

Uso de Oralstats para la caracterización prosódica de hablantes en distintos géneros discursivos

Adrián Cabedo Nebot

Universitat de València grupo Val.Es.Co.
adrian.cabedo@uv.es ORCID: <https://orcid.org/0000-0002-3881-9308>

Enviado: 28/09/2021; Aceptado: 16/11/2021; Publicado en línea: 28/06/2022

Cómo citar este artículo / Citation: Adrián Cabedo Nebot (2021). Uso de Oralstats para la caracterización prosódica de hablantes en distintos géneros discursivos. *Loquens*, 8(1-2), e079. <https://doi.org/10.3989/loquens.2021.e079>.

RESUMEN: Este artículo analiza 437 grupos entonativos, 2777 palabras y 12520 fonos de cuatro políticos españoles en cuatro géneros discursivos distintos (tertulia, rueda de prensa, mitin y entrevista). A partir de preceptos metodológicos procedentes de la Lingüística Forense, la Sociolingüística, la Fonética y la Lingüística Computacional, se ha desarrollado un conjunto de *scripts* en R, llamado Oralstats, que ha permitido realizar un análisis multivariable, dinámico e interactivo de los datos acústicos y textuales de los grupos entonativos, palabras y alófonos recogidos con PRAAT. El estudio exploratorio se basa en la visualización de las variables mediante árboles de decisiones, mapas de calor y diagramas de caja. En cuanto a los resultados generales, prácticamente todos los hablantes se comportan de manera similar por género discursivo; sin embargo, algunas marcas fónicas, como intensidad o tonos bajos en géneros confrontacionales, en los que se esperaba encontrar lo contrario, sí han permitido caracterizar a alguno de los hablantes analizados. En conclusión, este artículo sugiere vías computacionales para el análisis mecánico de datos fónicos y plantea la búsqueda de marcas lingüísticas individuales que combinen lo fónico con otros factores lingüísticos (por ejemplo, valores léxicos o morfosintácticos).

Palabras clave: idiolecto, géneros discursivos, prosodia, Oralstats

ABSTRACT: *Using Oralstats for prosodic characterisation of speakers in different discourse genres.* - This paper presents the prosodic and morphosyntactic analysis of 437 intonational phrases, 2777 words and 12520 phones registered from the speech of four Spanish politicians in four different discourse genres (TV gathering, press conference, political rally and interview). Based on methods from forensic linguistics, sociolinguistics, phonetics and computational linguistics, a set of R scripts called Oralstats has been developed, which allows to carry out a multivariate, dynamic and interactive analysis of acoustic and textual data aligned with PRAAT. By doing so, a new computational approach is suggested for finding individual forensic marks, combining both phonetic and linguistic factors, such as lexical or morphosyntactic values. Data exploration comes from the visualization of prosodic and morphosyntactic variables collected through decision trees, heatmaps and boxplots. The results show how speakers behave in a significantly similar way when it comes to differentiating themselves by genre; however, in confrontational genres some unexpected phonic marks were found, such as low intensity or low pitch, (the expected behavior being the opposite), thus enabling us to single out some of the speakers.

Keywords: idiolect, discourse genre, prosody, Oralstats

1. INTRODUCCIÓN

El objetivo de este trabajo es doble: por un lado, presentamos un trabajo exploratorio y descriptivo sobre el comportamiento prosódico y morfosintáctico de un mismo hablante en diferentes géneros discursivos orales. Por otro lado, y quizá de manera más relevante, exponemos una metodología computacional, basada en transformación de datos y en su posterior visualización, para poder realizar una caracterización lingüística amplia de hablantes. Esta caracterización, que utiliza criterios prosódicos como base, también extiende su alcance a otros factores, como el uso de bigramas, trigramas, la frecuencia de categorías morfosintácticas, la posición de algunas formas léxicas en el discurso, etc. Así pues, para poder acceder al primer objetivo, el de observar la adaptación prosódica de los sujetos a los géneros discursivos, hemos desarrollado una serie de scripts en R (R Core Team, 2020), bajo el nombre de *Oralstats* (Cabedo 2021), que, en general, permiten realizar un estudio exploratorio avanzado sobre características discursivas de los hablantes.

Algunos estudios han sugerido la viabilidad de este tipo de metodología cuantitativa para analizar variedades discursivas y géneros orales, como lo hace Berez (Berez, 2011, p. 231) para el Athna, una lengua del sur de Alaska:

Finally, on a methodological note, quantitative analysis can contribute meaningfully to our understanding of complex perceptual phenomena in real-time speech. Discourse-level prosodic units are emergent, and unlike much of phonetics and segmental phonology, the variables of which they are comprised are often below the level of the conscious awareness of speakers and listeners.

De manera inicial, podemos señalar que los hablantes modifican su manera de hablar según el contexto discursivo específico en el que se encuentran (Espinosa y Dabrowski, 2019). Habitualmente, el control de la elocución viene determinado por varios factores, como la toma de turno predeterminada (en una entrevista, por ejemplo) o el carácter monológico o dialógico de la intervención (Briz, 2019). Los contextos interactivos síncronos favorecen una expresión mucho menos controlada, sobre todo en situaciones donde es posible encontrar solapamientos de habla entre los hablantes; por el contrario, en formatos de intercambio menos inmediato, como en un debate parlamentario, el control sobre lo dicho es mucho mayor y puede tener en algunos hablantes una repercusión en su prosodia, precisamente en la modificación del tono (Romero Nieto, 2018).

Tengamos en cuenta, no obstante, que la manera de hablar de un ser humano, más allá del control elocutivo marcado por los parámetros del género discursivo, puede tener constantes relativamente estables que definen lo que conocemos por *idiolecto* (Bloch, 1948) o, una de las acepciones más frecuentes, el estilo discursivo de una persona. Este estilo, como hemos comentado previamente, es susceptible de ser modificado según las características

contextuales, pero cabe la posibilidad de plantearse si hay elementos que permanecen y que son emitidos de manera regular por el hablante (Espinosa y Dabrowski, 2019). De esta manera, lo que parte de un análisis estilístico del habla de una persona deviene en una especie de estudio forense e identificativo de ese ser humano.

Si extendemos la idea previamente expuesta, existen situaciones, y no solo en el ámbito forense, en los que hay muestras de habla que no son fáciles de caracterizar desde un punto de vista acústico, bien por las características deficientes de la grabación, que puede estar distorsionada, bien porque el hablante ha intentado camuflar su voz. En estos casos, bastante habituales (Cicres i Bosch, 2011), la identificación forense es mucho más complicada, pero todavía resta material presente en la transcripción que puede ser analizado y que no tiene por qué ser desestimado, ya que puede formar parte de esas constantes discursivas del sujeto (formas léxicas o gramaticales reiteradas, posiciones comunes de algunas palabras, duración amplia o breve de unidades fónicas...).

Por lo tanto, a partir de la selección de muestras de habla procedentes de cuatro políticos españoles, es objetivo de este estudio observar el comportamiento lingüístico de cada uno de ellos en oposición al resto, por una parte, pero también en oposición a sí mismos cuando participan en géneros discursivos distintos. Con la finalidad de facilitar esa tarea, se ha diseñado un sistema de análisis computacional en R (R Core Team, 2020), llamado *Oralstats*, que puede ser ampliado y perfeccionado en el futuro para estudiar una variación idiolectal mayor o incluir un número mayor de registros o unidades de análisis.

2. SOBRE GÉNEROS, VARIACIÓN IDIOLECTAL Y PROSODIA

Los géneros discursivos son constructos cognitivos que se manifiestan en distintas realizaciones textuales (Biber y Conrad, 2009). Se trata, por tanto, de moldes o arquitecturas textuales que promueven la utilización de recursos verbales y prosódicos de un modo particular. Los elementos lingüísticos concretos que se utilizan dependen a su vez del mayor o menor control sobre el discurso emitido (Briz, 2019; Briz y Albelda, 2013). Son géneros discursivos, por tanto, la entrevista, la conversación, el mitin político, el monólogo de humor, la tertulia discursiva, etc. (Cabedo, 2009; Biber y Conrad, 2009; Calvi, 2010).

Estos géneros incluyen características lingüísticas comunes que los definen como estereotipos comunicativos (Caalen-Haumont, 1993; McMenamin, 1993). Las características de estos géneros son puestas en funcionamiento por los hablantes, que los usan según el contexto comunicativo en el que participan. En este sentido, el hablante cumple con requisitos sociales en su manera de comunicarse, si bien es posible plantearse también la posibilidad de que haya un núcleo común lingüístico independiente de cualquier manifestación discursiva.

No se trata de un acercamiento nuevo para otras lenguas, como el francés, donde ya existen estudios que han encontrado comportamientos fónicos y morfosintácticos

distintos para géneros como narraciones conversacionales, noticiarios televisivos, conferencias y discursos políticos (Degand y Simon, 2009, pp. 94–95):

the different genres in our sample obey to different production rules. Conversational narration is the more informal genre in our comparison, characterized by a lower degree of complexity that results in a low articulation ratio and short syntactic units (...) Political address, on the other hand, demonstrates very formal and well-prepared scripted speech. Here the opposite tendency can be observed. The speakers deliver their message in a solemn way, speaking slowly (...). Radio news speech is a typical written-to-be-spoken type of discourse where we encounter most congruent units, reading being supported by typographic punctuation units. Conferences, then, show a high degree of mixed BDUs, a combination of syntax-bound and intonation-bound units

En español, el estudio fónico de géneros discursivos ha sido abordado por diferentes trabajos de investigación (Cabedo, 2007, 2009; Estellés y Albelda, 2014). En estudios como el de Estellés y Albelda (2014), en el que se relacionan marcas evidenciales, prosodia y géneros discursivos, se detecta amplia variedad asociada al comportamiento prosódico del debate parlamentario; esa misma variedad es causante de constituir diferentes grupos explicativos en el seno del mismo género discursivo:

there are instances of impolite meanings containing zero, one, two or three prosodic marks, whereas the other groups show a certain degree of coherence in the number of marked parameters exhibited. The underlying reason for this odd behaviour has to do with the fact that all impolite examples in our corpus are ascribed to one and the same genre, namely parliamentary debate. This genre is characterized, in the Spanish tradition, by its confrontational nature; interventions in the Spanish Parliament are meant to be impolite

En estos casos, la individualidad del hablante y su comunicación idiosincrásica trae a la mente conceptos como el *idiolecto*. Precisamente, una de las primeras definiciones sobre el término la aporta Bloch (1948, p. 7), que utiliza el término de manera explícita para referirse principalmente a características fonéticas del ser humano, aunque no exclusivamente, y señala lo siguiente:

The totality of possible utterances of one speaker at one time in using a language to interact with one other speaker is an idiolect... Our definition implies a) that an idiolect is peculiar to one speaker, (b) that a given speaker may have different idiolects at successive stages of his career, and (c) that he may have two or more different idiolects at the same time. (Bloch, 1948, p. 7)

Aunque la definición de Bloch (1948) implica en general la adaptación contextual de los requisitos lingüísticos del sujeto, el idiolecto se ha consolidado tradicionalmente en Sociolingüística como el habla individual de una persona y se ha percibido siempre como un elemento que debe superarse para poder caracterizar comunidades de habla. “Thus, while idiolects (or the speech of individuals) considered in isolation might seem random, the speech community as a whole behaved regularly.” (Romaine, 2000, p. 67); o en la misma línea, “within sociolinguistics, language variation patterns of individuals were not often studied except when aggregated to approximate and reveal the language variation patterns of the speech community” (Hazen, 2007, p. 73). De hecho, no es nada extraño que algunos de los mayores expertos en lingüística forense procedan del ámbito de la Sociolingüística (Coulthard, 2013, 2014, 2007; Coulthard y Johnson, 2010). Curiosamente, el objetivo sociolingüístico, como decíamos previamente, es prescindir de las marcas heterogéneas de los propios hablantes para buscar el núcleo común: “Findings such as these are vital for the task of identifying the unity of a language from among the enormous diversity of dialects (and idiolects) that we find in the world (...) this kind of orderly heterogeneity” (Guy, 2011, p. 183). Así pues, tanto la Lingüística Forense como el presente estudio pretenden mantener el foco en esa heterogeneidad discursiva, con la finalidad última de caracterizar el habla particular de un sujeto y, si cabe, poder identificar sus *marcas* constantes en el discurso hablado, sea cual sea el género en el que se inserte.

De esta manera, el concepto de *idiolecto* y de *autor* son de facto concomitantes. Determinar las características lingüísticas de un autor pueden tener, por tanto, una doble finalidad: por un lado, se trata de un análisis estilístico del modo de hablar de un ser humano; por otro lado, si esas características son estables o, incluso, si son poco habituales frente al resto de expresiones verbales en una comunidad de habla, podrían servir, en el marco forense, para poder identificar a ese hablante en un momento discursivo concreto. Aunque algunos trabajos señalan la importancia de las características lingüísticas de autor en los textos escritos (Chaski, 2012; Coulthard et al., 2011; McMenamin, 1993), la idea subyacente es que estas marcas lingüísticas pueden extrapolarse a las muestras orales.

Los requisitos que debe cumplir una variable lingüística para erigirse como marca de autoría son distintos según los estudios (Chaski, 2012; Coulthard et al., 2011; Grant, 2008; Grant y Baker, 2001), aunque estos estudios suelen priorizar valores léxicos o morfosintácticos, dado que se centran en textos escritos. En la práctica se han demostrado útiles elementos ortotipográficos, morfológicos, sintácticos, léxicos, pragmáticos y discursivos (Coulthard, 2007; Coulthard y Johnson, 2010; McMenamin, 1993). Las prácticas basadas en la identificación de *marcas* de autoría en los textos escritos han centrado su interés en marcas de distintos niveles lingüísticos, principalmente en las marcas léxicas y, de manera secundaria, en las marcas sintácticas (Spasova, 2009), entre otros tipos de marcas (por ejemplo, n-gramas de caracteres,

puntuación, etc.). Al mismo tiempo, en el marco de los estudios orales, se ha partido también de otros tantos factores, como el estudio de factores como la fonética segmental (Leemann et al., 2018), las pausas llenas (Cicres i Bosch, 2014), etc.; algunos autores incluso postulan la importancia de tomar varios rasgos como conjunto o clúster: “in forensic authorship attribution, the linguist’s task is to identify the writer’s habitual choices and define them as a set” (McMenamin, 2002, p. 158).

En relación con la prosodia de los géneros discursivos, existen ya trabajos que estudian las características prosódicas de diferentes géneros discursivos, aunque en muchas ocasiones se estudia géneros de una mayor preparación discursiva; por ejemplo, hay estudios sobre la radio o las noticias en televisión (de la Mota y Rodero, 2010; Rodero, 2012). También estudios fónicos sobre otros géneros discursivos, como la retransmisión deportiva, el discurso político o la conversación coloquial (Cabedo, 2007, 2009; Hidalgo, 2019). Estudios más recientes sobre la entrevista, la conversación y el monólogo de humor han determinado la importancia del grado de planificación discursiva en el tipo de factor prosódico alterado: “los [géneros] menos planificados presentan una mayor polarización, sobre todo centrada en el descenso de la intensidad y en el aumento de la velocidad de habla” (Cabedo, 2018, pp. 234–235).

3. METODOLOGÍA

La metodología del presente estudio está basada en el análisis cuantitativo de datos lingüísticos extraídos de material sonoro de habla y de sus respectivas transcripciones, así como de los datos prosódicos (f_0 e intensidad) extraídos de PRAAT (Boersma y Weenink, 2021). Si bien el sistema utilizado para la explotación de los datos (Oralstats, desarrollado en la Sección 3.2) permite analizar grandes cantidades de información, funciona mejor con cantidades medias, en una horquilla variable de 1000 a 10000 registros.

Al mismo tiempo, cabe señalar que la perspectiva de análisis es exploratoria; es decir, aunque algunas pruebas realizadas por el sistema arrojan valores de significación estadística, estos deben completarse en la medida de lo posible con un análisis valorativo por parte del investigador (López Serena, 2019). En este sentido, incidimos en el hecho de que, en modo similar a como sucede en algunos estudios de Lingüística Forense (Cicres i Bosch, 2011), los audios analizados en esta investigación proceden de entornos de habla espontánea y eso puede condicionar el comportamiento de algunas variables fonéticas específicas, como valores de f_0 o intensidad. No obstante, el sistema desarrollado para este estudio permite observar posibles datos anómalos o alterados mediante diagramas de caja y procedimientos de estadística descriptiva, así como filtrar casos extremos para las diferentes variables de análisis.

3.1. Datos recogidos

Con finalidad exploratoria, hemos recogido un conjunto de datos procedente de cuatro políticos españoles

(Pablo Casado, Rafael Hernando, Pablo Iglesias e Íñigo Errejón). Todos ellos han participado en eventos discursivos de distinta índole que han sido publicados posteriormente en Youtube. De esta manera, los fragmentos analizados proceden de mítines, tertulias, ruedas de prensa y entrevistas. Normalmente, se ha transcrito una breve porción de estos archivos, alrededor de un minuto o dos minutos por género y hablante.

La elección de figuras de la política española se debe a que precisamente su labor ciudadana motiva que participen en múltiples actos comunicativos, con diferentes requisitos y correspondientes expectativas discursivas. Al mismo tiempo, la voluntad pública de su trabajo, el hecho de querer conectar con la ciudadanía facilita que puedan encontrarse muestras de habla de variada tipología en plataformas audiovisuales como Youtube. Los archivos de vídeo procesados en Youtube suelen proceder de fragmentos extraídos de noticiarios o de las propias páginas web de los partidos políticos. Esto suele asegurar que la calidad del audio sea adecuada, dado que las grabaciones se han recogido mediante micrófono y muchas veces en entorno cerrado, a excepción de los mítines o de algunas ruedas de prensa. Por otro lado, en este estudio se han elegido cuatro géneros discursivos: dos de ellos, las tertulias y las entrevistas, son dialógicos, mientras que los otros dos, los mítines y las ruedas de prensa, son monológicos.

De cada archivo de vídeo se ha extraído un archivo de audio para ser posteriormente transcrito en grupos entonativos mediante el programa ELAN (Max Planck Institute for Psycholinguistics, The Language Archive, 2021). La identificación de los grupos entonativos se ha determinado básicamente mediante el uso de pausas superiores a 300 milisegundos, aunque también se ha prestado atención, en ocasiones, a otros criterios como el reajuste entonativo, siguiendo pautas establecidas previamente en la bibliografía (Cabedo, 2011; Hidalgo, 2019). Con la ayuda de PRAAT (Boersma y Weenink, 2021), las palabras y los diferentes alófonos de esta transcripción manual en grupos entonativos se han alineado de manera semiautomática al tiempo. También con el uso de PRAAT se han recogido, además, los datos acústicos de frecuencia fundamental e intensidad (archivos *PitchTier* e *IntensityTier*).

Así pues, para cada fragmento susceptible de análisis hemos dispuesto de diferente material audiovisual y textual:

1. Vídeo procedente de Youtube.
2. Audio extraído del vídeo.
3. Archivo eaf (del programa ELAN) con la transcripción alineada al audio.
4. Archivo TextGrid con la transcripción alineada al audio en grupos entonativos, palabras y alófonos.
5. Archivo tabulado extraído del archivo eaf. En él se incorporan varias columnas: hablante, tiempo de inicio, tiempo final, duración, texto y nombre del archivo.
6. Archivo Pitchtier de PRAAT con valores de pitch recogidos para cada 0.001 segundos y guardados en formato tabulado sin encabezado.

7. Archivo Intensitytier de PRAAT con valores de pitch recogidos para cada 0.001 segundos y guardados en formato tabulado sin encabezado.

Las frecuencias de los grupos entonativos recogidos y transcritos quedan reflejados en la Tabla 1:

Tabla 1. Grupos entonativos recogidos por géneros discursivos y hablante. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa

sp_g	time_used	time_max	ip	words	words_ip	phon
e_e	43500	67010	27	179	6.63	830
e_m	46400	152790	29	184	6.34	729
e_r	54560	93330	39	201	5.15	923
e_t	44710	181030	39	168	4.31	783
pc_e	52110	81040	22	210	9.55	952
pc_m	42460	1441060	22	144	6.55	648
pc_r	48050	174890	27	161	5.96	768
pc_t	52830	60910	28	226	8.07	913
pi_e	32290	61950	16	122	7.62	572
pi_m	55160	159200	37	137	3.70	679
pi_r	48870	61570	27	147	5.44	721
pi_t	51250	273740	30	179	5.97	817
rh_e	40590	93440	21	125	5.95	593
rh_m	56510	181130	21	181	8.62	834
rh_r	47520	65170	17	156	9.18	678
rh_t	74650	197590	35	257	7.34	1080
total	791460	1441060	437	2777	6.35	12520

En general, tal y como se observa en la Tabla 1, hay un cierto equilibrio entre los grupos entonativos recogidos. Debe tenerse en cuenta que los marcos de tiempo han sido los mismos para todas las muestras de los distintos géneros discursivos; es decir, las frecuencias de la Tabla 1 son los grupos entonativos emitidos por los hablantes durante dos minutos de grabación de cada archivo. De hecho, las propias frecuencias ya señalan posibles características idiosincrásicas de los hablantes. Por ejemplo, Rafael Hernando emite una media de 8.5 a 9 palabras por grupos entonativos para el mitin y la rueda de prensa, muy por encima de los mítines o las ruedas de prensa del resto de políticos. Ello indica, a falta de un análisis más específico, que sus enunciados son mucho más amplios. Curioso es también el dato en la dirección inversa de Pablo Iglesias, con una media de 3.7 palabras para el mitin político; es decir, sus grupos entonativos son habitualmente más breves.

A modo de ejemplificación de lo comentado sobre la duración de los grupos entonativos de Pablo Iglesias y

Rafael Hernando, podemos ver algunos ejemplos, con sus valores de duración medios en la Tabla 2:

Tabla 2. Muestra de los grupos entonativos de los mítines de Pablo Iglesias y de Rafael Hernando. Leyendas: p_i = Pablo Iglesias, r_h = Rafael Hernando; _m = mitin

sp_g	text	duration
pi_m	nunca más	0.92
pi_m	una España sin su gente nunca	1.81
pi_m	nunca más una España sin sus pueblos	2.06
pi_m	nunca más	0.69
pi_m	una españa sin ternura	1.47
pi_m	buenas tardes- madrid	1.24
rh_m	y este es el gran cambio al que algunos se niegan a reconocer	3.24
rh_m	el de pasar de la recesión a la recuperación	2.80
rh_m	en la locomotora del crecimiento	2.16
rh_m	este gran cambio es también el resultado de la suma de historias de superación de millones de españoles	5.49
rh_m	capaz de superar grandes dificultades cuando tiene un buen gobierno con la visión y la ambición	5.20
rh_m	de afrontar con valentía los problemas que es lo que hace nuestro partido en este caso con Mariano Rajoy al frente	6.48

3.2. Descripción general de Oralstats

Con el nombre de Oralstats (Cabedo 2021) designamos un conjunto de scripts desarrollados en R que permiten dos operaciones básicas: en primer lugar, la transformación de datos textuales procedentes de transcripciones (etiquetado morfosintáctico, transcripción de grupos entonativos, palabras o alófonos, etc.) y su combinación con datos de frecuencia fundamental, duración y de intensidad. Todos estos scripts se encuentran a disposición de cualquier investigador de manera gratuita en <https://github.com/acabedo/oralstats>.

Normalmente, las transcripciones proceden de la exportación en texto tabulado realizada desde el programa ELAN (Max Planck Institute for Psycholinguistics, The Language Archive, 2021) y los datos fónicos proceden del programa PRAAT (Boersma y Weenink, 2021). En segundo lugar, Oralstats permite una visualización de los datos en un formato exploratorio mediante gráficos básicos (círculos, diagramas de caja) y, también, mediante la aplicación de algunas técnicas de visualización estadística más avanzada, como mapas de calor o árboles de decisiones.

Las variables que genera el módulo de conversión de Oralstats son las siguientes:

1. Columnas de identificación id para los alófonos, palabras, grupos entonativos, archivos y hablantes.
2. Variables prosódicas: máximo, mínimo, mediana y media de f_0 , así como rango tonal (en Hz y St), mediana y media de intensidad, duración media y duración de las pausas previas y posteriores. También se calcula la diferencia de la unidad fónica analizada con la media de la unidad fónica previa y con la media del hablante. Además, se computa una versión básica y experimental de codificación TOBI y de las diferencias porcentuales según el análisis melódico del habla (Cantero Serena y Font Rotchés, 2009). Para estas últimas operaciones se toman como referencia valores de las vocales (primero, medio y final) y no de las sílabas, como se propone en sistemas de etiquetado fónico automático mucho más elaborados (Elvira-García et al., 2016).
3. Variables morfosintácticas: cantidad de verbos, sustantivos, adjetivos, numerales, adverbios y pronombres. Para ello se realiza un etiquetado morfosintáctico automático mediante la librería UDPipe (Straka, 2018).
4. Variables posicionales: primera y última palabra del grupo entonativo, así como tonemas, estructura acentual de la palabra (aguda o no aguda) y posición de la palabra en el grupo (pre cuerpo, cuerpo o final).

Con Oralstats pueden realizarse diferentes operaciones estadísticas; mientras que algunas proceden de una visualización específica de los datos (gráfico circular, diagrama de caja, mapa de calor...), otras aplican técnicas de discriminación entre grupos, como los árboles de decisiones, la prueba ANOVA o el análisis discriminante. Estas últimas técnicas ofrecen valores de significación estadística, pero nuestro objetivo se articula más como exploratorio que como validatorio. En contextos comunicativos orales, sobre todo con material de habla espontánea, parece más sensato considerar tendencias en los datos y no extrapolar validaciones generales sin precaución o cautela. En todo caso, un valor p por debajo del nivel de significación establecido siempre apuntará en direcciones que deben explorarse de modo cualitativo por parte del investigador. Lo cuantitativo y lo cualitativo deben necesariamente ir de la mano (López Serena, 2019).

Una versión accesible del entorno computacional generado y del material utilizado para este estudio (sin incluir los audios), puede encontrarse en el siguiente enlace: <<https://adrin-cabedo.shinyapps.io/oralstats-generes/>>. Por su parte, los scripts de transformación y visualización del programa se encuentran a disposición general en el siguiente repositorio de Github: <<https://github.com/acabedo/oralstats>>.

Actualmente, las versiones con las que se ha trabajado deben considerarse beta.

4. ANÁLISIS Y RESULTADOS

Esta sección de análisis de las muestras orales recogidas incluye varios subapartados: mapa de calor, árbol de decisiones y diagramas de caja, ya que consideramos que es el orden que debería seguirse para realizar un análisis lingüístico del material utilizado.

En Oralstats la sección de *Heatmap* permite visualizar un gráfico con la media de las variables numéricas que consideremos de interés para ser estudiadas y una variable de agrupación. En la parte inferior del gráfico, así mismo, se añade una tabla con la media, mediana y desviación típica de las variables que se proyectan en el mapa de calor. Al mismo tiempo, la opción de *heatmap* permite establecer una agrupación basada en un análisis de clúster. Por eso mismo, la visualización por color y el análisis de clúster facilitan detectar variables que pueden ser relevantes para ser incorporadas en un modelo de explicación multivariable, como el que se observa en la sección de Árbol de decisiones o análisis discriminante (Sección 4.4), y los diagramas de caja para cada factor o categoría que reflejan la mediana y los cuartiles de una variable numérica. Por tanto, en el análisis con el que se opera en Oralstats se recomienda empezar por técnicas de visualización amplias y posteriormente por técnicas de clasificación más específicas.

Muchas técnicas de visualización estadística, como los árboles de decisiones, tienen en la actualidad muchos seguidores en el ámbito científico. Como señala Gries (2019, p. 644): “Treebased methods have become a welcome alternative for data sets that defy regression-based methods especially in noisy and unbalanced corpus data, and that, in and of itself, is potentially a good thing.” En general, el uso de los árboles de decisiones y de técnicas de visualización similares parece por lo tanto adecuado para explorar y analizar relaciones entre variables de bases de datos orales, donde el equilibrio o balance entre los datos no se cumple siempre con la frecuencia esperada.

4.1. Mapa de calor

En esta sección presentamos dos mapas de calor generados mediante Oralstats para el análisis de los grupos entonativos de los hablantes, tanto cuando se considera a estos hablantes como seres individuales, sin influencia del género discursivo, como cuando son tomados como sujetos discursivos diferentes según el género discursivo en el que participan.

Los análisis de clúster establecen dos grandes grupos con sucesivas bifurcaciones o subgrupos. Los dos primeros oponen los mítines políticos de los cuatro hablantes y las tertulias de todos menos la de Pablo Iglesias frente al resto de géneros, es decir, las entrevistas y las ruedas de prensa. Por su parte, los subgrupos son los siguientes:

1. Entrevista y rueda de prensa de Errejón y rueda de prensa de Pablo Casado
2. Entrevistas de Pablo Iglesias y Pablo Casado.

Figura 1. Mapa de calor por hablante y género con variables prosódicas y morfosintácticas. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa

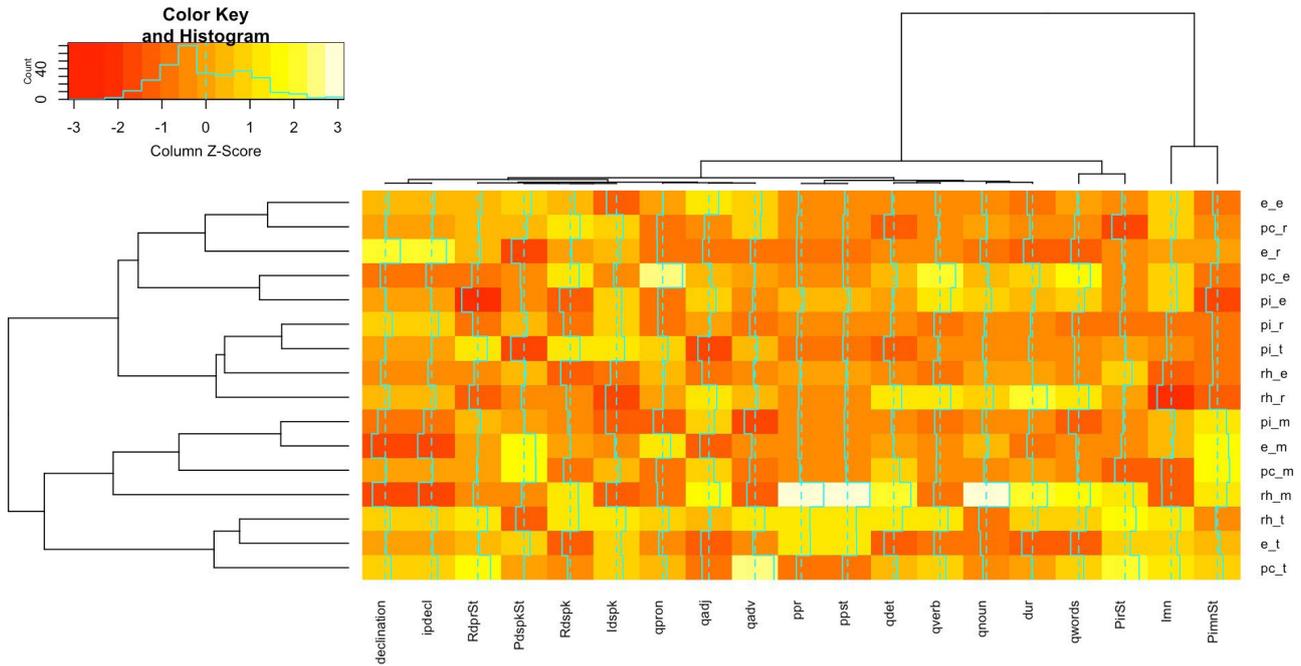
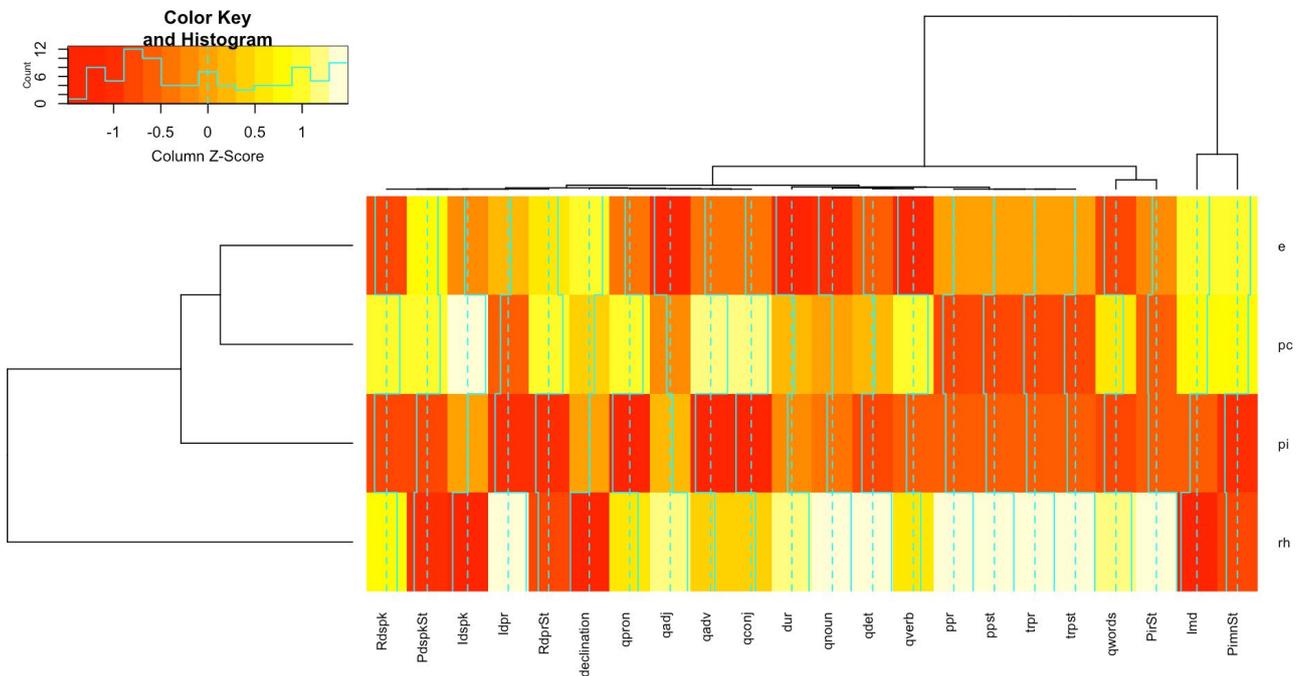


Figura 2. Mapa de calor por hablante con variables prosódicas y morfosintácticas. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa



3. Rueda de prensa y tertulia de Pablo Iglesias y rueda de prensa y entrevista de Rafael Hernando.
4. Mítines de todos los políticos analizados.
5. Tertulias de todos los políticos, menos la de Pablo Iglesias.

En cuanto a la agrupación de las variables más importantes, la intensidad y la media tonal tomada en semitonos, junto con la cantidad de palabras y el rango tonal, también en semitonos, resultan ser las variables más relevantes en la agrupación de los grupos y subgrupos que hemos comentado con anterioridad.

Los resultados de la visualización implican consideraciones totalmente esperables: los cuatro géneros se dividen según el mayor énfasis comunicativo; en el caso de los mítines, por la voluntad de ser enérgico en la comunicación con el público (Degand y Simon, 2009; Estellés y Albelda, 2014), mientras que en las tertulias, debido a su entorno de conflicto, el interviniente tiende a subir tono de voz e intensidad (Cabedo, 2007). En el caso de la entrevista y de la rueda de prensa, tanto la intensidad como las medias tonales son ciertamente inferiores. Como dato anecdótico, y quizá puede percibirse en él una marca idiolectal, llama la atención que Pablo Iglesias, a diferencia de los otros políticos, mantenga un tono de voz y una intensidad más bajas en la tertulia, donde el trasfondo confrontacional parece provocar lo contrario.

Cabe recordar que este análisis preliminar no tiene en cuenta variables categóricas y que, en este caso, variables como la configuración melódica (recogida en sistema TOBI) todavía no se aplica y, por tanto, no podemos conocer en este momento su alcance para generar un modelo adecuado de clasificación.

Con el mapa de calor por hablante, en la Figura 2, sí parecen configurarse singularidades idiolectales, aunque el análisis de clúster, activado por defecto en la prueba, detecta que Pablo Casado e Íñigo Errejón tienen más cercanías en sus discursos que Pablo Iglesias. Pueden verse a continuación casos de Pablo Iglesias extraídos en este caso de la entrevista y del mitin; en ellos se observan valores relativamente bajos, tanto de intensidad como de rango tonal:

Tabla 3. Grupos entonativos de mítines de Pablo Iglesias

phon	Pim-nSt	PirSt	Imn	file
buenos días Pepa	80.13	6.55	69.96	pi_e
yo creo que no debilita al gobierno de Pedro Sánchez la dimisión de Máxim Huerta o la dimisión de la	76.99	9.02	69.53	pi_e
antigua ministra de sanidad al contrario	77.14	5.87	69.85	pi_e
una España sin su gente nunca	88.48	7.54	67.61	pi_m
nunca más una España sin sus pueblos	89.48	6.17	68.77	pi_m
sonreíd	87.25	5.41	66.69	pi_m

Por su parte, Rafael Hernando presenta una mayor cantidad de elementos morfosintácticos en sus enunciados, si bien debemos tener en cuenta la variable de la duración, notablemente más alta que la del resto de políticos. Al mismo tiempo, los valores globales de intensidad y tono son también menores. Este es sin duda un dato curioso, dado que aunque el tono global sea menor, el rango

tonal es mucho mayor que para los otros políticos; los valores tonales amplios suelen implicar una mayor variación en los contornos melódicos empleados y un mayor énfasis en partes concretas del discurso.

4.2. Árbol de decisiones y análisis discriminante

En la sección de árbol de decisiones podemos elegir un conjunto de variables numéricas o categóricas que sirvan de predictores para caracterizar grupos de otra variable de entrada. De esta manera, en esta sección exploremos dos variables de entrada: la variable *archivo* que, en realidad, es la variable que cataloga a los datos por autor y por género; por otro lado, también analizaremos la variable *hablante* que aglutina los datos sin considerar el género discursivo y que, por tanto, permite observar el núcleo común denominador de un sujeto.

En el caso concreto de los dos árboles de decisiones que comentamos en este apartado, hemos usado un conjunto amplio de variables como predictores, aunque, como se vio en la sección del mapa de calor, realmente son dos variables (*media tonal* e *intensidad media*) las más concluyentes en la separación de datos. En este caso, además, hemos podido introducir dos variables categóricas que aportan configuraciones melódicas en sistema TOBI, tanto las anotaciones melódicas de las últimas sílabas tónicas de los enunciados, como los tonos de frontera. En todo caso, debemos recordar que este etiquetado no es tan detallado ni específico como el propuesto por otros autores para otros sistemas de etiquetado automático (Elvira-García et al., 2016).

En los dos árboles de las Figura 3 y 4, advertimos una singularidad que se observa en el eje de abscisas, donde aparecen los nombres abreviados de los hablantes y de los géneros discursivos. Dado que el gráfico es amplio y hay dieciséis categorías, el eje solo incluye algunas de ellas en la visualización (e_e, pc_r y pi_t). El eje de abscisas está ordenado alfabéticamente, por lo que las sucesivas líneas verticales se establecen siguiendo el presente orden: e_e, e_m, e_r, e_t, pc_e, pc_m, pc_r, pc_t, pi_e, pi_m, pi_r, pi_t, rh_e, rh_m, rh_r, rh_t. En general, se trata del nombre del hablante (e por Errejón, pc por Pablo Casado, pi por Pablo Iglesias y rh por Rafael Hernando) seguido del nombre del género discursivos (e por entrevista, m por mitin, r por rueda de prensa y t por tertulia). En la Figura 3, el orden es e, pc, pi y rh.

En el árbol expuesto en la Figura 3, observamos un núcleo amplio de 305 grupos entonativos en el nodo 6, caracterizados por una media tonal entre 80.13 y 90.97 semitonos y una intensidad media superior a 62.12 dB. El problema de esta clasificación, y que luego se contemplará aún mejor en el *análisis discriminante*, señala que la mayor parte de grupos entonativos se aglutinan bajo este marco de tono y de intensidad y, de esta manera, es difícil separarlos por categoría, es decir, discriminar mejor los géneros discursivos y los hablantes.

No obstante, sí hay datos que pueden resultar curiosos y llamativos; por ejemplo, en el nodo 2, con tonos bajos, aparecen unos diez grupos entonativos de la entrevista de

Figura 3. Árbol de decisiones con variables fónicas y morfosintácticas por hablante y género. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa

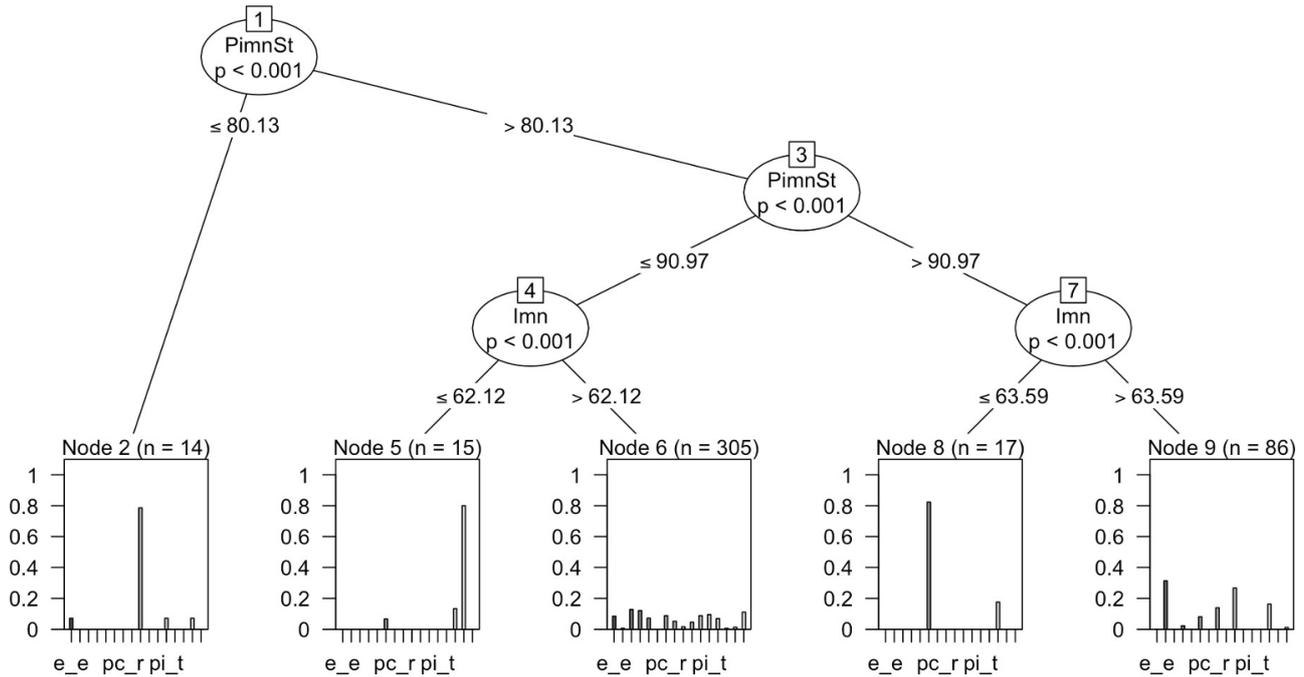
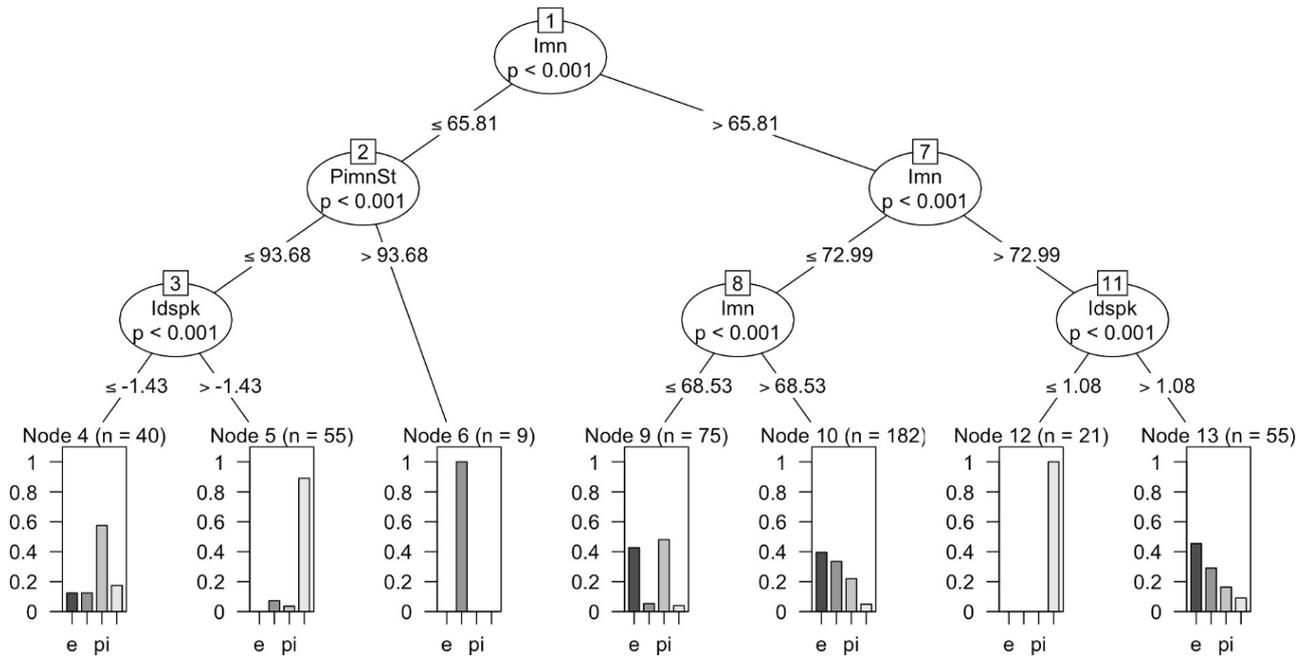


Figura 4. Árbol de decisiones con variables fónicas y morfosintácticas por hablante. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando



Pablo Iglesias; en el nodo 5, con una intensidad relativamente baja, inferior a 62.12 dB, se observan otros tantos grupos entonativos de Rafael Hernando para la rueda de prensa; en el nodo 8, los grupos entonativos presentan un tono alto, superior a 90.97 semitonos pero una intensi-

dad baja, inferior a 63.59 dB, y pertenecen en su mayoría al mitin de Pablo Casado; ello implica un aumento del énfasis tonal sin necesidad de aumentar la energía articuladora de la voz. Finalmente, hay un conjunto amplio de 86 casos, en el nodo 9, que incluyen valores altos de tono

e intensidad. Son sobre todo casos de mítines de Errejón y Pablo Iglesias, pero también de las tertulias de Pablo Casado.

En la Figura 4 se clasifican los hablantes sin atender a los géneros discursivos en los que participan. Los diferentes nodos generados pueden verse a continuación:

Tres variables son las más llamativas en la Figura 4, la *Intensidad media*, la *media de f0* (en semitonos) y la diferencia de intensidad de los grupos entonativos en relación con la media general del hablante. A diferencia de la Figura 3, en este árbol de decisiones los grupos entonativos se subdividen frecuentemente de un modo más repartido. En el nodo 6 se observan 9 casos de Pablo Casado con valores bajos de intensidad pero altos en tono: eso implica énfasis melódico pero energía articuladora disminuida. El nodo 3, y concretamente el nodo 5, con intensidad medio baja y valores tonales también medios, discriminan básicamente casos de Rafael Hernando, casi 50 grupos entonativos en los que la diferencia de intensidad en relación con su media es superior a -1.43 dB. Este factor indica básicamente o subidas de intensidad marcadas o un componente monótono, ya que los valores particulares de los grupos entonativos se asemejan a la media general. Por otro lado, el nodo 4 muestra casos de Pablo Iglesias. En este nodo se distribuyen grupos entonativos en los que la intensidad es más baja que la media del hablante. Es, por tanto, marca dialectal de Pablo Iglesias usar secuencias de habla marcadamente más bajas que sus secuencias habituales.

Siguiendo con la Figura 4, la parte derecha del gráfico distribuye valores más altos de intensidad, superiores a 65.81 dB. El nodo 10 es el más poblado. En él se distribuyen grupos entonativos de todos los políticos a excepción de valores residuales de Rafael Hernando. Son grupos con una intensidad media comprendida entre 68.53 y 72.99 dB. Muchos de los grupos de Rafael Hernando, sin embargo, aparecen en el nodo 12, con una media alta de intensidad superior a 72.99 dB y con una diferencia en relación con la media de 1 dB. Se trata esta de una subida menor, a diferencia de los casos del nodo 13, pertenecientes en su mayoría a Íñigo Errejón y Pablo Casado; en ellos hay una media superior a 72.99 dB y grupos entonativos con una diferencia de más de 1 dB en relación a la media. Este dato implica que algunos grupos entonativos suben marcadamente la intensidad probablemente para ser percibidos con mayor claridad o para disputar el espacio dialógico. Podemos encontrar ejemplos de ambos hablantes

Tabla 4. Grupos entonativos de Pablo Casado e Íñigo Errejón.
Legendas: pc = Pablo Casado, e = Íñigo Errejón

spk	phon	Imn
e	[esto]	76.16
e	[Eduardo]	77.10
e	[hablabas de las elecciones]	75.80
e	[pero cuando multiplicas]	76.75

en la Tabla 4; los corchetes indican que se trata de fragmentos confrontacionales en los que hay solapamiento o lucha por el turno con otro hablante:

En último término, para ampliar los resultados el árbol de decisiones, puede completarse el modelo de clasificación usando las mismas variables en el apartado de *análisis discriminante* de Oralstats. La diferencia es que mientras el árbol de decisiones restringe las variables de predicción a las más concluyentes, en el análisis discriminante podemos observar la importancia de cada variable dentro de la función de discriminación. En el presente caso, interesa estudiar sobre todo el porcentaje de clasificación de las funciones discriminantes generadas. En cuanto a la variable que separa por género y hablante, la clasificación correcta es de solo un 58 %; para la variable que únicamente identifica hablante, la clasificación correcta sube a un 73 %; las variables más priorizadas siguen siendo en ambos casos la media tonal y la media de intensidad.

Son en todo caso modelos de discriminación muy pobres y, seguramente, la información más concluyente que transmiten es que, dentro de la variación observable, hay un núcleo común amplio en cuanto a la configuración discursiva de los hablantes; dicho de otro modo, los hablantes, desde una perspectiva idiolectal, son más parecidos que distintos. Si se comparan a lo largo de géneros discursivos, donde muchas veces las convenciones discursivas están estereotipadas culturalmente, sus semejanzas aumentan y solo algún caso aislado, como el de Pablo Iglesias para las tertulias, establecen marcas de separación más amplias.

4.3. Diagramas de caja

En las secciones anteriores hemos podido observar que dos variables son muy importantes en la discriminación de grupos entonativos por hablante y género discursivo: concretamente, se trata de las variables de media tonal y de intensidad. Por lo tanto, en las siguientes líneas, analizaremos la visualización en forma de diagrama de caja que proporciona Oralstats para estas dos variables. Así mismo, el sistema ofrece también como resultado una prueba de contraste de la varianza (ANOVA) y, en este caso, podemos ver las diferencias significativas entre grupos.

En la Figura 5, los diagramas de caja muestran las medianas de intensidad para los distintos hablantes a lo largo de los cuatro géneros discursivos analizados. En general, la variación es patente, pero destacan los valores más elevados de Rafael Hernando en la tertulia y los más bajos de Pablo Casado en los mítines y de Rafael Hernando en las ruedas de prensa. Pablo Iglesias presenta los valores más bajos de intensidad para las tertulias discursivas, un discurso que, como manifiestan el resto de políticos, precisa de una energía articuladora amplia, sobre todo en secuencias de disputa o mantenimiento del turno de habla. Esta podría considerarse inicialmente una marca de autor, en línea similar a la que indica McMenamin (1993).

En la Figura 6 se observa el diagrama de caja por hablante y género de la variable *media de f0*, tomada en semitonos:

Figura 5. Valores medios de intensidad por hablante y género discursivo. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa

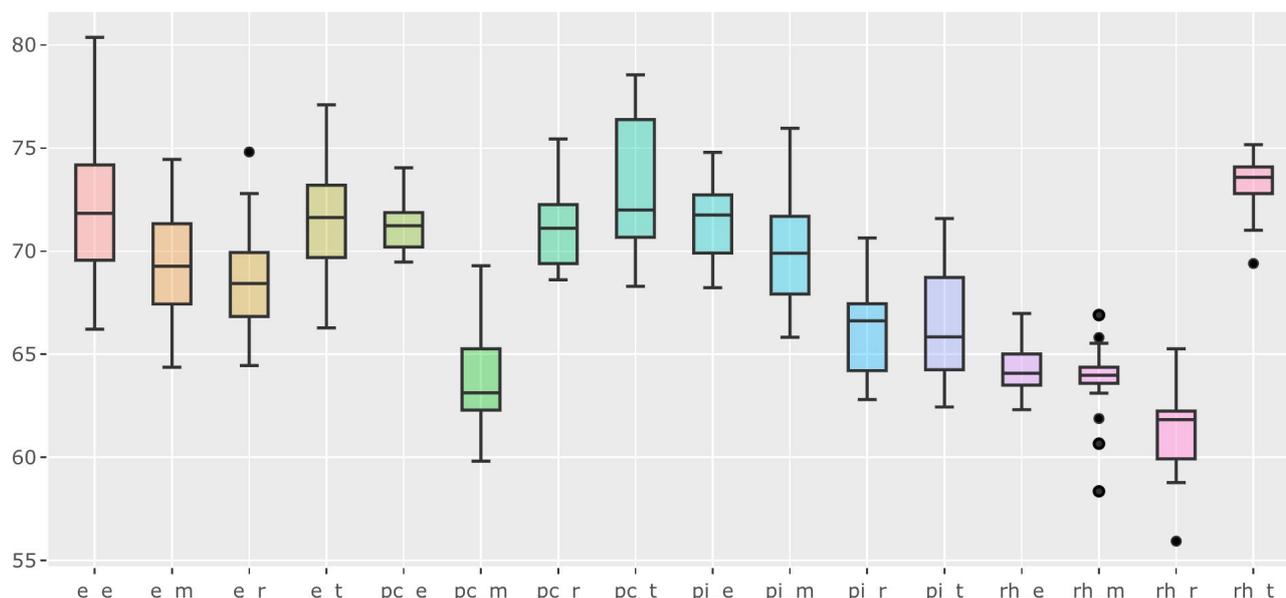
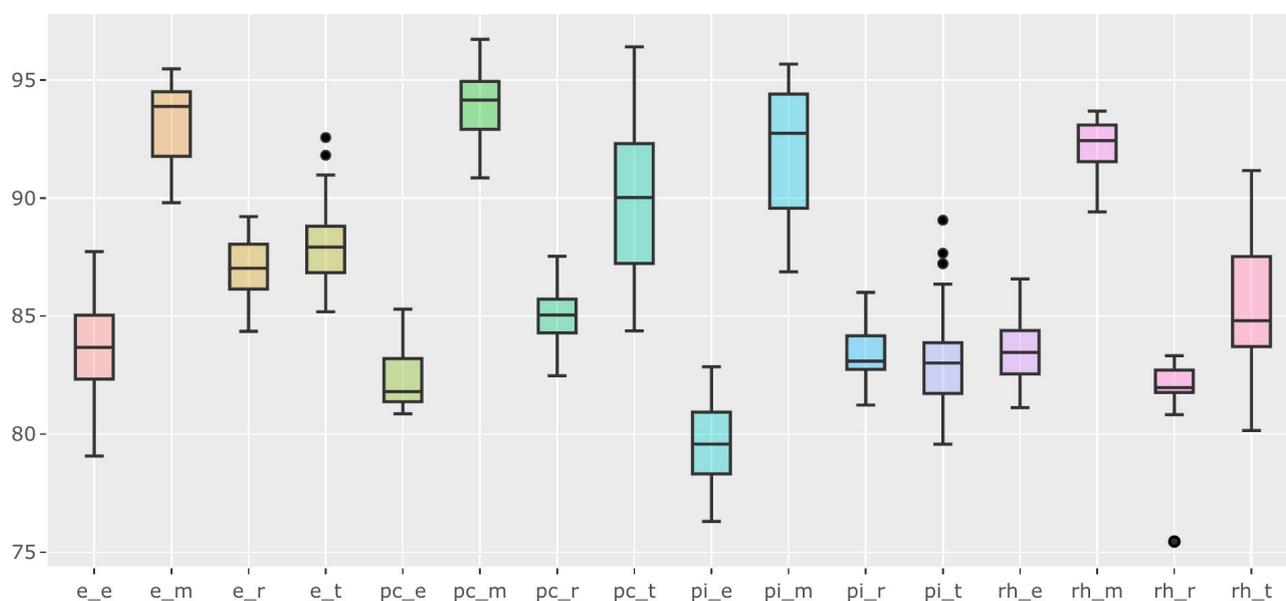


Figura 6. Valores medios de f_0 por hablante y género discursivo. Leyendas: e = Íñigo Errejón, p_c = Pablo Casado, p_i = Pablo Iglesias, r_h = Rafael Hernando; _e = entrevista; _m = mitin; _t = tertulia; _r = rueda de prensa



El gráfico anterior, combinado con los resultados de la prueba ANOVA, sugiere una posible agrupación con un tono medio alto para el grupo entonativo, compuesto por los mítines de todos los políticos y la tertulia de Pablo Casado. En contraste con la Figura 5 anterior, donde el mitin de Pablo Casado, tenía los valores de intensidad más bajos, en este caso hay una compensación inversa mediante el uso de un tono más amplio. Llama la atención el tono más bajo utilizado por Pablo Iglesias en relación con los demás políticos. También hay una zona media compues-

ta por el resto de géneros discursivos, a excepción de las tertulias de Pablo Casado e Íñigo Errejón y la rueda de prensa de este último, que se situarían en unos valores medio altos.

5. CONCLUSIONES Y VALORACIONES FINALES

A partir de todo lo expuesto en los apartados anteriores, podemos elaborar unas conclusiones generales sobre

el trabajo de investigación realizado. En general, se deduce que los hablantes analizados presentan algunas diferencias pero no de una manera excesivamente amplia ni precisa; una de las posibles explicaciones es la brevedad de las muestras, si bien se han tomado muestras breves de habla en un modo similar al que puede encontrarse para otras disciplinas, como la Lingüística Forense (Cicres i Bosch, 2011, 2014; Coulthard et al., 2011); en estas disciplinas las circunstancias pueden propiciar una recogida de audios breves o de audios con una posible malformación acústica producida por un contexto de habla real y espontáneo.

Para poder analizar de manera sistemática, variada y dinámica la base de datos acústica y textual, se ha generado un conjunto de scripts en R (R Core Team, 2020) llamado *Oralstats* que transforma los datos en una primera fase y que, en segundo término, permite visualizar en formato interactivo, mediante la adición o elisión de algunas variables, estos mismos datos en forma de árbol de decisiones, mapas de calor o diagramas de caja. La idea general es que la información de alófonos, palabras y grupos entonativos, procedente de una transcripción alineada en PRAAT (Boersma y Weenink, 2021), pueda servir de entrada a un análisis estadístico mediante la combinación de relaciones entre unidades inferiores y unidades mayores; por ejemplo, la segmentación en alófonos permite tomar datos de las vocales y obtener posteriormente valores de reajuste o subida tonal, mayor o menor duración..., y también otro tipo de catalogaciones, como la transformación en un sistema TOBI básico, siguiendo los pasos de otros modelos computacionales mucho más elaborados y depurados (Elvira-García et al., 2016). La base de datos generada, en el módulo de visualización de *Oralstats*, puede consultarse en el siguiente enlace: <<https://adrin-cabedo.shinyapps.io/oralstatsgenres/>>

La voluntad de este trabajo ha sido explorar, de manera computacional, en primer lugar, y valorativa, en último lugar, el comportamiento discursivo diferenciado de los mismos hablantes en distintos formatos o moldes discursivos. En el marco concreto de este artículo, se han analizado variables fónicas y morfosintácticas del habla de cuatro políticos españoles en muestras sonoras procedentes de cuatro géneros discursivos distintos (la tertulia discursiva, la rueda de prensa, la entrevista y el mitin). Los resultados generales, presentados en la Sección 4 en forma de diferentes técnicas de visualización estadística multivariable, indican que los géneros discursivos se diferencian entre ellos de manera más o menos esperable, es decir, los géneros confrontacionales, como el mitin o la tertulia, presentan valores medios de tono e intensidad más altos que otros géneros donde no se apela tanto a la atención del oyente o donde no hay polémica discursiva.

Sobre los políticos españoles analizados se han podido extraer datos curiosos sobre su correspondiente *idiolecto* (Bloch, 1948); en tal sentido, se apuntan marcas individuales que, no de manera sistemática ni privativa, afectan a algunos hablantes. Por ejemplo, llama la atención cómo Pablo Iglesias se sirve de un tono y una intensidad bajas en géneros como la tertulia, donde el resto de políticos

enfatan estos rasgos. También son relevantes los tonos y la intensidad medio altos de Rafael Hernando que se combinan con una mayor duración de los grupos entonativos y, como consecuencia, por una mayor cantidad de riqueza léxica en su interior, en forma de más cantidad de sustantivos, verbos, etc. Políticos como Íñigo Errejón o Pablo Casado, por su parte, suben la intensidad de manera notable en géneros confrontacionales como la tertulia y suben en ocasiones su media tonal para enfatizar aspectos concretos de su discurso, no necesariamente acompañados de incrementos de intensidad para estos casos.

Así pues, este artículo, a partir de un simple análisis de caso de cuatro hablantes en cuatro géneros discursivos distintos, plantea la necesidad de recurrir a una vertiente cuantitativa en primera instancia, es decir, a una mirada exploratoria sobre los datos de modo dinámico y ágil; posteriormente, esta visualización interactiva permite un estudio cualitativo de los datos.

Las vías de ampliación del presente trabajo pasan precisamente por explorar de qué manera esas variables determinantes o singulares pueden ampliarse y, de este modo, precisar su labor o tarea identificativa. Por ejemplo, consideramos que en el futuro puede ser enriquecedor añadir información posicional en los análisis realizados (lugar de la palabra en el grupo entonativo, posición más común de las subidas o descensos tonales sintomáticos en el nivel de la palabra...).

El uso concreto del lenguaje de programación R incrementa notablemente las posibilidades de construcción y explotación de una plataforma interactiva de consulta y análisis de los datos, tanto en construcción web como en la aplicación de pruebas estadísticas más complejas. El beneficio del sistema, por tanto, es que se trata de un entorno ampliamente escalable y que, en tal sentido, puede crecer en propiedades y recursos para realizar análisis discursivos más completos y sofisticados.

5. REFERENCES

- Berez, A. L. (2011). Prosody as a genre- distinguishing feature in ahtna: A quantitative approach. *Functions of Language*, 18(2), 210–236.
- Biber, D. y Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press. <https://doi.org/10.1075/jhp.13.1.07leh>
- Bloch, B. (1948). A set of postulates for phonemic analysis. *Language*, 24(1), 3. <https://doi.org/10.2307/410284>
- Boersma, P. y Weenink, D. (2021). PRAAT (Version 6.1.53) [Computer software]. <http://www.praat.org/>
- Briz, A. (2019). Español coloquial. En E. Ridruejo (Ed.), *Manual de lingüística española* (pp. 614– 637). Berlín: De Gruyter. <https://doi.org/10.1515/9783110362084-023>
- Briz, A. y Albelda, M. (2013). Una propuesta teórica y metodológica para el análisis de la atenuación lingüística en español y portugués. La base de un proyecto común (ES.POR.ATEN- UACIÓN). *Onomázein. Revista de Lingüística, Filología y Traducción*, 28, 288–319. <https://doi.org/10.7764/onomazein.28.21>
- Cabedo, A. (2007). Caracterización prosódica del estilo directo de habla en la conversación coloquial. *Estudios de Lingüística de la Universidad de Alicante*, 21, 53–64. <http://dx.doi.org/10.14198/ELUA2007.21.04>
- Cabedo, A. (2009). *La segmentación prosódica en español coloquial*. Valencia: Quaderns de Filología de la Universidad de Valencia.

- Cabedo, A. (2011). Hacia un modelo predictivo para la segmentación prosódica del discurso oral coloquial: MESTEL (Modelo Estadístico para la selección de Términos Entonativos Ligados). *Oralia: Análisis del discurso oral*, 14, 85-104.
- Cabedo, A. (2018). Atenuación con disminución prosódica significativa en géneros con distinto grado de planificación discursiva. *Spanish in Context*, 15(2), 219-237.
- Cabedo, A. (2021). Oralstats. A tool to visualize and explore transcriptions and phonic data. <https://github.com/acabedo/oralstats>
- Caelen-Haumont, G. (1993). Procesos cognitivos y adaptación de la codificación prosódica de los hablantes a las condiciones del discurso. *Cognitiva*, 5(2), 133-152.
- Calvi, M. V. (2010). Los géneros discursivos en la lengua del turismo: Una propuesta de clasificación. *Iberica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, 19, 9-32.
- Cantero Serena, F. J., y Font Rotchés, D. (2009). Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental*, XVIII, 17-32.
- Chaski, C. E. (2012). Author identification in the forensic setting. En L. M. Solan y P. M. Tiersma (Eds.), *The Oxford Handbook of Language and Law*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199572120.013.0036>
- Cicres i Bosch, J. (2011). Transcripció i autenticació de gravacions en contextos judicials. *Llengua, Societat i Comunicació: Revista de Sociolingüística de La Universitat de Barcelona*, 9, 26-32.
- Cicres i Bosch, J. (2014). Comparación forense de voces mediante el análisis multidimensional de las pausas llenas. *Revista Signos: Estudios de Lingüística*, 86, 365-384. <http://dx.doi.org/10.4067/S0718-09342014000300002>
- Coulthard, M. (Ed.) (2013). *Advances in Spoken Discourse Analysis*. Londres: Routledge. <https://doi.org/10.4324/9780203200063>
- Coulthard, M. y C. N. Conclin (Eds.) (2014). *An Introduction to Discourse Analysis*. Londres: Routledge. <https://doi.org/10.4324/9781315835884>
- Coulthard, M. (2007). *An Introduction to Forensic Linguistics*. Londres: Routledge. <https://doi.org/10.4324/9780203969717>
- Coulthard, M., Grant, T. y Kredens, K. (2011). Forensic linguistics. In R. Wodak, B. Johnstone y P. E. Kerswill (Eds.), *The SAGE Handbook of Sociolinguistics* (pp. 529-544). Londres: SAGE Publications Ltd. <https://doi.org/10.4135/9781446200957.n36>
- Coulthard, M. y Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Londres: Routledge. <https://doi.org/10.4324/9780203855607>
- De la Mota, C., y Rodero, E. (2010). La demarcación entonativa y el énfasis en la locución de los editores de boletines informativos radiofónicos. En Sociedad Española de Lingüística (Ed.), *XXXIX Simposio Internacional de la Sociedad Española de Lingüística* (pp. 1-69). <https://dialnet.unirioja.es/servlet/articulo?codigo=5412029>
- Degand, L. y Simon, A. C. (2009). Mapping prosody and syntax as discourse strategies: How basic discourse units vary across genres. En D. Barth-Weingarten, N. Dehé y A. Wichmann (Eds.), *Where Prosody Meets Pragmatics* (pp. 79-105). Leiden: Brill.
- Elvira-García, W., Roseano, P., Fernández Planas, A. M. y Martínez Celdrán, E. (2016). A tool for automatic transcription of intonation: Eti-ToBI a ToBI transcriber for Spanish and Catalan. *Language Resources and Evaluation*, 50(4), 767-792. <https://doi.org/10.1007/s10579-015-9320-9>
- Espinosa, G. E., y Dabrowski, A. E. (2019). ¿La práctica discursiva condiciona la prosodia? Evidencia de una conferencia y una entrevista del español rioplatense. *Repositorio Digital Institucional, Universidad Nacional del Comahue*. <http://rdi.uncoma.edu.ar/handle/123456789/15781>
- Estellés, M., y Albelda, M. (2014). Evidentials, politeness and prosody in Spanish: A corpus analysis. *Journal of Politeness Research: Language, Behaviour, Culture*, XXX(1), 29-62.
- Grant, T. (2008). Approaching questions in forensic authorship analysis. En J. Gibbons y M. T. Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 215-229). Amsterdam: John Benjamins. <https://doi.org/10.1075/aals.5.15gra>
- Grant, T. y Baker, K. (2001). Identifying reliable, valid markers of authorship: A response to Chaski. *Forensic Linguistics*, 8(1), 66-79. <https://doi.org/10.1558/sll.2001.8.1.66>
- Gries, S. (2019). On classification trees and random forests in corpus linguistics: Some words of caution and suggestions for improvement. *Corpus Linguistics and Linguistic Theory*, 16, 617-647. <https://doi.org/10.1515/cllt-2018-0078>
- Guy, G. R. (2011). Language, social class, and status. En R. Messtrie (Ed.) *The Cambridge Handbook of Sociolinguistics* (pp. 159-185). Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511997068.015>
- Hazen, K. (2007). The study of variation in historical perspective. En R. Bayley y C. Lucas (Eds.), *Sociolinguistic Variation* (pp. 70-89). Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511619496.005>
- Hidalgo, A. (2019). *Sistema y uso de la entonación en español hablado*. Santiago de Chile: Universidad Andrés Burtado.
- Leemann, A., Kolly, M.-J., Nolan, F. y Li, Y. (2018). The role of segments and prosody in the identification of a speaker's dialect. *Journal of Phonetics*, 68, 69-84. <https://doi.org/10.1016/j.wocn.2018.02.001>
- López Serena, A. (2019). *La lingüística como ciencia humana: Una incursión desde la filosofía de la ciencia*. Madrid: Arco/Libros-La Muralla.
- ELAN (version 6.2) [Computer software] (2021). Nimega: Max Planck Institute for Psycholinguistics. <https://archive.mpi.nl/tla/elan>
- McMenamin, G. R. (1993). Forensic stylistics. En M. Coulthard y A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics*. Londres: Routledge. <https://doi.org/10.4324/9780203855607.ch32>
- McMenamin, G. R. (2002). *Forensic Linguistics: Advances in Forensic Stylistics*. Nueva York: CRC Press.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rodero, E. (2012). A comparative analysis of speech rate and perception in radio bulletins. *Text & Talk*, 32(3), 391-411. <https://doi.org/10.1515/text-2012-0019>
- Romaine, S. (2000). *Language in Society: An Introduction to Sociolinguistics* (2ª edic.). Oxford: Oxford University Press.
- Romero Nieto, A. (2018). Prosodia y gestualidad y su relación con la expresión del humor en el debate parlamentario. *Cuadernos de Investigación Filológica*, 44, 45-66. <https://doi.org/10.18172/cif.3413>
- Spasova, M. (2009). *El potencial discriminatorio de las secuencias de categorías gramaticales en la atribución forense de autoría de textos en español* Tesis Doctoral. Barcelona: Instituto Universitario de Lingüística Aplicada, Universitat Pompeu Fabra. <https://www.tdx.cat/handle/10803/7512%3Bjsessionid%3D#page=165>
- Straka, M. (2018). UDPipe 2.0 prototype at CoNLL 2018 UD shared task. En *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies* (pp. 197-207). Bruselas: Association for Computational Linguistics. <https://doi.org/10.18653/v1/K18-2020>