

PROCESAMIENTO DEL ANÁLISIS MELÓDICO DEL HABLA EN EL CORPUS PRESEEA (NIVEL ALTO): DIFICULTADES DESDE UNA PROPUESTA COMPUTACIONAL

Adrián Cabedo Nebot

Universitat de València (España)

adrian.cabedo@uv.es

<https://orcid.org/0000-0002-3881-9308>

Recibido: 14/3/2024. Aceptado: 6/4/2024. Publicado: 31/7/2024

Cita recomendada:

Cabedo Nebot, A. (2024). Procesamiento del análisis melódico del habla en el corpus PRESEEA (nivel alto): dificultades desde una propuesta computacional. *Phonica* 20, 1-25. <https://doi.org/10.1344/phonica2024.20.2>

Resumen

Este trabajo explora el procesamiento computacional del Análisis Melódico del Habla (AMH) en un corpus de habla sociolingüístico, el corpus PRESEEA (nivel de instrucción alto). La investigación destaca la importancia de los entornos computacionales para el análisis fonético, centrándose en el estudio del tono, la intensidad, y la duración del habla y su relación con los grupos entonativos en español. Mediante herramientas avanzadas como ORALSTATS, se demuestra la eficacia de la metodología del Análisis Melódico del Habla (AMH) para el análisis fonético en grandes volúmenes de datos. A pesar de las ventajas de estas herramientas de procesamiento automático, el estudio también señala desafíos y limitaciones en la precisión y interpretación de los datos obtenidos. Los resultados enfatizan la complejidad de la prosodia y su variabilidad según factores sociolingüísticos, abriendo nuevas direcciones para futuras investigaciones en lingüística computacional y fonética.

Palabras clave: Análisis Melódico del Habla (AMH); Corpus PRESEEA; Prosodia; ORALSTATS; Español



© 2024. Los autores. Este artículo es de acceso abierto y está sujeto a la licencia de [Reconocimiento 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), de Creative Commons, que permite utilizar, distribuir y reproducir la obra original por cualquier medio sin restricciones, siempre y cuando se cite adecuadamente.

Processament de l'Anàlisi Melòdica de la Parla en el corpus PRESEEA (nivell alt): dificultats des d'una proposta computacional

Resum: Aquest treball explora el processament computacional de l'Anàlisi Melòdica de la Parla (AMP) en un corpus de parla sociolingüístic, el corpus PRESEEA (nivell d'instrucció alt). La investigació destaca la importància dels entorns computacionals per a l'anàlisi fonètica, centrant-se en l'estudi del to, la intensitat i la duració de la parla i la seva relació amb els grups entonatius en espanyol. Mitjançant eines avançades, com ORALSTATS, es demostra l'eficàcia de la metodologia de l'Anàlisi Melòdica de la Parla (AMP) per a l'anàlisi fonètica en grans volums de dades. Malgrat els avantatges d'aquestes eines de processament automàtic, l'estudi també assenyala desafiaments i limitacions en la precisió i la interpretació de les dades obtingudes. Els resultats emfasitzen la complexitat de la prosòdia i la seva variabilitat segons factors sociolingüístics, obrint noves direccions per a futures investigacions en lingüística computacional i fonètica.

Paraules clau: Anàlisi Melòdica de la Parla (AMP); Corpus PRESEEA; Prosòdia; ORALSTATS; Espanyol

Processing of the Melodic Analysis of Speech (MAS) in the PRESEEA corpus: a challenge from a computational perspective

Abstract: *This study explores the computational processing of the Melodic Analysis of Speech (MAS) in a corpus of sociolinguistic speech, the PRESEEA corpus (high level of education). The research highlights the importance of computational environments for phonetic analysis and focuses on the study of tone, intensity and duration of speech and their relationship with intonation groups in Spanish. Using advanced tools such as ORALSTATS, the efficacy of the Melodic Analysis of Speech (MAS) methodology for the phonetic analysis of large volumes of linguistic data is demonstrated. Despite the advantages of these computational tools, the study also points out challenges and limitations in the accuracy and automatic interpretation of prosody. The results emphasize the complexity of prosody and its variability according to sociolinguistic factors, opening new directions for future research in computational linguistics and phonetics.*

Keywords: *Melodic Analysis of Speech (MAS); PRESEEA corpus; Prosody; ORALSTATS; Spanish*

1. Introducción

La presente investigación pretende resaltar la utilidad de entornos de análisis computacionales para la obtención y análisis de datos fonéticos, con un interés particular en el estudio del tono, la intensidad y la duración, y en cómo estos se relacionan y adaptan a los grupos entonativos del español.

A través de este estudio, se pretende demostrar la eficacia de codificar configuraciones entonativas desde un modelo de análisis específico, como por ejemplo las que pueden extraerse al emplear el modelo teórico del Análisis Melódico del Habla (AMH), una metodología desarrollada inicialmente por Cantero (2002), enriquecida posteriormente mediante su colaboración con Font-Rotchés (Cantero y Font-Rotchés, 2009), y aplicada en investigaciones posteriores (Devís, 2011, 2012; Mateo, 2010; Torregrosa, 2015; Sola y Torregrosa, 2023, entre muchos otros). Este trabajo subraya, por tanto, el valor que puede desempeñar una herramienta computacional en la extracción y análisis de datos fonéticos dentro del contexto del AMH y, concretamente, al aplicarlos a ámbitos de grandes datos, como los que se detectan en algunos corpora lingüísticos.

En tal sentido, este artículo expone las posibilidades que un sistema computacional ofrece para el procesamiento de *big data*, tanto en el procesamiento de grandes volúmenes de datos como en la visualización de los resultados obtenidos. Si bien se utiliza el corpus PRESEEA Valencia (Gómez Molina, 2001) como caso de estudio, el propósito no es adentrarse en un análisis detallado de estratos sociolingüísticos en esta ocasión. En su lugar, el corpus PRESEEA sirve en esta investigación como ejemplo práctico para ilustrar la metodología computacional empleada en el análisis de datos mediante el uso de R Core Team (2023), y también para ofrecer una panorámica general sobre los procesos correctos de análisis, pero, al mismo tiempo, para incidir en las dificultades y deficiencias que procesos automáticos generan en los resultados finales, sobre todo cuando se encuentran o detectan patrones entonativos que no habían sido registrados inicialmente.

En cuanto a los datos que se analizan, el corpus PRESEEA (Proyecto para el Estudio Sociolingüístico del Español de España y de América) constituye una invaluable fuente de datos lingüísticos para la investigación sociolingüística del español hablado en diversas regiones de España y América Latina. PRESEEA es una herramienta esencial para los investigadores interesados en la sociolingüística y el análisis de las variedades regionales del español, ya que proporciona acceso a datos orales recopilados en una variedad de contextos sociales y situaciones cotidianas, reflejando de manera fidedigna el uso del español en la comunicación oral a través de diferentes estratos sociales, edades y contextos socioculturales. La iniciativa PRESEEA se destaca como un pilar fundamental para el análisis fonético del español, al ofrecer una plataforma sólida para la investigación sobre la variación y el cambio lingüístico en el ámbito hispanohablante. Gracias a este proyecto, ha sido posible recopilar y analizar de manera exhaustiva datos lingüísticos de diversas ciudades hispanohablantes, aspecto que facilita un contraste entre distintas variedades del español.

Actualmente, el corpus PRESEEA de la ciudad de Valencia recoge un número de entrevistas amplio y que no ha sido procesado previamente para facilitar el estudio fonético. Con la intención de realizar esa actividad, la metodología de esta investigación no solo subraya el valor de herramientas innovadoras como ORALSTATS (Cabedo, 2022) en el ámbito de la investigación lingüística, sino que además resalta la relevancia del Análisis Melódico del Habla (AMH) como un marco teórico fundamental para el estudio profundo de la entonación en el español. Al introducir prácticas de análisis de datos utilizando R, se proporciona una plataforma robusta para futuras investigaciones dedicadas a examinar patrones entonativos en extensos conjuntos de datos fonéticos.

Esta estrategia metodológica amplía significativamente las capacidades de los investigadores para procesar y analizar grandes cantidades de información, permitiendo un acercamiento más detallado y matizado a las variaciones de entonación y a su impacto en la comunicación verbal. Al emplear R, un entorno de software libre para la computación estadística y gráficos, los investigadores pueden aplicar técnicas de análisis cuantitativo y cualitativo avanzadas. Esto abre nuevas posibilidades para descubrir tendencias y patrones en la entonación que antes podrían haber pasado desapercibidos debido a las limitaciones de los métodos de análisis más tradicionales.

Además, el uso de programas como ORALSTATS (Cabedo, 2022) facilita la automatización del proceso de análisis, reduciendo el tiempo y el esfuerzo necesarios para examinar vastos volúmenes de datos. Esto no solo hace que la investigación sea más eficiente, sino que también aumenta la precisión de los resultados, minimizando el potencial error humano en el procesamiento de los datos. La capacidad de manejar grandes bases de datos con eficacia abre el camino hacia investigaciones más complejas y a gran escala, las cuales pueden arrojar luz sobre las sutilezas de la entonación en variadas comunidades de habla y contextos comunicativos.

2. Sobre la propuesta de análisis melódico o prosódico del habla

El análisis melódico del habla, que inicialmente consideraba solo el comportamiento de la melodía y del flujo tonal, ha sido ampliado más recientemente en lo que se conoce como *análisis prosódico del habla*, como ha sido denominado por Cantero (2019); en este caso, este tipo de análisis constituye una metodología rigurosa que busca explorar las complejidades de la entonación lingüística y que no solo se mantiene en los límites del tono o variación acústica de la frecuencia fundamental, sino que añade otros valores acústicos, como la intensidad o la duración de los pies rítmicos (duración existente desde el pico de intensidad de una vocal hasta el pico de la siguiente vocal).

El modelo inicial de análisis melódico se define en Cantero (2002), donde se sugiere que, para comprender adecuadamente la entonación, es crucial identificar los segmentos tonales relevantes en la cadena hablada. Esto implica una serie de pasos meticulosos, como la estandarización de los resultados y la relativización de los valores tonales para permitir una comparación precisa entre distintos hablantes y contextos lingüísticos. Esta estandarización, dentro del análisis prosódico del habla, se extiende a los valores de intensidad y de pie rítmico. En palabras de Cantero y Font-Rotchés (2007), «los valores absolutos del tono están condicionados por las características fisiológicas de los hablantes», lo que hace difícil comparar directamente los datos obtenidos de diferentes individuos. Por lo tanto, se vuelve imprescindible relativizar los valores tonales, convirtiéndolos en porcentajes o intervalos relativos, lo que facilita la descripción de las melodías y su comparación entre diferentes hablantes:

Para ello, se identifican las vocales y los segmentos tonales que en ella se producen, apoyándonos en el sonograma, y se recoge el valor medio -si es estable- o el valor central. En el caso de vocales tónicas puede producirse una inflexión y existir dos y hasta tres segmentos tonales inflexión circunfleja-; estos valores se toman de los extremos de la inflexión. (Mateo, 2010, p. 51)

Siguiendo la metodología propuesta, Cantero (2016) enfatiza la necesidad de definir con precisión las reglas y criterios que rigen el análisis melódico, también conocido como *análisis prosódico del habla*. Esta precisión metodológica es esencial para garantizar la fiabilidad y replicabilidad de los resultados obtenidos en la investigación fónica. Establecer un conjunto de normas bien definidas y universalmente aplicables asegura que los hallazgos no solo sean sólidos y válidos

dentro de un estudio particular, sino que también puedan ser corroborados y extendidos por otros investigadores en diferentes contextos lingüísticos y culturales.

La estandarización mencionada por Cantero (2002, 2019) juega un rol crucial en la construcción de una metodología coherente y sistemática, permitiendo que los análisis de prosodia se realicen con un marco común de referencia. Esto no solo facilita la comparación de resultados entre estudios diferentes, sino que también posibilita la acumulación de conocimientos sobre cómo la entonación afecta la comunicación en una diversidad de idiomas más allá del español. Tal uniformidad en el enfoque metodológico es indispensable para avanzar en el entendimiento de los patrones entonativos y su variabilidad, tanto en la dimensión intralingüística como en la interlingüística. Un ejemplo de patrón melódico se proyecta en la Figura 1, concretamente el patrón enunciativo:

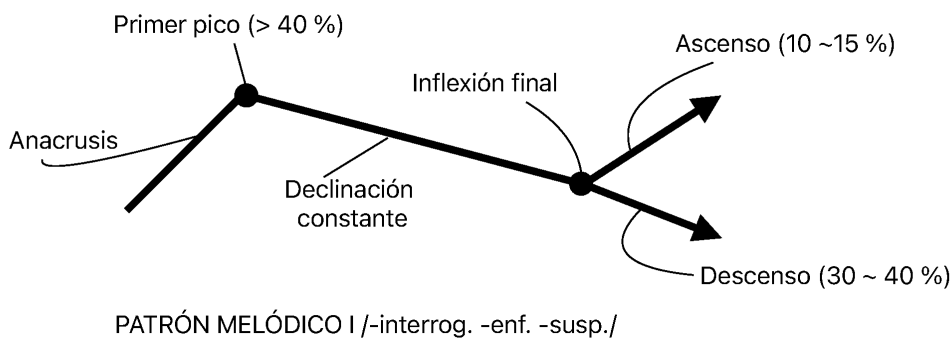


Figura 1. Imagen del patrón melódico I (enunciativo), basado en Cantero y Font-Rotchés (2007)

Como se deduce de lo anterior, el AMH se basa en principios sólidos y reglas que permiten un análisis riguroso y objetivo de la entonación lingüística. Una metodología de estas características es fundamental para avanzar en nuestro entendimiento de la prosodia y su papel en la comunicación cotidiana. En general, el AMH permite a un investigador de la prosodia alcanzar los siguientes objetivos:

1. Identificar patrones melódicos. El AMH permite identificar y describir patrones tonales, como tonos ascendentes, descendentes, suspendidos y circunflejos, que pueden indicar multitud de valores pragmáticos (Cantero y Font-Rotchés, 2007; Font-Rotchés y Mateo, 2013).
2. Analizar la estructura de la entonación. El modelo permite descomponer el discurso en unidades entonativas o grupos fónicos que pueden desglosarse en partes (*anacrusis*, *cuerpo* e *inflexión final*). La *anacrusis* se define como el conjunto de vocales previas al primer pico tonal, que suele consolidarse generalmente alrededor de la primera sílaba tónica del grupo entonativo. El *cuerpo* de la unidad entonativa se extiende desde el primer pico hasta el último pico o *núcleo*, que suele coincidir con la última sílaba tónica. Y la *inflexión final* se inicia en el núcleo y llega al final del contorno.
3. Comprender la intención comunicativa. El AMH puede revelar cómo las variaciones en la entonación reflejan la intención del hablante, como enfatizar información, expresar sorpresa o indicar incertidumbre, es decir, permite observar la transmisión de valores pragmáticos; un ejemplo de ello se ha podido ver al aplicarse al estudio de la cortesía y atenuación del español (Devís, 2011, 2012).

4. Investigación lingüística. Se utiliza en investigaciones lingüísticas para comprender las diferencias entonativas en diferentes dialectos o lenguas, así como para estudiar cómo la entonación varía en diferentes contextos comunicativos; por ejemplo, puede mencionarse el trabajo de Sola y Torregrosa sobre el inglés (Sola y Torregrosa, 2023).

3. Necesidad de *software* auxiliar en el análisis de la entonación

La necesidad de utilizar *software* en lingüística, especialmente en el dominio de la fonética, se ha hecho cada vez más evidente a medida que aumentaba el interés por el estudio del sonido y por su vinculación con la expresión de valores más o menos codificados en el nivel lingüístico. La aplicación de herramientas de *software* en disciplinas como la lingüística facilita el análisis de datos a una escala y con una precisión que los métodos manuales no pueden igualar (Llisterri, 1991; Mota et al., 2005; Estruch et al., 2007), particularmente en fonética, donde el análisis acústico detallado es esencial para comprender los sonidos del habla.

Herramientas de *software*, como Praat (Boersma y Weenink, 2022), permiten a los investigadores analizar ondas sonoras, extraer características fonéticas y visualizar patrones de habla de maneras que son cruciales para el análisis fonético avanzado. Por ejemplo, Mateo (2010) demostró cómo se pueden utilizar protocolos y scripts en Praat para la extracción de datos tonales y curvas estándar en el Análisis Melódico del Habla (AMH), subrayando la importancia del software en la realización de estudios fonéticos detallados. De hecho, en los propios scripts diseñados por Mateo (2010) se alerta sobre inconsistencias del propio Praat a la hora de ofrecer puntos tonales sobre los audios analizados.

De manera similar, la creación de corpora especializados, como *Glissando*, discutida por Garrido et al. (2013), se hace posible mediante el uso de herramientas computacionales. Estos corpora sirven como recursos fundamentales para estudios prosódicos multidisciplinares en idiomas como el español y el catalán, ilustrando cómo el *software* habilita la compilación, análisis y diseminación de datos lingüísticos a gran escala. Además, el trabajo de Garrido y Chica (2018) enfatiza el papel de los grandes corpora y herramientas computacionales en la descripción de la prosodia, presentándolo como un desafío emocionante para el futuro. El uso de programas informáticos en lingüística no solo aborda problemas pendientes, sino que también abre nuevas vías para la investigación al mejorar la capacidad del investigador para analizar datos lingüísticos de manera más completa.

El aspecto de programación, como discute Anthony (2020) en el contexto de la lingüística de corpus, es integral para el desarrollo y personalización de herramientas adaptadas a necesidades específicas de investigación lingüística. Esto enfatiza el papel del software no solo como una herramienta, sino también como un conjunto de habilidades que los lingüistas deben desarrollar para avanzar en su campo.

En cuanto a la metodología aplicada o los programas utilizados para el análisis prosódico, algunas propuestas establecen límites de delimitación tonal según fronteras de tipo perceptivo. En tal sentido, cabe mencionar aquí proyectos computacionales para analizar y catalogar la curva melódica, como sería el caso del script de Praat ETI-TOBI, diseñado y presentado por Elvira et al. (2016) para el reconocimiento de patrones entonativos en el habla, a partir de principios sólidos y reglas objetivas. Según Elvira et al. (2016), «se utiliza un umbral de 1.5 semitonos para determinar la significancia de los movimientos de F0», lo que se ha demostrado efectivo para la percepción de la entonación en español y otros idiomas. Además, Elvira et al. (2016) establecen reglas específicas para la asignación de etiquetas de tonemas, lo que permite etiquetar las *frases entonacionales* de manera sistemática y objetiva:

For being labelled as high, the value needs to be above the high third of the range of the utterance. If there is movement, the script follows a series of rules that go from the least restrictive to the most restrictive to determine which label has to be assigned. For example: if from the F0 valley to the F0 peak, there is a difference greater than 1.5 semitones, a L*? H label applies but, if there is also a difference greater than 1.5 between the start point and the end point of the stressed syllable, a L? H* label is written. If the previous two are true, but the target of the movement is in the posttonic syllable, the label will be L + >H* (Elvira et al. 2016: 777:778)

Así pues, la necesidad de utilizar software en lingüística, particularmente en fonética, radica en su capacidad incomparable para facilitar el análisis detallado, la visualización y el manejo de datos lingüísticos. Esta integración tecnológica no solo mejora la precisión y eficiencia de la investigación científica, sino que también expande los límites de lo que se puede lograr en la comprensión de las complejidades del lenguaje. En ese sentido, la importancia del software se extiende a la visualización de fenómenos fonéticos, como destacó Cabedo (2022), quien utilizó el módulo de visualización de la herramienta ORALSTATS para ilustrar la melodía con múltiples valores acústicos y de etiquetado. Esta capacidad de visualización es crucial para entender patrones fonéticos complejos y comunicar hallazgos de manera efectiva.

Precisamente ORALSTATS, que se ampliará específicamente en la sección 4.2, incorpora tanto codificación en sistema ToBI como incorporación en sistema del AMH, a partir de la incorporación de un trabajo computacional previo, desarrollado por Mateo (2010):

El siguiente paso es determinar cuántos valores tonales significativos hay en el segmento; en el caso del castellano y catalán se ha verificado en diversos experimentos perceptivos realizados en el Laboratorio de Fonética Aplicada de la Universidad de Barcelona que este umbral es del 10%, por ello verificamos la diferencia existente entre los valores mencionados y en función de dicha diferencia grabamos los valores: - Un valor: principal, media: si las diferencias entre los valores tonales son inferiores al 10%. - Dos valores, F0 inicial y F0 final: si las diferencias entre ellos son superiores al 10% y no lo son respecto al máximo y mínimo. - Dos valores, F0 inicial y F0 mínimo: si la diferencia entre F0 mínimo e inicial es superior al 10% pero no la diferencia entre mínimo y final. - Dos valores, F0 inicial y F0 máximo: si la diferencia entre F0 inicial y máximo es superior al 10%, pero no la diferencia entre los valores máximo y final. - Tres valores, F0 inicial, mínimo y final: si las diferencias entre los tres valores son superiores al 10% - Tres valores, F0 inicial, máximo y final: si las diferencias entre los tres valores son superiores al 10%. (Mateo, 2010:59)

En el campo del análisis prosódico, a pesar de los numerosos avances en la generación de sistemas capaces de transcribir la entonación, las implementaciones automáticas actuales aún no incorporan de manera efectiva la interpretación semántica de las secuencias entonativas en sus soluciones de *software*. Este déficit se puede atribuir a dos factores críticos. En primer lugar, existe una dificultad significativa en la adquisición de datos de tono vocal precisos y fiables, esenciales para la decodificación adecuada de la entonación. En segundo lugar, la separación conceptual y operativa entre los ingenieros informáticos, responsables de desarrollar la infraestructura técnica de estos sistemas, y los lingüistas, que investigan cuestiones como la prosodia, constituye un obstáculo considerable. Esta distancia se ve aumentada por la diversidad de modelos teóricos existentes sobre la prosodia, cuya naturaleza abstracta complica aún más la integración de estos conocimientos en aplicaciones tecnológicas concretas.

4. Metodología

Esta sección de metodología presenta el enfoque adoptado para el análisis del corpus y la codificación de la melodía del habla utilizando una herramienta informática específica, el programa ORALSTATS. Inicialmente, se describe la selección y composición del corpus

analizado, enfatizando su relevancia y representatividad para el estudio de la entonación en contextos lingüísticos específicos. Se proporcionan detalles sobre el origen, la naturaleza y el tamaño del corpus, así como los criterios utilizados para su selección. Posteriormente, se explica el proceso y las técnicas empleadas para la codificación de la melodía del habla, incluyendo una descripción exhaustiva del software utilizado para este fin.

4.1 Corpus utilizado

Para esta investigación se ha utilizado una muestra del corpus PRESEEA Valencia (Gómez Molina, 2001). Concretamente, se han utilizado 24 entrevistas del nivel sociocultural alto, estratificadas además por sexo y nivel de instrucción de la siguiente manera:

sexo	edad	entrevistas	palabras	vocales	hablantes
hombre	18-35	4	16607	30938	4
hombre	35-55	4	20223	38970	4
hombre	Superior 55	4	20861	41304	4
mujer	18-35	4	15332	28442	4
mujer	35-55	4	24886	46750	4
mujer	Superior 55	4	19876	37474	4

Tabla 1. Distribución de entrevistas, palabras y fonos por sexo y edad.

En la Tabla 1, de los 24 hablantes analizados, se observa que las 4 mujeres de 35-55 años son las hablantes con más registros sonoros, con 24886 palabras y 46750 vocales, lo cual sugiere un uso más variado o extenso de la lengua en comparación con otros grupos. Por otro lado, las mujeres de 18-35 años registraron el menor uso de palabras, con 15332.

Entre los hombres, los mayores de 55 años presentaron el uso mayor de fonos, con 41304, aunque no superan a las mujeres en el mismo rango de edad. Generalmente, la diferencia en el uso del lenguaje entre sexos varía según el rango de edad; por ejemplo, los hombres de 35-55 años y mayores usaron más palabras que las mujeres de 18-35 años, pero menos que las mujeres en su mismo rango de edad.

Metodológicamente, el estudio pretende mostrar una consistencia al realizar cuatro entrevistas por grupo y tener un número equilibrado de hablantes, lo que sugiere un esfuerzo centrado en la comparabilidad. No obstante, la variabilidad significativa en el número total de palabras y fonos refleja posiblemente diferencias inherentes en el estilo comunicativo entre los conjuntos de seres humanos estudiados.

En cuanto a la calidad de los audios recogidos, todos los audios del corpus fueron recogidos mediante micrófono unidireccional ubicado a una distancia equivalente tanto desde la posición del entrevistador como desde la posición de la persona entrevistada. Se trata de audios con una calidad adecuada para el estudio acústico, dado que hay pocos solapamientos; en ocasiones, pueden recogerse interrupciones de la grabación o golpes, pero estos han sido filtrados de manera auditiva previamente al procesamiento computacional de los datos acústicos.

4.2 Uso de ORALSTATS

ORALSTATS (Cabedo, 2022) ha sido desarrollado mediante el lenguaje de programación R Core Team (2023) y la sección de visualización incorpora la funcionalidad dinámica proporcionada por Shiny (Chang et al., 2022), una biblioteca gratuita también desarrollada en R. En resumen, ORALSTATS brinda un panel interactivo que permite a los investigadores explorar datos de manera sincrónica, utilizando filtros y diversos métodos estadísticos disponibles. Aunque la idea de esta herramienta no es nueva en los sistemas de gestión de R (Vnjis, 2016), permite poner datos en línea y analizarlos de manera dinámica, lo que resulta especialmente útil para la investigación y el análisis de datos fonéticos y prosódicos. La última versión, junto con unas breves indicaciones sobre su uso, puede encontrarse en el siguiente enlace (<https://acabedo.quarto.pub/oralstats/>).

ORALSTATS (Cabedo, 2022) es una herramienta de análisis que combina información de transcripciones de habla con valores de tono e intensidad obtenidos de Praat. En este sentido, funciona como una herramienta de transformación. Su módulo de creación crea tres conjuntos de datos: uno para los fonos, otro para las palabras (que incluye etiquetas POS) y otro para los grupos entonativos. Todos ellos incorporan información prosódica como rango de tono, media, mediana, intensidad media, velocidad del habla, duración y otros datos fónicos similares (rango tonal, inflexión melódica, etc.). Además, es posible generar otros conjuntos de datos con pequeñas modificaciones en el código, lo que puede resultar útil para investigadores interesados en unidades lingüísticas más amplias, como las cláusulas entonacionales o los paratonos. El sistema ha sido diseñado tanto para codificar de manera general la prosodia según modelos teóricos específicos, como ToBI o AMH, pero también modelos más generales, como el modelo interactivo-funcional para analizar la entonación, diseñado por Hidalgo (2019).

En cuanto a la transformación AMH, ORALSTATS simplifica la propuesta de Cantero (Cantero 2002; Cantero y Font-Rotchés, 2009), adoptando un enfoque básico que calcula la desviación porcentual entre los puntos medios de las vocales, pero también combinaciones de patrones micromelódicos dentro de estas, con la voluntad de detectar inflexiones melódicas internas dentro de las vocales que puedan ser parciales o que, incluso, generen patrones circunflejos. Para ello, el análisis descompone la vocal en cuartiles y calcula si hay desviaciones entre estas partes según si se cumplen ascensos superiores o inferiores a un 10 % de inflexión melódica, dato porcentual perceptivamente significativo (Mateo, 2010) y que equivale en cierta medida al 1.5 semitono establecido en otros estudios (Pàmies et al., 2002). Las diferentes casuísticas se esbozan en la Figura 2:

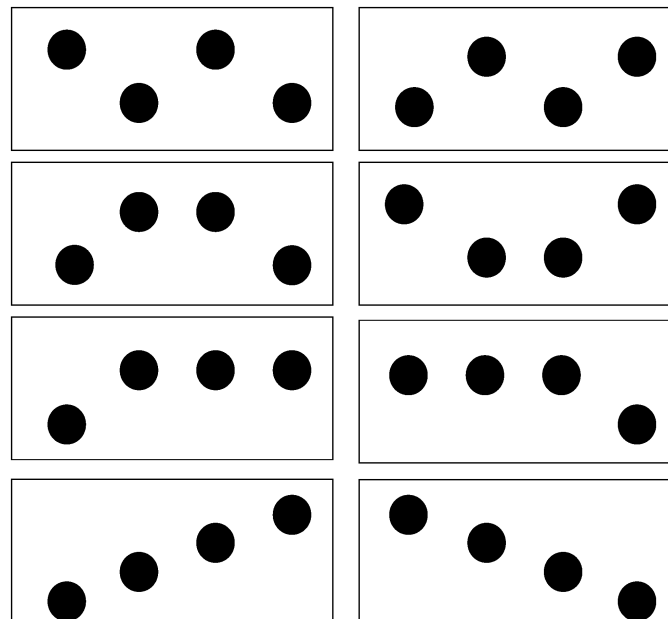


Figura 2. Variaciones micromelódicas dentro de la vocal.

Además del módulo de transformación y creación, ORALSTATS ofrece múltiples formas de visualizar datos, principalmente mediante gráficos (algunos específicos para información prosódica del proyecto PRESEEA). También permite aplicar diversos procedimientos estadísticos, como la prueba de chi-cuadrado, ANOVA, árboles de decisión o análisis de componentes principales, entre otros.

Finalmente, el módulo de visualización permite representar gráficamente múltiples variables fonéticas en una misma línea de tiempo. En estos gráficos, los investigadores pueden ajustar el tamaño de los puntos, los tipos de líneas y los colores para combinar información sobre tono, intensidad, duración y su pertenencia a una unidad lingüística, como las palabras dentro de un grupo fónico. Dentro de estas variables, se incluyen cruces complejos entre distintos aspectos de la entonación, como la frecuencia fundamental, la duración de las unidades y la intensidad del habla. Además, los rangos de valores tonales pueden variar significativamente entre hablantes y contextos lingüísticos, lo que agrega una capa adicional de complejidad al proceso de codificación.

Con la intención de agilizar la codificación melódica de un corpus según el modelo teórico del AMH, se ha elaborado una propuesta de codificación computacional. Esta propuesta tiene como objetivo principal analizar y etiquetar las características prosódicas, como los patrones tonales, en el discurso hablado. Sin embargo, es importante destacar que esta propuesta puede estar sujeta a ciertas limitaciones y posibles errores, dado que existen numerosas variables involucradas en el análisis de la entonación; concretamente, puede haber errores en la detección de puntos tonales desde Praat, aunque el trabajo de filtrado y recodificación previo de ORALSTATS, así como la interpolación de puntos tonales, minimiza esa posible incoherencia o falla metodológica.

Es fundamental reconocer que el análisis prosódico es una tarea desafiante y que las interpretaciones pueden ser subjetivas en algunos casos. Por lo tanto, es posible que el código propuesto inicialmente pueda contener errores o limitaciones en la representación de los patrones tonales y prosódicos en el habla. No obstante, esta porción de código se presenta como una base sólida para el análisis prosódico, y cualquier error identificado puede corregirse y mejorar en futuras versiones. La naturaleza adaptable y flexible del análisis prosódico en R permite ajustar y refinar el código a medida que se adquiere un mayor entendimiento de las características

prosódicas específicas de los datos y del idioma en estudio. En última instancia, este proceso iterativo busca mejorar la precisión y la robustez de la codificación prosódica según los principios del AMH.

Inicialmente, la codificación en patrones entonativos según el AMH procede de los requisitos establecidos en la bibliografía para el español (Cantero y Font-Rotchés, 2007), aunque puede ser modificada y extendida para otras lenguas. Por ejemplo, en términos computacionales, a modo de ejemplo, los patrones I y II se codifican según el siguiente código:

```
mutate(pattern = case_when ( ((between(anacrusis,-10,40) & anacrusis_displacement ==
"no") | is.na(anacrusis)) & (between(body_2,-80,10)|is.na(body)) & between(toneme, -40,
15) ~ "PI", ((between(anacrusis,-10,40) & anacrusis_displacement == "no") |
is.na(anacrusis)) & (between(body_2,-80,10)|is.na(body)) & toneme > 70 ~ "PII"
```

Además, a partir de las transcripciones, los archivos de tono e intensidad procedentes de Praat, se consigue codificar y crear gráficos de manera automática como los que siguen en las Figuras 3 y 4, que incluyen representaciones de los llamados patrones II (interrogativo) y I (enunciativo). Como se observa, ORALSTATS no solo proyecta en el gráfico la codificación de los puntos centrales de las vocales, sino que, en caso de haber una inflexión micromelódica interna, también se añade al gráfico, tanto en la curva melódica como en la tabla de valores que se adjunta al final. En todo caso, se ha intentado que el modelo de gráfico se asemeje al ideado y presentado por Cantero para el análisis prosódico del habla (Cantero, 2019).

Actualmente, como puede deducirse de los gráficos visibles en las figuras generadas desde Oralstats y que se muestran en este estudio (Figuras 3, 4, 10, 11 y 12), pueden existir ciertas *alucinaciones* (nombre técnico en computación que alude a situaciones en las que los programas sobregeneran información o alteran notablemente algún resultado), sobre todo en lo relacionado a la visualización de inflexiones melódicas internas, ya que el programa, en su versión actual, intenta que las vocales puedan dividirse en cuatro partes (cuatro cuartiles) para calcular melodía interna, aunque en ocasiones no siempre hay suficientes puntos tonales para poder realizar este reparto de manera realmente representativa.

En estas situaciones, por ejemplo, en las que se encuentran vocales con una duración breve o en las que Praat no ha generado puntos tonales para esas partes, ORALSTATS compensa con la media tonal de la vocal. Eso hace que algunos gráficos tengan una excesiva elongación. En todo caso, la computación de los patrones, en el momento actual, utiliza únicamente los puntos centrales de las vocales; es decir, si bien el patrón se computa con la centralidad de la vocal, los gráficos generados por ORALSTATS muestran inflexiones melódicas internas. La solución para que en el futuro los gráficos sean mucho más certeros es, como señala Mateo (2010), modificar un parámetro en Praat llamado *voice threshold* para recoger más puntos tonales (pero que incorporaría también otras posibles *alucinaciones* de Praat), o, como recogen los propios autores del programa (Boersma y Weenink, 2022), crear interpolaciones tonales.

Con la intención de mostrar la funcionalidad actual de ORALSTATS y de demostrar la complejidad que se mencionaba anteriormente, se proyectan los gráficos en dos versiones: la primera, que recibe la letra *a*, es el gráfico resultante de la exportación de ORALSTATS; la segunda, a la que se adjunta la letra *b*, es el gráfico extraído desde Praat (Figura 3b). Esta es una sugerencia de uno de los revisores anónimos del artículo y, realmente, mejora sustancialmente la ejemplificación de la herramienta.

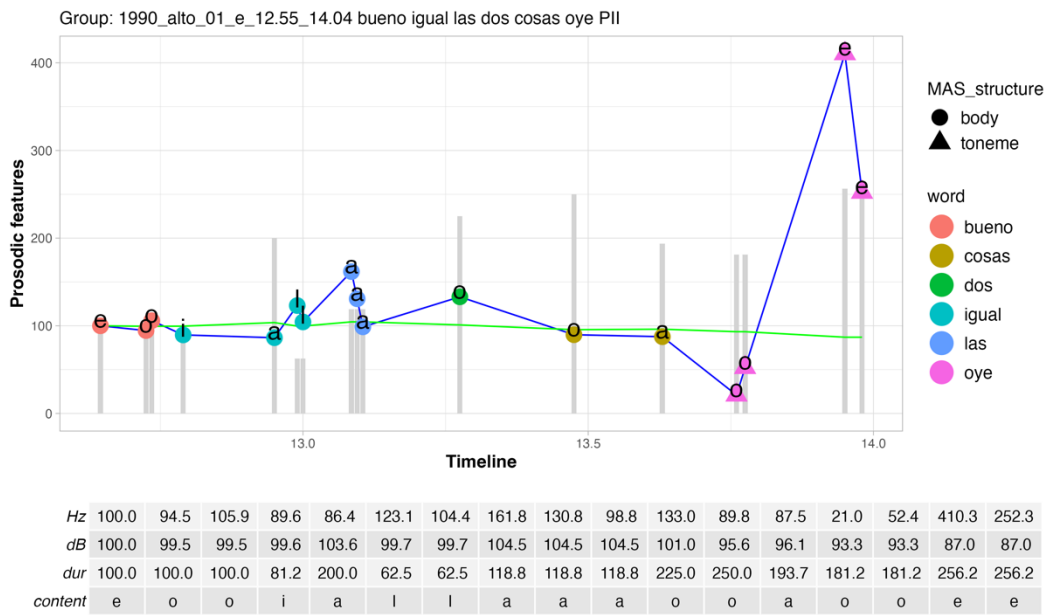


Figura 3a. Ejemplo codificado como patrón II en Oralstats (interrogativo). Enunciado: *Bueno igual las dos cosas oye*.

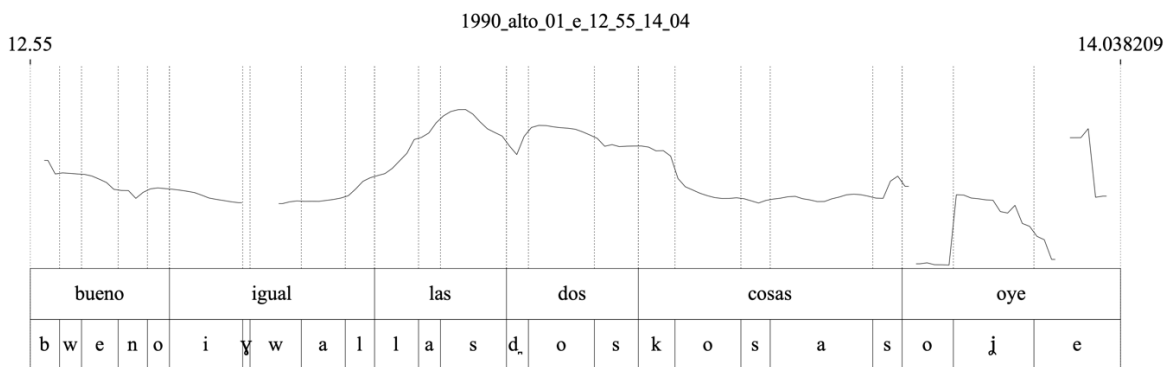


Figura 3b. Ejemplo extraído desde Praat. Enunciado: *Bueno igual las dos cosas oye*.

La Figura 3a presenta un gráfico que forma parte de un análisis melódico del habla, específicamente mostrando el patrón II (+interrogativo). En un eje de tiempo que va aproximadamente de 12.55 a 14.00 segundos, se visualizan puntos de datos de diferentes colores y formas que representan palabras individuales, cada uno siguiendo una trayectoria que refleja su configuración prosódica específica. Los colores y las formas corresponden a la leyenda de la derecha, identificando palabras como *bueno*, *cosas*, *dos*, entre otras. También se marcan elementos de la estructura del discurso como *body* y *toneme*, indicando partes fundamentales del patrón entonacional. Debajo del gráfico principal, se observan barras grises verticales que podrían indicar la duración de los pies rítmicos. A su vez, en la parte inferior se ofrece información detallada en cuanto a la frecuencia (Hz), intensidad (dB) y duración (dur) de los segmentos sonoros, y se muestra una serie de sílabas o fonemas correspondientes a los datos acústicos en la parte inferior. Estos valores son valores estandarizados, codificados según la propuesta teórica del AMH. Los valores originales quedan recogidos en una base de datos, pero no se proyectan en el gráfico.

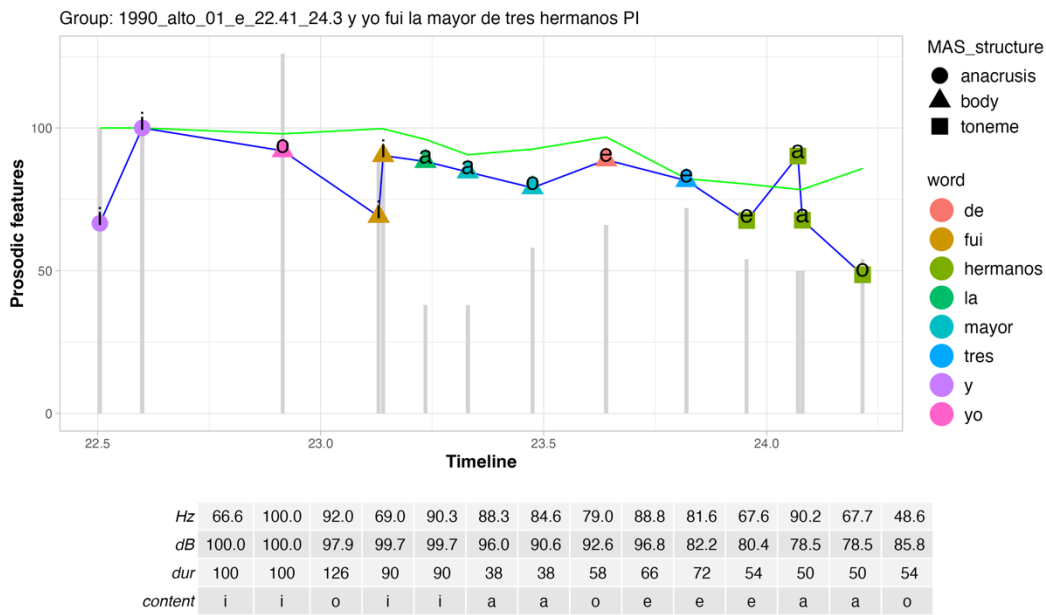


Figura 4a. Ejemplo codificado como patrón I (enunciativo). Enunciado: *Y yo fui la mayor de tres hermanos.*

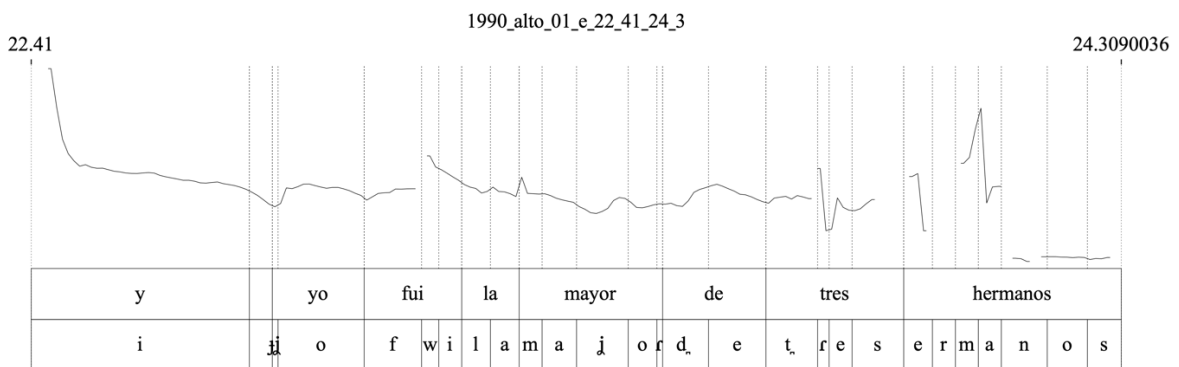


Figura 4b. Ejemplo extraído desde Praat. Enunciado: *Y yo fui la mayor de tres hermanos.*

La Figura 4a muestra el patrón I de un enunciado analizado mediante la metodología del Análisis Melódico del Habla (AMH), donde se visualiza la melodía de un grupo fónico enunciativo en un intervalo de tiempo de aproximadamente 1.5 segundos y en la Figura 4b el mismo enunciado extraído de Praat.

5. Resultados

En las secciones subsiguientes, presentaremos el análisis de la distribución de patrones tonales encontrados en el corpus. Más adelante, se abordará el apartado de los patrones no identificados, elementos que, a pesar de su presencia en la base de datos, han desafiado la clasificación automática convencional y no han recibido un patrón tonal según las reglas de computación establecida.

5.1 Distribución de patrones tonales

Los patrones tonales codificados por ORALSTATS se distribuyen como se muestra en la Figura 5:

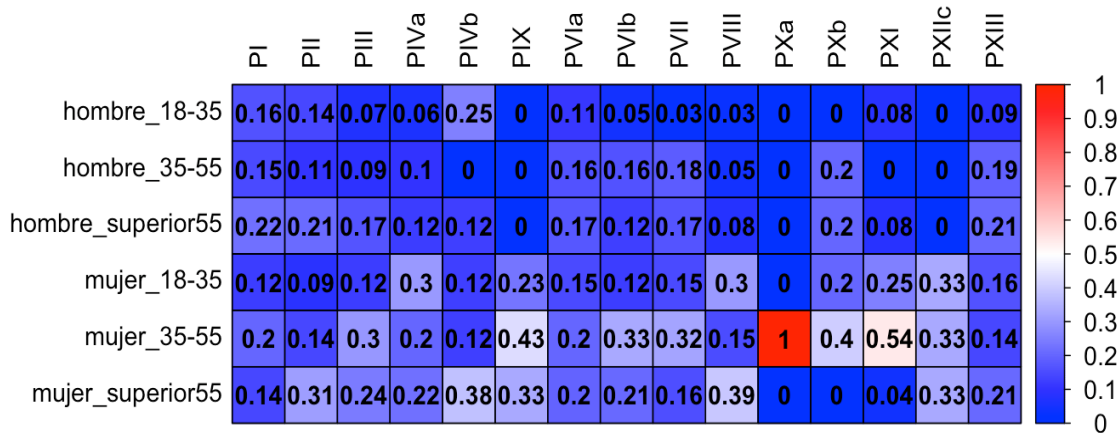


Figura 5. Distribución de patrones tonales por sexo y edad. La escala de 0 a 1 es porcentual.

Al agrupar los datos por sexo, podemos observar que tanto hombres como mujeres presentan el patrón tonal enunciativo (PI) como el más frecuente, con un 15-20 % de ocurrencias en hombres y en mujeres (en general, aunque no es un dato que se observe en la Figura 5, es el patrón más frecuente recogido en el corpus analizado). Sin embargo, hay diferencias notables en la diversidad y frecuencia de otros patrones tonales entre hombres y mujeres:

- **Diversidad de patrones:** las mujeres presentan una mayor variedad de patrones tonales diferentes en comparación con los hombres. Por ejemplo, se observan patrones únicos en mujeres como PIX, PXIc, PXa, que no aparecen en hombres.
- **Frecuencia de otros patrones:** algunos patrones, como PVIa y PVIb, muestran una mayor frecuencia en mujeres (20 a 30 % de frecuencia) en comparación con hombres (de 0 a 12 %, menos en el caso de los hombres jóvenes).

Este análisis sugiere que, mientras hay un patrón tonal dominante en el corpus analizado, existen diferencias notables en la diversidad y la utilización de patrones tonales entre distintos grupos de edad y sexos. Estas diferencias podrían estar influenciadas por una variedad de factores sociolingüísticos que deben ser aún explorados en detenimiento, quizá mediante la comprobación en otros niveles de instrucción, como el medio o el bajo.

Aún y con todo, en los datos recogidos, la catalogación automática de patrones ha dejado un total de 7543 grupos entonativos sin asignación de patrón. Dado que todos los grupos del estudio tienen computados valores como la anacrusis, el cuerpo, el tonema y otros factores, como la inflexión porcentual en relación a la vocal precedente o posterior, se elabora la sección 5.2, en la que un análisis multivariante y una técnica de clasificación estadística permiten elaborar una caracterización prosódica de esos grupos no catalogados, con la finalidad de poder etiquetarlos posteriormente o, incluso, de detectar nuevos patrones no definidos previamente en la bibliografía.

5.2 Patrones tonales no identificados

Con el total de 7543 grupos entonativos sin patrón detectado, se ha realizado un análisis de componentes principales (PCA, por sus siglas en inglés). Esta técnica estadística se utiliza para reducir la dimensionalidad de un conjunto de datos, mientras se conserva la mayor cantidad

posible de su variabilidad. Esto se logra identificando los ejes (componentes principales) que maximizan la varianza de los datos proyectados sobre estos ejes (Levshina, 2015).

Este procedimiento de análisis, por lo tanto, primero realiza un PCA para analizar los datos y reducir la variación, luego realiza un *clustering* jerárquico sobre los componentes principales extraídos del análisis. Finalmente, visualiza los clústeres resultantes para identificar agrupamientos o patrones naturales dentro del conjunto de datos. Todas las operaciones se han realizado con la librería de R FactoMineR (le et al. 2008) y con su paquete de visualización Factoextra (Kassambara y Mundt, 2020).

Las variables que se han tenido en cuenta para poder caracterizar a los grupos entonativos a los que no se les había podido asignar patrón melódico son las siguientes:

1. `standard_pitch`: tono estandarizado, medido en hercios (Hz).
2. `standard_intensity`: intensidad estandarizada, medida en decibelios (dB).
3. `standard_dur`: Duración estandarizada de los *pies rítmicos*, es decir, la duración que transcurre entre el centro de una vocal y la siguiente. No se trata en este caso de la duración estandarizada de las vocales, sino de la estandarización de la distancia existente entre las vocales.
4. `anacrusis`: se trata de las sílabas átonas precedentes a un primer pico tonal que, generalmente, coincide con el entorno de la primera sílaba tónica.
5. `body`: es la distancia porcentual entre el primer pico tonal y el último o tonema.
6. `anacrusis_displacement`: desplazamiento de la anacrusis, lo que podría implicar el ajuste o cambio en la posición de la anacrusis respecto a la estructura principal a la que introduce. Es habitual en español que el pico tonal, sobre todo en palabras paroxítonas, recaiga en la última sílaba y no directamente en la tónica (Alfano, Llisterri y Savy, 2008).
7. `accent`: se trata de diferenciar entre palabras oxítonas, paroxítonas o proparoxítonas.
8. `toneme`: inflexión melódica porcentual a partir de la última sílaba tónica.
9. `inflexion_percent_Hz_from_prev`: cambio porcentual en la frecuencia (Hz) desde la vocal anterior.
10. `inflexion_percent_Hz_to_next`: cambio porcentual en la frecuencia (Hz) hacia la vocal posterior.
11. `q1piHz_q2piHZ`, `q2piHz_q3piHZ`, `q3piHz_q4piHZ`: estas variables hacen referencia a las inflexiones melódicas, calculadas de forma porcentual, que suceden en los cuartiles de la duración de las vocales. Estos valores son útiles para capturar variaciones micromelódicas y, en general, patrones circunflejos.
12. `displacement`: desplazamiento de la inflexión melódica en el tonema.
13. `dur`: duración de la vocal tonemática.

Las trece variables anteriores se computan como entrada en el PCA. A modo de historial de los pasos de análisis estadístico, se incluyen las siguientes figuras:

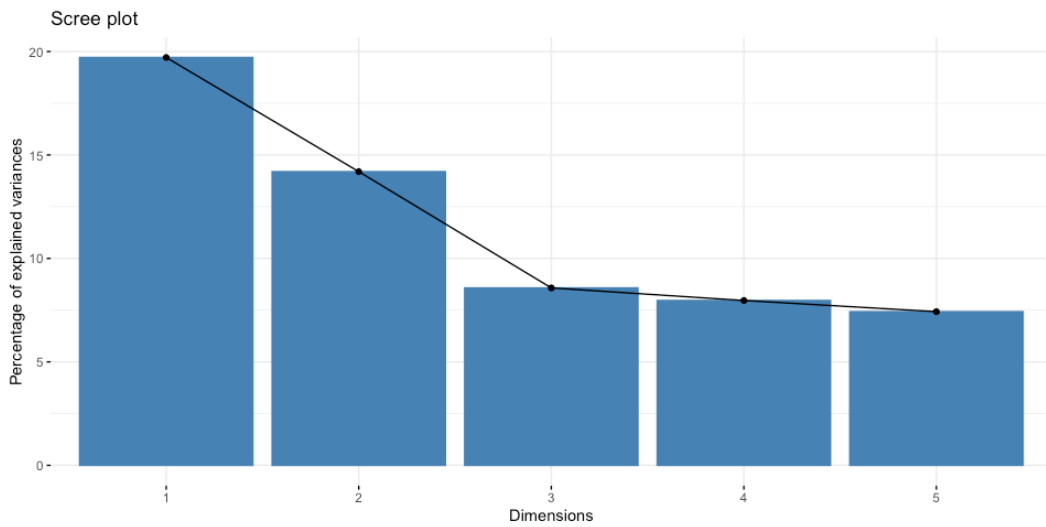


Figura 6. Eigenvalues o explicación porcentual de la variación por dimensión.

Con el análisis de componentes principales, las dimensiones pretenden reducir la variación uniendo variables que manifiestan correlaciones frecuentes. Generalmente, el sistema crea conjuntos de agrupaciones o dimensiones, en los que cada variable tiene un peso determinado. La acumulación de dimensiones explica la variación completa de una base de datos; generalmente, las más relevantes para los estudios son siempre las dos primeras (Glynn, 2014; Greenacre, 2007). De acuerdo con los datos proporcionados, las cinco primeras dimensiones explican un total de casi 60 %. La elección de cuántos componentes incluir en el análisis posterior depende de la cantidad de varianza total que se desea explicar y del equilibrio entre simplificación de los datos y retención de información.

En un conjunto de datos en el que se exploran más de cinco variables, es normal que las dimensiones, al menos las dos primeras, cubran solo un porcentaje menor de la varianza. En el caso de este estudio, las dos primeras dimensiones cubren en torno a un 40 %, como se observa en la Figura 6. En el marco de una investigación multivariable y con múltiples categorías, un valor como el obtenido puede considerarse una explicación de la variación bastante elevada.

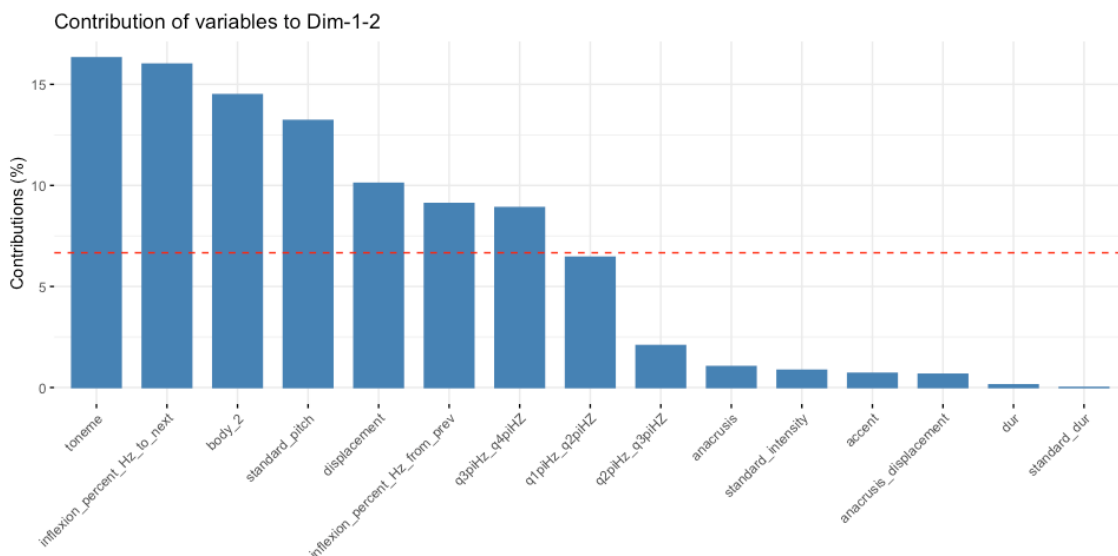


Figura 7. Contribución de las variables a las dos primeras dimensiones del PCA.

Variables de izquierda a derecha, es decir, de mayor a menor importancia en el modelo: *toneme*, *inflexion_percent_Hz_to_next*, *body*, *Standard_pitch*, *displacement*, *inflexion_percent_Hz_from_prev*, *q3piHz_q4piHz*, *q1piHz_q2piHz*, *q2piHz_q3piHz*, *anacrusis*, *Standard intensity*, *accent*, *anacrusis displacement*, *duration*, *Standard duration*

El análisis proporciona información sobre la contribución de diversas variables a las dos primeras dimensiones en un estudio de componentes principales o análisis factorial, como se observa en la Figura 7. La contribución de cada variable a una dimensión específica ayuda a entender cómo de importante es esa variable para explicar la variabilidad en esa dimensión. A continuación, se explica cómo contribuyen las variables a las Dimensiones 1 y 2:

Dimensión 1

- *Body* es la variable que más contribuye a esta dimensión, con un 20.10%.
- *Inflexion_percent_Hz_from_prev* sigue con un 12.43%, sugiriendo que las variaciones en la inflexión desde el sonido anterior también son significativas para la dimensión 1.
- *Q3piHz_q4piHZ* y *Toneme* tienen contribuciones considerables (14.65% y 9.15%, respectivamente).
- *Standard_pitch* contribuye con un 18.42%.
- Las variables restantes, como *Standard_intensity*, *Anacrusis* y *Dur*, tienen contribuciones menores, indicando una influencia menos significativa en esta dimensión.

Dimensión 2

- *Toneme* y *Inflexion_percent_Hz_to_next* dominan esta dimensión, con contribuciones del 26.26% y 25.52%, respectivamente.
- *Displacement* tiene una contribución significativa del 19.24%.
- *Body_2* y *Q2piHz_q3piHZ* tienen contribuciones más moderadas (6.70% y 4.95%, respectivamente).
- Las demás variables, como *Standard_pitch*, *Q1piHz_q2piHZ*, y *Q3piHz_q4piHZ*, tienen contribuciones menores; no son tan determinantes como las variables de mayor contribución.

En resumen, la Dimensión 1 se caracteriza principalmente por la contribución de variables relacionadas con características acústicas y variaciones melódicas, mientras que la Dimensión 2 está más influenciada por aspectos estructurales, como el desplazamiento tonal en el tonema. Estas contribuciones ayudan a interpretar las dimensiones en términos de las características más significativas de los datos analizados.

La cuantificación de la importancia de las variables por dimensión, a partir del análisis PCA, conlleva preguntarse de qué manera los grupos entonativos pueden aglutinarse según los parámetros que cumplan en esas variables. Una posible forma de actuar sería partir de un hipotético número de grupos preestablecido para determinar esas semejanzas o diferencias en los parámetros de las variables mencionadas. En esta investigación, sin embargo, ese número no se conoce, por lo que la aplicación del *análisis de clúster*, también realizado mediante la librería FactoMineR (Le et al., 2008), establecerá el número más adecuado de grupos que cumplan valores semejantes.

De esta manera, el análisis de clúster basado en un PCA se utiliza para simplificar conjuntos de datos mediante la reducción de dimensionalidad, lo que a su vez facilita la identificación de patrones o agrupaciones y mejora la eficiencia de los algoritmos de clúster. Al transformar variables posiblemente correlacionadas en componentes principales no correlacionados, PCA permite una visualización efectiva de los datos en espacios bidimensionales o tridimensionales, facilitando la identificación visual de agrupaciones o anomalías. Esta técnica mejora la

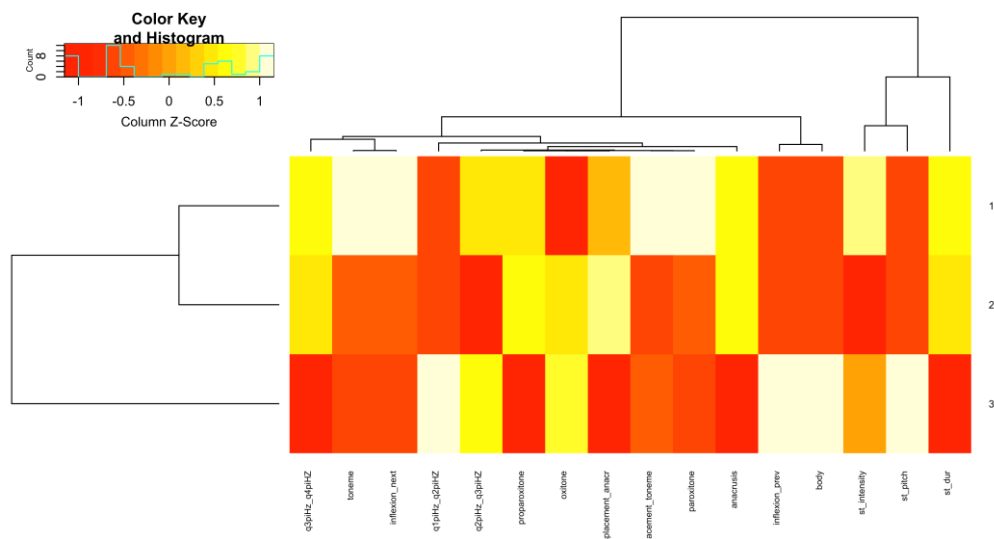


Figura 9. Mapa de calor con la caracterización de los grupos no clasificados por ORALSTATS.

Variables que aparecen de izquierda a derecha: *q3piHz_q4piHz*, *toneme*, *inflexion_next*, *q1piHz_q2piHz*, *q2piHz_q3piHz*, *proparoxitone*, *oxitone*, *displacement_anacrusis*, *displacement_toneme*, *paroxitone*, *anacrusis*, *inflexion_prev*, *body*, *standard intensity*, *Standard pitch*, *Standard duration*

La Figura 9 muestra un mapa de calor, visualización que se utiliza para proyectar la intensidad de los datos a través de colores y que permite identificar patrones, correlaciones y anomalías de manera rápida y efectiva dentro de una matriz o tabla de datos.

El *clúster 1* se caracteriza por un tono y una intensidad moderados, destacándose por tener la duración más larga entre los tres. En términos de anacrusis, este clúster muestra valores superiores en comparación con el Clúster 3, pero similares al Clúster 2. Presenta valores relativamente bajos en las variables *body* y *toneme* en comparación con el Clúster 3. Un aspecto distintivo de este clúster es su significativa inflexión porcentual desde el sonido previo hasta el siguiente, lo que indica una tendencia hacia cambios de tono más notorios. Además, muestra una alta proporción de palabras paroxítonas y una tasa muy alta de desplazamiento acentual en el tonema. Véase un ejemplo en las Figuras 10a y 10b.

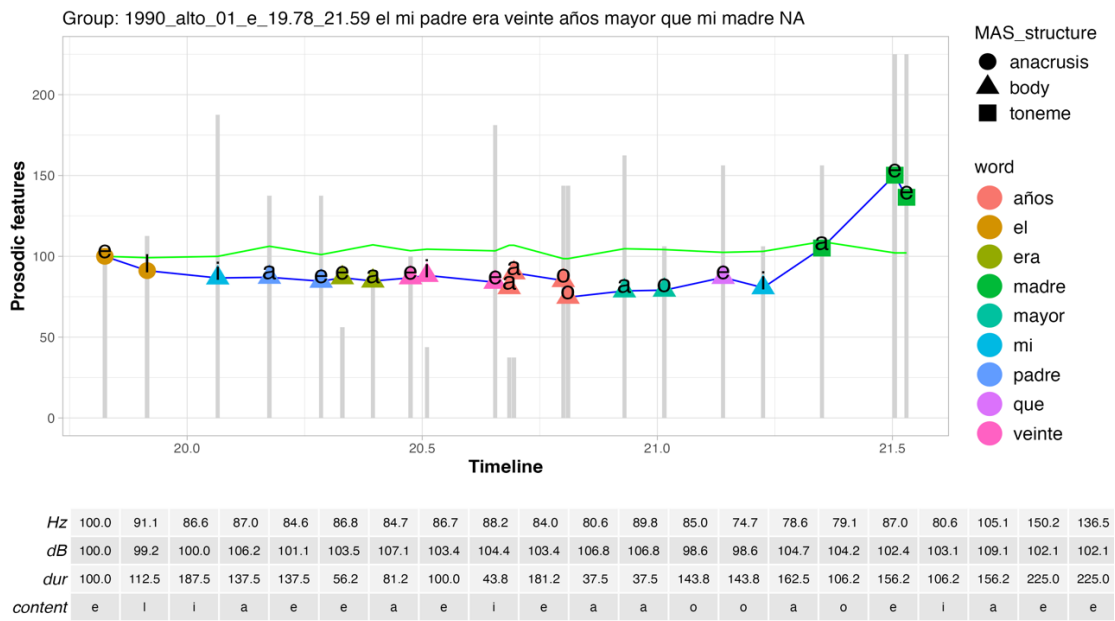


Figura 10a. Visualización del grupo entonativo clasificado como grupo 1. Enunciado: *El- mi padre era veinte años mayor que mi madre.*

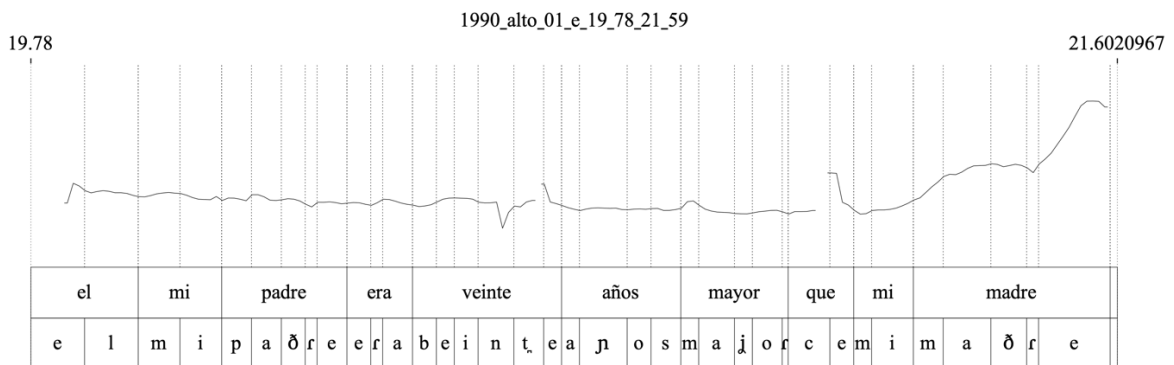


Figura 10b. Visualización del grupo entonativo clasificado como grupo 1 en Praat. Enunciado: *El- mi padre era veinte años mayor que mi madre.*

El *clúster 2*, por otro lado, presenta un tono e intensidad ligeramente inferiores al *clúster 1*. Su duración es menor que la del Clúster 1 pero mayor que la del *Clúster 3*. Comparte similitudes con el Clúster 1 en términos de anacrusis, pero se distingue por tener el valor más bajo en *body*. Este clúster muestra valores negativos para *toneme* y tiene valores menos extremos de inflexión porcentual, lo que sugiere menos variación de tono en comparación con el *clúster 1*. Es significativo por tener la proporción más baja de palabras paroxítonas y no registra instancias de desplazamiento acentual, diferenciándose notoriamente en términos de acentuación y desplazamiento fonético del Clúster 1. Véase un ejemplo en las Figuras 11a y 11b.

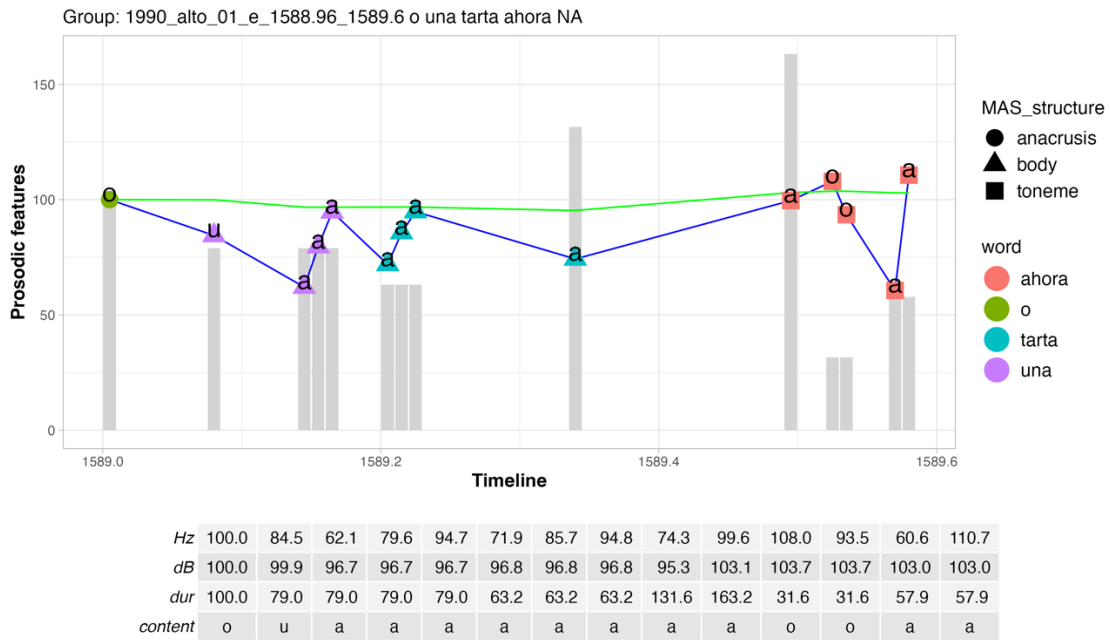


Figura 11a. Visualización del grupo entonativo clasificado como grupo 2. Enunciado: *O una tarta ahora.*

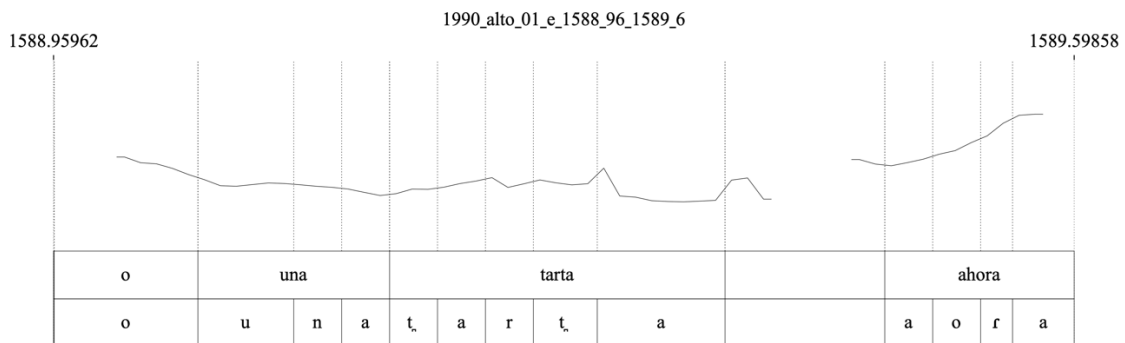


Figura 11b. Visualización del grupo entonativo clasificado como grupo 2 en Praat. Enunciado: *O una tarta ahora.*

Finalmente, el *clúster 3* se distingue significativamente por su tono mucho más alto en comparación con los otros dos clústeres. Aunque la intensidad y duración estándar son comparables a los otros clústeres, este tiene la duración más corta. Presenta un valor de *body* significativamente más alto, indicando propiedades de tiempo y físicas distintivas. Este clúster muestra un porcentaje de inflexión extremadamente alto desde el sonido previo, sugiriendo cambios de tono muy pronunciados. Posee la mayor proporción de palabras oxítonas.

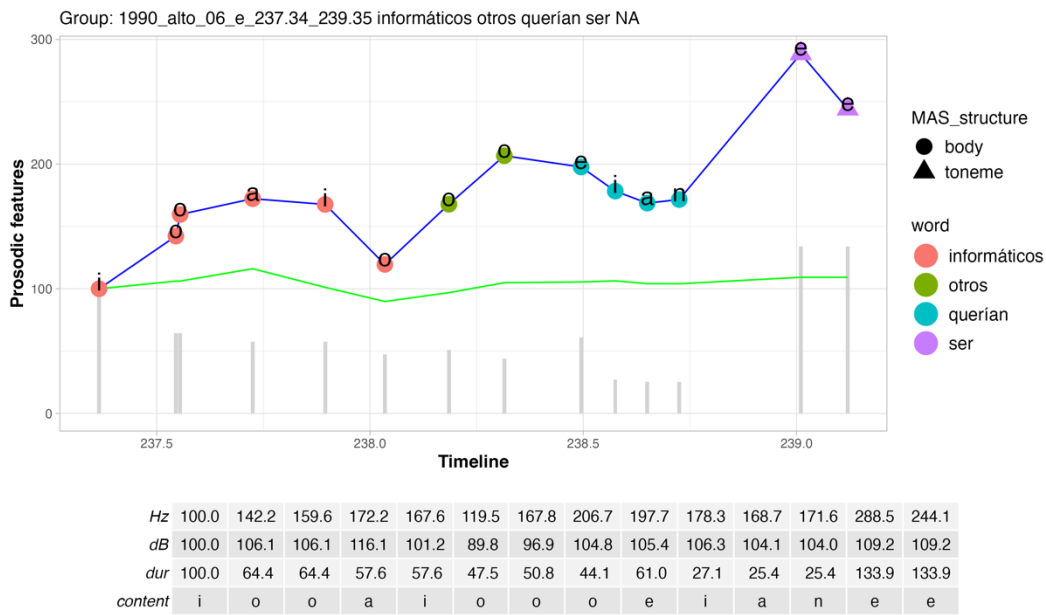


Figura 12a. Visualización del grupo entonativo clasificado como grupo 3. Enunciado: *Informáticos otros querían ser.*

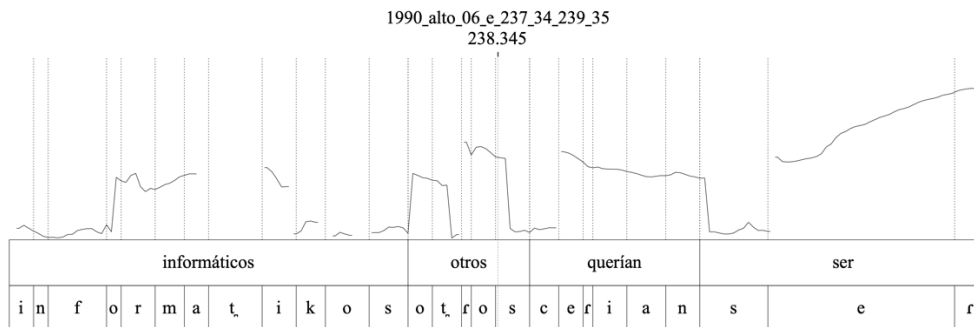


Figura 12b. Visualización del grupo entonativo clasificado como grupo 3 en Praat. Enunciado: *Informáticos otros querían ser*

En el caso de las Figuras 12a y 12b, por ejemplo, con una inflexión melódica continua y un descenso en la última parte de la vocal tonemática final, podría considerarse que patrones con estas características podrían integrar el llamado patrón XIII, que Font-Rotchés y Mateo (2013) registraron en alguno de sus estudios o, al menos, considerarse como una variación de este.

6. Conclusiones

El presente estudio ha demostrado la eficacia y relevancia del Análisis Melódico del Habla (AMH) aplicado al corpus PRESEEA Valencia (nivel de instrucción alto), mediante el uso de herramientas computacionales avanzadas para el análisis de datos fonéticos. A través de la implementación de ORALSTATS, hemos podido automatizar la extracción y el análisis de características prosódicas como el tono, la intensidad, y la duración de los segmentos fónicos, lo cual ha permitido una exploración detallada y sistemática de los patrones entonativos del español.

La utilización de *software* especializado ha evidenciado las ventajas de integrar tecnologías de análisis computacional en la investigación lingüística, no solo en términos de eficiencia y precisión, sino también en la capacidad de manejar grandes volúmenes de datos, fundamental en estudios sociolingüísticos y dialectales. Este enfoque computacional ha facilitado la identificación de patrones entonativos complejos y su clasificación automática, destacando la importancia de la entonación en la expresión de significados pragmáticos y emocionales en el habla. Aunque no se ha encontrado en forma de libre acceso un conjunto de datos etiquetados en patrones melódicos según AMH como base para realizar un estudio contrastivo, actualmente, ya hay hasta dos trabajos en los que, en conjunción con integrantes del grupo de AMH, se procederá a testar la viabilidad de la codificación automática del programa.

Los hallazgos de este estudio subrayan la complejidad y riqueza de la prosodia en el español hablado, evidenciando variaciones significativas en la entonación que reflejan factores sociolingüísticos como el sexo y la edad de los hablantes. La capacidad de documentar y analizar estas variaciones es crucial para una comprensión más profunda de la dinámica lingüística del español y sus dialectos.

Este trabajo también ha identificado desafíos inherentes al análisis automático de la prosodia, particularmente en lo que respecta a la precisión de la detección de patrones tonales y la interpretación semántica de las secuencias entonativas. Estos desafíos resaltan la necesidad de continuar refinando las herramientas computacionales y los modelos teóricos que sustentan el análisis prosódico, así como la importancia de una colaboración estrecha entre lingüistas y especialistas en tecnologías de la información.

Finalmente, la investigación deja abiertas algunas consideraciones. Se sugiere la expansión de este enfoque metodológico a otros corpus y variedades del español, así como la exploración de nuevas aplicaciones de la tecnología en el estudio de otros aspectos de la variación lingüística. Asimismo, se destaca la necesidad de un trabajo interdisciplinario que combine conocimientos lingüísticos con avances en inteligencia artificial y análisis de datos para avanzar en el estudio de la complejidad lingüística.

Este estudio, por tanto, confirma el valor del análisis computacional en la investigación lingüística y establece una base para futuras investigaciones en el análisis melódico del habla.

Referencias bibliográficas

- Alfano, I., Llisterri, J., y Savy, R. (2008). Las características acústicas y perceptivas del acento léxico en español y en italiano: los patrones acentuales paroxítonos. *Language Design: Journal of Theoretical and Experimental Linguistics*, 10, 23-30.
- Anthony, L. (2020). Programming for Corpus Linguistics. En Magali Paquot y Stefan Th. Gries (Eds.), *A Practical Handbook of Corpus Linguistics* (pp.181–207). Springer International Publishing. https://doi.org/10.1007/978-3-030-46216-1_9.
- Boersma, P., y Weenink, D. (2022). *Praat: Doing Phonetics by Computer*. <http://www.praat.org/>.
- Cabedo Nebot, A. (2022). *Oralstats* (1.3). <https://github.com/acabedo/oralstats>
- Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Edicions de la Universitat de Barcelona.
- Cantero, F. J. (2016). Corpus de habla espontánea para el estudio de la entonación. En A.M. Fernández Planas (Ed.), *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística* (pp. 153-160). Universitat de Barcelona.
- Cantero, F. J. (2019). Análisis prosódico del habla: más allá de la melodía. En M. R. Álvarez Silva, A. Muñoz Alvarado, y L. Ruiz Miyares (Eds.), *Comunicación Social: Lingüística, Medios*

- Masivos, *Arte, Etnología, Folclor y otras ciencias afines* (pp. 485–98). Santiago de Cuba: Ediciones Centro de Lingüística Aplicada.
- Cantero, F. J., y Font-Rotchés, D. (2007). Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia*, 13, 69-92. <http://hdl.handle.net/10347/6067>.
- Cantero, F. J., y Font-Rotchés, D. (2009). Melodic analysis of speech method (MAS) applied to Spanish and Catalan. *Phonica*, 5, 33–47.
- Chang, W., Cheng, J., Allaire, J.J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert A., y Borges, B. (2022). *Shiny: Web Application Framework for r*. <https://CRAN.R-project.org/package=shiny>.
- Devís, E. (2011). La entonación de (des)cortesía en el español coloquial. *Phonica*, 7, 4–44.
- Devís, E. (2012). Percepción de la cortesía atenuadora en el español coloquial. *Oralia: Análisis del discurso oral*, 15, 125–46.
- 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), 767–92. <https://doi.org/10.1007/s10579-015-9320-9>.
- Estruch, M., Garrido, J.M., Llisterri, J., y Riera, M. (2007). Técnicas y procedimientos para la representación de las curvas melódicas. *RLA: Revista de lingüística teórica y aplicada*, 45, 59–87.
- Font-Rotchés, D. y Mateo, M. (2013). Entonación de las interrogativas absolutas del español peninsular del sur en habla espontánea. *Onomázein*, 28, 256–275. <https://doi.org/10.7764/onomazein.28.17>
- Garrido, J. M., Escudero, D, Aguilar, L., Cardenoso, V., Rodero, E., de-la-Mota, C., González, C., et al. (2013). Glissando: A Corpus for Multidisciplinary Prosodic Studies in Spanish and Catalan. *Language Resources and Evaluation*, 47 (4), 945–71. <https://doi.org/10.1007/s10579-012-9213-0>.
- Garrido, J.M., y Chica, J. A. (2018). Pitch Range and Identification of Emotions in Spanish Speech: A Perceptual Study. *Estudios de Fonetica Experimental*, 27, 13-36.
- Glynn, D. (2014). Correspondence analysis: Exploring data and identifying patterns. En D. Glynn & J. A. Robinson (Eds.), *Human Cognitive Processing*, 43 (pp. 443-485). John Benjamins Publishing Company. <https://doi.org/10.1075/hcp.43.17gly>
- Gómez Molina, J. R. (2001). *El español hablado de Valencia. Materiales para su estudio I. Nivel sociocultural alto*. Anejo XLVI Cuaderno de Filología. Publicacions de la Universitat de València.
- Greenacre, M. (2007). *Correspondence Analysis in Practice*. Vol. 67. CRC Press. <https://doi.org/10.2307/1938672>.
- Hidalgo, A. (2019). *Sistema y uso de la entonación en español hablado*. Universidad Andrés Burtado.
- Kassambara, A., y Mundt, F. (2020). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. <https://CRAN.R-project.org/package=factoextra>.
- Le, S., Josse, J., y Husson, F. (2008). FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software*, 25 (1), 1-18. <https://doi.org/10.18637/jss.v025.i01>.
- Levshina, N. (2015). *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. John Benjamins Publishing Company.

- Llisterri, J. (1991). *introducción a la fonética: el método experimental*. Anthropos.
- Mateo, M. (2010). Scripts en Praat para la extracción de datos tonales y curva estándar. *Phonica*, 6, 91-111.
- Mota de la, C., Machuca, M. J., Ríos, A., Riera, M., y Llisterri, J. (2005). Corpus orales para el desarrollo de las tecnologías del habla en español. *Oralia: Análisis Del Discurso Oral*, 8, 289-328.
- Pàmies, A., Fernández Planas, A. M., Ortega Escandell, A., y Amorós, M. C. (2002). Umbrales tonales en español peninsular. *Actas del II Congreso de Fonética Experimental*. Universidad de Sevilla.
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org/>.
- Sola, A., y Torregrosa, J. (2023). Intonation of Wh-questions in Northern British English Spontaneous Speech. *IJES. International Journal of English Studies*, 23 (1), 1-25. <https://doi.org/10.6018/ijes.559521>
- Torregrosa, J. (2015). Una propuesta metodológica para el análisis melódico de enunciados orales espontáneos con partículas modales del alemán. *Revista de filología Alemana*, 23, 193-212.
- Vnjis, V. (2016). *Radiant, Business Analytics Using R and Shiny*. <https://vnjis.github.io/radiant/>.