

ESTUDIO ACÚSTICO DE LA DURACIÓN DE LAS VOCALES DEL HABLA EN USO A PARTIR DEL CORPUS PRESEEA-MEDELLÍN¹

Franklin Rodríguez Ordóñez

Universidad de Antioquia (Colombia) / Escuela Normal Superior Pedro Justo Berrío (Colombia)

franklin.rodriguez@udea.edu.co / frodriguez@enspedrojustoberrio.edu.co

Recibido: 15/08/2022 - **Aprobado:** 03/11/2022 - **Publicado:** 13/03/2023

DOI: doi.org/10.17533/udea.lyl.n83a05

Resumen: Los objetivos de este trabajo fueron establecer los valores de duración vocálica e identificar en qué medida diversas variables referentes al timbre, a la estructura de la sílaba, de la palabra y del enunciado se correlacionan con la conformación el parámetro acústico de la duración. A partir de una muestra tomada de entrevistas semidirigidas del corpus PRESEEA-Medellín, se pudo identificar un valor de duración vocálica intrínseca de 65,32 ms; además, los factores que generan variaciones en la duración están más correlacionados con las estructuras prosódicas del enunciado que con la estructura interna de la sílaba o de la palabra.

Palabras clave: duración de las vocales; fonética acústica; prosodia; español hablado en Colombia; español de América.

ACOUSTIC STUDY OF VOWEL LENGTH IN THE SPOKEN LANGUAGE FROM CORPUS PRESEEA-MEDELLÍN

Abstract: The purposes of this research were to establish duration values of vowels and identify how variables referring to the vowel category, to structure of the syllabic, the word, and the statement are related to the configuration of vowel length. From a corpus PRESEEA-Medellín sample, it was possible to identify an intrinsic vowel duration average of 65,32 ms; besides, the duration is more correlated to the prosodic structures of the statements than the internal structure of the syllables or the words.

Keywords: vowel duration; acoustic phonetics; prosody; spoken Spanish in Colombia; Spanish of the Americas.

1. Artículo producto de la ponencia presentada en el Tercer Encuentro del Seminario de Prosodia, organizado por el Laboratorio de Fonética de la Universidad de Antioquia y la Universidad Autónoma de Querétaro, realizado en los días 7 y 8 de octubre de 2021. A su vez, resultado de la tesis realizada para optar al título de Magíster en Lingüística de la Universidad de Antioquia en 2022.

1. Introducción

Las variantes dialectales del español en Colombia manifiestan cambios en diferentes niveles lingüísticos. En el plano de la fonética existen diversos fenómenos que pueden ser objeto de estudio. Se han realizado acercamientos a la comprensión del fenómeno de la duración vocálica, pero en función de parámetros prosódicos, y si bien la duración es uno de los parámetros de análisis de la entonación, hasta el momento no hay un estudio enfocado en el análisis de la duración del español en Colombia y su funcionamiento en contexto de habla en uso. De hecho, muchos estudios sobre la duración vocálica en Hispanoamérica han trabajado con análisis de corpus fijo y, a modo de conclusión, han sugerido continuar con investigaciones donde se hagan estudios acústicos del habla en uso con corpus amplios (Krohn, 2019, p. 222).

Por tales motivos, en este trabajo se describen cuáles son los valores de la duración vocálica y cómo estos valores cambian en función de los contextos donde se presentan en la variante del español de Medellín, Colombia. Para ello, se consideró necesario determinar las diferencias en los valores de duración de las vocales y si, además de los contextos vocálicos, hay otros factores que influyen en dichos valores.

Como antecedentes se encontró que Muñoz-Builes (2016) realizó un análisis prosódico de enunciados declarativos e interrogativos con expansión en sujeto de una informante de Medellín. En este estudio se concluye que, en general, en los sintagmas nominales de los enunciados declarativos, las vocales tónicas son más largas que en los interrogativos y las vocales tónicas de las expansiones del sintagma tienden a ser más largas que las del sujeto. Tanto en los sintagmas verbales como preposicionales, las diferencias de duración vocálica no son significativas entre los enunciados declarativos e interrogativos. Por lo tanto, en lo concerniente a la duración, en esta investigación se establece que las vocales acentuadas tienen una tendencia a ser más largas que la duración de las inacentuadas, independientemente de la modalidad declarativa o interrogativa en la que se presente.

Por otra parte, se encuentra un estudio similar realizado por Muñetón (2017), en el cual se concluye que sí existe una relación entre los parámetros acústicos que coocurren en la prosodia de la frase. No obstante, en relación a la duración, no se determina que este parámetro sea principal o secundario en la prosodia, pero sí importante en el contexto del análisis. Al igual que el estudio mencionado anteriormente, se identifica que hay una tendencia a presentarse una mayor duración vocálica en las sílabas tónicas, aunque la F0 y el pico tonal no siempre coinciden con la sílaba más larga, sino que hay un patrón de asociación entre estos tres factores para configurar el entorno prosódico, por lo que se centra también la observación en las sílabas adyacentes a la sílaba que contiene la vocal tónica.

En el estudio de variantes dialectales del español en países de habla hispana diferentes a Colombia, Marín (1995) hizo una caracterización acústica de la duración de las cinco vocales del español ibérico. A partir de los resultados en este estudio, se obtiene una agrupación de duración de las vocales que consta de tres bloques: /a/ 69,63 ms > /e/ 64,95 ms, /o/ 64,15 ms > /u/ 60,93 ms, /i/ 60,66 ms.

En un contexto similar de estudios sobre el habla en España, se encontró el estudio realizado por Cuenca (1996), en el cual se realiza un análisis instrumental de la duración de las vocales en español. Como resultados en

este trabajo se llega a conclusiones similares al estudio anterior. Además, se agrupan de igual manera las vocales según su valor de duración clasificándolas en tres grupos: /a/ 73,65 ms > /o/ 70,94 ms, /e/ 65,58 ms > /i/ 58,71 ms, /u/ 55,66 ms.

Otro estudio concerniente es el realizado por Krohn (2019), quien analizó la duración vocálica en el español del área metropolitana de Costa Rica. Los valores de duración identificados para la variante dialectal del español de Costa Rica que se encontraron son: /a/ 92 ms, /o/ 88 ms, /e/ 87 ms, /i/ 84 ms, /u/ 84 ms.

Otro trabajo que se orienta en la misma línea de estudios fonéticos sobre las vocales del español es el realizado por Pereira y Soto (2011), quienes analizaron la duración absoluta de las vocales de las zonas urbanas y rurales de la Región de Ñuble, en Chile. Se trata de una investigación donde se plantea una diferencia en la duración de las vocales de acuerdo al estrato social y relación de cercanía frente a la urbe. Según este trabajo, cuanto más cercano se está a la ciudad y a mayor estrato social se pertenece, menor es la duración vocálica del hablante. Se observan factores tanto articulatorios como extralingüísticos sobre la concepción del tiempo o familiaridad con la situación comunicativa.

Los estudios aquí presentados han sido un fundamento importante para plantear un diseño metodológico que permita alcanzar el propósito de esta investigación. Identificando las posibles variables a analizar, se logra tener una mejor idea de qué se puede observar más allá de solo los valores en bruto que se puedan obtener de la duración; incluso, surgió el cuestionamiento por los valores que se denominan intrínsecos o absolutos, puesto que la variabilidad del fenómeno tiende a exhibir cambios que seguramente requieran aplicación de métodos de análisis más sofisticados y precisos como, por ejemplo, con mediciones de la duración con redes neuronales profundas. Al respecto, Adi, Keshet y Goldrick (2015) exponen:

To obtain accurate data most researchers have relied on manual annotation. This approach is clearly not ideal: it is highly resource intensive and fundamentally subjective. To address these issues, recent phonetic studies have used computational methods to measure acoustic properties of speech automatically. These methods greatly reduce the resources required as well as minimizing the role of subjective judgments (p. 1).

2. El vocalismo, la duración y el habla de Medellín

2.1. La duración vocálica

El tiempo en el habla es una característica que llama la atención a toda persona estudiosa de la fonética. Se puede medir a nivel microestructural, del cual se obtiene la duración de cada sonido —como se ha aplicado en esta investigación— o desde un enfoque más métrico, del que se tienen a consideración segmentos más amplios, como la duración de la sílaba o la velocidad del habla. Los valores de duración pueden también analizarse buscando las cantidades absolutas en cada vocal o mirando las oposiciones y contrastes entre unas cantidades y otras. (Hidalgo & Quillis, 2012, p. 84). De cualquier manera, la variabilidad del fenómeno salta a la vista, porque puede cambiar según diversas condiciones físicas o psicológicas del hablante.

Se entiende la duración a partir de lo expuesto por Gil-Fernández (2007) quien la define como la cantidad de tiempo empleada en la emisión de un sonido. Eventualmente el término *duración* se usa de la misma manera haciendo referencia a la cantidad vocálica y a la longitud de un sonido del habla². Esta autora se refiere a la duración como el tercer atributo fundamental de los sonidos; sin embargo, vale la pena contrastarlo con la descripción de Cantero (2019) quien plantea que ninguno de los atributos tiene mayor relevancia, sino que todos están entrelazados conformando la red del *continuum* físico, así como lo expuesto por Muñetón (2017), quien tampoco atribuye una jerarquía a los parámetros de la estructura prosódica, sino que intenta demostrar cómo todos estos son relevantes y se interrelacionan.

Por su parte, Gil-Fernández (2007) plantea que la duración de los sonidos del habla se encuentra entre 30 y 300 ms, y sin importar la variabilidad que se presente en ellos, no pueden sobrepasar la franja comprendida entre los 10 y los 40 ms, ya que es el tiempo requerido por el oído humano para tener la capacidad de hacer la distinción de cada segmento del habla. Adicionalmente, se expone que la duración de cada sonido del habla está condicionada por los elementos contiguos en su realización; de esta manera, se explica que la tendencia de las vocales a ser menor cuando son cerradas o altas y a ser mayor cuando son abiertas es considerada casi un universal fonético, lo cual está determinado por una necesidad fisiológica, ya que las vocales abiertas necesitan una mayor extensión espacial para el movimiento articulatorio, lo que requiere mayor tiempo en el habla.

También se han estudiado lenguas donde los factores aerodinámicos producen cambios significativos; no obstante, se concluye que los tres factores que condicionan la duración del segmento del habla son las características articulatorias de los sonidos vecinos, la estructura segmental de la sílaba en la que se encuentra el sonido en cuestión y el lugar que ocupa esa sílaba dentro del enunciado (Gil-Fernández, 2007).

La duración es, además, un parámetro físico con una funcionalidad lingüística ligada a otros aspectos de la comunicación verbal, como el ritmo, el acento y la entonación. En ese sentido, Cantero (2019) señala que el habla no es una cadena de sonidos cohesionados de manera lineal como usualmente se considera, sino que, se trata de una red entramada entre sonidos vocálicos y consonánticos, donde las vocales tienen la cualidad de ser núcleos sonoros y las consonantes zona marginal, es decir, obstáculos en la producción oral con mayor o menor grado de perceptibilidad. También plantea que no todas las vocales son núcleo de una unidad suprasegmental y, por lo tanto, prosódicamente no son igualmente relevantes.

Lo anterior evidencia una jerarquía fónica donde las vocales átonas suelen centrar su identidad tímbrica o hasta elidirla, mientras que las tónicas rara vez se neutralizan o eliden, en tanto que en su tono recae la unidad de significado. En algunos casos, constituyen el núcleo del grupo fónico y a su vez el núcleo de la melodía del enunciado. De esta manera, se identifica que la duración es un parámetro acústico que puede tener funciones comunicativas en el plano segmental de la vocal y en el plano suprasegmental del enunciado.

2. Para efectos de este trabajo, se toma el término *duración* como el más representativo, debido a la forma de tratamiento que se le da a este término en los antecedentes consultados haciendo referencia al fenómeno en cuestión. No obstante, en los estudios más clásicos, como los de Navarro (1918) y Quilis (1993), el término usado es *cantidad*, pero en investigaciones posteriores este último está asociado al plano de la perceptibilidad, lo cual no se incluye en la dimensión de este trabajo. En relación al término longitud, no se encuentra referencia al fenómeno en ningún antecedente más que en el de Gil-Fernández (2007) y se puede interpretar el uso de este término como una adaptación del inglés, donde en todos los referentes se hace uso de la palabra *length*.

2.2. *El habla de Medellín*

Se puede hablar del español de Medellín como un habla regional, pero sin precisar hasta dónde llegan sus fronteras lingüísticas que podrían marcarse bien sea dentro de la misma área metropolitana o extenderse a isoglosas mucho más amplias en la región circundante, pues en Antioquia y dentro de la región que se clasifica como subdialecto antioqueño hay muchas otras distinciones de tipo segmental y suprasegmental que podrían resultar de interés para la investigación lingüística.

Por ejemplo, González-Rátiva (2022) identifica que en el habla de Medellín hay una variación fonológica con respecto a Antioquia, caracterizada por procesos de espirantización de /d/, pluralización de *haber* y la tendencia antihiática de /ea/, diferencia que parece estar encaminada a confluir más en una norma urbana andina, mientras que a nivel regional las adiciones vocálicas, la elisión de codas consonánticas, los refuerzos consonánticos y la aspiración de /f/ se relacionan más con una identidad antioqueña. Puede resultar interesante del mismo modo para los estudios de duración examinar si la variación de la duración vocálica se comporta conforme a las cualidades prosódicas de las variantes subdialectales, o si, por el contrario, este atributo evidencia semejanzas que se puedan categorizar en un nivel más amplio como a nivel de un superdialecto.

Sobre el habla de esta ciudad también se han investigado el uso particular de marcadores discursivos — *hágale, o sea*—, verbos de actitud proposicional, procesos fonológicos, entre otros. Al respecto, se encuentran algunos estudios (Castellano, 2013; García, 2016; Muñoz-Builes, 2021).

El habla de Medellín tiene cualidades únicas que la diferencian del habla de las regiones hispanohablantes donde se han realizado estudios de duración. Sobre el habla en la región se menciona que no es común el relajamiento vocálico típico de otras variantes dialectales de Colombia. Más bien, se mantiene un timbre normal en las vocales tónicas con tendencia al alargamiento y, en contraste, solo se da un relajamiento en posición final átona en el habla no cuidadosa. Por lo tanto, cuando se hace referencia al vocalismo de esta variante dialectal se describe como:

de timbre claro y definido, aunque en el habla rápida y descuidada sean corrientes algunas modificaciones del timbre vocálico, sobre todo cuando se trata de las vocales átonas /e/, /o/ que tienden a cerrarse o a convertirse en /i/, /u/ respectivamente, o a relajarse en posición final. En esta posición también la /a/ se ve afectada (Betancourt, 1993, p. 289).

3. Metodología

3.1. *Corpus*

Las entrevistas están contenidas en el corpus sociolingüístico de PRESEEA-Medellín, desarrollado en el marco del Proyecto para el Estudio Sociolingüístico del Español de España y América (González-Rátiva, 2008). Son un total de 119 entrevistas almacenadas en una base de datos que incluye archivos sonoros y de texto. Cada entrevista

tiene una duración promedio de 45 minutos y contienen ejes temáticos como el clima, el barrio, el vecindario, la vivienda, la ciudad, la gente que vive en Medellín, problemáticas de la ciudad, familiares y amigos, profesión y trabajo, esparcimiento, costumbres, deseo de mejora económica y relato de narraciones espontáneas a partir de las preguntas del entrevistador.

Debido al amplio tamaño del corpus, se seleccionaron 12 entrevistas, de las cuales se tomaron 83 enunciados como muestra para el análisis. De estos enunciados, se obtuvo un total de 1006 vocales que fueron objeto del análisis acústico. Del corpus se obtuvieron tres archivos: el registro sonoro, la transcripción de la entrevista en un formato simple y la misma transcripción de la entrevista detallada con etiquetas de pausas, tiempo, alargamiento, palabras cortadas, cita, habla simultánea, marcas de intervención del entrevistado y entrevistador, énfasis y vacilación.

3.2. Muestreo

Las muestras que se tomaron de las entrevistas consisten en segmentos enunciativos que conforman oraciones simples. Se entiende para esta investigación la noción de oración simple desde lo expuesto por Gili Gaya (1980) quien la define como una unidad lingüística que tiene un sujeto y un predicado y que depende de la calidad psicológica del juicio del hablante y de la naturaleza del predicado. Desde el plano pragmático, los segmentos que se seleccionaron corresponden a actos de habla ilocucionarios de tipo asertivo (Searle, 1990). Estos se comprenden como emisiones de secuencias de palabras dentro de oraciones, que, para este caso, pueden ser afirmativas o negativas, realizadas en contextos de entrevista, bajo condiciones específicas de la situación de la conversación y con la intención del hablante de dar información respondiendo a las diversas preguntas que el entrevistador le plantea.

Otra cualidad que contribuyó en la delimitación de los segmentos del muestreo es que los enunciados se aproximen a la estructura de foco informativo. Se entiende esta noción desde lo expuesto por Bravo (2008), quien lo define como el elemento que aparece en la posición final de la cláusula y, para que sean informativos, todos sus componentes deben ser información nueva, lo cual es difícil encontrar con estas restricciones, por lo que su aparición es de frecuencia reducida, dando como resultado un número limitado de segmentos para analizar.

Esta propiedad de la estructura informativa se tuvo en cuenta, ya que se trabajó con corpus de habla semiespontánea y, al ser un tipo de habla que posee variados elementos entonativos y emocionales, se procuró con estos criterios delimitar los tipos de enunciados buscando que fueran enunciados que no tuvieran marcas entonativas muy marcadas, ayudando a controlar también de esta manera el ritmo de habla. De ese modo, se procuró que fueran enunciados concluyentes, que no respondieran de inmediato a la pregunta con *sí* o *no*, que fueran enunciados sin muletillas, ni superlativos, ni pausas intermedias, que evitaran tener diminutivos o condicionales y que su extensión estuviera en un promedio de 15 sílabas.

3.3. Variables de análisis

3.3.1. Duración vocálica

Es la medida de la vocal obtenida del análisis espectral y cuya medida se expresa en un valor de milisegundos (ms).

3.3.2. Vocales del español /a/, /e/, /i/, /o/, /u/

Se tomó todo el grupo de las cinco vocales del español.

3.3.3. Estructura silábica

Hualde (1999) define la sílaba como un conjunto de segmentos agrupados en torno a un núcleo vocálico. A nivel fonético-articulatorio, este autor expresa que hay una sincronización precisa entre la consonante a principio de sílaba y la vocal siguiente, mientras que la consonante posnuclear tiene una realización menos coordinada con la vocal que la precede; aun así, se conserva una relación existente entre estos dos elementos en el plano de la duración.

La variable de estructura silábica corresponde al contexto silábico de cada vocal, el cual se presenta en modo de núcleos silábicos simples o complejos. Al igual que otras variables, la estructura silábica puede presentar más frecuencia de uso de algunas combinaciones que otras; por tal motivo, se requirió hacer agrupaciones de contextos silábicos que tuvieron poco rendimiento. De ese modo, la categoría CCVC y CVCC, cuyas ocurrencias fueron de 11 casos, se clasificaron dentro del grupo CVC.

Vale la pena resaltar que todas las divisiones silábicas se hicieron siguiendo un patrón fonológico. De hecho, en contextos de juntura de palabra se categorizó según esta propuesta y solo en los casos donde se elidió algún fonema y se asimiló la vocal a otra sílaba se permitió clasificar como la nueva estructura silábica generada. Por ejemplo, en el fragmento *posibilidad d'empleo* el contexto *d'em-* se clasificó como CVC por tener una elisión y una juntura claramente identificables.

A continuación, se presentan las estructuras silábicas consideradas en la investigación con un respectivo ejemplo de cada una:

Núcleos silábicos simples		Núcleos silábicos complejos	
Estructura	Ejemplo de contexto	Estructura	Ejemplo de contexto
V	O-fi-cial	C + Dip (CD)	Fue
CV	Pe-ro	C + Dip + C (CDC)	Piel
CCV	Pro-fe-sor		
VC	Ad-mi-tir		
CVC	Pis-ta		

Tabla 1. *Clasificación de grupos silábicos encontrados en el corpus e incluidos en el análisis*

3.3.4. Segmento anterior, posterior y sonoridad

Esta variable hace referencia a los segmentos consonánticos que se presentan en posición anterior o posterior de las vocales. Estos segmentos se agruparon por modo de articulación en: aproximantes, fricativas, nasales, líquidas y oclusivas. Adicionalmente, se encuentra dentro de esta variable la categoría de pausa, que ocurre cuando la vocal está al inicio o al final del enunciado. También surgió la categoría denominada semi, que surge cuando la vocal esta antecedida o seguida por semiconsonante o semivocal. Adicionalmente, como una categoría complementaria a esta, se registró la sonoridad de cada uno de los segmentos consonánticos adyacentes. En esta subcategoría solo se tuvieron tres divisiones: sordo (SR), sonoro (SN) y no aplica (NA). Esta última cuando el elemento vocálico estaba precedido o seguido de pausa.

3.3.5. Número de sílabas por palabra

Esta variable se estableció por el conteo fonológico de las sílabas de la palabra donde se presenta cada vocal. Las palabras variaron entre 1 y 6 sílabas, las cuales se clasificaron en mono, bi, tri, y polisílabas.

3.3.6. Tonicidad

Esta variable corresponde al acento léxico y en este caso cada vocal puede aparecer en contexto de sílaba tónica o en sílaba átona dentro de cada palabra. En este caso, no se está considerando ninguna relación de posición dentro del enunciado ni de función gramatical de la palabra, como sí se hace en las siguientes variables.

3.3.7. Tonicidad enunciativa

Hace referencia a la posición en que se presentaron las vocales dentro del enunciado con respecto a la tónica nuclear del segmento y la posición acentual dentro de cada palabra o sintagma en el conjunto oracional.

De ese modo, se obtuvo la siguiente clasificación por tipo de tonicidad de la sílaba dentro del enunciado: antepretónica, pretónica, tónica, postónica, pretónica nuclear, tónica nuclear y postónica final. Las vocales tónicas en esta categoría estuvieron enumeradas para identificar la posición que ocupa dentro de la longitud del enunciado.

3.3.8. Categoría léxica

Esta variable se incluyó teniendo en cuenta que la función sintáctica de la palabra dentro del enunciado podría llegar a tener alguna incidencia en los valores de duración. Esta característica se relaciona con la variable de acento; sin embargo, en este caso se etiquetó cada palabra según su función lexical o gramatical dentro de la oración.

En ese sentido, Hidalgo y Quilis (2012) causing the pole inequality relations between men and women. Therefore, in this study wanted to dismantle the detail view of some theories, both social and feminist about gender relations in the family. Each of these theories (structural functional, conflict and feminist plantean que las palabras, según su categoría sintáctica de tipo lexical o gramatical, adoptan rasgos de acentuadas o inacentuadas. Para ello, brindan un listado donde clasifican las palabras según su categoría gramatical y la acentuación, describiendo, por ejemplo, que los sustantivos, verbos, adjetivos, adverbios, formas interrogativas y algunos pronombres tienen sílabas portadoras de acento; mientras que palabras de tipo preposiciones, conjunciones, adversativas, causales, entre otras, no son portadoras de acento. Con base en esta clasificación expuesta, se clasificó el corpus según la categoría léxica o gramatical.

3.3.9. Cantidad de sílabas por enunciado

Hace referencia a la clasificación de los enunciados según su cantidad de sílabas. Esta selección de variable se decidió en tanto que los enunciados se encontraban en un rango promedio de 15 sílabas y se consideró identificar si la constitución de un grupo fónico podía tener alguna incidencia en la conformación de la duración.

3.4. Tratamiento de datos sonoros

Inicialmente, se seleccionaron las 12 entrevistas y se utilizó el programa Audacity (versión 2.3.3; Audacity Team, 2020) para segmentar los 83 enunciados. Luego, estos se analizaron y se buscó que conservaran de la manera más rigurosa posible las características descritas en el muestreo.

Posteriormente, se utilizó el programa de uso libre Praat (versión 6.1.55; Boersma, Paul, & Weenink, 2021) para realizar el análisis acústico. Los 83 enunciados se organizaron en archivos de tipo proyecto Praat, con el fin de aplicar modificaciones similares en su conjunto. En la segmentación de cada vocal, se tuvo en cuenta el punto de inicio, donde comienza la oscilación pronunciada de la onda sonora y donde termina el patrón de su configuración. No obstante, se encuentran algunos casos cuyas divisiones son difíciles de establecer y durante este

proceso se evidenció que hay muchos estudios fonéticos realizados sobre la base de análisis espectrográfico, donde los límites entre un fonema y otro son más perceptibles. Sin embargo, se comprobó que hay una carencia didáctica de algún material en lengua española que oriente el proceso de un análisis acústico sobre habla espontánea.

Una buena guía que se pudo encontrar y aborda de manera general la metodología del análisis de la duración se encontró en Wright y Nichols (2015). Estos autores sugieren las siguientes consideraciones a tener en cuenta durante la medición de la duración:

- Excluding the consonant release burst and aspiration depending on what you are trying to measure.
- Fricative noise is never included as vowel duration since fricative turbulence is only generated during the consonant's closure phase.
- When measuring around fricatives, look for the point where the fricative turbulence ends (or changes dramatically in intensity) and where higher formant structure becomes visible in the spectrogram (higher formants are F2, F3, and maybe F4).
- When excluding the following consonant's closure, look for the point where there is a marked drop in intensity together with a loss of energy in the higher formants.
- The end of the vowel is marked by three cooccurring events: 1) a dramatic change in amplitude in the waveform, 2) a change in the energy in the formants accompanied by a change in complexity in the waveform indicating a loss of energy in F2 and F3, 3) the onset of aperiodicity (p. 2).

En el análisis se identificó un patrón recurrente para la configuración de las vocales. Ese patrón comienza a manifestarse con una longitud de onda grande que progresivamente va modificándose de manera coordinada con su amplitud, se genera una extensión de esta amplitud y, a la misma vez, comienza a reducirse la longitud. Parece un patrón acústico coordinado entre incremento y reducción de la amplitud y longitud de onda al inicio y final de la configuración espectral de las vocales. Paralelamente, comienzan a revelarse los formantes y se estabilizan en su huella espectral, de tal forma que dejan una marca resaltada en negro en el espectrograma.

En la ventana del oscilograma de las vocales, la amplitud de la onda puede ser cambiante con altibajos dependiendo de la intensidad de la vocal y, hacia el final, la amplitud se reduce y la longitud nuevamente se amplía para dar paso a la conformación de un nuevo patrón espectral de la consonante adyacente.

Adicional a las sugerencias referenciadas y la observación descrita, para los casos donde se presentó juntura vocálica por final, por inicio de palabra, por diptongo o hiato, se estableció la segmentación a partir del punto medio donde se apreciaba el cambio de formantes de una vocal a otra. Además, en las ocurrencias donde el espectrograma no permitía ver esas transiciones, se tuvieron en cuenta los picos ascendentes y descendientes de las ondas en el oscilograma y así como, en algunas de ellas, la visualización de la barra de intensidad ayudó a identificar puntos de inicio o final de las vocales. Entretanto, en los casos más complejos, una observación de todos los aspectos descritos anteriormente fue necesaria para poder establecer las fronteras acústicas de las vocales objeto de estudio.

Una vez se marcó cada segmento vocálico en cada enunciado, se continuó con la toma de la duración siguiendo la metodología expuesta por Díaz y Correa (2014) y por Boersma y Weenink (2020), quienes sugieren realizar

la medición a partir de la resta del tiempo de inicio de un intervalo a su tiempo final. Este procedimiento se realizó de manera automatizada a través de dos *scripts* tomados de Román (2021) y adaptados por Muñoz-Builes (2021). El primero se aplicó con el fin de insertar un intervalo adicional de punto que marcara el centro de cada vocal segmentada. El segundo se aplicó para extraer los datos de duración de cada vocal en un archivo externo, que posteriormente se importó y colectó en una hoja de cálculo, de manera que fuesen llevados al programa de tratamiento estadístico.

3.5. Estadística empleada

Para el tratamiento estadístico de los datos se utilizó el programa de uso libre para análisis estadístico JASP (versión 0.16.0.0; JASP Team, 2022). En este, se obtuvieron los estadísticos descriptivos que permitieron obtener los valores de la media de la duración de cada vocal, así como las distribuciones de los valores dentro del conjunto obtenido. Luego, se continuó con el tratamiento de los datos a través de análisis de varianzas (ANOVA), dado que la estadística descriptiva no es suficiente para interpretar las posibles correlaciones a partir de las variables propuestas.

El procedimiento general en la parte estadística fue el siguiente: 1) obtención de los gráficos descriptivos, 2) obtención de la tabla con estadísticos descriptivos, 3) interpretación y descripción de dichos valores y 4) aplicación de la prueba ANOVA con verificación de supuestos y tamaño de efecto. En todos los casos se aplicó el método de corrección Scheffé para solucionar las diferencias de tamaño de la muestra de cada vocal.

En los casos de verificación de supuestos en los que la prueba fue homogénea a partir del test de Levene, se adoptó el resultado ANOVA bajo el condicionamiento de la potencia del tamaño de efecto η^2 cuadrado parcial (η^2_p) como el indicador de la probable correlación de las variables (Goss-Sampson, 2018). En los casos en que la prueba de Levene no fue homogénea, se aplicó la prueba no paramétrica Kruskal-Wallis, sin que se tenga en esta una medición del tamaño de efecto, en tanto que este tipo de prueba no permite ese cálculo de modo general.

4. Análisis de los resultados

4.1. Duración de las vocales

Los resultados permiten observar diversas manifestaciones acústicas del fenómeno lingüístico estudiado. El primer aspecto que se identifica es el rendimiento vocálico que se presentó en el grupo total de los enunciados. En ellos, se evidencia que las cantidades obtenidas de cada vocal son desiguales, aunque este factor no representa un problema para el tratamiento estadístico de los datos, dado que se trabajó con comparación de valores de la media de cada vocal y, en principio, los promedios no son dependientes de la cantidad de los datos.

La muestra quedó conformada por una cantidad de datos de la vocal /a/ con 29,8 %, la vocal /e/ con 28,7 %, la vocal /i/ con 11,8 %, la vocal /o/ con 24,1 % y la vocal /u/ con 5,4 % de la muestra total, respectivamente. Las frecuencias para cada una se pueden observar en la Tabla 3, así como los valores promedio de la duración, la desviación y asimetría que muestran los datos.

La duración vocálica obtenida a partir de los estadísticos descriptivos reveló que el habla de Medellín presenta una media de duración de mayor a menor en el siguiente orden a partir de una clasificación general de las vocales: /i/ 71,11 ms; /a/ 70,8 ms; /u/ 64,88 ms; /e/ 61,89 ms; y, por último, /o/ 59,27 ms.

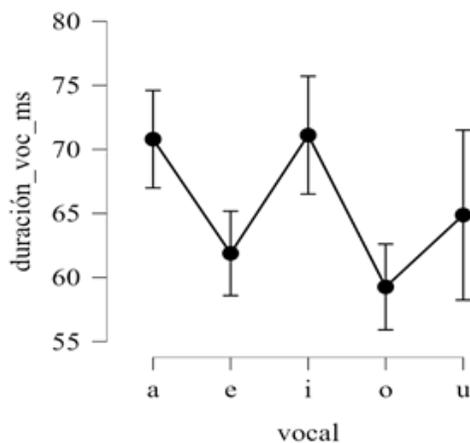


Figura 1. Gráficos descriptivos de la duración según la categoría vocálica general

duración_voc_ms					
	a	e	i	o	u
Válido	300	289	119	243	55
Moda	52.10	48.40	53.90	34.30	29.60
Mediana	66.20	58.50	68.00	51.40	63.70
Media	70.80	61.89	71.11	59.27	64.88
Curtosis	1.30	1.29	-0.20	2.27	2.07
Mínimo	9.80	15.00	23.20	15.10	29.60
Máximo	197.50	179.10	140.80	174.90	155.80

Tabla 2. Estadísticos descriptivos de la duración según categoría vocálica general

La figura anterior muestra la distribución de los valores de duración vocálica en el gráfico descriptivo. En este, se observa la variación en los valores promedio de la duración, destacándose la superioridad de /i/ e /a/ con respecto a las otras vocales. Estos datos contabilizan todas las vocales indistintamente de otras categorías que cumplan o no, aunque esta generalización puede ser controvertida, puesto que otros estudios han tratado de tomar la composición vocálica de estructura simple frente a la compleja por separado. No obstante, a pesar de la

ocurrencia de las vocales en contextos complejos, se procuró que la medición registrara únicamente el segmento vocálico, lo que no debería generar una distinción entre valores de duración a partir de este criterio más que en la variable de segmento anterior y posterior.

Tras un análisis ANOVA, se buscó conocer si existían diferencias que resultaran significativas de las medias obtenidas de las vocales, según su clasificación general. Los datos mostraron que la duración vocálica está afectada significativamente por la modalidad de tratamiento del análisis, $H = 32.21$, $p < .001$. Una vez que se determinó la existencia de diferencias estadísticamente significativas, se aplicaron las pruebas *post hoc* para identificar entre cuáles vocales se manifiestan los valores de significancia, encontrando que estos se dan entre la vocal /a/ con respecto a las vocales medias /e/ y /o/, al igual que la /i/ con respecto a la /o/.

Las diferencias significativas de la duración en las vocales según la categoría vocálica sugieren dos ideas concernientes. Por un lado, la duración en el habla espontánea de esta variante dialectal es más amplia cuando se producen en puntos articulatorios extremos, como es el caso de la vocal /a/ con relación a /i/, mientras que es más reducida para puntos articulatorios medios, como en las vocales /e/ y /o/. Por otro lado, la significancia en el contraste de las medias sugiere que la configuración articulatoria incide en los valores de la duración vocálica, aunque los valores promedio para la duración de las vocales no se comporten en un orden jerárquico de las más abiertas a las más cerradas. La vocal /u/ parece tener un comportamiento acústico de la duración más variado que oscila entre los valores de la media de las otras vocales.

Conociendo estos resultados preliminares de la duración vocálica a partir de su estructura interna, se concluye que la duración de las vocales para la variante dialectal del español de Medellín se manifiesta en orden de mayor a menor para las vocales en extremo anterior alto /i/ y extremo central bajo /a/, seguido de la vocal en extremo posterior alto /u/ y terminando con las vocales medias posterior y anterior /e/ y /o/, como se muestra en la siguiente figura:

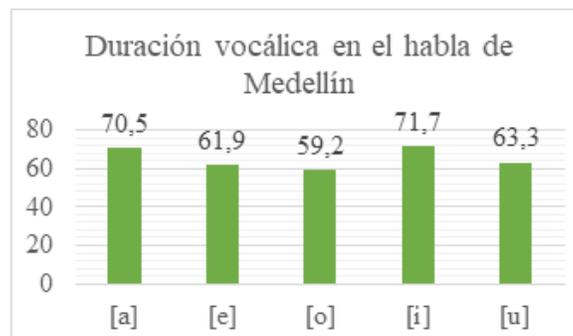


Figura 2. Duración vocálica en el español de Medellín a partir del corpus PRESEEA

Esos valores se corroboran cuando se miran desde un plano aislado de cada vocal y de las cinco vocales en conjunto, donde los valores de duración se comportan de modo similar. Una vez obtenidos estos, se puede obtener el valor de duración intrínseca de las vocales con el promedio de la media estadística de cada vocal, el cual

resultó ser de 65,32 ms para la variante dialectal del español de Medellín. Esta cifra también se corrobora con los resultados obtenidos a partir de la variable de núcleo silábico simple.

A partir de los datos obtenidos, se puede afirmar entonces que los 65,32 ms se constituyen en el valor intrínseco de duración y sirve de referencia para las vocales de esta variante dialectal. Con esta medición se obtiene un referente descriptivo sobre el cual se puede aplicar el umbral perceptivo de duración del 36 % descrito por Pamies y Planas (2015), a partir del cual se pueden diferenciar variaciones entre vocales breves y largas dentro de esta variante del español.

4.2. Núcleo silábico

Teniendo en cuenta que se busca describir valores de referencia en el comportamiento acústicos de las vocales, se presentan los descriptivos y los valores ANOVA de manera diferenciada para cada vocal que conforma los núcleos silábicos simples. En estos, los datos se comportan del mismo modo que la agrupación general de vocales y parecen presentarse diferencias significativas en los valores de la media con un tamaño de efecto medio. Sin embargo, las pruebas *post hoc* clasifican las diferencias entre pares de /a/ en relación a /o/ y entre /i/ con relación a /o/. Este sutil contraste podría reforzar la idea de que los cambios en la duración vocálica están influenciados por la composición de la vocal y se manifiestan principalmente entre puntos articulatorios en extremos opuestos.

duración_voc_ms					
	a	e	i	o	u
Válido	276	252	115	210	53
Moda	52.10	65.30	53.90	45.30	29.60
Mediana	66.20	59.55	70.70	51.00	63.70
Media	70.52	63.62	71.98	58.71	65.43
Curtosis	1.04	1.05	-0.21	1.73	2.07
Mínimo	9.80	15.00	23.20	15.10	29.60
Máximo	180.90	179.10	140.80	159.60	155.80

Tabla 3. Estadísticos descriptivos según categoría vocálica y núcleos simples

De otro modo, se analizan los resultados a partir de la variable de núcleo silábico agrupando todas las vocales. Los datos descriptivos evidencian una recurrencia mayor de los núcleos simples, al igual que un valor de duración mayor con respecto a los núcleos silábicos complejos. Sin embargo, ambos grupos presentan una alta desviación en los valores de duración, lo cual sugiere una distribución anormal de los datos. Realizando la observación de los gráficos descriptivos en la Figura 5 se evidencia la amplia diferencia de los valores de la media, pero por la variación que se presenta en los valores de duración de los núcleos complejos, que incluso alcanza el percentil de duración de los núcleos simples, se procede a continuar con un análisis de varianza no paramétrico.

En de la Figura 3 se puede observar que las medias puntúan en valores de duración distantes entre 60 ms para los núcleos complejos y 65 ms para los núcleos simples. Estos valores son más estables en estos últimos con respecto a los núcleos complejos que presentan más variación. De hecho, con una sobreposición en los quintiles que alcanzan los valores de duración de los núcleos simples. Estos datos sugieren que cuando las vocales ocurren en contexto de diptongos su duración tiende a reducirse a diferencia de cuando ocurren en contextos de núcleo vocálico simple:

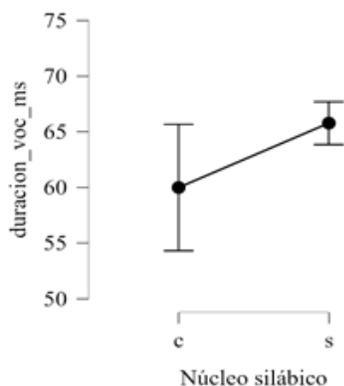


Figura 3. Gráficos descriptivos de la duración según el núcleo silábico

duración_voc_ms		
	complejo	simple
Válido	107	899
Moda	26.50	48.10
Mediana	53.50	61.30
Media	59.99	65.78
Curtosis	5.72	1.20
Mínimo	20.40	9.80
Máximo	197.50	180.90

Tabla 4. Estadísticos descriptivos de la duración según el núcleo silábico

Los resultados del análisis de la varianza ANOVA de la duración vocálica, según los núcleos silábicos muestran valores $F = 3.70$, $p < 0.05$, con un tamaño de efecto bastante pequeño de $\eta^2_p = < 0.01$. Estos valores sugieren unas diferencias significativas entre los valores de duración de los núcleos complejos y simples, pero estas diferencias no se pueden corroborar estadísticamente, debido al tamaño de efecto bastante bajo de los resultados. Por lo tanto, los datos revelan que, estadísticamente, en menor grado, la composición de los núcleos silábicos incide sobre los valores de la duración de las vocales para el análisis expuesto.

4.3. Estructura silábica

Las observaciones iniciales con respecto a esta variable sugieren un contraste de valores de duración entre estructuras silábicas abiertas con respecto a las cerradas, y entre estructuras silábicas de núcleos complejos frente a las estructuras silábicas con núcleos simples. Es llamativo el hecho de que una estructura silábica compleja tipo CCV sea la que presente mayor duración, dado que tiene ocurrencia con dos consonantes que le preceden y se podría pensar que, a mayores segmentos dentro de la estructura, se da una mayor duración de ellos. No obstante, los resultados no permiten tener una evidencia clara, debido a la variabilidad en cada una de las estructuras silábicas. Adicionalmente, lo que sí muestra esta variable es otro patrón correspondiente a la apertura de la sílaba. De este modo, si la sílaba es abierta, posee mayor duración que cuando es cerrada.

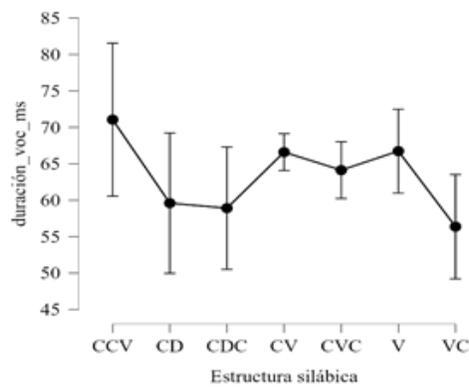


Figura 4. Gráficos descriptivos de la duración según la estructura silábica

duración_voc_ms							
	CCV	CD	CDC	CV	CVC	V	VC
Válido	41	53	36	534	158	96	76
Moda	18.70	26.50	62.20	48.10	71.50	35.40	22.60
Mediana	65.80	51.10	54.45	61.30	61.05	65.05	48.30
Media	71.06	59.60	58.92	66.60	64.13	66.74	56.38
Curtosis	-0.84	5.29	5.37	1.36	0.25	1.74	3.18
Mínimo	18.70	17.20	24.10	9.80	19.30	13.90	15.00
Máximo	139.30	197.50	155.20	180.90	139.40	177.10	179.10

Tabla 5. Estadísticos descriptivos de la duración según estructura silábica

La prueba ANOVA para esta variable revela los valores de $F = 2.30$, $p < 0.03$, con un tamaño de efecto pequeño de $\eta_p^2 = < 0.01$. Estos valores obtenidos, si bien muestran diferencias significativas entre las estructuras de sílaba abierta y cerrada que se pueden corroborar en las pruebas *post hoc*, su tamaño de efecto no le concede potencia a la aplicación de la prueba, por lo que las diferencias en las medias de esta variable no permiten afirmar que la estructura silábica tenga incidencia en la configuración de los valores de la duración de las vocales.

4.4. Segmento anterior

En la variable del segmento anterior, que hace referencia a los elementos que preceden a la vocal, se encontró que esos elementos pueden estar constituidos por segmentos consonánticos, semiconsonánticos y de pausa. Los gráficos reflejan una duración de la vocal, que se incrementa cuando está precedida por aproximantes y líquidas, registrando valores de 73,83 ms y 70,61 ms, respectivamente, mientras que, en oposición las vocales sin segmento previo, parecen ser las de menor duración, con 56,97 ms.

Es necesario mencionar que para este estudio los valores de duración se manifestaron con una tendencia de mayor a menor duración, en el orden aproximantes, líquidas > nasales, fricativas > oclusivas > africadas. La posición contigua a las pausas y semiconsonantes resultan mantener unos valores de duración medios:

duración_voc_ms									
	Afric.	Aprox.	Fric.	Liq.	Nas.	Oclus.	Pausa	Semi	Vocal
Válido	17	46	144	144	169	257	23	85	109
Moda	15.10	54.70	46.10	35.00	27.60	48.10	13.90	26.50	35.40
Mediana	47.20	60.95	60.15	69.55	62.20	60.00	70.40	53.50	46.10
Media	59.34	73.83	65.70	70.61	66.11	64.42	64.33	60.38	56.97
Curtosis	1.65	0.48	1.67	0.15	2.14	1.37	-0.02	5.68	2.35
Mínimo	15.10	29.50	17.20	18.70	24.30	9.80	13.90	20.40	15.00
Máximo	148.40	173.80	180.90	170.00	177.10	164.00	95.80	197.50	179.10

Tabla 6. Estadísticos descriptivos según el modo de articulación del segmento anterior

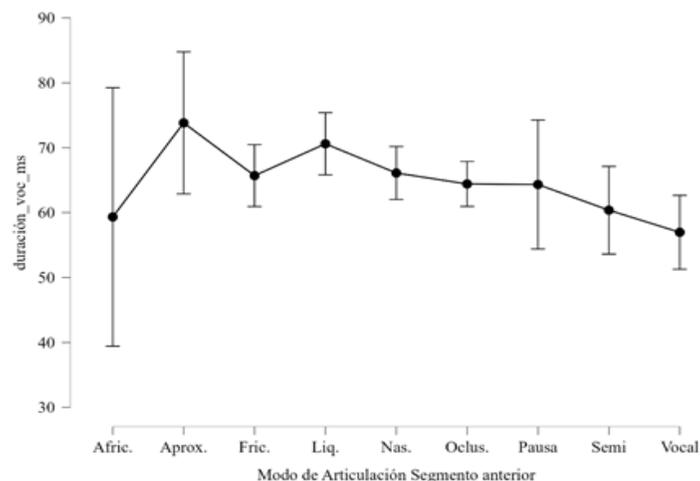


Figura 5. Gráficos descriptivos de la duración según el modo de articulación del segmento anterior

La prueba ANOVA para esta variable muestra valores $F = 2.62$, $p < 0.001$, $\eta_p^2 = 0.02$. Estos valores indican que hay niveles de significancia entre los grupos de esta variable, aunque el tamaño de efecto sea pequeño. La prueba *post hoc* muestra que las diferencias entre los valores de duración que resultan significativas son las que se presentan cuando la vocal está precedida de consonante, frente a los casos cuando se encuentra precedida de semiconsonante u otra vocal en juntura de palabra.

4.4.1. Sonoridad del segmento anterior

En relación a la sonoridad que se presentó en el segmento anterior, los gráficos muestran que la duración vocálica es mayor si el segmento anterior es sonoro y su nivel disminuye cuando está acompañada en su límite anterior por un segmento sordo. Estos resultados se asemejan a lo expuesto por Marín (1995), quien también describe una duración mayor para los segmentos sonoros:

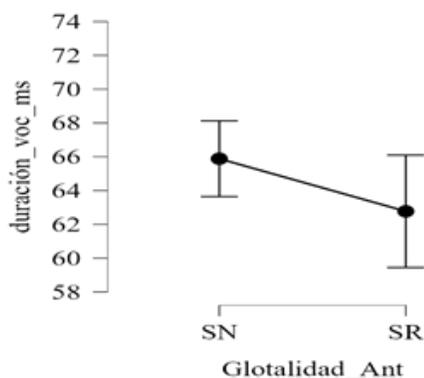


Figura 6. Gráficos descriptivos de la duración según la sonoridad del segmento anterior

duración_voc_ms		
	SN	SR
Válido	692	279
Moda	48.40	65.00
Mediana	61.40	57.20
Media	65.88	62.77
Curtosis	1.72	1.26
Mínimo	15.00	9.80
Máximo	197.50	164.00

Tabla 7. Estadísticos descriptivos según sonoridad del segmento anterior

Si bien los estadísticos descriptivos señalan una significancia en las diferencias de las duraciones entre la sonoridad de los segmentos prevocálicos, la prueba ANOVA y el tamaño de efecto sugiere que los valores de estas

medias no son significativos reportando los valores de $F = 2.21$, $p < 0.14$, $\eta_p^2 = 0.01$. Dicho de otro modo, la sonoridad del segmento previo a la vocal puede incidir en la configuración de la duración de la vocal, aunque no es un factor que se pueda corroborar estadísticamente.

4.5. Segmento posterior

Los segmentos que anteceden a las vocales en este trabajo reflejan una distribución en valores de duración similar a la de los segmentos en posición anterior y, de igual manera, se asemejan a los resultados revelados por Marín (1995) para los segmentos consonánticos posteriores. Los resultados sugieren que las aproximantes, africadas y líquidas tienen la tendencia a alargar los valores de duración con respecto al resto del grupo consonántico. Las pausas y la posición, seguida de vocales en juntura de la palabra posterior, revelan la tendencia a reducir la duración. Para el caso de las pausas a las que se refiere, corresponde a segmentos donde se hacía el recorte del audio hacia la parte final; diferente de situaciones de habla en las que la pausa se da en medio de todo un fragmento que se decide analizar y podría presentar valores de duración diferentes.

duración_voc_ms								
	Afric.	Aprox.	Fric.	Liq.	Nas.	Oclus.	Pausa	Vocal
Válido	11	52	208	141	234	183	54	111
Moda	29.60	55.70	38.80	40.70	48.40	48.10	87.00	35.00
Mediana	75.00	73.50	64.45	59.40	60.05	62.60	46.80	47.80
Media	74.65	80.92	65.83	67.37	64.48	64.48	61.77	55.31
Curtosis	-0.59	0.29	1.41	2.03	2.50	-0.18	1.05	0.36
Mínimo	29.60	20.40	15.00	20.80	13.90	18.70	9.80	15.30
Máximo	113.70	173.80	174.90	197.50	180.90	136.80	164.00	131.60

Tabla 8. Estadísticos descriptivos de la duración según el modo de articulación del segmento posterior

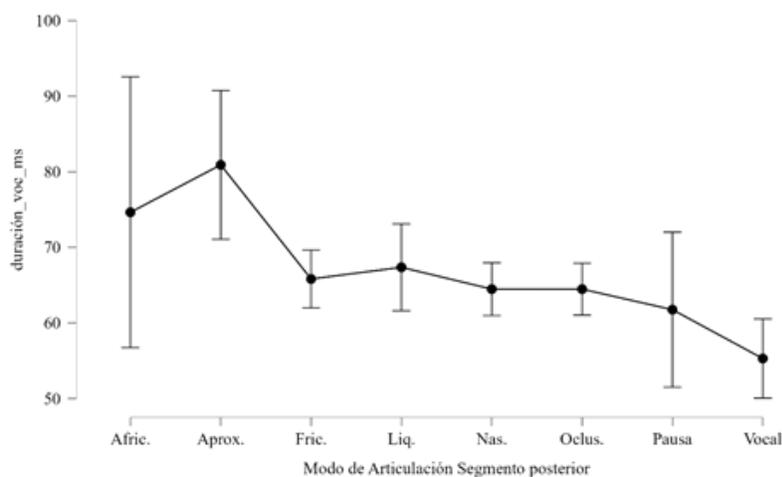


Figura 7. Gráficos descriptivos de la duración según el modo de articulación del segmento posterior

La prueba ANOVA para esta variable revela que sí hay valores significativos de diferencia en las medias estadísticas $H = 35.70$, $p < 0.001$. La comprobación del supuesto de homogeneidad muestra que se viola este supuesto, por lo que se recurre a la prueba no paramétrica, sin que se pueda reportar un tamaño de efecto. Esto quiere decir que sí hay incidencia en la duración de las vocales según el modo de articulación del segmento posterior, aunque esta conclusión no se puede extrapolar en la interpretación de otros datos que procuren describir el mismo fenómeno.

4.5.1. Sonoridad del segmento posterior

Para la variable de la sonoridad del segmento posterior los resultados revelan, al igual que en los segmentos anteriores, que hay mayor duración cuando las vocales están seguidas de segmentos sonoros. Sin embargo, en la prueba ANOVA los resultados no revelan que haya una significancia en las diferencias de los valores de duración a partir de este principio, $H = 0.57$, $p < 0.45$. Esto refuerza la idea de que, a efectos del diseño de esta investigación, se manifiesta una relación entre los modos de articulación y la duración, aunque estadísticamente no se refleja en la sonoridad de cada uno de esos segmentos.

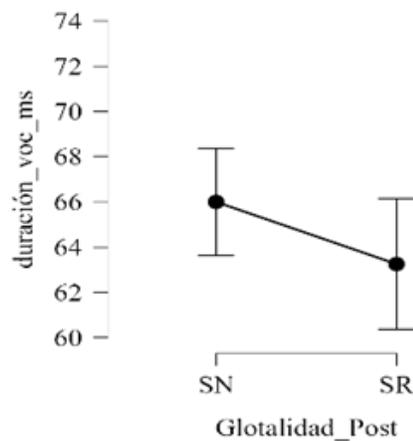


Figura 8. Gráficos descriptivos de la duración según sonoridad del segmento posterior

duración_voc_ms		
	SN	SR
Válido	630	308
Moda	49.40	38.80
Mediana	61.20	59.10
Media	66.00	63.25
Curtosis	1.54	1.83
Mínimo	13.90	15.00
Máximo	197.50	174.90

Tabla 9. Estadísticos descriptivos de la duración según la sonoridad del segmento posterior

4.6. Número de sílabas por palabra

Para la variable de número de sílabas que contiene la palabra donde ocurre la vocal, se obtuvo que la duración más alta se manifiesta en los bisílabos, con 67.16 ms, en contraste con la menor duración que se presenta en los monosílabos, con 60.39 ms, donde también hay una reducción entre trisílabos y polisílabos. Estos resultados no muestran una tendencia clara en el comportamiento de la duración a partir de esta variable, ya que, por ejemplo, los valores mínimos se asemejan entre palabras con cantidades de sílabas opuestas, es decir, que hay una menor duración, tanto en los monosílabos como los polisílabos, por lo que esta variabilidad no permite identificar un patrón de comportamiento acústico que permita hacer alguna aseveración al respecto.

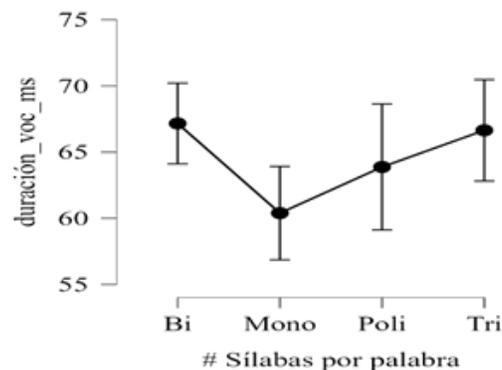


Figura 9. Gráficos descriptivos de la duración según el número de sílabas por palabra

duración_voc_ms				
	Bi	Mono	Poli	Tri
Válido	388	236	140	230
Moda	38.80	48.20	49.50	49.00
Mediana	63.40	54.50	60.25	60.95
Media	67.16	60.39	63.88	66.65
Curtosis	0.98	2.93	2.24	1.81
Mínimo	9.80	13.90	22.10	20.80
Máximo	177.10	180.90	173.80	197.50

Tabla 10. Estadísticos descriptivos de la duración según el número de sílabas por palabra

La verificación de supuestos revela el cumplimiento del supuesto de homogeneidad. Se obtienen los resultados de la prueba ANOVA mostrando los valores $F = 2.96$, $p < 0.03$, $\eta_p^2 = 0.01$. Estos resultados sugieren que, si bien hay valores de significancia en los datos de la media de duración según la cantidad de segmentos dentro de la palabra, el tamaño de efecto no permite aseverar que los resultados obtenidos en la prueba puedan demostrar que este

factor es representativo para la configuración de los valores de duración.

4.7. Tonicidad

Esta variable es la más comentada académicamente y referenciada como factor que determina la duración vocálica. Los gráficos descriptivos muestran que las diferencias de las medias son bastante marcadas, destacándose la preponderancia de la duración de las vocales tónicas con respecto a las átonas. En promedio, una vocal tónica registra una duración de 74 ms, mientras que una vocal átona registra 58 ms, como se puede observar en los gráficos descriptivos siguientes:

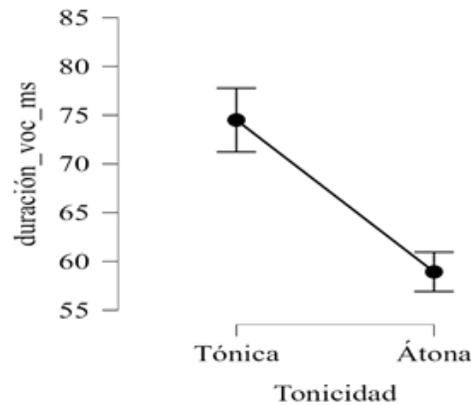


Figura 10. Gráficos descriptivos de la duración según la tonicidad

duración_voc_ms		
	Tónica	Átona
Válido	386	608
Moda	62.20	48.10
Mediana	69.35	54.70
Media	74.50	58.93
Curtosis	0.65	2.43
Mínimo	17.20	9.80
Máximo	197.50	179.10

Tabla 11. Estadísticos descriptivos de la duración según la tonicidad

La comprobación de supuestos para esta variable resulta no ser homogénea, por lo que se recurre a hacer pruebas no paramétricas. La prueba ANOVA muestra un resultado de $H = 57.56$, $p < 0.001$, reflejando una amplia correlación marcada entre la tonicidad y la duración de la vocal. De hecho, según los resultados se constituye en un factor con una fuerte incidencia sobre los valores de duración obtenidos.

4.8. Tonicidad enunciativa

Esta variable de cierta manera se complementa con la anterior, debido a que los gráficos descriptivos marcan también la prominencia de la duración en las vocales en que recae la tonicidad del enunciado. La posición de vocal tónica nuclear refleja una mayor duración de la vocal, con 96,33 ms, con respecto a la ocurrencia en otras posiciones del enunciado. También, se resalta el comportamiento acústico de la tercera tónica con una media de 75,17 ms, porque marca un incremento inicial en la curva entonativa que vuelve y toma fuerza hacia la parte final del enunciado.

Esto refleja que gran parte del comportamiento de la duración vocálica en el habla en uso está incidido por la prosodia del enunciado. De hecho, según los resultados, esta variable es la que tiene mayor incidencia en los valores de duración.

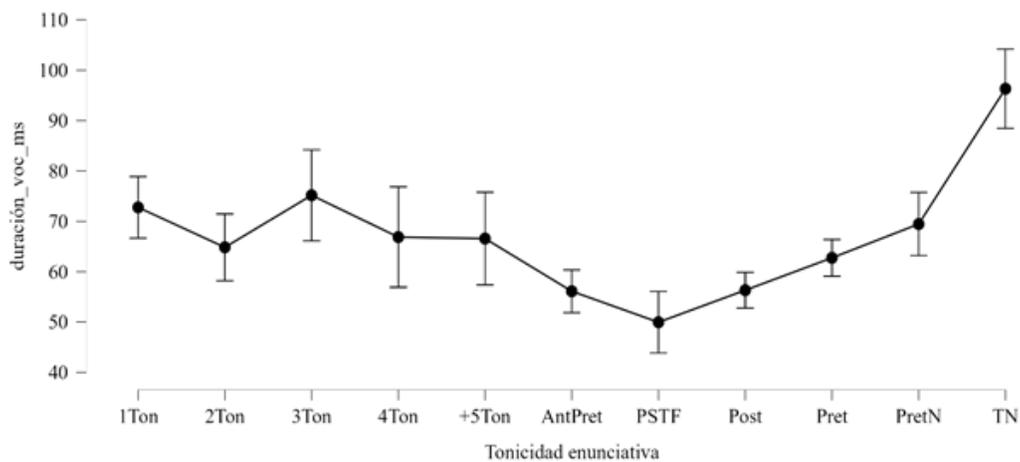


Figura 11. Gráficos descriptivos de la duración según la tonicidad enunciativa

duración_voc_ms					
	1Ton	2Ton	3Ton	4Ton	+5Ton
Válido	80	75	56	39	42
Moda	24.40	62.20	74.50	57.60	76.10
Mediana	69.50	62.20	66.75	58.70	68.90
Media	72.77	64.85	75.17	66.87	66.58
Curtosis	2.52	1.28	1.25	-0.86	0.81
Mínimo	24.40	17.20	20.40	22.90	26.80
Máximo	180.90	150.80	177.10	132.30	155.20

	AntPret	PSTF	Post	Pret	PretN	TN
Válido	133	63	192	160	75	79
Moda	32.50	87.00	43.40	49.00	48.40	35.00
Mediana	52.70	45.00	51.10	63.90	64.00	97.90
Media	56.12	49.97	56.33	62.75	69.46	96.33
Curtosis	6.51	0.98	3.61	0.42	1.07	-0.09
Mínimo	15.30	9.80	18.70	13.90	22.30	35.00
Máximo	179.10	125.40	174.90	137.20	150.80	197.50

Tabla 12. Estadísticos descriptivos de la duración según la tonicidad enunciativa

La prueba ANOVA para esta variable presenta los resultados $H = 134.35$, $p < 0.001$, corroborando que hay un alto nivel de significancia estadística de la correlación entre la variable de duración vocálica y tonicidad enunciativa.

4.9. Categoría léxica

En relación a la categoría léxica, los resultados mostraron una mayor duración de las vocales cuando se encuentran en palabras de contenido. En promedio, este tipo de palabras presentaron duraciones de 66 ms, con respecto a los 59 ms de las vocales en contexto de palabra gramatical. Este factor se complementa con los resultados de la duración en la tonicidad, pues son las palabras con contenido léxico más utilizadas y las portadoras del tono dentro de los enunciados.

De igual manera, se evidencia lo expuesto por Cantero (2019) en cuanto a la relevancia de las vocales en la configuración del significado a partir del tono, es decir, que se manifiesta una correlación entre la duración, el tono y el significado.

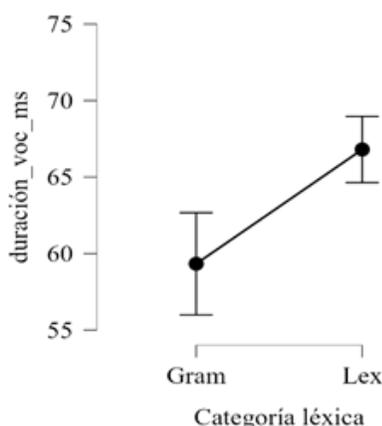


Figura 12. Gráficos descriptivos de la duración según la categoría léxica

duración_voc_ms		
	Gram	Lex
Válido	228	755
Moda	35.40	47.70
Mediana	53.95	62.30
Media	59.33	66.80
Curtosis	3.95	1.22
Mínimo	13.90	9.80
Máximo	180.90	197.50

Tabla 13. *Estadísticos descriptivos de la duración según la categoría léxica*

Los resultados de la prueba ANOVA para esta variable muestran los valores de $H = 10.39$, $p < 0.01$, revelando que, si bien los resultados parecen significativos para describir la correlación entre las variables, su nivel de significancia no es mayor de lo que permite explicar el fenómeno a través de valores expresados por otras variables como la de tonicidad enunciativa.

4.10. Número de sílabas por enunciado

En esta variable el comportamiento acústico de la duración de las vocales se manifestó con un incremento en los enunciados que estaban constituidos por 11 o por 6 sílabas, con duraciones de 77 ms y 75 ms, respectivamente. Enunciados con otra cantidad de sílabas tuvieron duraciones vocálicas que oscilaron entre los 51 y 69 ms. Estos resultados señalan que los cambios de extensión en los valores de duración de las vocales requieren de determinada cantidad de sílabas para constituir grupos fónicos donde el incremento de los valores sea posible, y ese incremento se posibilita en promedio cada seis sílabas.

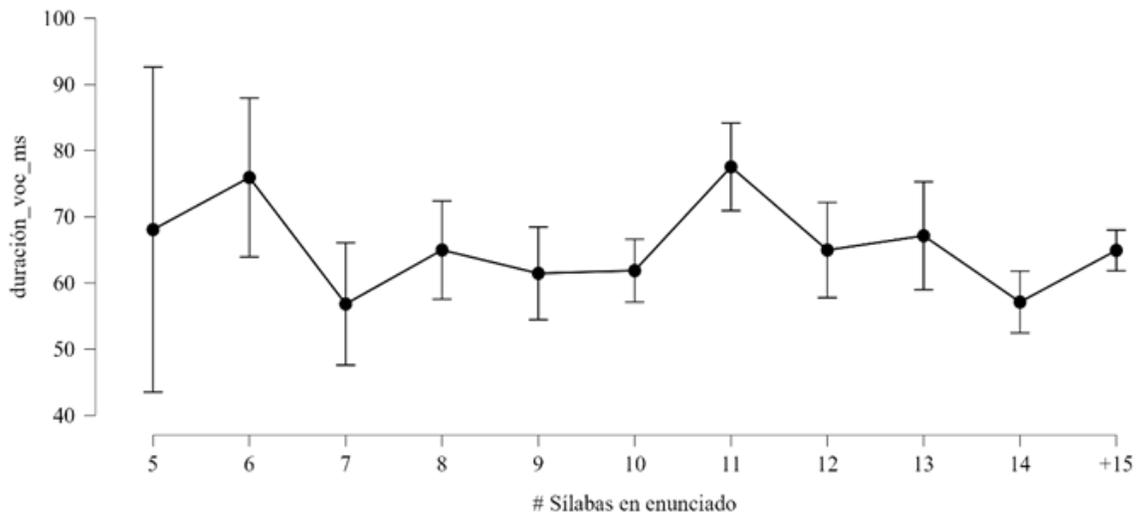


Figura 13. Gráficos descriptivos de la duración según el número de sílabas por enunciado

duración_voc_ms					
	5	6	7	8	9
Válido	10	22	28	43	36
Moda	22.30	72.60	17.20	46.10	27.20
Mediana	69.80	72.50	51.90	62.20	60.55
Media	68.07	75.95	56.83	64.98	61.46
Curtosis	-1.52	-0.64	2.92	2.10	2.70
Mínimo	22.30	37.80	17.20	20.40	27.20
Máximo	116.70	128.50	130.60	146.80	129.00

	10	11	12	13	14	+15
Válido	130	105	73	65	110	361
Moda	49.00	40.60	61.30	47.70	65.00	23.40
Mediana	56.20	71.20	58.30	57.10	54.50	61.10
Media	61.87	77.54	64.98	67.14	57.12	64.93
Curtosis	4.31	-1.37e-4	2.18	2.19	0.60	1.38
Mínimo	15.00	21.20	13.90	21.60	15.10	9.80
Máximo	197.50	162.70	180.90	177.10	136.50	179.10

Tabla 14. Estadísticos descriptivos de la duración según el número de sílabas por enunciado

Los resultados del ANOVA para esta variable muestran los valores de $H = 10.39$, $p < 0.001$. Estos resultados sugieren que hay niveles de significancia estadística que explican la correlación entre la cantidad de sílabas dentro de los enunciados y la duración que se puede dar en las vocales dentro de ellos. Esa correlación evidenciada entre estos factores parece surgir en la medida que se requiere la estructura enunciativa y, por ende, su patrón prosódico para generar cambios en la duración de las vocales.

5. Conclusiones

La duración vocálica en el habla de Medellín registra un valor de duración intrínseca de 65,32 ms y, según el timbre de cada vocal, los valores de duración se manifiestan de mayor a menor en el orden de /i/, /a/ > /u/ > /e/, /o/, demostrando que hay una correlación entre los valores de la duración y las propiedades de cada vocal que se determinan por el punto y modo de articulación de extremos entre alto-bajo y anterior-posterior.

Por otro lado, las categorías de tonicidad a nivel de la palabra y, sobre todo, a nivel del enunciado, estadísticamente reflejan mayor influencia en la composición de los valores de duración de las vocales. La duración vocálica en este estudio evidencia, pues, que tiene una fuerte correspondencia con la estructura prosódica del enunciado.

De modo general, las variables consideradas en esta investigación que más se correlacionan con la duración vocálica se pueden clasificar en el siguiente orden de importancia según los resultados estadísticos obtenidos: tonicidad enunciativa, tonicidad, cantidad de sílabas por enunciado, categoría vocálica y segmento posterior.

En esta investigación, la duración vocálica evidenció valores que son susceptibles a su grado de apertura, aunque el habla en uso no responde en sentido estricto al patrón, en el que la mayor duración se presenta en la vocal baja y se reduce hacia las vocales altas. También, el tamaño de los enunciados parece demostrar que las configuraciones de duración amplias se posibilitan cuando se conforman determinados grupos fónicos.

5.1. Recomendaciones

El estudio de la duración permite analizar los patrones fonéticos de la lengua desde otra óptica, por lo que continuar ampliando este tipo de trabajos a otras variantes dialectales del español puede resultar de interés como un posible complemento a los estudios de prosodia y a la comprensión de las variaciones que se dan en los dialectos del español de Colombia.

Disponer de los corpus lingüísticos existentes, como el PRESEEA, agiliza etapas de la investigación y se constituye en un aprovechamiento del recurso, lo cual puede posibilitar análisis de datos más grandes, en la medida en que se tecnifiquen otros procesos que se dan en etapas complementarias de los análisis, como en la segmentación de las vocales en Praat.

Por otro lado, es importante procurar una consistencia en los criterios de selección de los enunciados cuando se trabaja con el habla en uso, ya que son susceptibles de subjetividades en el proceso de selección. Lo mismo puede suceder con los parámetros que determinan las fronteras acústicas en la segmentación vocálica sobre los espectros de voz. Llegar al nivel de aplicar una asistencia computarizada que posibilite una segmentación más objetiva ayudaría a revelar resultados con un nivel de meticulosidad más fino.

Finalmente, contribuir en la creación de manuales, tanto para el tratamiento acústico de los datos como para los procedimientos estadísticos en lengua española, se hace imperante, dado que los insumos disponibles brindan un aporte *grosso modo*, pero aún hay carencia de materiales más especializados como fuentes de consulta que permitan esclarecer múltiples dudas que surgen en estos procesos de investigación fonética.

Referencias bibliográficas

1. Adi, Y., Keshet, J., & Goldrick, M. (2015). Vowel duration measurement using deep neural networks. In *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)* (pp. 1-6). IEEE.
2. Audacity Team (2.4.2). (2020). [Software]. <https://www.audacityteam.org/>
3. Betancourt, A. (1993). Lengua y región. *Thesaurus*, XLVIII(2), 255-291. http://cvc.cervantes.es/lengua/thesaurus/pdf/48/TH_48_002_023_0.pdf
4. Boersma, P. & Weenink, D. (2020). *Praat* (6.1.21). [Software]. <http://www.fon.hum.uva.nl/praat/>
5. Bravo, G. R. (2008). La identificación de los tópicos y los focos. *Nueva Revista de Filología Hispánica (NRFH)*, 56(2), 363-401. <https://doi.org/10.24201/nrfh.v56i2.969>
6. Castellano Ascencio, M. D. (2013). Cortesía verbal y fórmulas de tratamiento nominales en el habla de Medellín. *Lingüística y Literatura*, 33(62), 123–139.
7. Cantero, F, J. (2019). Análisis prosódico del habla : más allá de la melodía. En M. R. Álvarez Silva, A. Muñoz Alvarado & L. Ruiz Miyares (Eds.), *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología , Folclor y otras ciencias afines . Volumen II* (pp. 485-498). Ediciones Centro de Lingüística Aplicada.
8. Cedeño N.& Morales-Front A. (1999). *Fonología generativa contemporánea de la lengua española*. En P. Prieto i Vives & J. I. Hualde (Eds.). Georgetown University Press. <http://dx.doi.org/10.15581/008.17.1.125-126>
9. Cuenca, H, M. (1996). Análisis instrumental de la duración de las vocales en español. *Philologia Hispalensis*, 1(11), 295-307. <https://doi.org/10.12795/ph.19961997.v11.i01.20>
10. Díaz Romero, C. E. & Correa Duarte, J. A. (2014). Manual de análisis acústico del habla con Praat. Series Minor 49. *Revista de Investigación Lingüística*, 17, 245–248. <https://revistas.um.es/ril/article/view/224041>
11. García Z., C. (2016). Hágale, arcador conversacional en el habla coloquial de Medellín. *Lingüística y Literatura*, 37(69), 315-338. <https://doi.org/10.17533/udea.lyl.n69a14>
12. Gil-Fernández, J. (2007). *Fonética para profesores de Español*. Arco.
13. Gili, G. S. (1980). *Curso superior de sintaxis española*. VOX.
13. González-Rátiva, M. C. (2008). Corpus sociolingüístico PRESEEA-Medellín. [Base de datos] <https://comunicaciones.udea.edu.co/corpuslinguistico/?opcion=6>
14. González-Rátiva, M. C. (2022). Procesos fonológicos convergentes en el español hablado en Antioquia, a partir del Atlas Lingüístico de Antioquia y del Corpus PRESEEA-Medellín. *Anuario de Letras. Lingüística y Filología*, 10(1), 219-254. <https://doi.org/10.19130/iifl.adel.2022.10.1.2877x37>
15. Goss-Sampson, M. A. (2018). *Análisis estadístico con JASP: una guía para estudiantes*. FUOC.
16. Hidalgo, A. & Quilis, M. (2012). La voz del lenguaje: fonética y fonología del español, Valencia, Tirant Humanidades. *Estudios de Fonética Experimental*, 411-418. <https://www.raco.cat/index.php/EFE/article/view/275539>
17. JASP(2020). *JASP* (0.13.1). [Software]. <https://jasp-stats.org/>

18. Krohn, S, H. (2019). Duración vocálica en el español de la Gran Área Metropolitana de Costa Rica. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 45(1), 215-224. <https://doi.org/10.15517/rfl.v45i1.36736>
16. Marín G, R. (1995). La duración vocálica en español. *ELUA. Estudios de Lingüística Universidad de Alicante*, 10, 213-226. <https://doi.org/10.14198/elua1994-1995.10.11>
17. Muñetón, M. A. (2017). Asociación de la F0, duración e intensidad en el habla de una mujer de Medellín (Colombia) en función de la modalidad oracional y sus sintagmas. *RLA. Revista de Lingüística Teórica y Aplicada*, 55(1), 53-72. <https://doi.org/10.4067/s0718-48832017000100053>
18. Muñoz-Builes, D. M. (2016). Declarativas e interrogativas con expansión en sujeto: análisis prosódico de una informante de Medellín. *Lingüística y Literatura*, 37(69). <https://doi.org/10.17533/udea.lyl.n69a02>
19. Muñoz-Builes, D. M. (2021). Variación diastrática de la entonación femenina en el español colombiano de Medellín. *Lengua y Habla*, 25, 201-222. https://www.academia.edu/63378623/Variación_diastrática_de_la_entonación_femenina_en_el_español_colombiano_de_Medellín
20. Navarro, T. (1921). Manual de pronunciación española. *Bulletin Hispanique*, 23 (1), 69-76.
21. Pamies, A. & Planas. A. M. (2015). Sobre la percepción de la duración vocálica en español. *Actas del V Congreso Andaluz de Lingüística General*. Granada Lingüística.
22. Pereira, D. I., & Soto Barba, J. (2011). Duración absoluta de las vocales del español urbano y rural de la provincia de Ñuble. *Boletín de Filología*, 46(1), 153-161. <https://doi.org/10.4067/s0718-93032011000100005>
23. Quilis, A. (1993). *Tratado de fonética y fonología españolas*. Gredos.
24. Román, D. (2021). *Scripts para Praat*. http://www.domingo-roman.net/scripts_para_praat.html
25. Searle, J. (1990). *Actos de habla*. Cambridge University Press.
26. Wright, B. R. & Nichols, D. (2015). *Measuring Vowel Duration*. UW Phonetics/Sociolinguistics Lab Wiki. <https://depts.washington.edu/phonlab/resources/measuring-duration.pdf>