel perception and production in Spanish and Portuguese: European versus Latin American dialects. *Journal of the Acoustical Society of America, Express Letters*, 131(2), EL119-EL125.
- Chládkova, K., Escudero, P. y Boersma, P. (2011). Context-specific acoustic differences betwee Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America*, 130(1), 416-428.
- Chosmky, N. y Halle, M. (1969). *The sound patterns of English*. Harper & Row Publishers.

- Clopper, C. y Turnbull, R. (2018). Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors. En F. Cangemi, M. Clayars, O. Niebuhr, B. Schuppler y M. Zellers, *Rethinking reduction: interdisciplinary perspectives on conditions, mechanism, and domains for phonetic variation* (pp. 25-72). Walter de Gruyter.
- Coloma, G. (2005). Una versión alternativa de “El viento norte y el sol” en español. *Revista de investigación lingüística* (18), 191-212.
- Correa, J. A. (2014). *Manual de análisis acústico del habla con Praat*. Instituto Caro y Cuervo.
- Correa, J. A. (2017). Caracterización acústica de la reducción vocálica en el español hablado en Bogotá (Colombia). *Estudios de Fonética Experimental*, 26(1), 63 - 91.
- Correa, J. A. y Rodríguez, L. C. (2018). Reducción fonética de la secuencia consonántica /-st-/ en el español de Bogotá. *Estudios Filológicos*, 193-214.
- Crothers, J. (1978). Typology and universals of vowel systems. En J. Greenberg (ed.), *Universals of Human Language* (vol. 2, pp. 93-152). Stanford University Press.
- Cuervo, R. J. (1867). *Apuntaciones críticas sobre el lenguaje Bogotano*. Arnulfo M. Guarín.
- de Carvalho, F. (2010). Vowel acoustics in Pirahã. *Revista de Estudos da Linguagem*, 18(1), 11-33.
- de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America*, 97(1), 491-504.
- Delattre, P. (1969a). An acoustic and articulatory study of vowel reduction in four languages. *IRAL-International Review of Applied Linguistics in Language Teaching*, 7(4), 295-325.
- Delattre, P. (1969b). Coarticulation and the locus theory. *Studia Linguistica*, 23(1), 1-26.
- Delattre, P., Liberman, A. y Cooper, F. (1955). Acoustic Loci and Transitional Cues for Consonants, 769-773. *The Journal of the Acoustical Society of America*, 27.

- Delforge, A. M. (2008). Unstressed vowel reduction in Andean Spanish. En L. Colantoni y J. Steele (eds.) *Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 107-24.
- DiCanio, C., Nam, H., Amith, J. y Castillo García, R. (2015). Vowel variability in elicited versus spontaneous speech: evidence from Miztec. *Journal of Phonetics*, 48, 45-59.
- Diehl, R., Lindblom, B., Hoemeke, K. y Fahey, R. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics*, 24(2), 187-208.
- Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish. *The Journal of the Acoustical Society of America*, 83(5), 1863-1875.
- Engstrand, O. (1992). Systematicity of phonetic variation in natural discourse. *Speech Communication* (11), 337-346.
- Engstrand, O. y Krull, D. (1988). On the systematicity of phonetic variation in spontaneous speech. *Phonetic experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)*, 8, 34-47.
- Ericsson, C. y Ericsson, A. (2001). Gender differences in vowel duration in read Swedish: Preliminary results. *Fonetik 2001, XIVth Swedish Phonetics Conference*, (pp. 34-37). Lund.
- Ernestus, M. y Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39(3), 253-260.
- Espejo, M. B. (2012). El español bogotano de ayer y de hoy. En C. Patiño Rosselli y B. Leongómez (Edits.), *El lenguaje en Colombia* (pp. 197-210). Academia Colombiana de la Lengua e Instituto Caro y Cuervo.
- Fabricius, A., Watt, D. y Johnson, D. E. (2009). A comparison of three speaker-intrinsic vowel format frequency normalization algorithms for sociophonetics. *Language Variation and Change*, 21, 413-435.
- Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scalings. *STL- Quarterly Progress and Status Report*, 7(4), 22-30.
- Fant, G. (1970). *Acoustic theory of speech production*. Mouton.
- Fant, G. (1975). Non-uniform vowel normalization. *STL- Quarterly Progress and Status Report*, 16(1-3), 1-19.

- Ferguson, S. y Kewley-Port, D. (2007). Talker differences in clear and conversational speech: acoustic characteristics of vowels. *Journal of Speech, Language and Hearing Research*, 50(5), 1241-1255.
- Fleischer, J. y Schmid, S. (2006). Zurich German. *Journal of the International Phonetic Association*, 36(2), 243-253.
- Flemming, E. (2005). *A phonetically-based model of phonological reduction*.
<https://n9.cl/z2hpp>
- Flemming, E. (2009). The phonetics of schwa vowels. En D. Minkova (ed.), *Phonological weakness in English. From old to present-day English*. Palgrave MacMillan.
- Flórez, L. (1951). *La pronunciación del español en Bogotá*. Instituto Caro y Cuervo.
- Flórez, L. (1973). *Las «apuntaciones críticas» de Cuervo y el español bogotano cien años después*. Instituto Caro y Cuervo.
- Fourakis, M. (1991). Tempo, stress, and vowel reduction in american english. *The Journal of the Acoustical Society of America*, 90(4), 1816-1827.
- Gay, T. (1977). Effect of speaking rate on vowel formant movements. *STL-Quarterly Progress and Status Report*, 18(2-3), 8-30.
- Gay, T. (1978). Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical Society of America*, 63(1), 223-230.
- Gendrot, C. y Adda-Decker, M. (2007). Impact of duration and vowel inventory size on formant values of oral vowels: an automated formant analysis from eight languages. En J. Trouvain y W. Barry (ed.), *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken.
- Gendrot, C., Adda-Decker, M. y Santiago, F. (2019). Acoustic realization of vowels as a function of syllabic position. A crosslinguistic study with data from French and Spanish. En M. Gibson y J. Gil, *Romance Phonetics and Phonology* (pp. 77-88). Oxford University Press.
- Gil, J. (2007). *Fonética para profesores del español: de la teoría a la práctica*. Arco Libros, S. L.
- Godínez, M. (1978). A comparative study of some romance vowels. *UCLA Working Papers in Phonetics*, 41, 3-19.
- Goldman, J.-P. (2011). EasyAlign: an automatic phonetic alignment tool under Praat. *InterSpeech*. Firenze, Italy.

- González de la Calle, P. (1963). *Contribución al estudio del bogotano*. Instituto Caro y Cuervo.
- Gósy, M. y Horváth, V. (2010). Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association*, 40(2), 135-161.
- Gries, S. T. (2015). The most underused statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora*, 10(1), 95-125.
- Gussenhoven, C. (1992). Dutch. *Journal of the International Phonetic Association*, 22(1-2), 45-47.
- Harmegnies, B. y Poch-Olivé, D. (1992). A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication*, 11, 429-437.
- Harrington, J. (2010). *Phonetic Analysis of Speech Corpora*. Oxford: Willey Blackwell.
- Harrington, J. y Cassidy, S. (1999). *Techniques in Speech Acoustics*. Kluwer Academic Publishers.
- Harris, K. (1978). Vowel duration change and its underlying physiological mechanism. *Language and Speech*, 21(4), 354-361.
- Harris, K., Hoffman, H., Liberman, A., Delattre, P. y Cooper, F. (1958). Effect of Third-Formant Transitions on the Perception of the Voiced Stop Consonants. *The Journal of the Acoustical Society of America*, 30(2), 122-126.
- Hawkins, S. (2018). Preface. En M. a. *Interdisciplinary Perspectives on Conditions, Rethinking Reduction: (pp. V-VII)*. Walter de Gruyter.
- Hawkins, S. y Smith, R. (2001). Polysp: a polysystemic, phonetically-rich approach to speech understanding. *Rivista di Linguistica*, 1(13), 99-189.
- Hayward, K. (2000). *Experimental Phonetics: An introduction*. Longman.
- Healy, K. (2019). *Data visualization. A practical introduction*. Princenton University Press.
- Heijmans, L. y Gussenhoven, C. (1998). The Dutch dialect of Weert. *Journal of the International Phonetic Association*, 28(1-2), 107-112.
- Herrmann, F., Cunningham, S. y Whiteside. (2014). Speaker sex effects on temporal and spectro-temporal measures of speech. *Journal of the International Phonetic Association*, 59-74.

- Holliday, N. y Martin, S. (2017). Vowel categories and allophonic lowering among Bolivian Quechua–Spanish bilinguals. *Journal of the International Phonetic Association*, 1-24.
- Hualde, I. y Colina, S. (2014). *Los sonidos del español*. Cambridge University Press.
- International Phonetic Association. (1999). *Handbook of the International Phonetic Association*. Cambridge University Press.
- International Phonetic Association. (1949). *The Principles of the International Phonetic Association*. Department of Phonetics, University College of London.
- International Phonetic Association. (1989). Report on the 1989 Kiel Convention: International Phonetic Association. *Journal of the International Phonetic Association*, 19(2), 67-80.
- International Phonetic Association. (1993). Council actions on revisions of the IPA. *Journal of the International Phonetic Association*, 23(1), 32-34.
- Iskarous, K., Mooshammer, C., Hoole, P., Recasens, D., Shadle, C., Saltzman, E. y Whalen, D. (2013). The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance. *Journal of the Acoustical Society of America*, 134(2), 1271-1281.
- Jakobson, R., Fant, G. y Halle, M. (1969). *Preliminaries to speech analysis. The distinctive features and their correlates*. The MIT press.
- Jaworski, S. (2009). Inertial and non-inertial lenition processes. *Poznań Studies in Contemporary Linguistics*, 45(1), 103-129.
- Johnson, K. (2003). *Acoustic and Auditory Phonetics*. Blackwell Publishing.
- Johnson, K., Ladefoged, P. y Lindau, M. (1993). Individual differences in vowel production. *The Journal of the Acoustical Society of America*, 94(2), 701-714.
- Jones, D. (1922). *An outline of English phonetics*. G. E. Stechert & Co.
- Jones, D. (1944). Some thoughts on the phoneme. *Transactions of the Philological Society*, 43(1), 19-135.
- Joos, M. (1948). Acoustics phonetics. *Language*, 24(2), 1-136.
- Keating, P., Cho, T., Fougeron, C. y Hsu, C. (2004). Domain-initial articulatory strengthening in four languages. En J. Local, R. Ogden y R. Temple, *Phonetic Interpretation: Papers in Laboratory Phonology VI* (pp. 145-163). Cambridge University Press.

- Kendall, T. y Thomas, E. R. (2010). *Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, version 1.1*. Obtenido de [Software Resource: <http://ncslaap.lib.ncsu.edu/tools/norm/>]
- Koenig, L. (2006). Towards a physical definition of the vowel systems of languages. En V. Yngve y Z. Wasik (Edits.), *Hard-science Linguistics* (pp. 49-66). Bloomsbury Publishing.
- Kohler, K. (1999). German. *Journal of the International Phonetic Association*, 20(1), 48-50.
- Kohler, K. (2001). Foreword by the President of the International Phonetic Association. *Journal of the International Phonetic Association*, 31(1), V-VI.
- Kohler, K. J. (1996). Developing a research paradigm for sound patterns of connected speech in the languages of the world. *AIPUK*, 31, 227-233.
- Kondo, Y. (1994). Targeless schwa: is that how we get the impression of stress-timing in English? *Proceedings of the Edinburgh Linguistics Department Conference*. The University of Edinburgh .
- Koopmans-van Beinum, F. (1983). Systematics in vowel systems. En M. von den Broecke, V. van Heuven y W. Zonneveld (Edits.), *Sound structures. Studies for Antonie Cohen* (págs. 159-172). Fortis publications.
- Koopmans-van Beinum, F. y Harder, J. (1982). Word classification, word frequency, and vowel reduction. *Proceedings from the Institute of Phonetic Sciences of the University of Amsterdam*, 7, 61-69.
- Krull, D. (1989). Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)* (10), 87-108.
- Kuehn, D. P. y Moll, K. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4, 303-320.
- Kul, M. (2015). Spech rate plays marginal role in processes of connected speech. En T. S. 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, (pp. 0252. 1-5). Glasgow.
- Labov, W. (1972). Some principles of linguistic methodology. *Language in society*, 1(1), 97-120.
- Lacheret-Dujour, A. (1991). Phonological variations in read speech, reduction phenomena and speaker classes: Do allophonic choices represent speaking style?

- En J. Llisterra y D. Poch (ed.), *ESCA Workshop on the Phonetics and Phonology of Speaking Styles*, (pp. 38.1-10). Barcelona.
- Ladefoged, P. (1967). *Three Areas of Experimental Phonetics*. Oxford University Press.
- Ladefoged, P. (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Blackwell Publishing.
- Ladefoged, P. y Johnson, K. (2010). *A course in phonetics* (6ª edición ed.). Thomson Wadsworth.
- Lass, R. (2009). On schwa: Synchronic prelude and historical Fugue. En D. Minkova (ed.), *Phonological weakness in English from old to present-day English* (pp. 47-77). Palgrave Macmillan.
- Laver, J. (1994). *Principles of phonetics*. Cambridge University Press.
- Laver, J. (1996). *The gift of speech: papers in the analysis of speech and voice*. Edinburgh University Press.
- Lehiste, I. (1976). Suprasegmental Features of Speech. En N. Lass (ed.), *Contemporary Issues in Experimental Phonetics* (pp. 225-239). Academic Press, INC.
- Lengeris, A. (2012). Phonetic vowel reduction in Standard modern Greek. En Z. Gavriilidou, A. Efthymiou, E. Thomadaki y P. Kambakis-Vougioklis (ed.), *Selected papers of the 10th International Conference of Greek Linguistics*. Democritus University of Trace.
- Levshina, N. (2015). *How to do Linguistics with R*. John Benjamins Publishing Company.
- Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-187.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of Acoustical Society of America*, 35(11), 1773-1781.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. En W. Hardcastle y A. Marchal (Edits.), *Modeling, Speech Production and Speech* (pp. 403-439). Kluwer.
- Lindblom, B. y Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *The Journal of the Acoustical Society of America*, 50(4), 1166-1179.
- Lindblom, B., Sussman, H. y Agwuele, A. (2009). A duration-dependent account of coarticulation for hyper and hypoarticulation. *Phonetica*, 188-195.

- Lipski, J. (1990). Aspects of Ecuadorian Vowel Reduction. *Hispanic Linguistics*, 4, 1-19.
- Lipski, J. (1994). *Latin American Spanish*. Longman.
- Llisterri, J. (1991). *Introducción a la fonética: el método experimental*. Editorial Antropos.
- Llisterri, J. (1992). Speaking styles in speech research. *ELSNET/ESCA/SALT Workshop on Integrating Speech and Natural Language*. Dublin.
- Llisterri, J. y Poch-Olivé, D. (1992). *Speech Communication*, 11(4-5), 321-322.
- Llisterri, J., Machuca, M., de la Mota, C. R. y Ríos, A. (2003). The perception of lexical stress in Spanish. En M. J. Solé, D. Recasens y J. Romero (ed.), *15th International Congress of Phonetic Sciences (ICPhS-15)*, (pp. 2023-2026). Barcelona.
- Lobanov, B. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America*, 49(2), 606-608.
- Lope Blanch, J. M. (1963). En torno a las vocales caedizas del español mexicano. *Nueva Revista de Filología Hispánica*, 17(1-2), 1-19.
- Luce, P. y Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- Machuca Ayuso, M. y Poch-Olivé, D. (2016). Dinámica de las vocales del español en contacto con el catalán. *Oralia*, 19, 153-175.
- Maddieson, I. (1984). *Patterns of Sounds*. Cambridge University Press.
- Mairano, P., Santiago, F. y Romano, A. (2015). Cross-linguistic differences between accented vs. unaccented vowel durations. En T. S. Consortium (ed.), *Proceedings of the 18th International Congress of Phonetic*, (pp. 0748.1-5). Glasgow.
- Marín Gálvez, R. (1994-1995). La duración vocálica en español. *ELUA. Estudios de Lingüística*, 10, 213-226.
- Martín-Butragueño, P. (2014). Vocales en contexto. En E. Herrera y R. Barriga (eds.), *Lenguas, estructuras y hablantes: estudios en homenaje a Thomas C. Smith Stark* (págs. 971-992). El Colegio de México.
- Martínez Celdrán, E. y Fernández Planas, A. M. (2007). *Manual de Fonética Española. Articulaciones y Sonidos del Español*. Ariel.
- Meunier, C. y Espesser, R. (2011). Vowel reduction in conversational speech in French: The role of lexical factors. *Journal of Phonetics*, 39, 271-278.

- Milder, V. (2013). Experimental and quasi-experimental Research in Clinical Linguistics and Phonetics. En *Research Methods in Clinical Linguistics and Phonetics: A practical guide* (pp. 28-47). Willey-Blackwell.
- Montes, J. J. (1982). El español de Colombia: propuesta de clasificación dialectal. *Thesaurus*, XXXVII, 23-92.
- Montes, J., Figueroa, Y., Mora, S., Aparicio, R., Espejo, M. y Duarte, G. (1998). *El español hablado en Bogotá: análisis previo de su estratificación social*. Instituto Caro y Cuervo.
- Moon, S.-J. y Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America*, 96(1), 40-55.
- Mooshammer, C. y Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, 38(2), 117-136.
- Moosmüller, S., Schmid, C. y Brandstätter, J. (2015). Standard Austrian German. *Journal of the International Phonetic Association*, 45(3), 339-348.
- Munson, B. y Solomon, N. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language and Hearing Research*, 47, 1048-1058.
- Nadeu, M. (2013). *The effects of lexical stress, intonational pitch accent, and speech rate on vowel quality in Catalan and Spanish* [tesis doctoral, University of Illinois at Urbana-Champaign].
- Nadeu, M. (2014). Stress- and speech rate-induced vowel quality variation in Catalan. *Journal of Phonetics*, 46, 1-22.
- Nakagawa, S. y Schielzeth, H. (2013). A general and simple method for obtaining R from generalized linear mixed models. *Methods in Ecology and Evolution*, (4), 133-142.
- Navarro Tomás, T. (1916). Cantidad de las vocales acentuadas. *Revista de Filología Española*, 3, 387-408.
- Navarro Tomás, T. (1918). *Manual de pronunciación española*. Centro de Estudios Históricos.
- Nicoladis, K. (2003). Acoustic variability of vowels in Greek Spontaneous Speech. *15th International Congress of Phonetic Sciences* (pp. 3221-3224). Barcelona.

- Niebuhr, O. y Michaud, A. (2015). Speech data acquisition: the underestimated challenge. *KALIPHO- Kieler Arbeiten zur Linguistik und Phonetik*, 3, 1-42.
- Nord, L. (1986). Acoustic studies of vowel reduction in Swedish. *STL-Quarterly Progress and Status Report*, 27(4), 19-36.
- O'Rourke, E. (2010). Dialect Differences and the Bilingual Vowel Space in Peruvian Spanish. En M. Ortega-Llebaria (ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 20-30). Cascadilla Proceedings Project.
- Oostendorp, M. (2014). Schwa in phonological theory. En L. Cheng y R. Sybesma (eds.), *The Second Glot International State-of-the-Article Book. The Latest in Linguistics*. De Gruyter Mouton.
- Ortega, J., González, J. y Marrero, V. (2000). Ahumada: A large corpus in Spanish for speaker characterization and identification. *Speech Communication*, 31(2), 255-264.
- Ortega-Llebaria, M. y Prieto, P. (2010). Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech*, 54(1), 73-97.
- Patiño Roselli, C. (2004). Aspectos del lenguaje en Colombia. *Conferencia dictada en Bogotá, noviembre de 2003 como parte del ciclo: Los maestros y maestras piensan a Colombia* (pp. 1-22). Universidad Nacional de Colombia.
- Pätzold, M. y Simpson, A. (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. (A. Simpson, K. J. Kohler y T. Rettstadt, Edits.) *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, 215-247.
- Peirce, C. S. (24 de Junio de 2020 de 1905). *The Nature of Science*. Grupo de estudios Peirceanos. <https://n9.cl/t5ccf>
- Peters, J. (2010). The Flemish-Brabant dialect of Orsmaal-Gussenhoven. *Journal of the International Phonetic Association*, 40(2), 239-246.
- Pfützing, H. y Niebuhr, O. (2011). Historical development of phonetic vowel systems- The last 400 years. *The 17th International Congress of Phonetic Sciences*, (pp. 160-163). Hong Kong.
- Piroth, H. y Skupinski, P. (2011). On the realization of schwa in two varieties of standard German. *The 17th International Congress of Phonetic Sciences*, (pp. 1606-1609). Hong Kong.

- Pluymaekers, M., Ernestus, M. y Baayen, H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America*, 118(4), 2561-2569.
- Pluymaekers, M., Ernestus, M. y Baayen, H. R. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62, 146-159.
- Poch-Olivé, D., Dhainaut, K. y Harmegnies, B. (2003). La adaptación del locutor al interlocutor: dinámica del sistema vocálico del español en función de las características de los participantes en la conversación. *ELUA*, 17, 519-530.
- Poch-Olivé, D., Harmegnies, B. y Martín Butragueño, P. (2008). Influencia del estilo de habla sobre las características de las realizaciones vocálicas en el español de México. *Actas del XV Congreso Internacional ALFAL*. Montevideo.
- Poch-Olivé, D., Huet, K. y Harmegnies, B. (2001). Variabilidad acústica y estatus fonológico de la vocal central en las lenguas románicas. En F. Sánchez Miret (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica* (pp. 263-270). De Gruyter.
- Quilis, A. y Esgueva, M. (1983). Realización de los fonemas vocálicos españoles en posición fonética normal. En M. Esgueva y M. Cantarero (eds.), *Estudios de fonética I* (pp. 137-252). Consejo Superior de Investigaciones Científicas.
- R Core Team. (2014). *R: a language and environment for statistical computing*. (R. F. Computing, Productor). <http://www.R-project.org/>
- Rathcke, T., Suart-Smith, J., Torsney, B. y Harrington, J. (2017). The beauty in a beast: Minimising the effects of diverse recording quality on vowel formant measurements in sociophonetic real-time studies. *Speech Communication*, 86, 24-41.
- Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, 97-114.
- Recasens, D. (2014). *Fonètica i fonologia experimentals del català. Vocals i consonants*. Institut d'Estudis Catalans.
- Recasens, D. y Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*, 48, 645-666.
- Recasens, D. y Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America*, 125(4), 2288-2298.

- Rehor, C. y Pätzold, M. (1996). The phonetic realization of function words in German spontaneous speech. *AIPUK*, 31, 5-10.
- Rischel, J. (1992). Formal linguistics and real speech. *Speech communication*, 11(4-5), 379-392.
- Roettger, T. (2018). Researcher degrees of freedom in phonetic research. *Journal of phonetics*, 10(1).
- Romanelli, S. y Menegotto, A. C. (2018). Características acústicas de las vocales tónicas y átonas del español rioplatense. Efectos del estilo de habla y del contexto consonántico. *Signo y Seña*, 157-179.
- Ronquest, R. (2013). An acoustic examination of unstressed vowel reduction in Heritage Spanish. En C. e. Howe (ed.), *Selected Proceedings of the 15th Hispanic Linguistics Symposium* (pp. 157-171). Cascadilla Proceedings Project.
- Rosales, M. A. (2010). El español de Nicaragua. En M. Á. Quesada Pacheco (ed.), *El español hablado en América Central: nivel fonético* (pp. 137-154). Iberoamericana/Vervuert.
- Ryalls, J. y Lieberman, P. (1982). Fundamental frequency and vowel perception. *The Journal of the Acoustical Society of America*, 72(4), 1631-1634.
- Sadowsky, S. (2012). Vocales de referencia del castellano de Chile. *V Jornadas Nacionales de Fonética*. Temuco.
- Sapir, S., Raming, L. y Spielman, J. (2010). Formant centralization ratio: a proposal for a new acoustic measure of disarthric speech. *Journal of Speech, Language, and Hearing Research*, 53, 114-125.
- Schützler, O. (2015). Transforming acoustic data: A comparison of methods, using multi-dimensional scaling. En S. Schmid, D. Stephan, A. Volker y M. J. Kolly (Edits.), *Trends in Phonetics and Phonology: Studies from German-speaking Europe* (pp. 35-47). Peter Lang AG.
- Schwartz, J., Boë, L., Vallée, N. y Abry, C. (1997). Major trends in vowel system inventories. *Journal of Phonetics*, 25(3), 233-253.
- Schwegler, A. y Correa, J. (2018). Languages in contact. The case of Colombia. En E. Nuñez, *Biculturalism and Spanish in contact. Sociolinguistic case studies* (pp. 145-175). Routledge.
- Sessarego, S. (2012a). Unstressed vowel reduction in Cochabamba, Bolivia. *Revista Internacional de Lingüística Iberoamericana*, X(2), 213-227.

- Sessarego, S. (2012b). Vowel weakening in Afro-Yungueño: Linguistic and social considerations. *Revista Brasileira de Estudos do Contato Linguístico*, 22(2), 279-294.
- Simpson, A. (2001a). Does articulatory reduction miss more patterns than it accounts for? *Journal of the International Phonetic Association*, 1(31), 29-39.
- Simpson, A. (2001b). Dynamic consequences of differences in male and female vocal tract dimensions. *The Journal of the Acoustical Society of America*, 109(5), 2153-2164.
- Simpson, A. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics*, 30(3), 417-435.
- Simpson, A. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621-640.
- Simpson, A. (2013). Spontaneous speech. En M. J. Jones y R. A. Knight (eds.), *Bloomsbury companion to phonetics* (pp. 155-170). Bloomsbury.
- Simpson, A. y Ericsson, C. (2007). Sex-specific differences in f0 and vowel space. En J. Trouvain y W. J. Barry (ed.), *16th International Congress of Phonetic Sciences*, (págs. 933-936). Saarbrücken.
- Singmann, H., Bolker, B., Westfall, J. y Aust, F. (2016). *afex: Analysis of factorial experiments. R package version 0.16-1*. <https://n9.cl/k501b>
- Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17(1-2), 3-45.
- Stevens, K. (1998). *Acoustic Phonetics*. The MIT Press.
- Stevens, K. y House, A. (1955). Development of a quantitative description of vowel articulation. *The Journal of the Acoustical Society of America*, 27(3), 484-493.
- Stevens, K. y House, A. (1963). Perturbation of vowel articulations by consonantal context: an acoustical study. *Journal of Speech, Language, and Hearing Research*, 6(2), 111-128.
- Stevens, M. y Harrington, J. (2014). The individual and the actuation of sound change. *Loquens*, 1(1), 1-10.
- Šturm, P. (2015). International Phonetic Congresses: The shift in research practices and areas of interest over 44 years. En T. S. 2015 (ed.), *International Phonetic Congresses (ICPhS). Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK

- Sussman, H., Dalson, E. y Gumbert, S. (1998). The effect of speaking style on a locus equation characterization of stop place of articulation. *Phonetica*, 55, 204-225.
- Swerts, M., Kloots, H., Gillis, S. y Schutter, G. (2003). Vowel reduction in spontaneous Spoken Dutch. En T. I. Technology (ed.), *ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, (pp. MAO4 1-4). Tokyo.
- Torreira, F. y Ernestus, M. (2010). The Nijmegen corpus of casual Spanish. En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis y D. Tapias (eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)* (pp. 2981-2985). Language Resources Association (ELRA).
- Torreira, F. y Ernestus, M. (2011). Realizations of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology*, 2, 331-353.
- Trautmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, 88(1), 97-100.
- Trautmüller, H. (1997). *Auditory scales of frequency representation*. <https://n9.cl/64f5>
- Trautmüller, H. (2005). *Speech considered as modulated voice*. <https://n9.cl/zfm4u>
- Trouvain, J. (2003). *Tempo variation in speech production. Implications for speech synthesis* [tesis de doctorado, Universidad de Saarland]. Saarbrücken.
- Tsao, Y.-C., Weismer, G. y Iqbal, K. (2006). The effect of intertalker speech rate variation on acoustic vowel space. *The Journal of the Acoustical Society of America*, 119(2), 1074-1082.
- van Bergen, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, 12(1), 1-23.
- van Bergem, D. R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication*, 14(2), 143-162.
- van Bergen, D. R. (1995). Perceptual and acoustic aspects of lexical vowel reduction, a sound change in progress. *Speech Communication*, 16(4), 329-358.
- van Son, R. J. y Pols, L. C. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America*, 88(4), 1683-1693.
- van Son, R. J. y Pols, L. C. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America*, 92(1), 121-127.

- van Son, R., Bolotonova, O., Lennes, M. y Pols, L. (2004). Frequency effects on vowel reduction in three typologically different languages (Dutch, Finnish, Russian). *Interspeech*, 5, 1277-1280.
- Verhoeven, J. (2005). Belgian Standard Dutch. *Journal of the International Phonetic Association*, 35(2), 243-247.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. y Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7(6), 632 - 638.
- Wagner, N. (2019). Reduced speech: All is variability. *WIREs Cognitive Science*, 10(4), 1-7.
- Wagner, P. y Trouvain, J. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, 48, 1-12.
- Warner, N. (2011). Methods for studying spontaneous speech. En A. C. Cohn, C. Fougerson y M. Huffman (eds.), *The Oxford handbook of laboratory phonology* (pp. 621-633). Oxford University Press.
- Weirich, M. y Simpson, A. (2013). Investigating the relationship between speaker fundamental frequency and acoustic vowel space size. *The Journal of the Acoustical Society of America*, 134(4), 2965-2974.
- Weirich, M. y Simpson, A. (2014). Articulatory vowel spaces of male and female speakers. *Proc. 10th International Seminar on Speech Production*, 449-452.
- Weirich, M. y Simpson, A. (2015). Impact and interaction of accent realization of speaker sex on vowel length in German. En S. Schmid, V. Dellwo, A. Leemann y M.-J. Kolly (eds.), *Trends in Phonetics and Phonology: Studies from German-speaking Europe* (pp. 109-123). Peter Lang.
- Whiteside, S. (1996). Temporal-based acoustic-phonetic patterns in read speech: some evidence for speaker sex differences. *Journal of the International Phonetic Association*, 26(1), 23-40.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.
- Winter, B. (2011). Pseudoreplication in phonetic research. *Proceedings of the International Congress of Phonetic Science*, (pp. 2137-2140). Hong Kong.
- Winter, B. (2015). The other N: The role of repetitions and items in the design of phonetic experiments. *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK.

- Winter, B. (2020). *Statistics for linguistics. An introduction using R*. Routledge.
- Wissing, D. y Wikus, P. (2014). Evaluating vowel normalisation procedures: a case study on Southern Sotho Vowels. *Southern African Linguistics and Applied Languages Studies*, 32(1), 97-111.
- Wright, R. (2004). Factors of lexical competition in vowel articulation. En J. Local, R. Ogden y R. Temple (eds.), *Papers in Laboratory Phonology VI* (pp. 26-50). Cambridge University Press.
- Xu, Y. (2010). In defense of lab speech. *Journal of phonetics*, 38(3), 329-336.
- Zharkova, N. (2007). Quantification of coarticulatory effects in several Scottish English phonemes using ultrasound. *QMU Speech Science Research Centre Working Paper, WP-13*.
- Zipf, G. (1929). Relative Frequency as a Determinant of Phonetic Change. *Harvard Studies in Classical Philology*, 40, 1-95.

A.Anexo: Textos usados en la tarea de lectura

Texto 1

El joyero Federico Vanero ha sido condenado por la audiencia de Santander a ocho meses de arresto mayor y cincuenta mil pesetas de multa por un delito de compra de objetos robados. La vista oral se celebró el miércoles pasado y, durante ella, uno de los fiscales, Carlos Valcárcel³, pidió para el joyero tres años de prisión menor y una multa de cincuenta mil pesetas. Gracias a las revelaciones de Vanero de hace dos años y medio se llegó a descubrir la existencia de una sospechosa mafia policial en España, parte de la cual se vio envuelta en el llamado “caso el Nani”.

Bruyninckx, M., Harmegnies, B., Llisterri, J. y Poch, D. (1994). Language-Induced voice quality variability in bilinguals. *Journal of Phonetics*, 22(1), 19-31.

Texto 2

Hay algo ahí, en el aire, que cambia el sentido de las cosas. Ese viento suave vuela, te toca la cara, mientras cuentas las hojas de los árboles. El agua corre buscando los campos. Al abrir las puertas de mi casa pienso: este país, una mañana más. A mi edad, comienzan a faltarme las fuerzas, ya casi no soy joven, y la muerte de mi mujer en la guerra me pesa mucho. Cuando el cuerpo llega a esa hora, la ciencia de los doctores no logra detener el paso del tiempo. De niño, allá en mi tierra, solía pasarme los días revolviendo de un lado a otro. Poco a poco, los coches de la ciudad fueron llamando mi atención. Mi madre decía que tuviera cuidado, pero yo me creía muy mayor, así que no

³ El apellido Valcárcel no se conoce en Colombia, así que en el texto presentado a los participantes cambiamos este apellido por Velásquez para no ocasionar errores de lectura.

tenía ni interés ni tiempo para mi propio signo. Pero sigo, es cierto, cuántas cosas buenas encontré entre su gente. Si cuento los queridos veranos de entonces, no son siete, ni nueve, ni veinte. Debe ser que soy niño de nuevo en este cuerpo triste.

Ortega, J., González, J. y Marrero, V. (2000). Ahumada: A large corpus in Spanish for speaker characterization and identification. *Speech Communication*, 31(2), 255-264.

Texto 3

El viento norte y el sol discutían sobre cuál de ellos era el más fuerte, cuando pasó un extraño viajero envuelto en unas ropas muy abrigadas. Convinieron en que quien antes lograra obligarlo⁴ al transeúnte a quitarse el abrigo sería considerado más poderoso. El viento (...) sopló con gran furia, pero cuanto más soplaba, más se ceñía el hombre su ropa al cuerpo. Entonces se dio por vencido, y el sol empezó a brillar con mucha fuerza. Inmediatamente el viajero se despojó de su abrigo; y así ya quedó claro que el sol tenía superioridad respecto del viento.

Coloma, G. (2005). Una versión alternativa de “El viento norte y el sol” en español. *Revista de investigación lingüística*, 18, 191-212

Texto 4

El municipio de Buenaventura cubre parte de la ruta conocida como la «vía al mar». Su población es de tez oscura y muy culta. En los últimos años ocupa la atención de los medios porque su población marcha en procura de hacer valer sus derechos: quieren un mayor impulso financiero para obtener más cupos escolares, poner la lupa sobre los corruptos, brindar ayudas a los adultos mayores y juventudes, capturar a los miembros de los grupos armados, acabar con las torturas, entre otros.

⁴ Esta construcción es propia del español argentino, y no resulta ‘natural’ para un hablante colombiano. Así que los participantes leyeron *quien antes lograra obligarlo al transeúnte*.

B. Anexo: Temas de discusión utilizados para obtener habla conversacional

Instrucciones y temas de discusión

Discuta cada uno de los siguientes temas con su interlocutor, expresen sus posiciones sobre el tema, identifiquen los problemas y lleguen a un acuerdo:

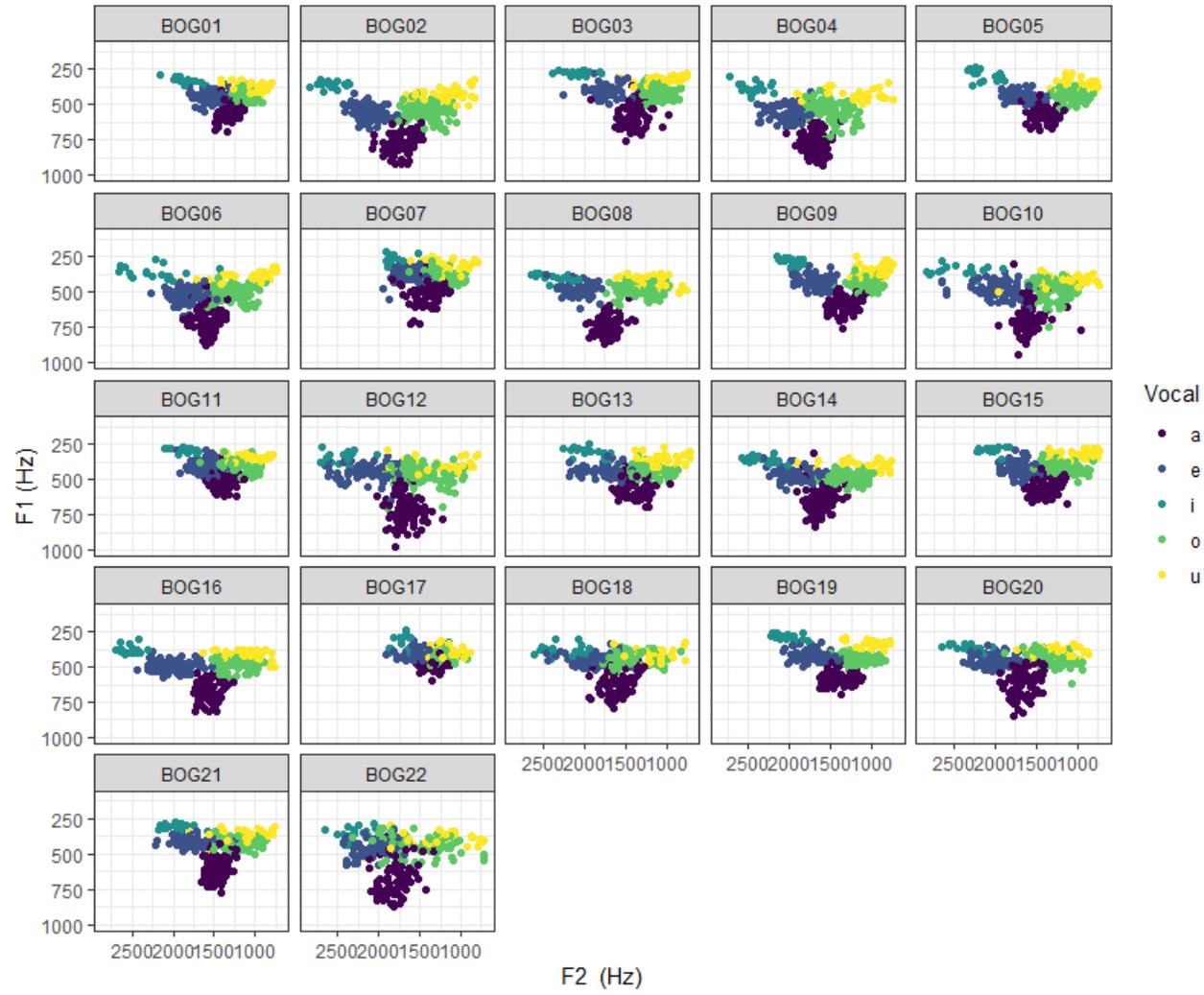
- 1) La senadora Vivian Morales afirma que la adopción deber ser solamente para parejas conformadas por un hombre y una mujer.
- 2) En los periódicos y revistas se afirma que los miembros de la generación *millennial* (nacidos entre las décadas de 1980 y 2000) le dicen “no” al matrimonio. De acuerdo con sus experiencias, ¿creen que esta afirmación es correcta?
- 3) Durante años, la Policía ha registrado decenas de casos de agresiones sexuales en el sistema de transporte Transmilenio. Frente a esta situación, las autoridades han retomado la idea de asignar vagones exclusivos para mujeres o ubicar a las pasajeras en una zona específica de los buses.
- 4) El Comité Noruego del Nobel le otorgó el año pasado el Premio Nobel de la Paz al presidente de Colombia, Juan Manuel Santos. Sin embargo, hoy muchos critican este reconocimiento.
- 5) El domingo 22 de enero se reabrieron las puertas de la plaza de toros de la Santamaría, luego de cuatro años desde que se prohibieran las corridas en la ciudad de Bogotá. El Distrito entregó el escenario remodelado para estas actividades. Las obras duraron ocho meses y se invirtieron \$ 8.871 millones de pesos.

C. Anexo: Valores medios y desviaciones estándar de la duración vocálica en función la tonicidad, la velocidad de elocución y el sexo del hablante

	Tónica	Velocidad	Duración (ms)	Átona	Velocidad	Duración (ms)
Mujeres	i	lenta	91.84 (21.15)	i	lenta	79.89 (35.04)
		normal	67.14 (13.97)		normal	65.47 (25.05)
		rápida	57.23 (12.97)		rápida	56.86 (16.37)
	e	lenta	101.63 (36.84)	e	lenta	70.16 (23.74)
		normal	78.31 (23.55)		normal	57.29 (14.01)
		rápida	66.76 (17.21)		rápida	50.34 (12.84)
	a	lenta	118.93 (31.47)	a	lenta	79.14 (26.36)
		normal	95.74 (24.98)		normal	61.38 (16.44)
		rápida	73.52 (24.74)		rápida	55.14 (13.54)
	o	lenta	105.83 (31.28)	o	lenta	67.26 (21.22)
		normal	88.90 (23.37)		normal	53.66 (14.63)
		rápida	69.99 (23.02)		rápida	50.21 (12.64)
u	lenta	99.26 (28.77)	u	lenta	77.19 (27.87)	
	normal	75.78 (24.30)		normal	56.90 (21.03)	
	rápida	61.25 (20.99)		rápida	51.72 (17.54)	
Hombres	i	lenta	81.84 (17.51)	i	lenta	67.00 (21.90)
		normal	62.20 (16.06)		normal	64.66 (24.73)
		rápida	49.61 (9.98)		rápida	49.95 (13.28)
	e	lenta	87.80 (30.17)	e	lenta	62.48 (14.39)
		normal	65.80 (19.56)		normal	50.24 (11.03)
		rápida	57.14 (14.17)		rápida	45.57 (12.44)
	a	lenta	95.09 (22.36)	a	lenta	69.46 (19.64)
		normal	78.43 (20.04)		normal	57.76 (14.27)
		rápida	61.92 (18.29)		rápida	50.66 (11.68)
	o	lenta	91.51 (26.47)	o	lenta	62.39 (17.26)
		normal	72.43 (22.05)		normal	51.33 (12.87)
		rápida	57.70 (18.24)		rápida	46.83 (11.39)
u	lenta	84.46 (18.77)	u	lenta	67.05 (20.27)	
	normal	70.68 (17.74)		normal	57.00 (17.59)	
	rápida	54.25 (14.97)		rápida	52.75 (15.19)	

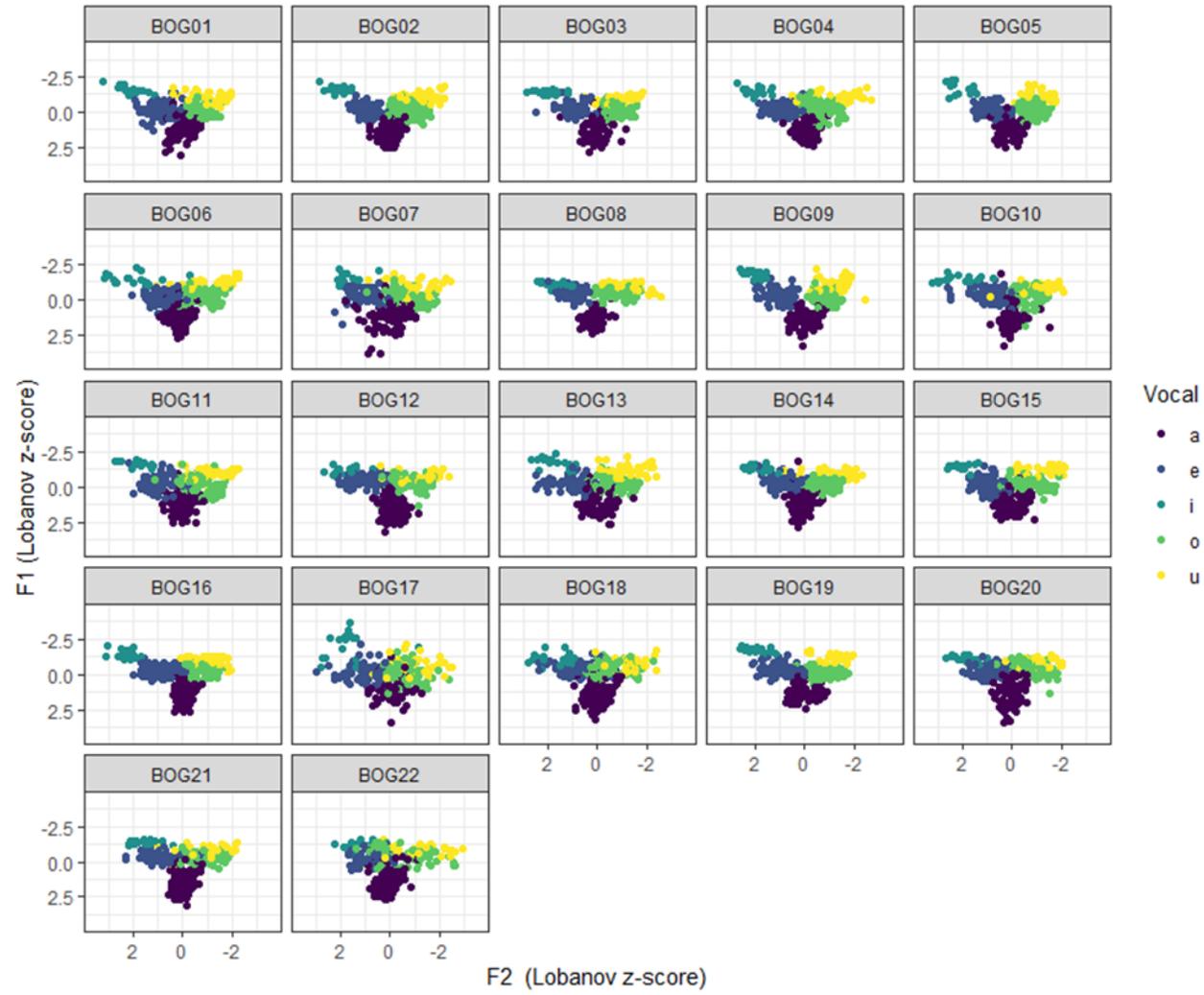
D.Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (Hz) en función del timbre, la velocidad de elocución y el sexo del hablante

	Vocal	Velocidad	F1 (Hz)	σ	F2 (Hz)	σ
Mujeres	i	lenta	353.05	34.96	2523.86	224.54
		normal	376.64	38.38	2422.07	207.59
		rápida	384.81	31.57	2366.38	226.02
	e	lenta	485.80	63.32	2097.76	210.28
		normal	503.66	62.28	2015.46	168.93
		rápida	484.16	63.44	1957.36	204.49
	a	lenta	708.60	104.93	1650.18	151.77
		normal	717.46	91.36	1655.27	147.31
		rápida	659.89	103.94	1650.32	155.33
	o	lenta	482.52	59.95	1326.20	235.43
		normal	494.21	63.01	1386.11	211.54
		rápida	479.92	66.70	1424.43	219.97
	u	lenta	399.09	38.25	1109.45	271.15
		normal	409.78	38.78	1174.11	289.38
		rápida	413.93	43.60	1276.23	310.04
Hombres	i	lenta	288.37	35.93	2037.00	158.93
		normal	300.76	31.78	1962.93	195.26
		rápida	315.94	37.60	1883.84	157.23
	e	lenta	412.55	42.61	1743.04	169.83
		normal	421.66	43.82	1677.74	150.55
		rápida	420.45	42.99	1622.15	146.60
	a	lenta	568.48	77.82	1394.09	127.96
		normal	557.60	69.27	1379.72	133.37
		rápida	529.74	68.74	1365.27	139.53
	o	lenta	416.67	37.80	1129.34	146.34
		normal	424.33	40.27	1175.39	140.49
		rápida	423.31	40.84	1214.81	147.26
	u	lenta	335.02	36.14	987.41	187.32
		normal	351.00	37.58	1053.83	212.75
		rápida	358.76	38.62	1104.66	167.20



E. Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (normalizadas con el método de Lobanov) en función del timbre, la velocidad de elocución y el sexo del hablante

	Vocal	Velocidad	F1	σ	F2	σ
Mujeres	i	lenta	-1.32	0.30	1.99	0.49
		normal	-1.13	0.34	1.75	0.48
		rápida	-1.07	0.30	1.63	0.50
	e	lenta	-0.29	0.39	0.98	0.44
		normal	-0.15	0.40	0.79	0.36
		rápida	-0.30	0.39	0.65	0.44
	a	lenta	1.46	0.73	-0.07	0.33
		normal	1.52	0.63	-0.05	0.31
		rápida	1.07	0.69	-0.07	0.34
	o	lenta	-0.31	0.37	-0.84	0.52
		normal	-0.22	0.41	-0.70	0.47
		rápida	-0.33	0.42	-0.61	0.49
	u	lenta	-0.96	0.29	-1.36	0.62
		normal	-0.88	0.31	-1.20	0.67
		rápida	-0.84	0.35	-0.96	0.72
Hombres	i	lenta	-1.61	0.40	2.01	0.40
		normal	-1.47	0.37	1.75	0.55
		rápida	-1.31	0.44	1.52	0.43
	e	lenta	-0.25	0.44	1.07	0.48
		normal	-0.14	0.44	0.85	0.42
		rápida	-0.15	0.45	0.68	0.42
	a	lenta	1.40	0.70	-0.04	0.38
		normal	1.30	0.64	-0.09	0.40
		rápida	1.00	0.66	-0.13	0.43
	o	lenta	-0.20	0.38	-0.87	0.45
		normal	-0.12	0.41	-0.72	0.41
		rápida	-0.12	0.43	0.59	0.43
	u	lenta	-1.08	0.35	-1.33	0.59
		normal	-0.90	0.37	-1.11	0.65
		rápida	-0.82	0.39	-0.94	0.50



F. Anexo: Valores medios y desviaciones estándar de las frecuencias de F1 y F2 (Datos en hercios (Hz) y normalizados con el método de Lobanov) en función del timbre y el sexo del hablante en habla espontánea

Valores medios y desviaciones estándar en hercios

	Vocal	Tonicidad	F1 (Hz)	σ	F2 (Hz)	σ	Duración	σ
Mujeres	i	Átona	374.92	50.95	2213.31	286.88	50.99	11.56
		Tónica	379.04	49.94	2351.19	259.09	64.89	21.49
	e	Átona	515.68	73.94	1845.40	243.26	58.50	19.98
		Tónica	544.76	70.91	1993.13	230.50	76.96	29.11
	a	Átona	671.59	84.88	1624.31	163.48	66.09	23.02
		Tónica	743.40	76.55	1660.25	150.96	89.92	31.71
	o	Átona	503.49	55.48	1367.79	202.24	58.88	20.54
		Tónica	538.07	60.00	1253.67	195.72	78.87	30.59
	u	Átona	424.57	55.87	1305.52	280.01	58.20	21.54
		Tónica	415.15	39.91	1103.24	245.51	75.66	31.67
Hombres	i	Átona	307.13	43.97	1939.41	155.53	49.22	15.56
		Tónica	315.39	41.79	1985.91	152.47	66.89	25.80
	e	Átona	410.90	54.74	1629.14	197.62	54.49	18.08
		Tónica	424.24	45.96	1674.39	150.41	67.70	25.73
	a	Átona	523.30	73.17	1412.22	144.78	61.13	18.09
		Tónica	575.89	70.63	1441.43	120.63	77.80	28.05
	o	Átona	408.37	49.49	1187.92	179.53	55.56	20.45
		Tónica	431.26	49.91	1083.74	147.34	69.30	28.13
	u	Átona	330.21	38.75	1191.92	190.97	47.41	15.75
		Tónica	353.38	36.89	1004.99	194.81	68.92	26.87

Valores medios de la frecuencia de F1 y F2 normalizados con el método de Lobanov

	Vocal	Tonicidad	F1 (Lobanov)	F2 (Lobanov)
Mujeres	i	Átona	-1.29	1.45
		Tónica	-1.30	1.74
	e	Átona	-0.22	0.47
		Tónica	-0.01	0.86
	a	Átona	1.00	-0.08
		Átona	1.47	-0.02
	o	Átona	-0.31	-0.74
		Tónica	-0.06	-1.02
	u	Átona	-0.92	-0.88
		Tónica	-1.00	-1.39
Hombres	i	Átona	-1.44	1.68
		Tónica	-1.28	1.86
	e	Átona	-0.29	0.66
		Tónica	-0.13	0.85
	a	Átona	0.93	-0.04
		Tónica	1.44	-0.06
	o	Átona	-0.32	-0.78
		Tónica	-0.09	-1.10
	u	Átona	-1.12	-0.76
		Tónica	-0.94	-1.36

Departament de Filologia Espanyola
Facultat de Filosofia i Lletres
Programa de Doctorat en Filologia Espanyola

UAB
Universitat
Autònoma
de Barcelona