

ESTUDIO PILOTO

Análisis melódico del habla como herramienta distintiva para el perfil idiolectal de hablantes

Adrián Cabedo NEBOT 
Universitat de València (UV)

RESUMEN

Este artículo presenta un análisis de 1218 grupos entonativos y 8432 palabras de 5 hablantes de español peninsular en dos tipos de discursos interactivos orales: la entrevista y el podcast. Los grupos entonativos son secciones naturales de habla delimitadas por pausas o inflexiones melódicas. La elección de estos dos tipos de discurso se basa en la menor cantidad de superposición en este tipo de formato discursivo, la gradación de espontaneidad (más planificada en la entrevista), y la calidad del audio recogido, que proviene de videos de Youtube. Se ha utilizado la metodología del análisis melódico del habla (AMH) para clasificar estos grupos entonativos en 4 tonemas básicos (enunciativo, interrogativo, suspendido y enfático) con hasta 12 variantes diferentes. Los resultados muestran que los patrones de distribución son similares tanto entre hablantes como por género, aunque algunos factores acústicos de la estructura melódica, como la duración de la primera sílaba tónica en la anacrusis, pueden ser significativos para discriminar a los hablantes.

ABSTRACT

This paper presents an analysis of 1218 intonational phrases and 8432 words from 5 speakers of Peninsular Spanish in samples extracted from two interactive oral discourse genres: the interview and the podcast. Intonational phrases are natural units of speech located between pauses or marked pitch modulations. The choice of discourse genres is justified by the lower amount of overlap in this type of discourse format, the gradation of spontaneity (more planned in the interview), and, finally, by the



OPEN ACCESS

- Monique Leite Araújo (UnB)
- Sara Recio Pineda (UB)

CLASIFICADO POR

- Manuella Carnaval (UFF)
- Yolanda Congosto (US)
- Evandro Oliveira Santos (UFAL)

DATAS

- Recibido: 29/10/2022
- Aceito: 20/12/2022
- Publicado: 31/12/2022

COMO CITAR

NEBOT, Adrián Cabedo. (2022). Análisis melódico del habla como herramienta distintiva para el perfil idiolectal de hablantes. *Revista da Abralín*, v. 21, n. 2, p. 48-70, 2022.

quality of the audio collected, which comes from Youtube videos. Specifically, the methodology of the speech melody analysis (AMH) has been applied to classify these intonational phrases according to the 4 basic tones established by the model (assertive, interrogative, suspended, and emphasized) with up to 12 different variants. The main results indicate that patterns are distributed both in the same way among speakers and by genre, although some acoustic factors of the melodic structure, such as the duration of the first tonic syllable in the anacrusis, may be significant in discriminating speakers.

PALABRAS CLAVE

Entonación. Análisis melódico del habla. Idiolecto. Hablantes.

KEYWORDS

Intonation. Melodic Analysis of Speech. Idiolect. Speakers.

RESUMEN PARA NO ESPECIALISTAS

Este artículo estudia de qué manera las entonaciones de las personas pueden contribuir a identificarlas, por un lado, o a caracterizarlas, por otro. Se supone que cada persona tiene o puede tener un modo particular de hablar y que puede depender de elementos como la velocidad de habla, las entonaciones que se utilizan, etc. Para poder hacer esta investigación se ha utilizado un programa específico llamado Oralstats.

Introducción

El propósito de este artículo consiste en explorar la configuración melódica de múltiples grupos entonativos emitidos por los mismos hablantes en géneros discursivos distintos; al mismo tiempo, el objetivo es determinar en qué medida esas configuraciones pueden utilizarse como un elemento de análisis importante en la identificación de hablantes o, al menos, para formar parte de una caracterización lingüística y/o estilística de los sujetos analizados.

De este modo, para computar los valores prosódicos de las unidades de habla recogidas, y para permitir realizar con posterioridad un estudio fonético y estadístico más completo, se ha utilizado el sistema de análisis prosódico conocido como análisis melódico del habla (AMH), introducido por CANTERO (2002) y desarrollado y ampliado en muchos trabajos posteriores (CANTERO; FONT-ROTCHÉS, 2007; CANTERO; FONT-ROTCHÉS, 2009; MATEO, 2010; CANTERO, 2015, 2019).

El etiquetado prosódico en patrones fonológicos ha permitido realizar distintas pruebas de análisis y gestión estadística de los datos, de modo que puede decirse cuántos patrones se usan por qué hablantes y en qué momento. Así pues, el objetivo general del artículo supone realizar una primera aproximación hacia la rentabilidad de incorporar un modelo de análisis melódico del habla al estudio forense de hablantes o, al menos, al de su caracterización idiolectal.

1. Prosodia e identificación de hablantes

Tradicionalmente, el concepto de variación lingüística ha sido de gran interés para los lingüistas, dado que el estudio de la realidad discursiva pronto sugirió la idea de que los elementos que se comparten en una comunidad discursiva pueden estar afectados por variación tanto entre hablantes como intrahablantes (FERNÁNDEZ, 1990; CORVALÁN; ENRIQUE-ARIAS, 2017).

En tal caso, las propuestas que incluyen al individuo como centro de estudio establecen que hay motivos para pensar en un factor o un grupo de ellos que permita caracterizar, y por ende posteriormente identificar, el habla de los sujetos. Esa variación enunciativa no tiene por qué ser siempre constante, aunque muy probablemente existirán rasgos que serán continuos a lo largo de la vida adulta del sujeto. En ese contexto, es en el que el concepto de “idiolecto” o variación “idiolectal” cobra sentido:

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).

Generalmente, los hablantes presentan múltiples variaciones lingüísticas en sus emisiones, si bien algunas de ellas, más dependientes de factores anatómicos o fisiológicos, se consideran mucho más relevantes en la identificación o singularización de hablantes, proceso enmarcado en el ámbito de la Lingüística forense y, más concretamente, en la Fonética forense (CICRES I BOSCH, 2003; DELGADO *et al.*, 2009; LEEMANN *et al.*, 2018; DELGADO, 2020).

Algunos casos de factores idiolectales que se han señalado en inglés para tareas de identificación de hablantes serían, entre otros, la glotalización de las oclusivas sordas a final de palabra, la velarización de la /l/ postvocálica e incluso la importancia de la coarticulación (NOLAN, 2009). En el caso del español, y también en cuanto a fonética segmental, algunos especialistas del ámbito policial establecen multitud de parámetros. A modo de ejemplo, citamos solo algunos parámetros segmentales que se tienen en consideración (DELGADO *et al.*, 2009), como la aspiración, la duración de la oclusión, la asimilación, la nasalidad, la fricación, la reducción, las estridencias captadas en el espectro, etc.

También se consideran factores importantes en la discriminación de hablantes valores asociados a la frecuencia fundamental, correlato acústico del tono, así como cuestiones anatómicas, como la altura de los sujetos investigados; en general, individuos más altos se asocian a tonos más graves, mientras que individuos más bajos se relacionan con tonalidades más agudas (JESSEN, 2020).

En la comparación forense de voces, es necesario analizar varios parámetros acústicos que tienen una relación directa con la forma de las cavidades (nasal, faríngea y oral) de resonancia del trato vocal, por un lado, y con la frecuencia de vibración de las cuerdas vocales (...) que presenten una alta variabilidad entre hablantes distintos y una variabilidad baja en las diversas realizaciones de un mismo hablante; que sean resistentes a los intentos de enmascaramiento y de imitación; que tengan una alta frecuencia de ocurrencia en las muestras de análisis; que sean robustos en la transmisión (es decir, que la transmisión no altere o elimine sus propiedades acústicas), y que sean relativamente fáciles de extraer y analizar. [la traducción es nuestra] (CICRES I BOSCH; GAVALDÀ, 2014, p. 64–65).

Aun así, hay otras ideas procedentes de lo anteriormente expuesto que fomentan otro acercamiento que pudiera resultar interesante. En ocasiones, los valores fónicos pueden ser compartidos por la comunidad de habla (lógico en todo caso); sin embargo, en esos valores compartidos pueden aparecer momentos en los que los hablantes, por varios factores adscritos al contexto o al estilo de habla, pueden presentar alteraciones notables que lo alejen tanto del grupo social de hablantes al que pertenece como a sí mismo en otras situaciones de habla. El uso de esas divergencias también pueden ser tenidas en cuenta en el ámbito del estudio forense, aunque se recalca habitualmente la dificultad de encontrar elementos plenamente idiosincrásicos:

The features that are used for speaker classification are by definition not speaker-specific but rather characteristics of sets of speakers. Features such as a high-pitched voice or a fast speech rate, on the other hand, are individual features. That does not mean that speakers are uniquely characterised by these individual features; there is still a long way to go until we know whether, and if so, how speakers can be completely individualised by speech evidence alone (...). But it means that information about pitch and speech rate is important in the forensic task of distinguishing individual speakers. (JESSEN, 2020, p. 381).

En el marco de la Lingüística forense, y concretamente de la identificación de hablantes, el papel de la prosodia y de la fonética en general ha resultado ser de un valor indudable. No obstante, el objeto de estudio para la diferenciación de hablantes ha resultado estar casi siempre focalizada en aspectos vinculados a la fonética de segmentos aislados, es decir, a variaciones alofónicas de fonemas particulares y, en sobre ellos, en cuestiones acústicas muy restringidas (DELGADO, 1998, 2012; DELGADO *et al.*, 2009). Este modo de actuación está justificado si se tiene en cuenta que la configuración anatómica del tracto vocal puede tener una repercusión inmediata en propiedades acústicas observables, como los valores formánticos, y que también puede tener un reflejo en interpretaciones auditivas como el timbre (CICRES I BOSCH, 2003; NOLAN, 2009; CICRES I BOSCH; GAVALDÀ, 2014). En otras palabras, los sonidos segmentales parecen estar afectados más directamente por valores dependientes de la propia y genuina identidad del sujeto que los produce.

No obstante, la configuración melódica puede ser tomada también como un elemento de análisis en la caracterización de esos mismos hablantes. Estos patrones tonales, dentro del marco de la fonética suprasegmental, pueden ser utilizados como tendencias que se asocian, con una cierta regularidad, a la producción específica de los hablantes. Es en esta línea exploratoria en la que situamos el valor de análisis del presente trabajo.

Por tanto, el objetivo de este artículo es investigar la configuración melódica de múltiples grupos entonativos producidos por los mismos hablantes en géneros discursivos diferentes, y al mismo

tiempo, determinar en qué medida estas configuraciones pueden ser utilizadas como un elemento importante en la identificación de hablantes o al menos como parte de una caracterización lingüística y/o estilística de los sujetos analizados.

La metodología utilizada en este artículo implica la aplicación de la metodología de análisis melódico del habla (AMH) para clasificar los grupos entonativos en 4 tonemas básicos (enunciativo, interrogativo, suspendido y enfático) con hasta 12 variantes diferentes. Los pormenores específicos de la metodología se desarrollan en la Sección 2.

2. Metodología

La metodología de este estudio es de base cuantitativa, es decir, a partir de la recogida de un corpus de estudio, que definimos en la sección 2.1., se realiza un análisis de datos mediante programas de explotación estadística, como *Oralstats* (CABEDO, 2021), cuyas prestaciones comentamos más extensamente en la sección 2.3. El análisis estadístico, por su parte, presenta visualizaciones de tipo exploratorio que incluyen mapas de calor, tablas cruzadas o diagramas de caja. Si bien este estudio presenta un carácter exploratorio y no pretende una representatividad poblacional depurada a partir de los resultados obtenidos, técnicas como la prueba ANOVA, utilizada en la sección 3.3., sí se basan en un formato de validación inferencial.

2.1 Corpus de estudio

El corpus analizado en esta investigación está compuesto por 8 archivos de audio procedentes de dos géneros discursivos (entrevista y podcast) extraídos de la plataforma Youtube; concretamente, la persona que publica el material en su canal es el *youtuber* Pico de Oro/Padre Zorro¹. La selección de los archivos analizados y no de otros es meramente eventual y se justifica en general por dos motivos: la presencia de géneros discursivos distintos y la calidad del material sonoro recogido. La elección de la temática y de los hablantes, todos ellos usuarios de la variedad peninsular del español, ha sido aleatoria, es decir, podrían haberse seleccionado otros hablantes y otros temas. Aun así, el requisito básico, más allá de la elección concreta de las personas, es que fueran hombres de entre 25 a 45 años y que hubieran participado en ambos formatos, es decir, tanto en la entrevista como en el podcast.

Los géneros discursivos analizados incluyen, en primer lugar, un formato más formal, en forma de entrevista discursiva, y registrada en Youtube con la etiqueta “5 preguntas a”; en este formato el *youtuber* realiza una entrevista a un deportista (normalmente, culturista) y le formula cinco preguntas, aunque no

1 A fecha de octubre de 2022, Pico de Oro, también conocido como Padre Zorro, es un *youtuber* con 132000 suscriptores. Su temática es de entretenimiento y se centra en el análisis semanal de la actualidad del mundo del fitness, sobre todo del culturismo. Los videos que sube regularmente a su canal cubren varios géneros discursivos, desde monólogos editados hasta entrevistas o podcasts. El enlace actual al canal es <https://www.youtube.com/channel/UC4oGsbaR87FQ1b3C04ZxVwQ>

se prohíbe realizar más interacciones si van surgiendo durante el intercambio. En la última parte de la entrevista, el entrevistado tiene la opción de realizar una pregunta al entrevistador.

Por otro lado, el segundo género discursivo que hemos recogido procede de un formato bastante más distendido en cuanto a la interacción discursiva, etiquetado como *podcast*. En él no hay preguntas prefijadas de antemano, aunque los participantes, habitualmente dos o tres personas pertenecientes al mundo del fitness, intercambian opiniones y pareceres sobre temas actuales de la comunidad: una competición reciente, el éxito de algún deportista famoso, polémicas generadas entre los mismos competidores, etc.

Aunque se trata de formatos similares que recogen habla espontánea, no leída ni planificada previamente (al menos, no planificada en las respuestas; sí en las preguntas de las entrevistas), es cierto que, en el formato de entrevista discursiva, como comentábamos, se realizan cinco preguntas al entrevistado que fomentan un posterior diálogo; es por ello por lo que este género discursivo se puede considerar algo más formal o ritualizado que el del *podcast*. La idea de recoger estos dos géneros ha sido la de observar posibles diferencias asociadas al formato discursivo en cuanto al uso de diferentes patrones melódicos. Aunque ambos formatos recogen habla espontánea, la entrevista es un poco más formal o ritualizada debido a que se hacen cinco preguntas al entrevistado que incentivan un diálogo posterior. Además, la elección de la entrevista y el *podcast* se justifica por la menor cantidad de solapamientos en este tipo de formatos discursivos, así como por la gradación de espontaneidad (que es mayor en el *podcast* y más planificado en la entrevista). La calidad del audio recogido en ambos formatos también es un factor importante, ya que los datos provienen de videos de Youtube.

Así pues, se ha obtenido un corpus compuesto por 5 hablantes (4 de ellos se repiten por cada género analizado y el *youtuber* aparece en todos ellos), 1218 grupos entonativos y 8432 palabras. En realidad, el total de grupos entonativos es mayor, pero se han eliminado una gran cantidad por varios motivos: la presencia de solapamiento, la ininteligibilidad de alguna parte del grupo entonativo (se daña entonces la información acústica obtenida), imposibilidad de recoger suficiente material acústico en la porción de audio analizada...

Para cada archivo, como puede observarse en la Tabla 1, se ha recogido una media de unos 5 minutos de habla. En la Tabla 1, se describe, además, la cantidad de grupos entonativos que entran en los análisis, el tiempo total de habla y la proporción de tiempo de habla ocupado por el hablante en relación con esos 5 minutos. Los cuatro hablantes, distintos al *youtuber* Pico de oro/Padre Zorro, se han nombrado con los nombres que generalmente utilizan en las redes sociales. El nombre suele incluir una referencia parcial al nombre del deportista entrevistado (*angelreal*, por ejemplo), o el apellido (*Estepa*, *Dufort* y *Castellano*). El *youtuber* recibe una reducción de su nombre en internet: *pzorro* es equivalente a *Padre Zorro* o *Pico de oro*. Así, como puede verse en la Tabla 1, la columna del archivo tiene una misma referencia (*5pangelreal*, por ejemplo) y luego dos referencias a los hablantes que participan en él: *5pangelreal* y *5ppzorro*.

filename	spk2	grupos	palabras	tiempo_habla	tiempo_total	tiempo_prop
5pangelreal	5pangelreal	103	624	131.07	283.563	46.22253
5pangelreal	5ppzorro	64	414	95.02	300.436	31.62737
5pcastellano	5ppzorro	68	511	118.02	318.746	37.02635
5pcastellano	5prcastellano	97	806	148.94	328.691	45.31308
5pdufort	5pdufort	42	374	80.27	299.484	26.80277
5pdufort	5ppzorro	95	748	185.28	292.744	63.29079
5pestepa	5pestepa	83	474	135.55	294.427	46.03858
5pestepa	5ppzorro	57	396	88.64	251.559	35.23627
podangelreal	podangelreal	43	369	86.64	297.144	29.15758
podangelreal	Podpzorro	109	714	179.43	299.082	59.99358
poddufort	Poddufort	120	737	176.63	307.133	57.50929
poddufort	Podpzorro	36	222	52.41	304.152	17.23152
podestepa	Podestepa	86	488	133.01	310.330	42.86083
podestepa	Podpzorro	72	462	103.80	302.816	34.27824
poscastellano	Podpzorro	86	621	145.73	298.072	48.89087
poscastellano	podrcastellano	57	472	101.42	299.784	33.83103

TABLA 1 - Distribución por hablante y archivo de grupos entonativos, palabras, tiempo de habla de la grabación (en segundos) y porcentaje de tiempo de habla en el total de la porción de grabación analizada. Leyendas: 5p (5 preguntas), pod o pos (podcast).

Fuente: Elaboración propia.

Solo como comentario general sobre los datos que se observan en la Tabla 1, aunque no forman parte del interés de esta investigación, se registran casos en los que algunos hablantes participan muy poco, como Padre Zorro en el podcast con Dufort, o, curiosamente, Dufort en relación con la entrevista. Al mismo tiempo, algunos hablantes manifiestan un dominio bastante amplio del total de la duración del fragmento analizado; así es el caso de Padre Zorro en gran parte de los archivos recogidos. Este último dato no deja de ser lógico y esperable, dado que es el propio youtuber el que se articula como mantenedor del intercambio comunicativo.

2.2 Alineación y tratamiento de los datos

Los grupos entonativos recogidos en el corpus de esta investigación han sido tokenizados y divididos en fonemas con ayuda de la opción de división y alineamiento de sonidos que ofrece el programa PRAAT (BOERSMA; WEENINK, 2022). Esta manera de proceder permite generar archivos de textos (*TextGrids*) que tienen tres líneas o capas de análisis:

1. Línea con la transcripción de los grupos entonativos.
2. Línea de palabras tokenizadas alineadas al audio.
3. Línea de fonos alineados al audio.

La primera línea de transcripción, la de los grupos entonativos, se realiza manualmente, mientras que las otras dos líneas, las de la tokenización y la fonemización se realizan de manera semiautomática, es decir, a la división realizada por PRAAT se aplica una segunda capa de revisión para comprobar, sobre todo, que las vocales tónicas han sido correctamente ubicadas en el marco temporal al que pertenecen. Esta tarea es importante para generar posteriormente los patrones melódicos de modo automático, aspecto que abordaremos a continuación en la sección 2.3.

2.3 Herramienta de análisis

El análisis cuantitativo se ha realizado con la herramienta *Oralstats*²(CABEDO, 2021), un entorno de análisis desarrollado con el lenguaje de programación R y que utiliza una de sus librerías, en concreto Shiny (CHANG *et al.*, 2022), para permitir transformar y analizar datos procedentes de transcripciones de habla, normalmente recogidos en formato de grupos de entonación (secuencias de habla entre pausas de habla o inflexiones melódicas muy pronunciadas (QUILIS; CANTARERO; ESGUEVA, 1993). *Oralstats* permite cruzar esa información con datos fónicos, mediante la importación de datos acústicos de frecuencia fundamental, medidos en hercios, e intensidad, medidos en decibelios.

En la versión 1.3 del programa, se permite realizar una codificación de la curva melódica mediante dos sistemas de etiquetado tonal: el sistema ToBI³ (PIERREHUMBERT, 1980) y el análisis melódico del habla⁴ (CANTERO, 2002; CANTERO; FONT-ROTCHÉS, 2009). Se trata en todo caso de etiquetado automático que requieren siempre de una última supervisión manual por parte de investigador, no porque el etiquetado realizado no sea consistente con unos parámetros computacionales previamente incorporados, sino porque en ocasiones hay anomalías que solo se detectan con posterioridad al proceso de computación; por ejemplo, el archivo de valores tonales o de intensidad puede estar corrupto o puede tener alguna malformación que conlleve interpretaciones erróneas por parte del programa.

Finalmente, los cómputos estadísticos o los gráficos exploratorios que integran este artículo (prueba ANOVA, mapa de calor, tablas y diagramas de caja) se han generado también con el programa *Oralstats*.

2 El script completo está disponible de forma gratuita en <https://github.com/acabedo/oralstats/> y el utilizado específicamente para transformación en el modelo del análisis melódico del habla se encuentra también disponible de manera gratuita en https://github.com/acabedo/abralin_1, así como la base de datos en formato xlsx (Excel).

3 En el caso del etiquetado ToBI, mientras que el script de computación de *Oralstats* es algo básico para realizar la enorme variedad de patrones tonales que se pueden obtener (no siempre bien establecidos en la bibliografía), se recomienda el uso de otros scripts mucho mejor diseñados, como el caso de *EtiToBI* (ELVIRA-GARCÍA *et al.*, 2016).

4 Para un script de procesado del análisis melódico del habla, puede usarse también el desarrollado por MATEO (2010).

3. Resultados

En esta sección incluiremos los análisis y resultados procedentes del material transformado y computado por *Oralstats*. En la primera sección (3.1), se describen los patrones tonales generales que se han recogido a partir de las configuraciones del análisis melódico del habla; en la siguiente sección (3.2), esos patrones se presentan en relación con los hablantes y con el género; finalmente, en la última sección de resultados (3.3), se incluye un dato que ha sido significativamente distinto en la identificación de los hablantes de esta investigación, la duración de la primera vocal tónica del grupo entonativa como marca para la caracterización de algunos hablantes.

3.1 Descripción de los datos

Los patrones entonativos para el español han sido definidos convenientemente por la bibliografía (CANTERO SERENA; FONT-ROTCHÉS, 2007). Se trata de cuatro tonemas generales (enunciativo, interrogativo, suspendido y enfático) y hasta 17 variantes: 1 del tonema enunciativo, 4 del interrogativo, 3 del suspendido y 9 del enfático. Estas variantes toman en consideración sobre todo las distintas modulaciones de la estructura melódica; por ejemplo, el comportamiento de la primera sílaba tónica, según si el pico se alinea con esta sílaba o con la siguiente, la declinación tonal y, finalmente, el porcentaje de inflexión tonal a partir de la última sílaba tónica del enunciado.

Los datos recogidos en esta investigación, en general, se alinean con los patrones entonativos recogidos por CANTERO y FONT-ROTCHÉS (2007). Los tonemas básicos recogidos son cuatro, que simplifican otras propuestas anteriores (CANTERO, 2002): el neutro, entre un 15 % de inflexión tonal para los tonos descendentes y un 30 % de inflexión para tonos ascendentes; el suspendido, con una subida entre 15 y 70 %; el interrogativo, con una inflexión superior al 70 %; y el patrón enfático, con un descenso final superior al 30 %, y habitualmente con desplazamientos tonales en la primera sílaba tónica o con modulaciones complejas en el cuerpo del grupo entonativo.

En cuanto a la información codificada mediante el cómputo del análisis melódico del habla, la distribución obtenida se presenta en la Tabla 2 para los tonemas:

tonemes	Frecuencia
desconocido	383
Other	289
enunciativo	234
suspendido	199
enfático	102
interrogativo	11

TABLA 2 - Distribución de los tonemas más frecuentes etiquetados con el sistema del análisis melódico del habla
Fuente: Elaboración propia.

En la Tabla 2 podemos observar que los tonemas que más presencia tienen en el corpus son los enunciativos y los suspendidos, seguidos de los enfáticos y, circunstancialmente, de los interrogativos. Estos últimos tienen poca presencia en el corpus recogido, dado que en el modelo AMH los tonemas interrogativos incluyen ascensos tonales superiores a 70 %⁵; la mayor parte de ascensos, inferiores a ese porcentaje, han quedado registrados habitualmente como tonemas suspendidos. Los tonemas etiquetados como desconocido significa que esos grupos entonativos han sido descartados del análisis por varios motivos: están compuestos por elementos paralingüísticos como pausas llenas, vacilaciones, risas, chasquidos, etc., la información acústica no se ha podido recoger adecuadamente en la zona del tonema, etc.

Aun así, hay 289 grupos que se han etiquetado como “other”, es decir, que no han cumplido los parámetros establecidos por la computación del programa Oralstats para los patrones entonativos; curiosamente, en ellos el tonema tiene una inflexión tonal pequeña, entre -20 y 20 de descenso y ascenso porcentual, combinado con un amplio número de desplazamiento tonal en la primera sílaba tónica de la zona de la anacrusis. Posiblemente, estos casos podrían considerarse como un patrón enunciativo con desplazamiento tonal en la anacrusis, con lo que el número de tonemas enunciativos aún ascendería notablemente.

Los tonemas de la Tabla 2 pueden dividirse en los 17 patrones o variantes entonativas que comentábamos anteriormente. Esos patrones se disponen en la Tabla 3:

MAStag	Frecuencia
desconocido	374
other	289
PI	234
PV	151
PXIIa	64
PVIa	35
PVIII	29
PVIb	13
PIX	9
PII	8
PXI	5
PIII	3
PVII	3
PXa	1

TABLA 3 - Distribución de las variantes melódicas más frecuentes etiquetadas a partir del análisis melódico del habla

Fuente: Elaboración propia.

En la Tabla 3 se describen con mayor precisión los datos que se habían presentado en la Tabla 2. En general, más allá de los datos no procesados (374 casos), se observan 289 que han sido catalogados como “other”; de nuevo, recordamos que este patrón puede considerarse un subtipo del

5 El tonema interrogativo también puede recoger inflexiones con un porcentaje menor, 40-60 %, aunque según la bibliografía (CANTERO; FONT-ROTCHÉS, 2007) parece que, en estas variantes, se requiere también que exista un desplazamiento tonal en el primer pico en la zona de la anacrusis.

patrón enunciativo, pero con un desplazamiento tonal en la anacrusis. Por su parte, el enunciativo PI es el más usado (234), seguido del suspendido PV (151); posteriormente, el tercero más usado es el enfático PXIIa (64 casos), en este patrón hay un desplazamiento en la anacrusis, una declinación plana y un final tonal descendente. También se detectan 35 casos del PVIa, un patrón suspendido sin desplazamiento tonal en la anacrusis y con un ascenso entre el 15 y el 70%. Finalmente, otro patrón enfático, el PVIII, es el siguiente más utilizado, con 29 casos; este patrón se caracteriza por no tener declinación tonal, por lo que el primer pico y el tonema quedan unidos.

Todos los cruces estadísticos realizados para determinar la relación entre la variable que integraba hablante y género, por un lado, y el tonema (o sus variantes), por otro, han resultado negativos, con valores p superior a 0.05 en la prueba de la chi-cuadrado. De este modo, no se ha encontrado una relación identitaria entre los patrones melódicos registrados y su asociación a individuos particulares en diferentes géneros discursivos. Al menos, no en este corpus de estudio.

Por lo tanto, dado que la distinción entre el género *5 preguntas a* y el género *podcast* no generan diferencias estadísticas entre sí, podemos permitirnos aludir sencillamente al hablante y combinar los tonemas de un género con los del otro. De esta manera, en tablas como la Tabla 4 y la Tabla 5, los hablantes se consideran de manera unitaria, sin distinguir entre géneros.

Spk	enfático	Enunciativo	interrogativo	suspendido	<NA>
angelreal	5	31	1	20	89
Dufort	16	37	3	21	85
Estepa	23	47		35	64
Pzorro	49	98	6	108	326
rcastellano	9	21	1	15	108

TABLA 4 - Distribución de los tonemas por hablante
Fuente: Elaboración propia.

En la Tabla 4 se perciben varios elementos llamativos; en primer lugar, los datos de Pzorro siempre serán superiores al resto, dado que su aparición es constante en los 8 archivos de audio procesados. Al mismo tiempo, destaca que los contornos interrogativos tengan una aparición casi residual en nuestro corpus. Puede deberse a mucho factores, como el hecho de que hay un control elocutivo mayor, pero no tanto por una presunta formalidad del acto de habla, que no existe en realidad, sino porque la relajación de la situación (por lo menos, en el podcast) permite, en el caso de que el hablante lo decida, un mayor control de lo que se dice.

En cuanto a modulaciones entonativas, dos de los deportistas registrados en las grabaciones, concretamente Angelreal y RCastellano, apenas muestran casos de patrones enfáticos. Su discurso, por lo tanto, es algo más lineal en cuanto a la expresión fónica.

Por lo demás, si desglosamos los tonemas de la Tabla 4, nos encontramos con los valores que aparecen en la Tabla 5:

spk	Other	PI	PII	PIII	PIX	PV	PVIa	PVIb	PVII	PVIII	PXa	PXI	PXIIa
angelreal	29	31		1	1	15	4	1		3			2
dufort	44	37	1	2	1	15	4	2	1	6		1	8
estepa	34	47				29	4	2	2	8			13
pzorro	152	98	6		6	80	21	7		10	1	4	34
rcastellano	30	21	1		1	12	2	1		2			7

Tabla 5. Distribución de los patrones melódicos por hablante
Fuente: Elaboración propia.

Si atendemos a los valores de la Tabla 5, puede comprobarse efectivamente que, a nivel proporcional, para la cantidad total de grupos entonativos que aporta cada hablante del corpus, la distribución entre los patrones tonales está bastante equilibrada; los valores más poblados son PI, PV, PVIb y PXIIa; mientras que la menor frecuencia en otros patrones parece constante, a excepción de Padre Zorro, que recoge valores de PII, PIX, PVIa y PXI; debe tenerse en cuenta que, como moderador de los actos interactivos en los que participa, se presupone que su comportamiento melódico también debe ser amplio y variado.

A modo de ejemplificación de estos patrones, aportamos gráficos extraídos directamente de PRAAT en los que se puede ver la curva melódica. Se trata de dos patrones pertenecientes al youtuber Padre Zorro (enunciativo, PI, y suspendido, PV) y de un patrón suspendido, emitido en uno de los enunciados del deportista/culturista Castellano.

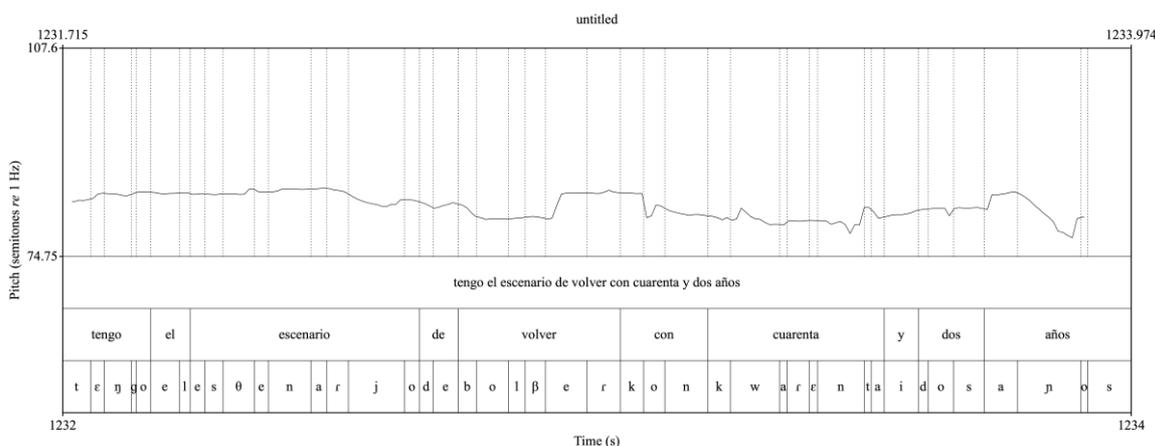


FIGURA 1 - Imagen de la curva melódica del grupo entonativo "tengo el escenario de volver con cuarenta y dos años", emitido por el youtuber Padre Zorro.
Fuente: Elaboración propia.

Esta curva melódica de la Figura 1 ha sido catalogada como PI por el sistema automático de Oralstats. Anacrusis ausente o, al menos, el grupo comienza con una sílaba tónica, luego se aprecia una declinación constante y un ascenso de un 10.2 % en la sílaba tónica del tonema, seguido de una inflexión descendente de más del 30 % en la sílaba postónica.

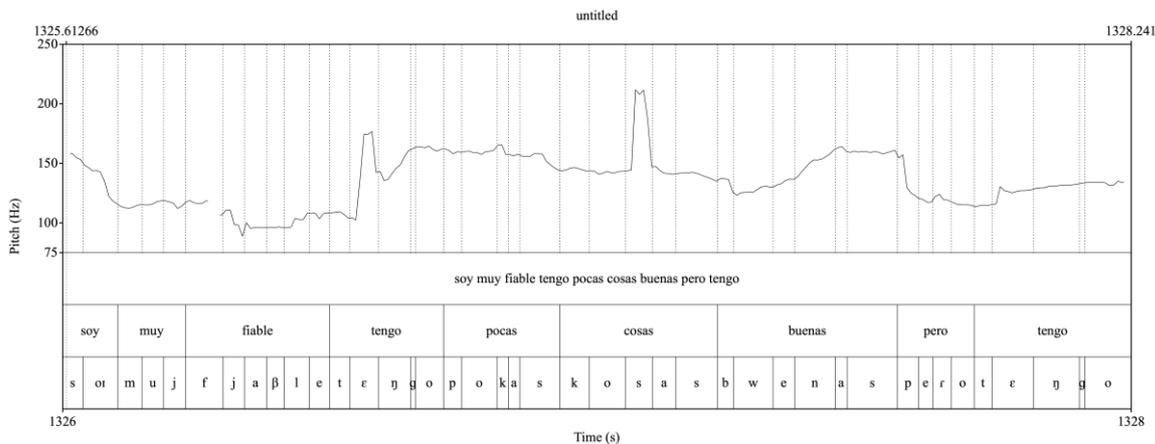


FIGURA 2 - Curva melódica del grupo entonativo "soy muy fiable tengo pocas cosas buenas pero tengo", emitido por el youtuber Padre Zorro.
Fuente: Elaboración propia.

La curva de la Figura 2 ha sido catalogada como PV por el sistema automático de Oralstats. Declinación constante, con una interferencia acústica en cosas por un pequeño golpe del micrófono, desde la primera sílaba tónica del grupo hasta la última sílaba tónica de la palabra "tengo", que no presenta una inflexión tonal especialmente llamativa.

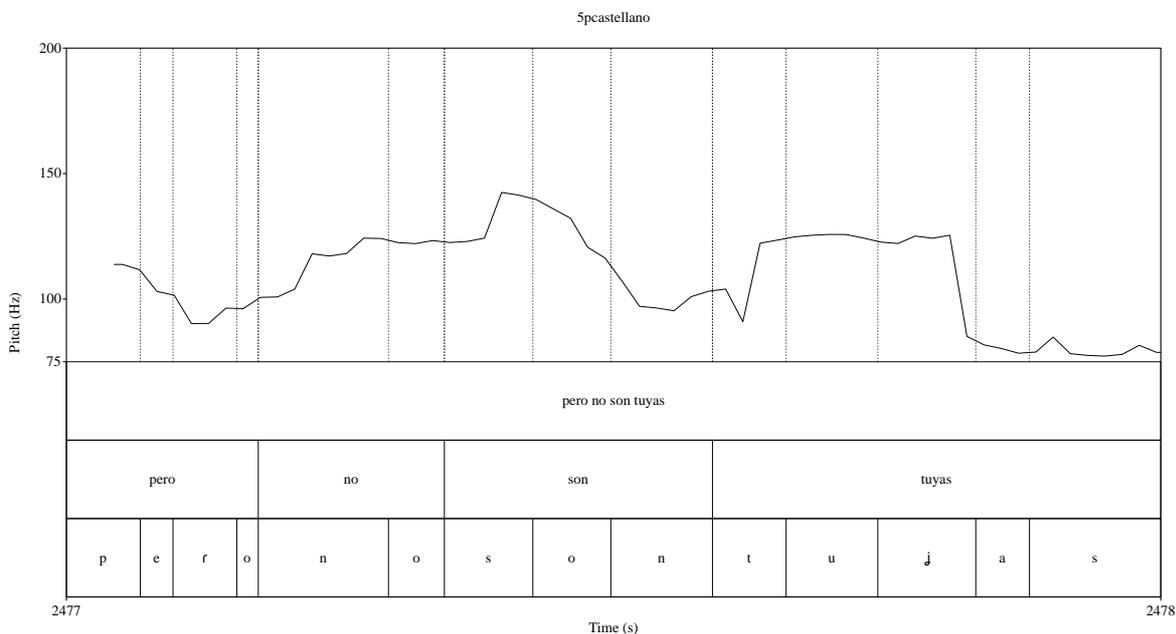


FIGURA 3 - Curva melódica del grupo entonativo "pero no son tuyas", emitido por el deportista Castellano
Fuente: Elaboración propia.

La curva melódica de la Figura 3 ha sido catalogada como PXIIa por el sistema automático de Oralstats. Hay un desplazamiento de la sílaba tónica en el anacrusis a la siguiente sílaba, también

tónica, y finalmente sigue una inflexión descendente que comienza en la sílaba tónica del tonema. En el caso de la Figura 3, puede debatirse efectivamente sobre si hay o no desplazamiento en “no”, ya que aunque la subida tonal culmina en el principio de la sílaba de la palabra “son”, también es cierto que esta se trata en todo caso de una sílaba también tónica. Curiosamente el descenso tonal comienza en la [o] de la palabra “son”. La declinación tonal es plana, puesto que Oralstats establece supera un 5 % de desviación (negativa o positiva) para clasificar las inflexiones del cuerpo del tonema como inclinaciones o declinaciones; en caso contrario, la declinación se considera plana.

La Figura 1, la Figura 2 y la Figura 3 no solo muestran los patrones entonativos que los clasifican (PI, PV y PXIIa), sino también el modo de proceder del módulo de etiquetado de Oralstats para el análisis melódico del habla. Si bien no todos los etiquetados tienen por qué ser correctos, es cierto que el sometimiento de los datos a una misma estructura computacional de transformación asegura que aquello que se compare sea homogéneo estadísticamente. Aun así, como puede percibirse, el etiquetado automático de valores tonales supone un problema complejo que puede ir modificándose paulatinamente a partir de un aprendizaje semiautomático por parte de Oralstats.

El etiquetado automático de los valores tonales es un problema complejo que puede mejorar gradualmente a partir de un aprendizaje semiautomático realizado por el sistema Oralstats. La subjetividad humana puede influir en la correcta identificación de los valores tonales, pero el hecho de que todos los datos se sometan a la misma estructura computacional de transformación asegura que aquello que se compara sea homogéneo estadísticamente. A pesar de esto, es importante tener en cuenta que no todos los valores tonales etiquetados por el sistema pueden ser necesariamente correctos. Por lo tanto, el proceso de etiquetado automático requiere una supervisión humana constante para garantizar su precisión.

3.2 Hablantes y configuración acústica por género

En la sección anterior, los diferentes patrones entonativos establecidos según la metodología del análisis melódico del habla indicaban que el comportamiento de los hablantes era similar en todas las muestras analizadas; sin embargo, también es interesante plantearse de qué manera los hablantes pueden comportarse de modo distinto o, al menos, ligeramente diferente sin tener que recurrir necesariamente a valoraciones estadísticamente significativas, esto es, los elementos que hemos comentado en la sección 3.1. pueden explorarse nuevamente desde una visualización simple de la información. Para ello, es para lo que Oralstats permite realizar, entre otros gráficos, mapas de calor.

Los mapas de calor permiten representar visualmente distintas variables numéricas en el eje de coordenadas, por un lado, y, por otro, en el eje de ordenadas, los grupos o categorías en los que se proyectan estas variables cuantitativas. Dado que los valores de las variables pueden contener valores con unidades de medida distinta, se relativizan previamente antes de desplegarse en el mapa. En la paleta de colores, los colores más claros de las celdas indican gran cantidad en esa variable, mientras que los colores más oscuros señalan ausencia de valores o, incluso, un componente negativo en esa celda. Un ejemplo de mapa calor puede verse en la Figura 4:

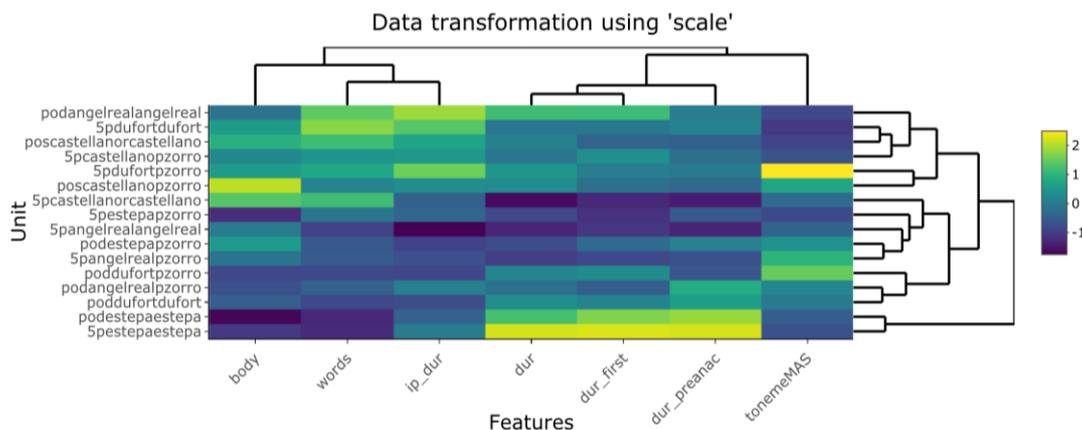


FIGURA 4 - Mapa de calor de las variables numéricas del estudio con disposición de cada una de las manifestaciones orales de los hablantes por entrevista o podcast analizado

Fuente: Elaboración propia.

Un mapa de calor generado por *heatmap* en R es un tipo de visualización de datos que utiliza una matriz de colores para representar valores numéricos en una tabla. Los valores se asignan a diferentes colores en una escala que va desde colores claros, valores altos, a colores oscuros, valores bajos. Esta escala se puede personalizar para destacar las diferencias en los valores. La idea detrás de un mapa de calor es que cuanto más oscuro sea un color, más alto es el valor correspondiente. En R, los mapas de calor se pueden crear utilizando el paquete *heatmap* o una función similar. Estos mapas son útiles para identificar patrones o tendencias en grandes conjuntos de datos numéricos y para comparar valores en una tabla bidimensional.

En la Figura 4, se observan varios factores que cabe recalcar:

1. Estepa, para el podcast y la entrevista, presenta valores muy altos en la duración de la vocal tónica del tonema, pero también en la primera vocal tónica del grupo entonativo y en la duración de la primera sílaba átona del grupo. Contrasta con los valores de Castellano y de Angelreal en la entrevista, que son muchísimos más bajos.
2. El tonema más elevado, interrogativo, se percibe en la entrevista a Dufort y lo enuncia el youtuber Padre Zorro.
3. En el dendograma se crean tres grupos a partir de la exploración. El primero lo componen Angelreal, Castellano y Padre Zorro para el podcast, y Dufort y Padre Zorro para las entrevistas. El segundo grupo es el más numeroso y está compuesto por Castellano, Angelreal en la entrevista, y Dufort en su podcast, así como Padre Zorro en casi todos los podcasts. Finalmente, Estepa se diferencia del resto en el tercer grupo, compuesto solo por él mismo tanto en su podcast como en su entrevista.

Los valores proyectados en la Figura 4 pueden observarse numéricamente en la Tabla 6:

spk3	dur	dur_first	dur_preatac	body	tonemeMAS	words	ip_dur
5pangelrealangelreal	0.08	0.06	0.05	3.49	1.84	5.86	1.21
5pangelrealpzorro	0.08	0.06	0.06	-0.68	13.45	6.02	1.43
5pcastellanopzorro	0.10	0.09	0.06	1.41	-0.18	6.71	1.60
5pcastellanorcastellano	0.07	0.06	0.05	6.98	2.87	7.85	1.44
5pdufortdufort	0.10	0.08	0.06	9.87	-3.39	8.66	1.90
5pdufortpzorro	0.11	0.08	0.06	1.22	25.86	7.21	1.80
5pestepaestepa	0.13	0.12	0.08	-2.63	-0.54	5.31	1.54
5pestepapzorro	0.09	0.06	0.06	-1.45	-1.46	6.66	1.50
podangelrealangelreal	0.11	0.10	0.07	-1.00	-1.39	7.90	1.86
podangelrealpzorro	0.10	0.07	0.07	-1.33	6.95	6.09	1.53
poddufortdufort	0.10	0.08	0.07	0.31	5.13	5.70	1.41
poddufortpzorro	0.10	0.09	0.06	1.89	17.68	5.79	1.37
podestepaestepa	0.12	0.11	0.08	-7.14	1.10	5.31	1.48
podestepapzorro	0.09	0.07	0.07	4.86	8.80	5.98	1.36
poscastellanopzorro	0.10	0.08	0.06	11.51	11.52	6.63	1.58
poscastellanorcastellano	0.10	0.08	0.06	5.28	-2.16	7.81	1.71

Tabla 6. Medias de las variables acústicas de duración y de inflexión melódica del corpus de análisis para cada hablante por género.

Fuente: Elaboración propia.

Uno de los factores que más llaman la atención en la Tabla 6, sobre todo en lo relativo a la inflexión tonal en el tonema, es que hay hablantes que utilizan una amplia variedad de inflexiones tonales. Los que mantienen el seguimiento de los intercambios comunicativos, como es el caso del youtuber Padre Zorro, utilizan tonemas ascendentes generalmente, aunque no en unos límites suficientes como para etiquetarlos como patrones interrogativos. Se trata muchas veces de patrones suspendidos y continuativos que fomentan la continuación por parte de los interlocutores.

Así pues, aunque no se ha articulado como una variable estadísticamente significativa, si volvemos sobre los datos de la Tabla 6 o de la Figura 4, sí se apuntan datos relevantes como, por ejemplo, una cierta neutralidad en la mayoría de los grupos entonativos de casi todos los hablantes, a excepción del youtuber Padre Zorro que, por su papel de dinamizador y mantenedor del intercambio, utiliza tonemas ascendentes bastante más marcados que los realizados por el resto de interlocutores. Otra de las conclusiones que se deduce, y que se extiende en la Sección 3.3, es que pueden existir otros factores acústicos que, incluso siendo estadísticamente significativos, pueden colaborar en la caracterización de hablantes.

3.3 Elementos identificativos

En las secciones anteriores (3.1. y 3.2), hemos podido ver que los patrones entonativos generados a partir del análisis melódico del habla no discriminan a los sujetos de nuestro corpus en la medida que esperábamos. No obstante, el AMH incluye un factor relevante en cuanto a la constitución de los grupos entonativos que sistematiza. El modelo, en este sentido, permite la recopilación de datos acústicos que se asocian con diferentes partes de los enunciados que pretende explicar. De esta

manera, las sílabas más importantes se sitúan en la zona de la anacrusis (inflexión tonal hasta la primera sílaba tónica), desde esta hasta el tonema (declinación o inclinación en el cuerpo del grupo entonativo) y, finalmente, el tonema, que incluye la última sílaba tónica y su posterior inflexión.

Por lo tanto, no solo disponemos de una etiqueta de análisis (interrogativo, enunciativo, PI, PXIIa, etc.), sino también los datos de duración e inflexión tonal asociados a las sílabas tónicas de las zonas que comentábamos: la anacrusis y el tonema. A partir de estos datos acústicos, pueden realizarse pruebas estadísticas de contraste inferencial entre grupos; en el caso de la presente investigación, como señalábamos en la sección de Metodología, solo buscamos una visión exploratoria a partir de los datos, aunque algunas pruebas estadísticas sí permiten realizar labores de representatividad poblacional de los datos y, así, reducir su grado de aleatoriedad (LEVSHINA, 2015; GRIES, 2016). Un ejemplo de ello puede verse en la Tabla 6, que recoge los valores de la prueba ANOVA:

spk	Diff	p.adj
dufort-angelreal	9.432014	0.616
estepa-angelreal	42.952176	0.000
pzorro-angelreal	2.231150	0.994
rcastellano-angelreal	-9.496798	0.620
estepa-dufort	33.520162	0.000
pzorro-dufort	-7.200864	0.632
rcastellano-dufort	-18.928812	0.032
pzorro-estepa	-40.721027	0.000
rcastellano-estepa	-52.448974	0.000
rcastellano-pzorro	-11.727948	0.172

TABLA 6 - Resultados de la prueba ANOVA que contrasta los valores de duración de la primera vocal tónica con los hablantes del corpus.

Fuente: Elaboración propia.

Como puede verse en la Tabla 6, la prueba ANOVA, mediante el contraste post hoc HSD Tukey⁶, sí encuentra un factor discriminatorio entre los sujetos comparados. Se trata, en este caso, de los valores de duración asociados a la vocal de la primera sílaba tónica del contorno. Los datos de significación estadística o valor p inferiores a 0.05, con un nivel de significación del 95 %, señalan diferencias en 5 de los 10 cruces posibles entre los hablantes. Concretamente, el deportista que más se separa del resto, con una duración superior, es Estepa; también otros dos deportistas, Castellano y Dufort, se diferencian entre ellos. La distribución numérica de la variable duración de la vocal de la primera sílaba tónica puede verse mejor en los diagramas de caja de la Figura 7:

⁶ La prueba ANOVA es una prueba de análisis de la varianza de una variable que pretende comparar los valores medios de una variable, por ejemplo, duración, para distintos grupos, por ejemplo, hablantes; si los grupos presentan valores muy similares se considera que, para esa variable, no constituyen grupos realmente distintos, mientras que si lo contrario sucede, puede concluirse que esa variable diferencia significativamente los grupos con valores más distintos entre sí. Puede encontrarse más información sobre esta prueba en obras recientes sobre estadística en lingüística (LEVSHINA, 2015).

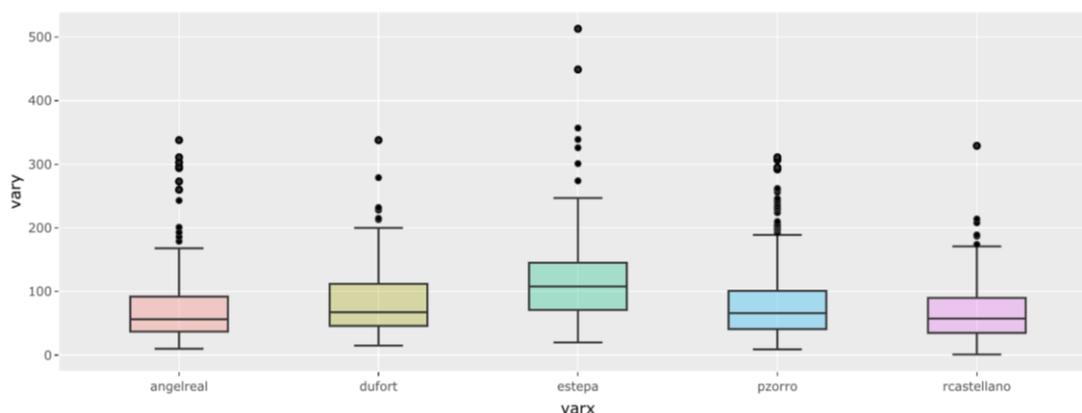


FIGURA 7 - Diagramas de caja para la variable duración de la vocal de la primera sílaba tónica por hablante.

Fuente: Elaboración propia.

En la Figura 7 se puede observar que la variable duración de la vocal tónica de la primera sílaba tónica del enunciado, situada en la anacrusis, presenta distintos valores según el emisor. El participante con valores más altos sería Estepa (0.12 segundos de media, 0.11 segundos mediana), mientras que Dufort presentan una duración algo menor (sobre 0.09 segundos de media y 0.07 de mediana). Por su parte, Padre Zorro, con 0.08 de media y 0.07 de mediana, y Angelreal, con 0.08 de media y 0.06 de mediana, presentan valores similares; finalmente Castellano, con 0.07 de media y 0.06 de mediana, es el hablante que presenta una duración menor.

La distribución, en este caso, señala que hay hablantes que enfatizan esa primera sílaba tónica, mientras que otros, por el contrario, la ralentizan. Curiosamente, el fenómeno de la duración sí ha resultado significativo para la discriminación de hablantes en la fonética segmental, como era el caso de la duración de las sibilantes, o del *voice onset time* de las oclusivas, por señalar solo un par de ejemplos (CICRES I BOSCH, 2003; DELGADO *et al.*, 2009; NOLAN, 2009); en el caso del presente estudio, ha sido interesante encontrar que una sílaba, ubicada en una parte estructural del grupo entonativo, puede presentar valores en un rango de distribución tan significativamente distinto.

De esta manera, se demuestra que el análisis melódico del habla no solo es un sistema de estudio entonativo que se base en el tonema, sino que, a diferencia de otros modelos como el ToBI, donde esta parte queda algo más descuidada o desatendida en los estudios, sí realiza una disposición estructural que atiende a la realidad de las unidades entonativas que se analizan, dando un papel importante tanto al principio, al medio, como, también hay que decirlo, al final ocupado por el tonema.

Algunos artículos científicos, como este que presentamos, se enfocan en presentar la metodología utilizada y los resultados obtenidos, sin profundizar en la discusión de los mismos. En este caso, la sección de discusión puede ser omitida porque el objetivo principal es describir el proceso utilizado para obtener los resultados. De todos modos, no existen en la bibliografía enfoques que vinculen el análisis melódico del habla a patrones idiolectales como marca de identificación de hablantes, por lo que, al ser una línea de investigación que se abre en este estudio, deberán ser los próximos estudios e investigaciones quienes determinen futuras discusiones entre ellos.

4. Valoraciones finales

En este artículo hemos presentado una primera cala exploratoria para determinar en qué medida el sistema de transformación y codificación porcentual proporcionado por el análisis melódico del habla puede contribuir a la identificación de hablantes. Aunque se trata de un sistema claramente destinado a la comparación de patrones melódicos completos, es cierto que su división en tres partes estructurales (anacrusis, cuerpo y tonema) también permite tener en consideración otros valores relevantes, como el desplazamiento tonal o su duración.

En los datos analizados (5 hablantes, ocho archivos, 16 muestras de habla distintas procedentes de dos géneros discursivos, 1218 grupos de entonación, 8432 palabras y 33 minutos de habla efectiva sin pausas), se han encontrado algunas regularidades, si bien no en aquellos factores de la curva melódica que, a priori, parecían destinados a ser mucho más importantes de lo que al final han resultado ser. En general, los resultados obtenidos en esta investigación de corpus, son los siguientes:

1. Los 12 patrones melódicos encontrados (PI, PII, PIII, PIX, PV, PVIa, PVIb, PVII, PVIII, PXa, PXI, PXIIa) se reparten de modo similar tanto entre los hablantes como entre los géneros discursivos.
2. Esos 12 patrones melódicos corresponden a los cuatro tonemas establecidos por Cantero (CANTERO; FONT-ROTCHÉS, 2007) y también se distribuyen de manera proporcionalmente equivalente entre los 5 hablantes, independientemente del género discursivo. El tonema más generalizado es el enunciativo (234 casos), seguido del suspendido (199), el enfático (102) y el interrogativo (solo 11 casos).
3. La duración de la primera sílaba tónica, justo en la sección de la anacrusis, sí ha resultado ser estadísticamente significativa para diferenciar a muchos de los hablantes entre sí, y dejando manifiesto que también se configura como una marca idiosincrásica en algún caso, sobre todo en las muestras de habla del youtuber/culturista Sergio Estepa, quien se diferencia del resto mediante una duración más amplia de esa sílaba tónica de la anacrusis.

De esta manera, esta investigación ha explorado la pertinencia del sistema de análisis melódico del habla para el estudio de diferencias o similitudes particulares entre hablantes. El estudio deja abierto un interesante campo de estudio en el que pueden efectuarse una serie de contribuciones epistemológicas. En primer lugar, la aplicación de métodos de computación avanzados, como el efectuado mediante el uso del programa *Oralstats* (CABEDO, 2021), permite analizar de un modo económico y ágil grandes cantidades de datos; dado que el proceso de codificación y transformación es homogéneo para todos los datos, se asegura que la posterior comparación estadística entre ellos

(y más si se relaciona con valores de identificación de hablantes) sea coherente. En todo caso, es aún un programa en fase de maduración y, por ello, el margen de mejora es todavía amplio.

En segundo lugar, el estudio forense de hablantes, incluso en el marco de la estilometría oral, puede beneficiarse de las sólidas estructuras de análisis de un modelo como el análisis melódico del habla (AMH). Al mismo tiempo, y a partir de los resultados obtenidos de material sonoro real y espontáneo, pueden añadirse otros factores acústicos que puedan ofrecer alternativas de análisis y que mejoren en lo práctico las valoraciones concretas que se generen. Ya en la actualidad el propio análisis melódico del habla ha introducido nuevos factores, como la intensidad o los factores de duración (CANTERO, 2019). De este modo, mientras que análisis generales de estudio de la entonación como el ToBI se centran en los valores de tono (PIERREHUMBERT, 1980; ESTEBAS; PRIETO, 2008; PRIETO; RIGAU, 2011), otros sistemas, como el AMH, presentan una perspectiva de estudio algo más flexible y escalable en términos computacionales, esto es, pueden incorporar progresivamente a su modelo nuevas capas que potencian el análisis tonal de base; estas capas, por lo tanto, añadirán factores como la intensidad, la duración o la velocidad de habla.

En tercer lugar, el análisis melódico del habla, a nuestro juicio, es perfecto para el estudio fonético de hablantes, dentro de la perspectiva forense, ya que, como señalábamos previamente, es un sistema basado en valores acústicos con una base objetivable y medible. A nivel computacional, facilita la programación de estos valores y, posteriormente, permite realizar operaciones de cómputo estadístico, como las que hemos presentado en este estudio, en las que el objetivo ha sido determinar en qué medida los hablantes pueden diferenciarse o no entre ellos según los patrones melódicos utilizados y el comportamiento acústico de cada una de las partes que los componen.

Agradecimientos

Agradecemos enormemente la revisión realizada por los revisores de este artículo, ya que han permitido mejorar sustancialmente el resultado final.

Informaciones complementarias

Evaluación y respuesta de los autores

Evaluación: <https://doi.org/10.25189/rabralin.v21i2.2103.A>

Respuesta de los autores: <https://doi.org/10.25189/rabralin.v21i2.2103.R>

Conflicto de intereses

El(los) autor(es) no tiene(n) intereses contrapuestos que declarar.

Enlace a preimpresión

<https://doi.org/10.5281/zenodo.7263620>

Protocolo y pre-registro de investigación

He revisado los estándares de Equator Network. No encontré ningún estándar aplicable a las humanidades. Esta investigación no fue pre registrada.

Declaración y disponibilidad de los datos

Los datos, códigos y materiales que respaldan los hallazgos de este estudio están disponibles de forma abierta en https://github.com/acabedo/abralin_1. Los dos archivos disponibles en este repositorio incluyen un módulo especial para Oralstats (<https://github.com/acabedo/oralstats>), diseñado para codificar el modelo Melodic Analysis of Speech, y el dataset/dataframe de este estudio en formato xlsx. Otros materiales, como audios, videos y transcripciones de PRAAT están disponibles a pedido del autor correspondiente.

REFERENCIAS

BLOCH, B. A Set of Postulates for Phonemic Analysis. *Language*, v. 24, n. 1, p. 3-46, 1948. DOI <<https://doi.org/10.2307/410284>>.

BOERSMA, P.; WEENINK, D. Praat: doing phonetics by computer [Computer program]. Disponible em <<http://www.praat.org/>>. Acceso em: outubro 2022.

CABEDO NEBOT, A. Oralstats. Disponible em <<https://github.com/acabedo/oralstats>>. Acceso em: outubro 2021.

CANTERO SERENA, F. J. *Teoría y análisis de la entonación*. Barcelona: Edicions de la Universitat de Barcelona, 2002.

CANTERO, F. J. De la fonética del habla espontánea a la fonología de la complejidad. *Normas: revista de estudios lingüísticos hispánicos*, n. 5, p. 9-29, 2015.

CANTERO, F. J. Análisis prosódico del habla: más allá de la melodía. In: ÁLVAREZ, M. R.; MUÑOZ, A.; RUIZ, L. (Ed.). *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines*. Santiago de Cuba: Ediciones Centro de Lingüística Aplicada, 2019. p. 485-498.

CANTERO, F. J.; FONT-ROTCHÉS, D. Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia*, n. 13, p. 699-715, 2007.

CANTERO, F.J.; FONT-ROTCHÉS, D. Protocolo para el análisis melódico del habla. **Estudios de fonética experimental**, v. 18, nov. 2009. Disponible em: <<https://raco.cat/index.php/EFE/article/view/140087>>.

CHANG, W. *et al.* Shiny: Web Application Framework for R. 2022. Disponible em: <<https://CRAN.R-project.org/package=shiny>>. Acceso em: outubro 2022.

CICRES I BOSCH, J. Estudi pilot d'identificació de parlants amb finalitat forense. **Papers de l'IULA: Sèrie Monografies**, n. 7, 2003.

CICRES I BOSCH, J.; GAVALDÀ, N. La lingüística forense: la llengua com a evidència. **Revista de llengua i dret**, n. 61, p. 60-71, 2014.

CORVALÁN, C. S.; ENRIQUE-ARIAS, A. **Sociolingüística y pragmática del español**. Washington: Georgetown University, 2017.

DELGADO, C. Pasaporte vocal: utilidad de la estratificación del uso lingüístico. **Ciencia policial: revista del Instituto de Estudios de Policía**, n. 40, p. 57-89, 1998.

DELGADO, C. *et al.* Identificación forense de locutores (I.F.L.): categorización de parámetros acústicos y fonarticulatorios del español. **Revista Española de Lingüística**, v. 39, n. 1, p. 33-60, 2009.

DELGADO, C. La pericia de identificación de habla: el papel fundamental del experto. In: GARAYZÁBAL, E.; JIMÉNEZ, M.; REIGOSA, M. **Lingüística Forense: La lingüística en el ámbito legal y policial**. Madrid: Euphonia Ediciones, 2012. p. 199-212.

DELGADO, C. Comparación forense del habla: El cambio conceptual. **Ciencia policial: revista del Instituto de Estudios de Policía**, n. 158, p. 61-74, 2020.

ELVIRA-GARCÍA, W. *et al.* A Tool for Automatic Transcription of Intonation: Eti_ToBI a ToBI Transcriber for Spanish and Catalan. **Language Resources and Evaluation**, v. 50, n. 4, p. 767-792, 2016. DOI <<https://doi.org/10.1007/s10579-015-9320-9>>.

ESTEBAS, E.; PRIETO, P. La notación prosódica del español: una revisión del Sp- ToBI. **Estudios de fonética experimental**, n. 17, p. 263-283, 2008.

FERNÁNDEZ, F. M. **Metodología sociolingüística**. Madrid: Editorial Gredos, 1990.

GRIES, S. **Quantitative corpus linguistics with R: A practical introduction**. Routledge; 2016.

JESSEN, M. Speaker profiling and forensic voice comparison. In: COULTHARD, M. **The Routledge Handbook of Forensic Linguistics**. Routledge. 2020, p. 378-395. DOI <<https://doi.org/10.4324/9780429030581-31>>.

LEEMANN, A. *et al.* The Role of Segments and Prosody in the Identification of a Speaker's Dialect. **Journal of Phonetics**, v. 68, p. 69-84, 1 maio 2018. DOI <<https://doi.org/10.1016/j.wocn.2018.02.001>>.

LEVSHINA, N. **How to do Linguistics with R: Data exploration and statistical analysis**. London: John Benjamins Publishing Company, 2015. DOI <<https://doi.org/10.1075/z.195>>.

MATEO, M. Scripts en praat para la extracción de datos tonales y curva estándar. **Phonica**, n. 6, 2010.

NOLAN, F. **The Phonetic Bases of Speaker Recognition**. Cambridge: Cambridge University Press, 2009.

PIERREHUMBERT, J. **The Phonology and Phonetics of English Intonation**. Phd Thesis, 1980.

PRIETO, P.; RIGAU, G. Prosody and pragmatics. In: PAYRATÓ, L.; COTS, J. **The Pragmatics of Catalan**. Berlin-Boston: De Gruyter Mouton, 2011. p. 17-48. DOI <<https://doi.org/10.1515/9783110238693.17>>.

QUILIS, A.; CANTARERO, M.; ESGUEVA, M. El grupo fónico y el grupo de entonación en español hablado. **Revista de Filología Española**, v. LXXIII, p. 55-65.