

Género discursivo y prosodia en Las Palmas

comparativa PRESEEA-AMERESCO

Raúl González; Adrián Cabedo

Universitat de València

Table of contents I

- 1 Introducción y objetivos
- 2 Las Palmas de Gran Canaria
- 3 Modelos entonativos utilizados
- 4 La prosodia en el español canario
- 5 Metodología
- 6 Resultados principales
- 7 Conclusiones
- 8 Referencias

Section 1

Introducción y objetivos

Introducción y objetivos

La **prosodia** es un componente fundamental de la variación lingüística y de la **organización del discurso**, especialmente sensible a factores contextuales y sociales. Aunque la **entonación del español canario** ha sido descrita previamente, existen todavía **pocos estudios comparativos** que analicen de forma sistemática el efecto del **género discursivo** a partir de corpus orales de habla espontánea. Esta comunicación examina la **variación prosódica en la ciudad de Las Palmas** mediante la comparación entre **entrevistas semidirigidas** y **conversaciones coloquiales**, con el objetivo de evaluar **cómo el género discursivo condiciona los patrones prosódicos** y su distribución.

Comparar la prosodia del español de Las Palmas en dos **géneros discursivos** contrastivos: entrevistas semidirigidas (PRESEEA) y conversaciones coloquiales (Ameresco).

- **Caracterizar las diferencias prosódicas** entre géneros a partir de medidas acústicas (F0, intensidad, duración) y del **inventario de patrones entonativos** (TOBI, AMH).
- **Identificar los tonos y configuraciones entonativas más frecuentes** en cada género, evaluando su distribución absoluta y relativa.
- **Evaluar la capacidad discriminativa de la prosodia** mediante modelos de clasificación (Random Forest); qué variables contribuyen más a distinguir entrevista y conversación.

Section 2

Las Palmas de Gran Canaria

Las Palmas de Gran Canaria (I)



Figure 1: Las Palmas

Las Palmas de Gran Canaria (II)

- Situada en el noreste de la isla de Gran Canaria (Islas Canarias, España).
- Aproximadamente a 150 km al oeste de la costa africana en el Océano Atlántico.
- Fundada el 24 de junio de 1478.
- Aproximadamente 380–384 mil habitantes (INE/estimaciones recientes, 2024–2025), ciudad más poblada de las Islas Canarias.
- Con área metropolitana superan los 600–700 mil habitantes incluyendo municipios cercanos.
- Capital de la isla de Gran Canaria y de la provincia de Las Palmas.

Section 3

Modelos entonativos utilizados

Sistema AMH (Análisis melódico del habla [Cantero y Font 2007])

Análisis acústico-perceptivo del contorno completo

Estructura del grupo fónico:

- ① **Anacrusis** (opcional): hasta +40% antes del primer pico
- ② **Primer pico**: primera vocal tónica
- ③ **Cuerpo**: declinación entre primer pico y última tónica
- ④ **Inflexión final (tonema)**: segmento más saliente

Medición:

Cambios porcentuales entre centros vocálicos (normaliza diferencias individuales)

Patrones MAS principales

Clasificación en 3 categorías:

Neutros

- Patrón I: descenso moderado
- Patrón V: suspensión

Interrogativos

- Patrón II: ascenso >70%
- Patrón III: ascenso 40-60%

Enfáticos

- Patrón VII-XII: diversos tipos
- Patrón IX: descenso >30%

Total: 16 patrones descritos en la literatura

El Modelo ToBI (Tones and Break Indices)

- ToBI es un sistema de consenso creado para etiquetar emisiones de habla, marcando eventos entonativos fonológicamente contrastivos basados en el modelo métrico-autosegmental (Hualde & Prieto, 2015; Sp-TOBI, Estebas Vilaplana & Prieto, 2008).
- Se fundamenta en que la prominencia (acentos tonales) y tonos de frontera son los dos aspectos clave de los sistemas entonativos (Hualde & Prieto, 2015).
- **Estructura de niveles:** el sistema establece cuatro capas de etiquetado alineadas con la señal de voz: palabras, tonos, índices de ruptura (break indices) e información miscelánea (Hualde & Prieto, 2015).
- **Nivel tonal:** describe los contornos como secuencias de unidades tonales distintivas, representadas mediante objetivos Altos (H) y Bajos (L) y sus combinaciones (Hualde & Prieto, 2015, Estebas Vilaplana & Prieto, 2008).
- **Índices de Ruptura:** utiliza índices numéricos para indicar los grados de disyunción o separación prosódica entre las palabras (Hualde & Prieto, 2015).

Section 4

La prosodia en el español canario

La prosodia en el español canario

El español de Canarias se caracteriza por una prosodia que lo vincula estrechamente con las hablas del Caribe y el sur peninsular, distanciándose de los patrones estándar del centro de España (Madrid) o México. A continuación, se presenta una comparativa de los tonos y configuraciones **Sp_ToBI** más frecuentes en la variedad canaria

Relación del español canario con otras variedades

- **Vínculo atlántico:** El español de Canarias comparte con las variedades caribeñas el rasgo de la **ausencia de declinación** o el uso de “**upstep**” (elevación tonal) en el acento final de las interrogativas para diferenciarlas de las declarativas.
- **Alineación tonal:** La preferencia por los L^*+H) en posición prenuclear es una de las marcas de identidad más fuertes que agrupan a Canarias con el Caribe, diferenciándolos del sustrato común peninsular.
- **Interrogación circunfleja:** Mientras que en la mayor parte de España la pregunta “sube” (final ascendente), en Canarias la pregunta “baja” tras un pico alto (**circunflejo**), una característica que a menudo suena similar al habla de Buenos Aires o Caracas para los hablantes de otras regiones.

Referencias generales

- Cabrera Abreu, M., & Vizcaíno Ortega, F. (2010). Canarian Spanish intonation. En P. Prieto & P. Roseano (Eds.), Transcription of intonation of the Spanish language (pp. 87–121). Lincom Europa.
- Díaz Cabrera, I. C. (2013). Contribución al atlas prosódico de Canarias (AMPER-Can): declarativas e interrogativas de La Gomera (Islas Canarias) [Tesis doctoral]. Universidad de La Laguna.
- Dorta, J. (2007). La entonación canaria y su relación con las variedades caribeñas. En Temas de dialectología (pp. 141-175). Instituto de Estudios Canarios.
- Dorta, J. (Ed.). (2013). Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela. La Página Ediciones.
- Dorta, J. (Ed.). (2018). La entonación declarativa e interrogativa en cinco zonas fronterizas del español: Canarias, Cuba, Venezuela, Colombia y San Antonio de Texas. Peter Lang.
- Dorta, J., & Díaz, C. (2018). Proximidad y distancia prosódica desde el punto de vista contrastivo entre Canarias y Venezuela. *Dialectología*, (21), 27-60.

Tonos y Patrones: Variedad Canaria vs. Otras Variedades

Contexto / Función	Patrón Canario (ToBI)	Similitudes Dialectales	Diferencias Dialectales
Acento Prenuclear (Declarativas)	L*+H (ascenso muy tardío).	Muy similar al español caribeño , donde el ascenso se confina totalmente a la sílaba postónica.	En el español peninsular y otras variedades, el patrón común es L+<H* (pico desplazado pero con inicio en la tónica)
Preguntas Sí/No (No marcadas)	L%* o L%* (Contorno circunflejo o descendente).	Compartido con el Caribe, Buenos Aires y el noroeste de España (Cantabria/Bilbao).	Se opone radicalmente al Madrid y México , donde el patrón neutral es ascendente (H%*)
Preguntas Sí/No (Marcadas/Sesgadas)	H%* (ascenso final).	Es el patrón neutral en Madrid, México, Chile y la zona andina de Ecuador	En Canarias, este patrón no es neutro; indica contraexpectativa o incredulidad .
Preguntas con Pronombre (Wh-)	Caída final desde un pico nuclear prominente.	Similar al estándar del inglés y a la mayoría de los dialectos del español (Madrid, Caribe, Buenos Aires).	En Madrid, también pueden terminar en ascenso (H%*) para indicar interés o cortesía.
Declarativas (Neutrales)	L%* (interpolación caída).	Común en Madrid, Cantabria, Ecuador y México	El Caribe a veces presenta una caída nuclear específica desde la pretónica (L%*).

Section 5

Metodología

Pipeline de procesamiento (I): transcripción y alineación

Paso 1: Transcripción automática

- Herramienta: **Whisper** (OpenAI)
- Conversión de audio a texto con alta precisión
- Resultado: transcripción ortográfica completa

Paso 2: Alineación forzada

- Herramienta: **Montreal Forced Aligner (MFA)**
- Alineación palabra por palabra y fonema por fonema
- Generación de archivos TextGrid con tres niveles:
 - Phones (fonemas individuales)
 - Words (palabras)
 - Intonation Phrases (frases entonativas)

Paso 3: Análisis acústico con PRAAT

Extracción de parámetros por vocal y por frase:

- **F0**: frecuencia fundamental (Hz)
- **Intensidad**: energía acústica (dB)
- **Duración**: longitud temporal (ms)
- **Rango tonal**: diferencia entre F0 máximo y mínimo

Pipeline de procesamiento (III): anotación prosódica

Paso 4: Etiquetado automático con Oralstats

- Aplicación Shiny desarrollada en R
- Implementa sistema similar a **Eti_ToBI** (basado en reglas)
- Versión beta: cubre patrones básicos

Etiquetado dual:

- 1 **Patrones MAS**: 16 tipos (PI, PII, PIII, PIVa, PIVb, PV, etc.)
- 2 **Patrones ToBI**: acentos tonales y tonos de frontera

Umbrales configurables:

Semitonos (ToBI) y porcentajes (MAS)

Variables extraídas del corpus

- **Unidad (phones_output):** observación a nivel de *segmento/fono* dentro de una palabra.
 - Claves: word_id, archivo, hablante, palabra, fono, tiempo_inicio, tiempo_fin, duracion.
 - Prosodia: f0_media, f0_min, f0_max, int_media, cuantiles q1_*...q4_*.
 - Entonación: tobi_accent, tobi_full, ip_id, ip_seccion.
- **Unidad (words_output):** observación a nivel de *palabra*.
 - Claves: word_id, archivo, hablante, palabra, tiempo_inicio, tiempo_fin, duracion, ip_id.
 - Prosodia agregada por palabra: f0_media, int_media, cuantiles q1_*...q4_*.
- **Género (genero):** conversacion si archivo contiene LP; entrevista si contiene LASP.

Métodos estadísticos empleados

1. Estadística descriptiva

- Tablas de contingencia (frecuencias y porcentajes)
- Distribución de patrones por género discursivo

2. Pruebas de chi-cuadrado (χ^2)

- Evalúa asociación entre variables categóricas
- Análisis de residuos (Pearson) para identificar patrones específicos
- Visualización mediante correlation plots

3. Random Forest (clasificación supervisada)

- 500 árboles de decisión
- Importancia de variables: Mean Decrease Accuracy

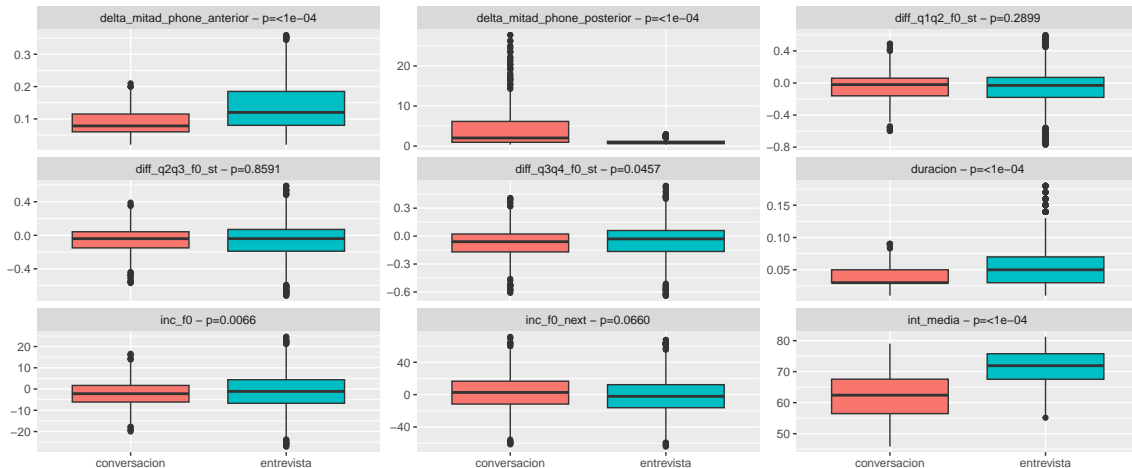
Características generales de la base de datos

Género	Nº de hablantes	Nº de archivos	Nº de vocales analizadas	Nº de tonemas (llanas)	Nº de palabras	Duración total (h:min)
Conversación	6	2	3 818	266	3 981	0:57
Entrevista	2	18	13 456	696	11 920	2:43
General	8	20	17 274	962	15 901	3:41

Section 6

Resultados principales

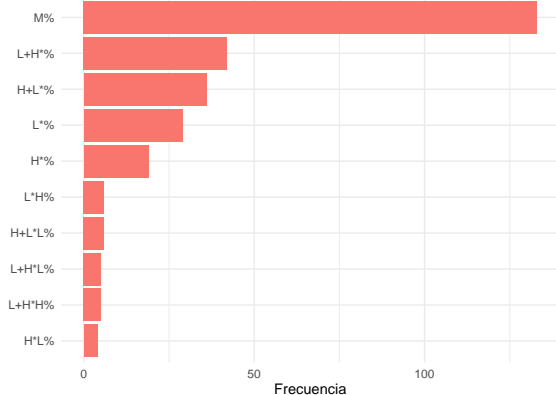
Diagramas de caja del tonema



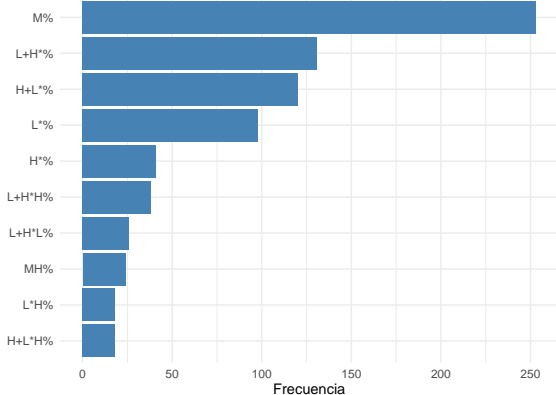
Tonos más frecuentes por género

Frecuencia de tonos ToBI (Top 10 por género)

Top 10 ToBI – Conversación



Top 10 ToBI – Entrevista

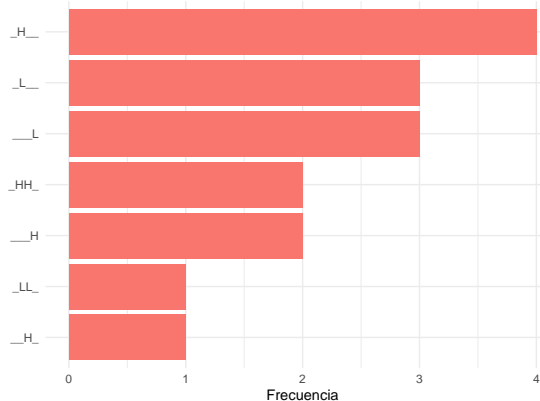


Inflexiones internas en vocal y géneros discursivos (II)

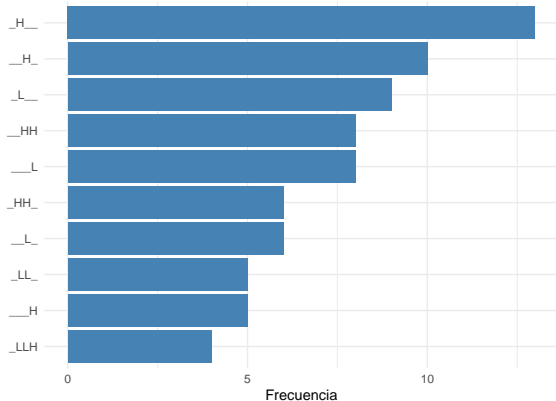
Solo casos con alguna inflexión interna (se excluyen vocales sin inflexión interna)

Frecuencia de inflexiones internas en la vocal (Top 10 por género)

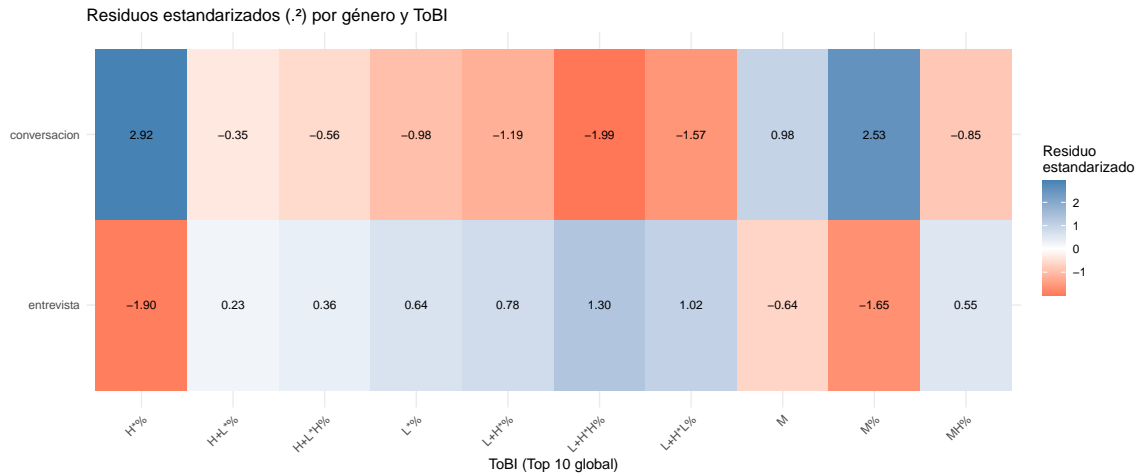
Top 10 ToBI – Conversación



Top 10 ToBI – Entrevista



Asociación tonos y géneros discursivos



Clasificación con randomForest

	entrevista	conversacion	class.error
entrevista	575	53	0.0843949
conversacion	140	99	0.5857741

	Real	Pred_entrevista	Pred_conversacion	Class.error
1 entrevista		575	53	0.0843949
2 conversacion		140	99	0.5857741

Importancia de variables

Importancia de variables (Mean Decrease Accuracy)

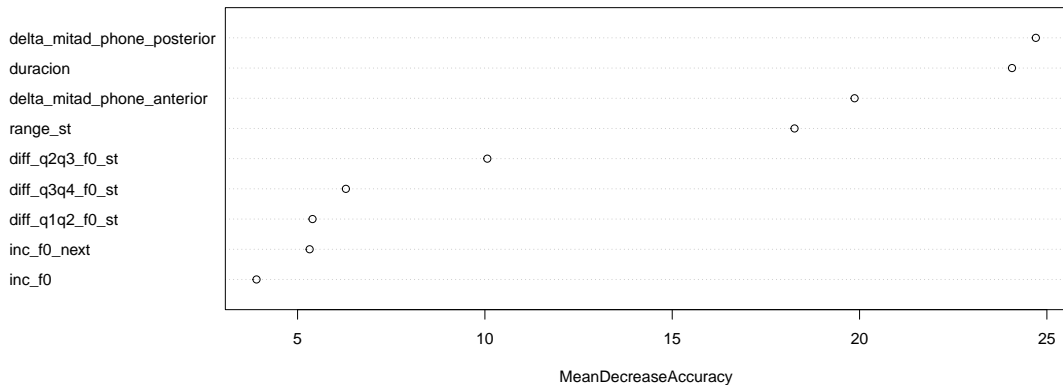


Gráfico M%

IP (conversacion) con tonema que termina en: M%

conversacion – sabes lo cuando te vas a una terraza a tomar algo y hay el un como una

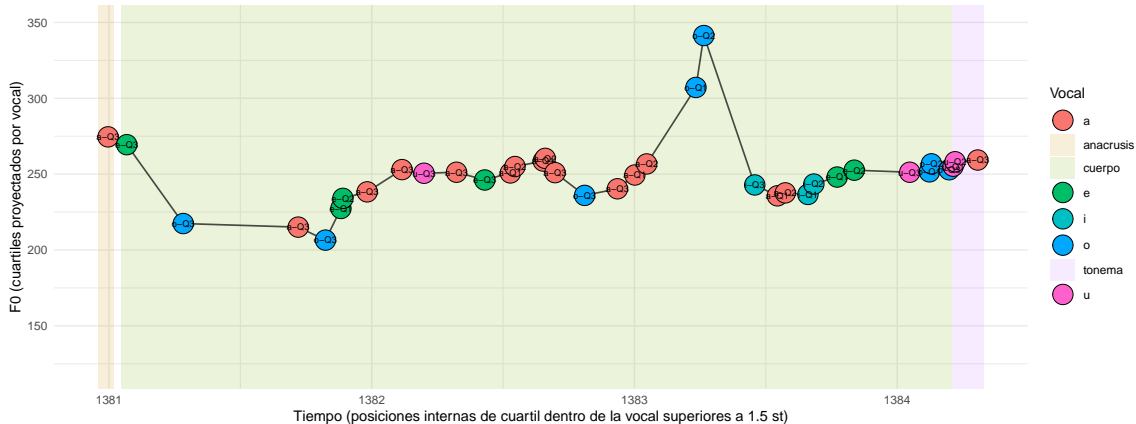


Gráfico L+H%

IP (entrevista) con tonema que termina en: L+H*H%

entrevista – y bueno que es una casa nueva y tal que se estaba construyendo

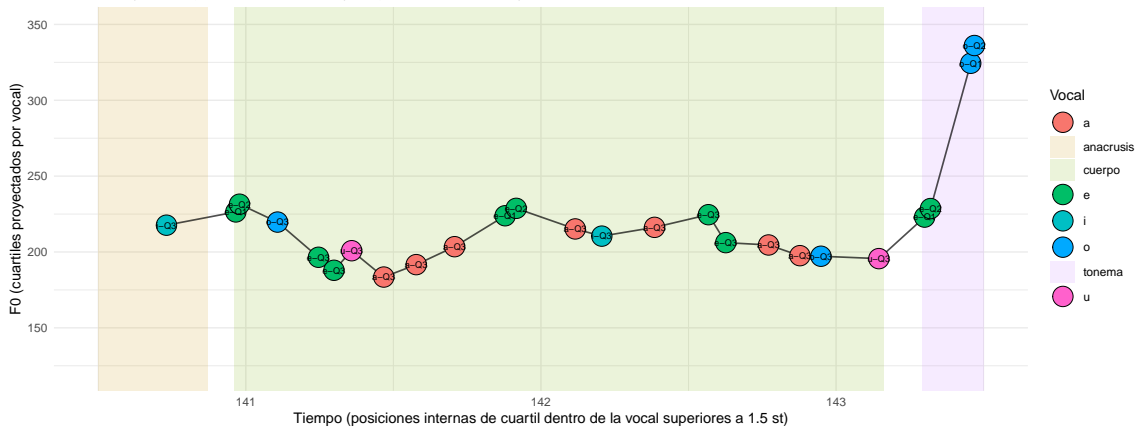


Gráfico H+L*%

IP (entrevista) con tonema que termina en: H+L*%

entrevista – cuando antiguamente los eran habituales entre los niños

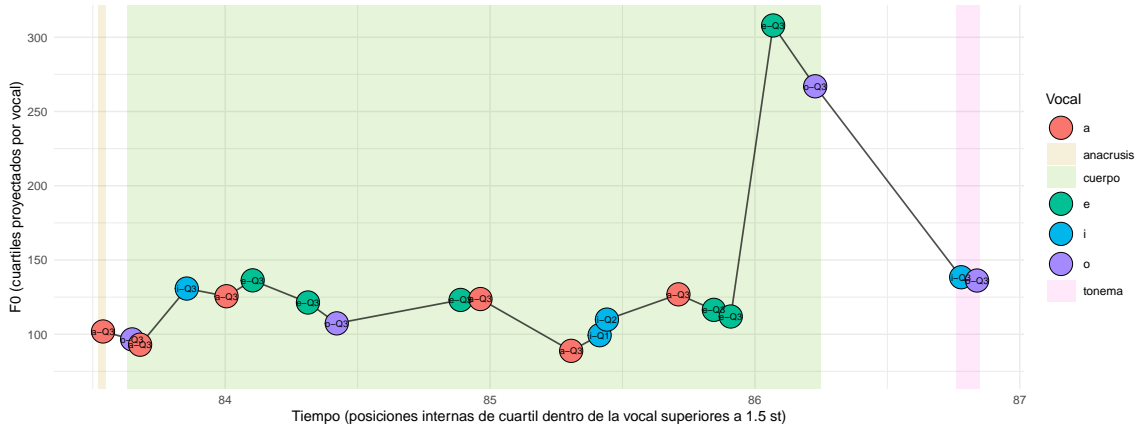
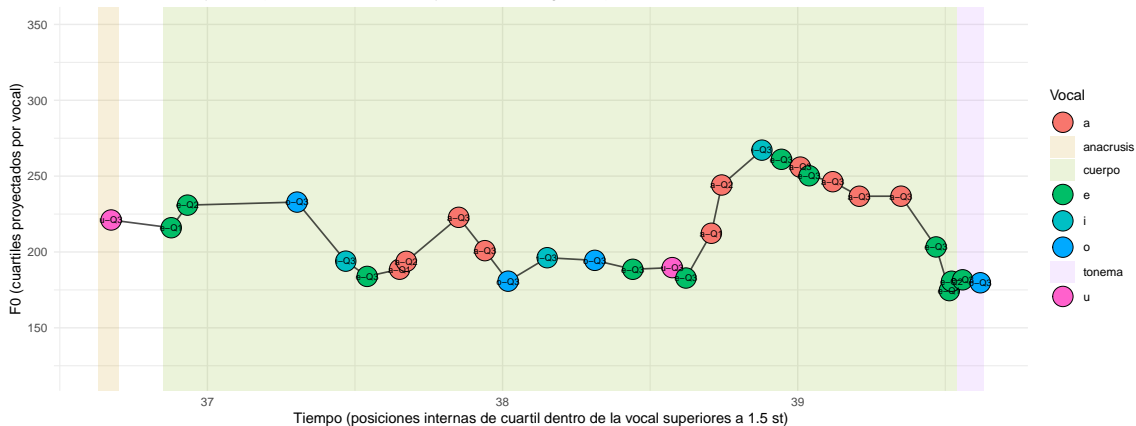


Gráfico L%

IP (entrevista) con tonema que termina en: L*%

entrevista – sus remos y demás para los niños de su edad y de la edad más grande que ellos



Section 7

Conclusiones

Conclusiones

- **Diferencias por género en el tonema.** En el tonema, el género discursivo se asocia significativamente con **inc_f0** ($p = 0.0066$), **duración** ($p < 1e-04$), **intensidad media** ($p < 1e-04$) y valores temporales de transición/centro vocálico (**mitad_phone_anterior/posterior**, $p < 1e-04$). En cambio, algunas microvariaciones internas no discriminan (p. ej., **diff_q2q3_f0_st**: $p = 0.8591$).
- Inventario ToBI resultante: **predominio de M** y alta presencia de bitonales. Frecuencias globales: M 4559, L+H* 2854, H+L* 2613, L* 1632, H* 766. La dominancia de “_____” en TOBI_interno (8426) sugiere que, bajo umbral 1.5 st, la mayor parte de vocales no muestra inflexión interna marcada.
- Clasificación: buena en entrevista, débil en conversación. El Random Forest acierta entrevista con error bajo (~ 0.084), pero conversación con error alto (~ 0.586), coherente con el desbalance (conversación $\sim 0:57$ vs entrevista $\sim 2:43$) y/o mayor heterogeneidad de conversación.

Futuras investigaciones

- Ampliar la base de datos con más hablantes y variedades dialectales.
- Incluir análisis de otros géneros discursivos.
- Observar las características sociolingüísticas de los hablantes.
- Integrar análisis de otros niveles prosódicos (e.g., ritmo, pausas...).
- Explorar técnicas avanzadas de machine learning para clasificación y predicción.
- Realizar estudios comparativos con otras variedades del español y lenguas relacionadas.

Section 8

Referencias

Referencias

- Cabrera Abreu, M., & Vizcaíno Ortega, F. (2010). Canarian Spanish intonation. En P. Prieto & P. Roseano (Eds.), Transcription of intonation of the Spanish language (pp. 87–121). Lincom Europa.
- Díaz Cabrera, I. C. (2013). Contribución al atlas prosódico de Canarias (AMPER-Can): declarativas e interrogativas de La Gomera (Islas Canarias) [Tesis doctoral]. Universidad de La Laguna.
- Dorta, J. (2007). La entonación canaria y su relación con las variedades caribeñas. En Temas de dialectología (pp. 141-175). Instituto de Estudios Canarios.
- Dorta, J. (Ed.). (2013). Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela. La Página Ediciones.
- Dorta, J. (Ed.). (2018). La entonación declarativa e interrogativa en cinco zonas fronterizas del español: Canarias, Cuba, Venezuela, Colombia y San Antonio de Texas. Peter Lang.
- Dorta, J., & Díaz, C. (2018). Proximidad y distancia prosódica desde el punto de vista contrastivo entre Canarias y Venezuela. *Dialectología*, (21), 27-60.