

Propuesta para la evaluación de métodos de segmentación y medición de parámetros acústicos

(de las consonantes aproximantes espirantes
del castellano)



UCL

Mauricio Figueroa y Bronwen G. Evans

m.figueroa.12@ucl.ac.uk, bronwen.evans@ucl.ac.uk

Speech, Hearing and Phonetic Sciences

Division of Psychology and Language Sciences

Aproximantes de /b, d¹, g/

Introducción

Alófonos aproximantes en distribución complementaria [sic]
(p. ej., Barlow, 2003; Parrell, 2011)

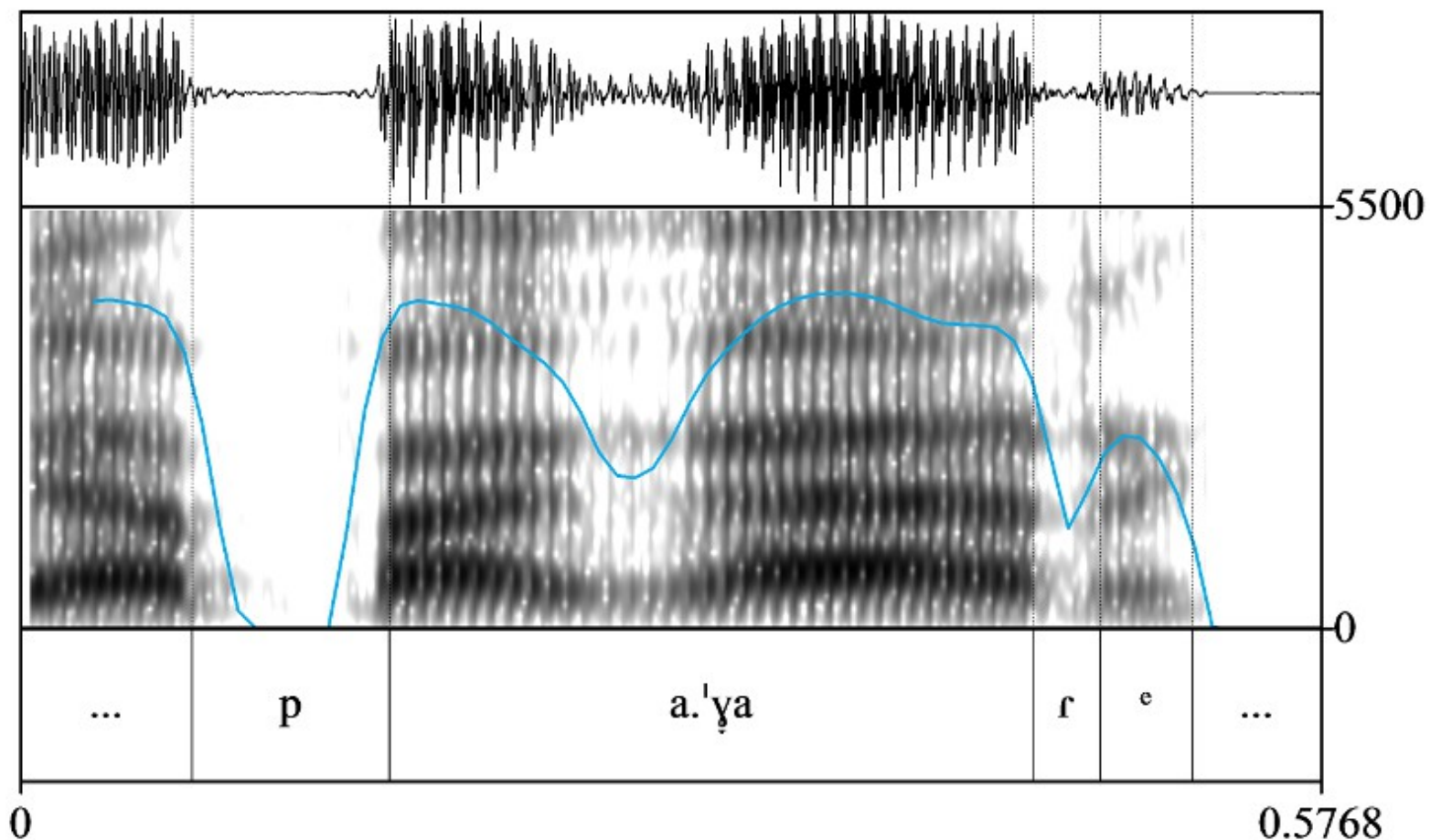
Forman parte de un **continuo** de oclusión hasta elisión
(p. ej., Hualde, Simonet, Shosted y Nadeu, 2010; Carrasco, Hualde y Simonet, 2012; Simonet, Hualde & Nadeu, 2012; Figueroa y Evans, 2014)

Pueden ser subclasificadas: aproximantes espirantes vocálicas (AV), **abiertas (AA)** y cerradas (AC)
(p. ej., Pérez, 2007; Martínez-Celdrán y Regueira, 2008; Martínez Celdrán, 2013)

¹ En el castellano chileno, las variantes de /d/ son habitualmente postdentales.

Aproximantes de /b, d, g/

Un ejemplo: AA



La palabra “pagar”, grabada en una lectura de listas de palabras. El contorno de intensidad se muestra como una línea continua (escala 30-65 dB).

Problema metodológico

¿Cómo segmentarlas?

Estas consonantes son difíciles de segmentar
(Turk, Nakai y Sugahara, 2006)

No siempre resulta claro cómo ha procedido la segmentación
(p. ej., Martínez Celdrán, 1984; Cepeda, 1991; Colantoni y Marinescu, 2010; Hualde, Shosted y Scarpace, 2011)

Predominan dos métodos complementarios: (a) segmentación manual y (b) uso de máximos y mínimos de intensidad

[a] p. ej., Martínez Celdrán, 1984; Kingston, 2008; Colantoni y Marinescu, 2010; Martínez Celdrán, 2013 (normalizadas)

[b] p. ej., Cole, Hualde e Iskarous, 1999; Ortega-Llebaría, 2003; Martínez-Celdrán y Regueira, 2008; Hualde et al., 2010; Colantoni y Marinescu, 2010

Problema metodológico

Dos métodos complementarios

Segmentación manual: uso de claves visuales y auditivas
(p. ej., Kingston, 2008; Colantoni y Marinescu, 2010; Figueroa y Evans, 2014)

Uso de máximos y mínimos de intensidad para mediciones basadas en la intensidad como correlato de constricción:

Proporción de intensidad (IntRatio¹):

intensidad en mínimo de C / intensidad en máximo de V2

Diferencia de intensidad (IntDiff_A², IntDiff_B³):

intensidad en máximo de V2 – intensidad en mínimo de C

intensidad en máximo de V1 – intensidad en mínimo de C

Velocidad de cambio (MaxVel⁴, MinVel⁵):

máxima velocidad de la curva de intensidad a partir de las primeras diferencias (intervalo 0,001 ms) entre el mínimo de C y el máximo de V2

mínima velocidad de la curva de intensidad a partir de las primeras diferencias (intervalo 0,001 ms) entre el mínimo de C y el máximo de V2

¹ p. ej., Parrell, 2010; Carrasco et al., 2012; ² p. ej., Hualde et al., 2010; Parrell, 2010; ³ p. ej., Martínez-Celdrán y Regueira, 2008; ⁴ p. ej., Parrell, 2010; Hualde et al., 2010; Kingston, 2008; ⁵ p. ej., Kingston, 2008.

Problema metodológico

¿Cómo se han evaluado estos métodos?

Búsqueda de la mejor correlación entre mediciones basadas en intensidad con mediciones de articulometría electromagnética (EMA) para /b/ a través de regresiones lineales simples. Mejor resultado: **IntRatio y mediciones sin filtro paso bajo (< 20 Hz)**.

(Parrell, 2010)

Estudio del efecto del segmento precedente en el grado de constricción de /d/ a través de IntDiff¹, inclinación espectral² y MaxVel. Mejor predictor en modelos de efectos mixtos: probablemente **MaxVel**.

(Hualde et al., 2010)

¿Alternativas?

¹ Filtro de paso Hann entre 500 y 10,000 Hz. ² Diferencia de energía en la consonante entre una banda de frecuencias bajas (50-500 Hz) y una de frecuencias altas (500-5000 Hz).

Evaluación de métodos

Uso de Regresión Logística Multinomial (RLM)

La RLM permite evaluar la membresía de dos o más niveles de una variable dependiente a partir de un conjunto de variables independientes, dado un nivel base de la variable dependiente.

(Field, Miles y Field, 2013)

Dicho más sencillamente: **permite determinar cuán posible es predecir el valor de una variable dependiente a partir de otras independientes.**

Asunciones: (a) Cada caso (dato) es único; (b) la variable dependiente no debe poder ser predecida de manera perfecta por alguna de las variables independientes; y (c) es preferible que las variables no estén altamente correlacionadas entre sí.

Estudio de producción

Características del estudio

Diez adultos, hablantes nativos del castellano chileno, formaron parte de la muestra.

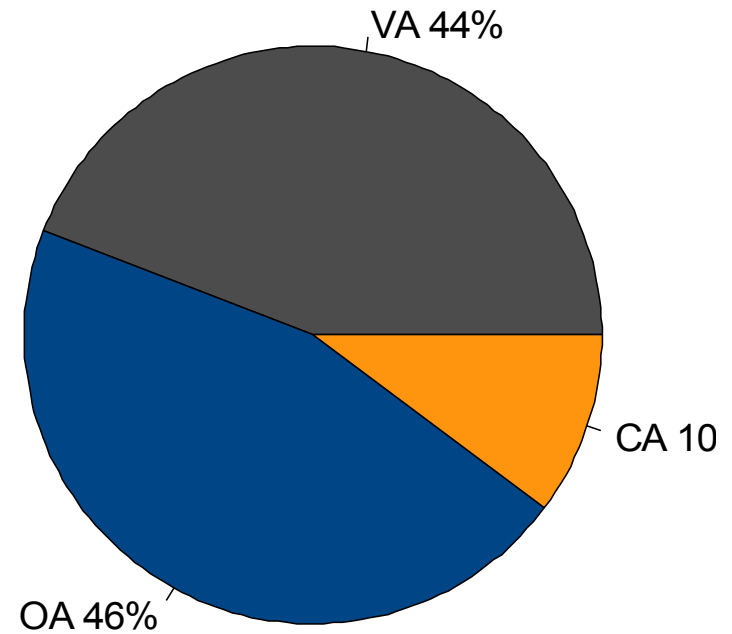
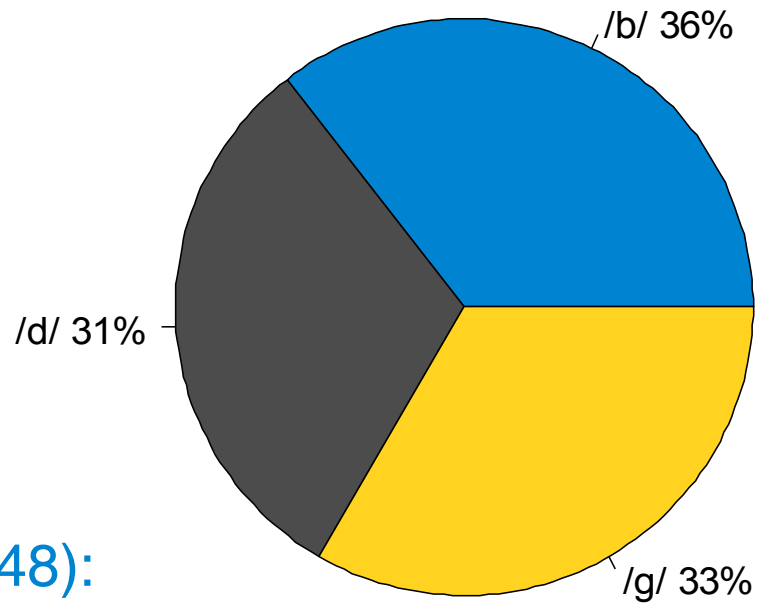
Tres estilos de habla fueron elicitados mediante: listas de palabras, lectura de textos breves y conversación semidirigida.

(Labov, 1972)

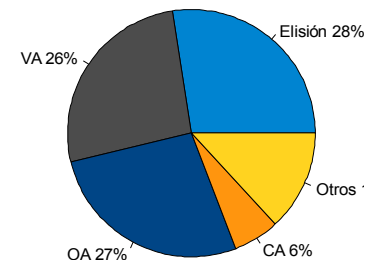
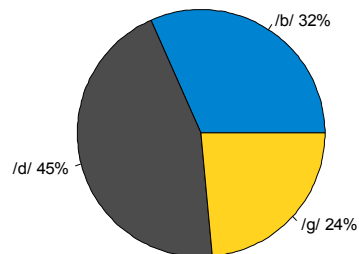
Las grabaciones fueron realizadas en la cabina insonorizada del Laboratorio de Fonética de la Pontificia Universidad Católica de Chile (micrófono de condensador cardioide AT3035; interfaz portable Mbox 2; Protools LE 7.4.2 para Mac).

Estudio de producción

Sobre los datos



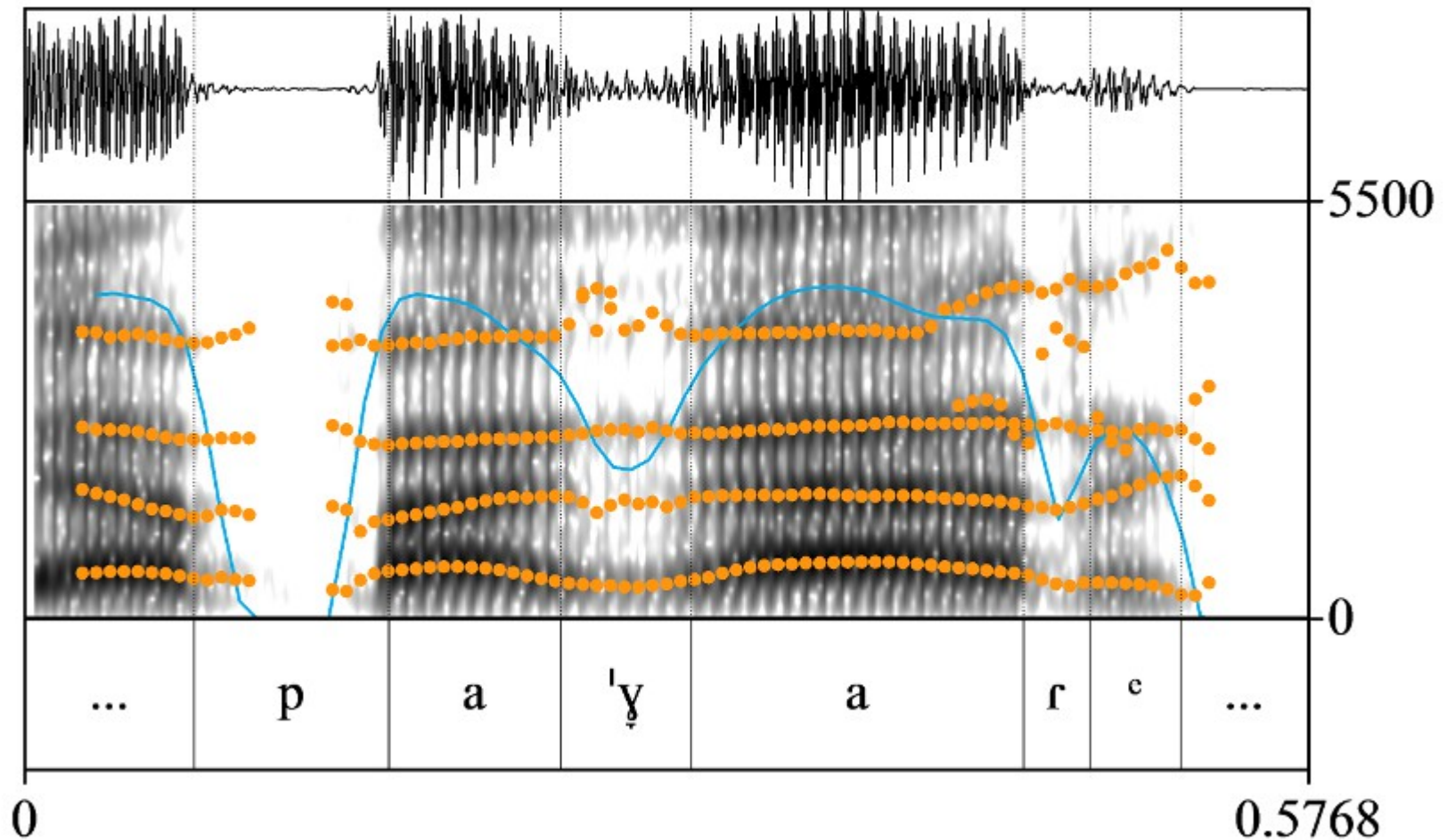
Subset (n = 4948):
solo aproximantes



Todos los datos (n = 9339): incluye
elisiones, fricativas y oclusivas

Método I

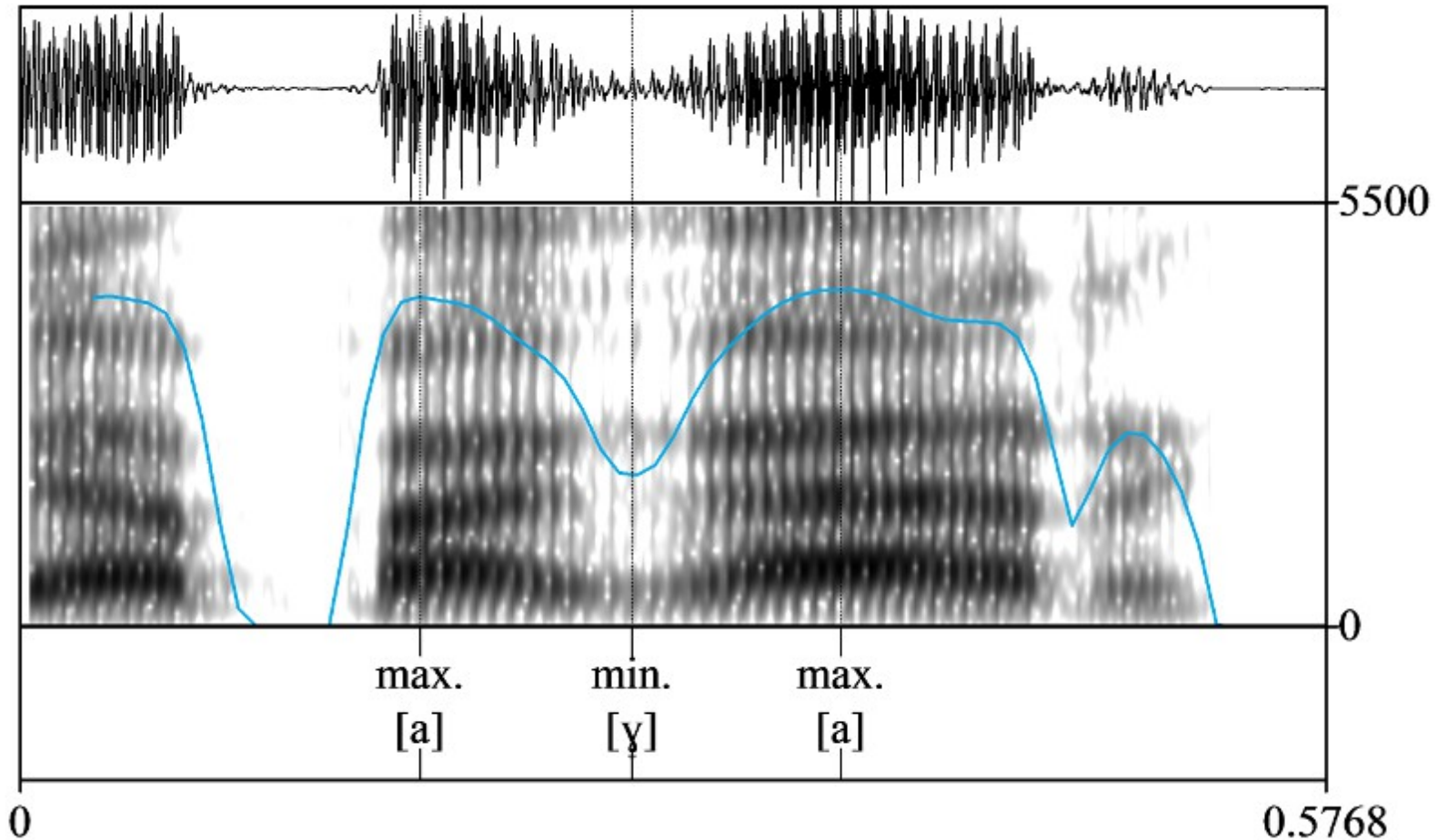
Segmentación manual



Uso de claves visuales en forma de onda, espectrograma, curva de intensidad y formantes. Localización de fronteras dentro de transiciones. Mediciones de (a) **duración** y (b) **intensidad** y **formantes** (F1 y F2) en el 50% interior de cada segmento.

Método II

Detección de máximos y mínimo



Identificación automatizada en *Praat* (Boersma y Weenink, 2014) de máximos y mínimo de intensidad. Mediciones acústicas de [intensidad](#) y [formantes](#) (F1 y F2) realizadas aquí. [Los cálculos de medidas de intensidad adicionales se basan en estos datos.](#)

Regresión Logística Multinomial

¿Qué tan buenos predictores?

Cuatro análisis de regresión logística multinomial fueron realizados en R (`mlogit`) para evaluar la capacidad predictiva de conjuntos de variables:

Método I.A – Segmentación manual: duración, intensidad, F1 y F2 para los tres segmentos de la secuencia.

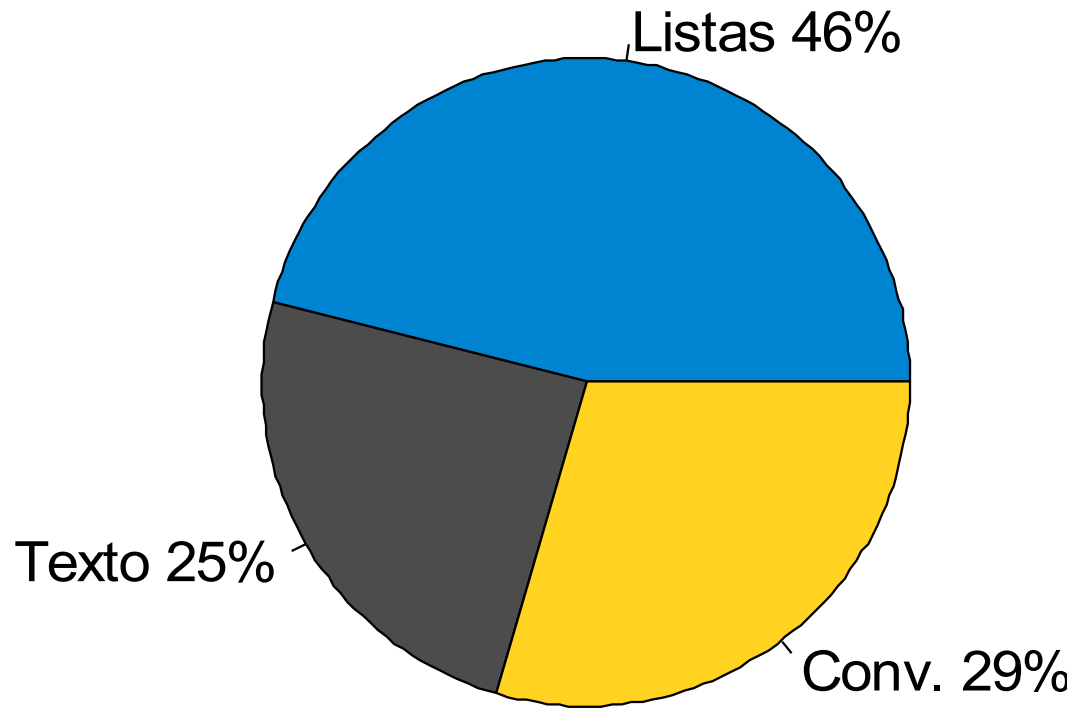
Método I.B – Segmentación manual: intensidad, F1 y F2 para los tres segmentos de la secuencia (¿y la duración?).

Método II – Identificación automatizada: mediciones de intensidad, F1 y F2 para los tres segmentos de la secuencia en los puntos máximos y mínimo de intensidad.

Mediciones de Intensidad: IntRatio, IntDiff_A, IntDiff_B, MaxVel y MinVel.

Regresión Logística Multinomial

¿Listas de palabras afectando valores formánticos?



Regresión Logística Multinomial

Método I.A: Resultados

	/b/				/d/			
	95% CI for odds ratio				95% CI for odds ratio			
	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)
(intercept)	0.24(0.58)	0.41	1.27	3.93	0.53(0.57)	0.56	1.70	5.18
dur_v1	4.37(1.27)***	6.50	78.84	955.85	8.67(1.23)***	526.35	5823.72	64436.21
dur_app	-5.94(2.79)*	0.00	0.00	0.63	-10.60(2.74)***	0.00	0.00	0.01
dur_v2	3.94(1.11)***	5.81	51.37	454.46	7.78(1.07)***	292.64	2383.40	19411.32
int_v1	-0.11(0.01)***	0.87	0.89	0.91	-0.11(0.01)***	0.87	0.89	0.91
int_app	0.08(0.01)***	1.06	1.09	1.11	0.06(0.01)***	1.04	1.06	1.08
int_v2	0.00(0.01)	0.97	1.00	1.02	0.01(0.01)	0.98	1.01	1.03
f1_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)*	1.00	1.00	1.00
f1_app	0.01(0.00)***	1.01	1.01	1.01	0.00(0.00)***	1.00	1.00	1.01
f1_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_app	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00

Frecuencias (n = 4948): /g/ = 0.33347, /b/ = 0.35509, y /d/ = 0.31144. Estimados con intervalos de confianza de un 95% y *odds ratios*. * $p < .05$, ** $p < .01$, *** $p < .001$.

Log-Likelihood: -4674.1

McFadden R²: 0.13902

Likelihood ratio test: chi-square = 1509.5, $p < 2.22e-16$

: variabilidad no explicada

: tamaño de efecto

: si $p < .05$, el modelo explica cantidad significativa de variabilidad

Regresión Logística Multinomial

Método I.B (-duración): Resultados

	/b/				/d/			
		95% CI for odds ratio				95% CI for odds ratio		
	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)
(intercept)	0.22(0.56)	0.42	1.24	3.70	0.75(0.55)	0.72	2.12	6.18
int_v1	-0.11(0.01)***	0.88	0.90	0.92	-0.10(0.01)***	0.88	0.90	0.93
int_app	0.08(0.01)***	1.06	1.09	1.11	0.05(0.01)***	1.04	1.06	1.08
int_v2	-0.01(0.01)	0.97	0.99	1.02	0.00(0.01)	0.98	1.00	1.02
f1_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)**	1.00	1.00	1.00
f1_app	0.01(0.00)***	1.01	1.01	1.01	0.00(0.00)***	1.00	1.00	1.00
f1_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_app	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00

Frecuencias (n = 4948): /g/ = 0.33347, /b/ = 0.35509, y /d/ = 0.31144. Estimados con intervalos de confianza de un 95% y *odds ratios*. * $p < .05$, ** $p < .01$, *** $p < .001$.

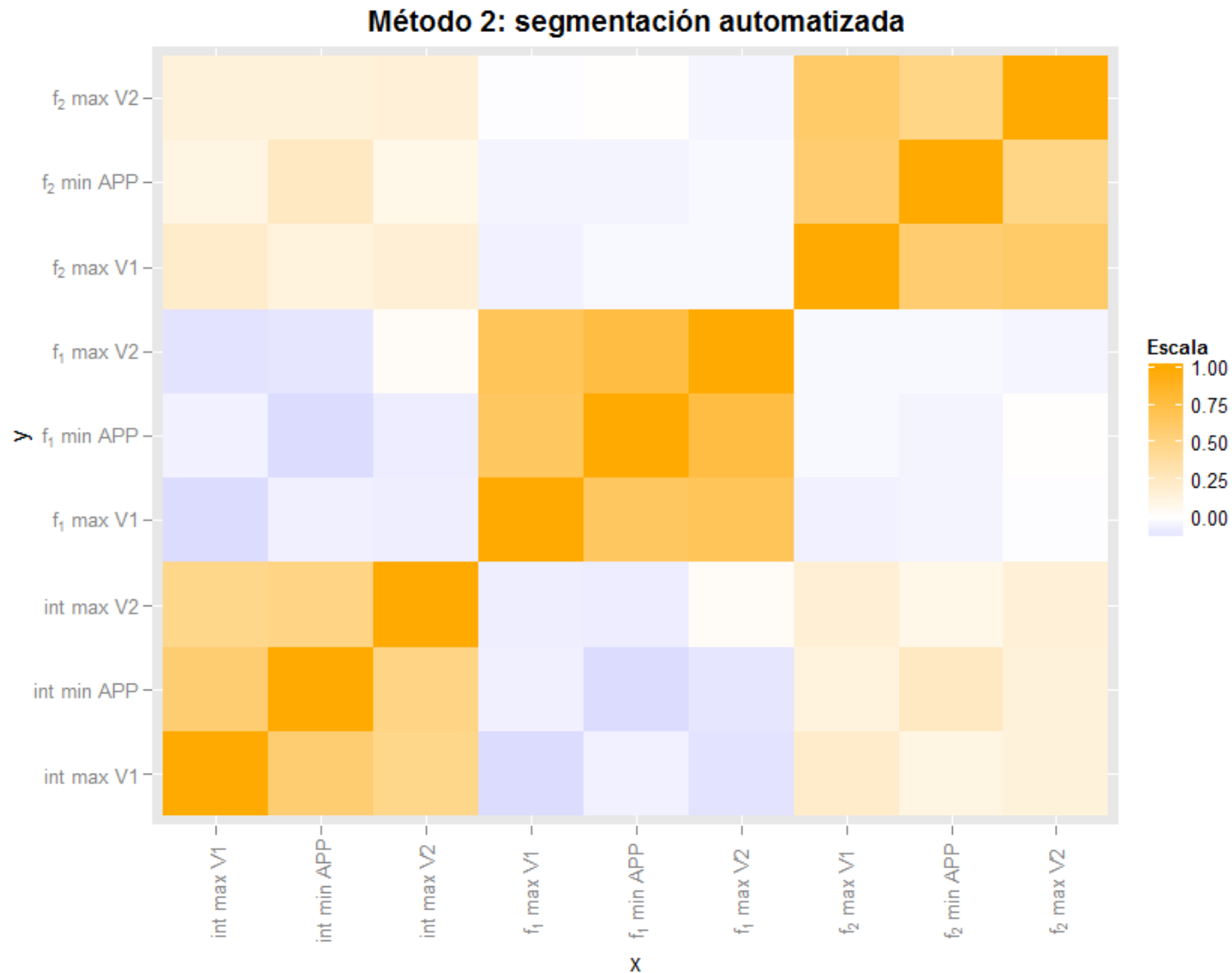
Log-Likelihood: -4743.6

McFadden R²: 0.12622

Likelihood ratio test: chi-square = 1370.4, $p < 2.22e-16$

Regresión Logística Multinomial

Método II: evaluación de multicolinearidad



Regresión Logística Multinomial

Método II: Resultados

	/b/				/d/			
	Estimate (SE)	95% CI for odds ratio			Estimate (SE)	95% CI for odds ratio		
		Lower (2.5%)	Odds Ratio	Upper (97.5)		Lower (2.5%)	Odds Ratio	Upper (97.5)
(intercept)	0.74(0.57)	0.69	2.09	6.33	0.88(0.56)	0.80	2.41	7.24
int_max_v1	-0.12(0.01)***	0.87	0.89	0.91	-0.09(0.01)***	0.89	0.91	0.93
int_min_app	0.08(0.01)***	1.06	1.08	1.10	0.04(0.01)***	1.03	1.04	1.06
int_max_v2	0.00(0.01)	0.97	1.00	1.02	0.00(0.01)	0.98	1.00	1.03
f1_max_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)**	1.00	1.00	1.00
f1_min_app	0.00(0.00)***	1.00	1.00	1.01	0.00(0.00)***	1.00	1.00	1.00
f1_max_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_max_v1	0.00(0.00)	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_min_app	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00
f2_max_v2	0.00(0.00)***	1.00	1.00	1.00	0.00(0.00)***	1.00	1.00	1.00

Frecuencias (n = 4948): /g/ = 0.33347, /b/ = 0.35509, y /d/ = 0.31144. Estimados con intervalos de confianza de un 95% y *odds ratios*. * $p < .05$, ** $p < .01$, *** $p < .001$.

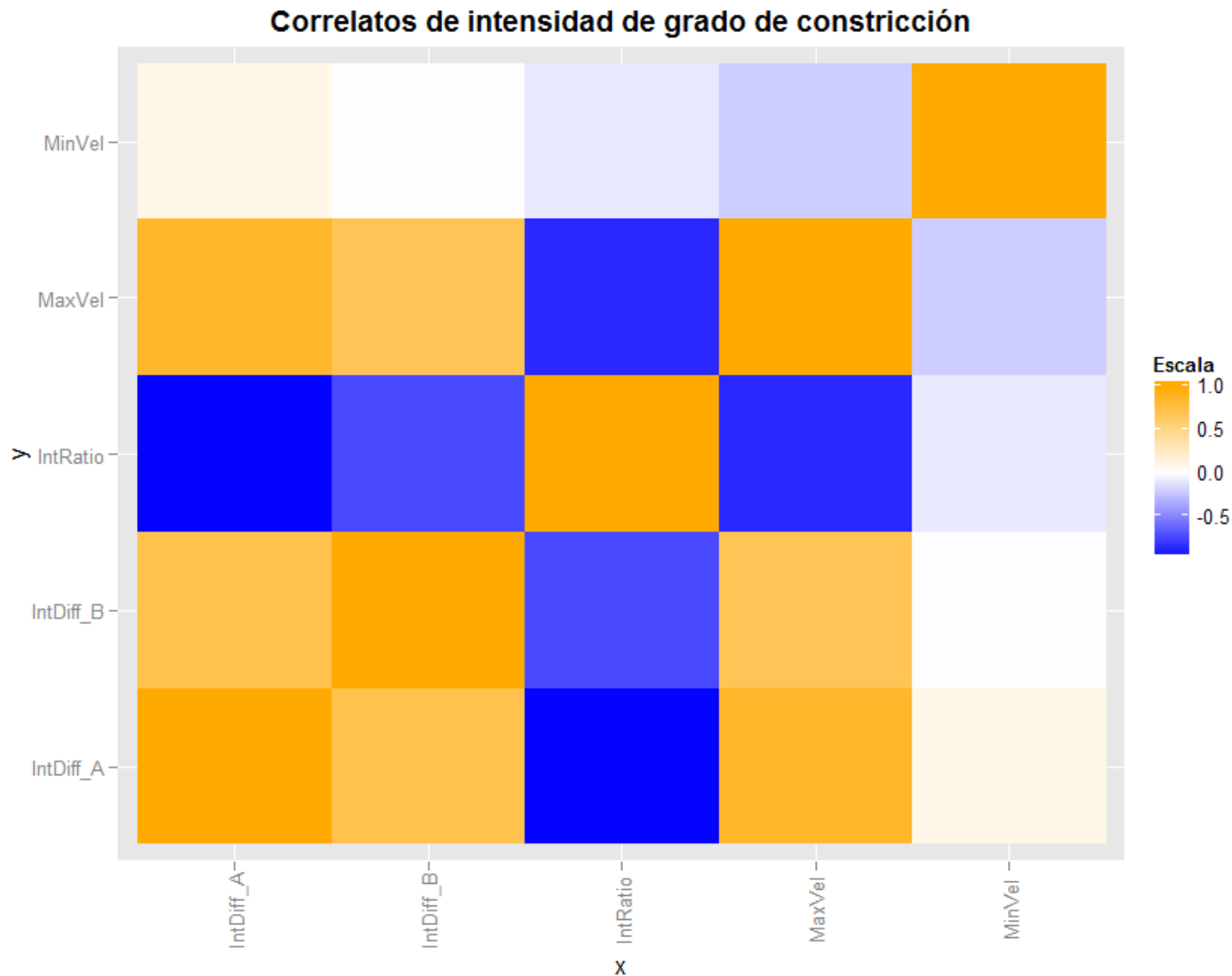
Log-Likelihood: -4891.7

McFadden R²: 0.098942

Likelihood ratio test: chi-square = 1074.3, $p < 2.22e-16$

Regresión Logística Multinomial

Mediciones de intensidad: multicolinearidad



Regresión Logística Multinomial

Mediciones de intensidad: Resultados

	/b/				/d/			
		95% CI for odds ratio				95% CI for odds ratio		
	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)	Estimate (SE)	Lower (2.5%)	Odds Ratio	Upper (97.5)
(intercept)	2.80(3.12)	0.04	16.52	7508.81	3.37(3.07)	0.07	29.09	11937.34
IntRatio	-2.12(3.12)	0.00	0.12	54.17	-2.96(3.07)	0.00	0.05	21.12
IntDiff_A	-0.14(0.06)*	0.78	0.87	0.97	-0.19(0.05)***	0.74	0.83	0.92
IntDiff_B	-0.10(0.01)***	0.89	0.91	0.93	-0.08(0.01)***	0.90	0.92	0.94
MaxVel	3.18(0.43)***	10.48	24.12	55.50	4.53(0.42)***	40.63	92.94	212.59
MinVel	1.66(0.70)*	1.33	5.25	20.78	4.29(0.79)***	15.67	73.26	342.44

Frecuencias (n = 4948): /g/ = 0.33347, /b/ = 0.35509, y /d/ = 0.31144. Estimados con intervalos de confianza de un 95% y *odds ratios*. * $p < .05$, ** $p < .01$, *** $p < .001$.

Log-Likelihood: -5273.9

McFadden R²: 0.028541

Likelihood ratio test: chi-square = 309.89, $p < 2.22e-16$

Conclusiones A

Mediciones acústicas de distintos métodos

Los valores acústicos obtenidos mediante la segmentación manual y en los puntos máximos y mínimo de intensidad (detectados automatizadamente) son **igualmente capaces** de predecir los valores de la categoría fonema.

La elección de uno u otro método debe obedecer, entonces, a otro tipo de consideraciones:

Método I – Segmentación manual. **Ventajas:** ecológicamente válido (se mide una unidad lingüística), incluye valores de duración. **Desventajas:** se ingresa variabilidad a través de errores de segmentación; consume mucho tiempo.

Método II – Identificación automatizada. **Ventajas:** puede programarse y en consecuencia puede realizarse muy rápidamente; altamente objetivo. **Desventajas:** no se mide una unidad lingüística; no permite medir duración propiamente tal.

Conclusiones B

Mediciones basadas en intensidad

La **proporción de intensidad (IntRatio)** no es capaz de predecir la **categoría fonema** (cf. Parrell, 2010). Los mejores predictores son diferencia de intensidad B (**IntDiff_B**) y velocidad máxima de cambio (**MaxVel**), que muestran mejores valores de significancia para /b/.

De todas maneras, las mediciones de diferencia de intensidad A (**IntDiff_A**) y velocidad mínima de cambio (**MinVel**) también son capaces de predecir con significatividad estadística tanto /b/ como /d/.



Referencias

- Barlow, J. A. (2003). The stop-spirant alternation in Spanish: Converging evidence for a fortition account. *Southwest Journal of Linguistics*, 22(1), 51–86.
- Boersma, P. y David W. (2014). *Praat: doing phonetics by computer (Versión 5.4)* [Programa computacional]. URL: <http://www.praat.org/>
- Carrasco, P., Hualde, J. I., y Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, 69(3), 149–79.
- Cepeda, G. (1991). *Las consonantes de Valdivia*. Valdivia: Imprenta America Ltda.
- Colantoni, L., y Marinescu, I. (2010). The scope of stop weakening in Argentine Spanish. En M. Ortega-Llebaria (Ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 100–114). Austin: Cascadilla Proceedings Project.
- Cole, J., Hualde, J. I., e Iskarous, K. (1999). Effects of prosodic and segmental context on/g/-lenition in Spanish. En O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of the Fourth International Linguistics and Phonetics Conference* (pp. 575–589). Thesalónica.
- Field, A., Miles, J. y Field, Z. (2013). *Discovering Statistics Using R*. Sage: Gran Bretaña.
- Figueroa, M. y Evans, B. G. (2014). Are acoustic cues sufficient for the identification and discrimination of Spanish approximants? Evidence from a production study. En *LabPhon 14 – The 14th Conference on Laboratory Phonology*. Tokio.
- Hualde, J. I., Shosted, R., y Scarpace, D. (2011). Acoustics and articulation of Spanish /d/ spirantization. En *Proceedings of the International Congress of Phonetic Sciences XVII* (pp. 906–909). Hong Kong.
- Hualde, J. I., Simonet, M., Shosted, R., y Nadeu, M. (2010). Quantifying Iberian spirantization: acoustics and articulation. En *LSRL 40* (pp. 26–28). Seattle.
- Kingston, J. (2008). Lenition. En L. Colantoni y J. Steele (Eds.), *3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 1–31). Toronto: Cascadilla Press.
- Labov, W. (1983). *Modelos Sociolingüísticos*. Madrid: Cátedra.

Referencias

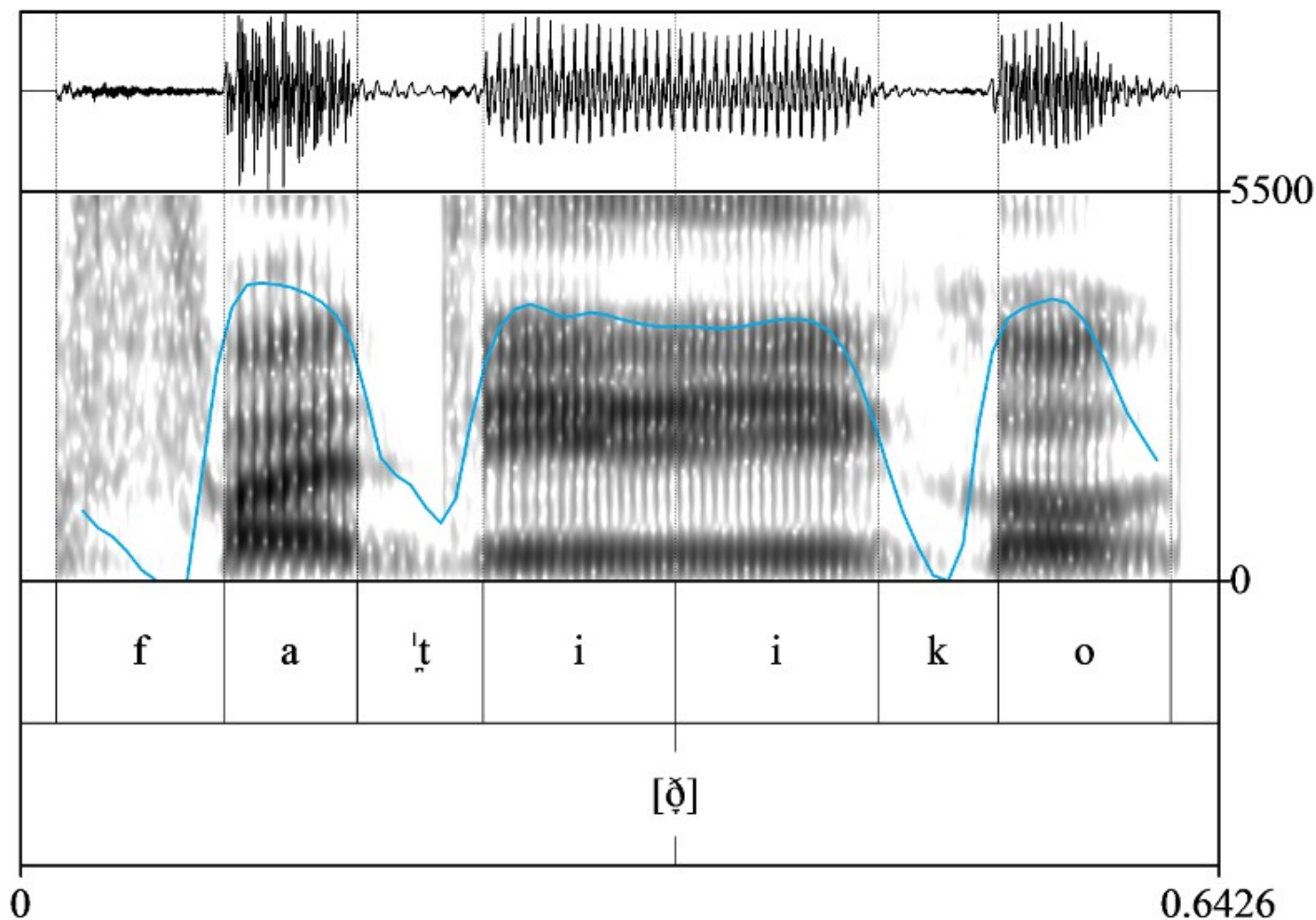
- Martínez Celdrán, E. (1984). Cantidad e intensidad en los sonidos obstruyentes del castellano: hacia una caracterización acústica de los sonidos aproximantes. *Estudios de Fonética Experimental*, 1, 73–129.
- Martínez Celdrán, E. (2013). Caracterización acústica de las aproximantes espirantes en español. *Estudios de Fonética Experimental*, XXII, 11–35.
- Martínez-Celdrán, E., y Regueira, X. L. (2008). Spirant approximants in Galician. *Journal of the International Phonetic Association*, 38(01).
- Ortega-Llebaria, M. (2003). Effects of Phonetic and Inventory Constraints in the Spirantization of Intervocalic Voiced Stops: Comparing two Different Measurements of Energy Change. En *15th ICPHS Barcelona* (pp. 2817–2820). Barcelona.
- Parrell, B. (2010). Articulation from acoustics: estimating constriction degree from the acoustic signal. En *2nd Pan-American/Iberian Meeting on Acoustics*. Cancún, México.
- Parrell, B. (2011). Dynamical account of how /b, d, g/ differ from /p, t, k/ in Spanish: Evidence from labials. *Laboratory Phonology*, 2(2), 423–449.
- Pérez, H. (2007). Estudio de la variación estilística de la serie /bdg/ en posición intervocálica en el habla de los noticieros de la televisión chilena. *Estudios de Fonética Experimental*, XVI, 227–259.
- R Development Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Simonet, M., Hualde, J., y Nadeu, M. (2012). Lenition of /d/ in spontaneous Spanish and Catalan. En *Interspeech*. Portland, Oregon.
- Turk, A., Nakai, S., y Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. *Methods in Empirical Prosody Research*, 3, 1–28.

Anexos

Otros materiales de relativo interés y de cuestionable relevancia.

Aproximantes de /b, d, g/

Ejemplo de elisión



Palabra “fatídico”, grabada en una tarea de lectura de listas de palabras. El contorno de intensidad se muestra como una línea continua (escala 30-65 dB).

Regresión Logística Multinomial

Todos los modelos

```
mlogit(formula = phoneme_code ~ 1 | duration_v1 + duration_app +  
  duration_v2 + intensity_50_v1 + intensity_50_app + intensity_50_v2  
  + formant_f1_50_v1 + formant_f1_50_app + formant_f1_50_v2 +  
  formant_f2_50_v1 + formant_f2_50_app + formant_f2_50_v2, data =  
  data_a_wide, reflevel = "g", method = "nr", print.level = 0)
```

```
mlogit(formula = phoneme_code ~ 1 | intensity_50_v1 +  
  intensity_50_app + intensity_50_v2 + formant_f1_50_v1 +  
  formant_f1_50_app + formant_f1_50_v2 + formant_f2_50_v1 +  
  formant_f2_50_app + formant_f2_50_v2, data = data_b_wide, reflevel  
  = "g", method = "nr",  
  print.level = 0)
```

```
mlogit(formula = phoneme_code ~ 1 | intensity_at_max_v1 +  
  intensity_at_min_app + intensity_at_max_v2 +  
  formant_f1_at_max_int_v1 + formant_f1_at_min_int_app +  
  formant_f1_at_max_int_v2 + formant_f2_at_max_int_v1 +  
  formant_f2_at_min_int_app + formant_f2_at_max_int_v2, data =  
  data_c_wide, reflevel = "g", method = "nr", print.level = 0)
```

```
mlogit(formula = phoneme_code ~ 1 | IntRatio + IntDiff_A +  
  IntDiff_B + MaxVel + MinVel, data = data_d_wide, reflevel = "g",  
  method = "nr", print.level = 0)
```