

UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# **Análisis acústico de las pausas llenas como parámetro en la comparación forense de voces**

**Luis Felipe Bautista Peña**

Universidad Nacional de Colombia  
Facultad de Ciencias Humanas, Departamento de Lingüística  
Bogotá, Colombia  
2021

# **Análisis acústico de las pausas llenas como parámetro en la comparación forense de voces**

**Luis Felipe Bautista Peña**

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de:

**Lingüista**

Director:

Mg. Lingüística, John Freddy Chaparro Rojas

Línea de Investigación:

Fonética Forense

Universidad Nacional de Colombia

Facultad de Ciencias Humanas, Departamento de Lingüística

Bogotá, Colombia

2021

## Resumen

Este estudio, enmarcado en la fonética forense, explora el potencial de las pausas llenas como parámetro en la comparación forense de voces (CFV), en el español hablado en Bogotá (Colombia). Para este fin, se empleó la mensajería instantánea (MI) como fuente para la obtención de datos, logrando recolectar muestras de habla de 12 hablantes femeninos, con características sociolingüísticas homogéneas. Se analizaron las características acústicas de la estructura formántica (F1, F2 y F3) y la duración de la pausa llena vocálica [e:], teniendo en cuenta que este fue el segmento con mayor frecuencia de aparición. Los datos de las mediciones acústicas fueron procesados mediante una prueba estadística paramétrica (ANOVA) para los formantes y una no paramétrica (Kruskal-Wallis) para la duración. Según los resultados se pueden considerar potenciales principalmente los valores provenientes de la frecuencia central de F2 y F3 puesto que según los valores de F1 y las medidas de la duración, la variación intrahablante es elevada, impidiendo anotar diferencias entre las hablantes.

**Palabras clave:** Fonética forense, comparación forense de voces, pausas llenas, mensajería instantánea.

## **Abstract**

This study, within the field of forensic phonetics, explores the strength of filled pauses as a parameter in forensic voice comparison (FVC) in the Spanish spoken in Bogota (Colombia). Instant messaging (IM) was used as source for data. Speech samples from 12 female speakers with homogeneous sociolinguistic characteristics were collected. The acoustic characteristics of the formant structure (F1, F2 and F3) and the duration of the full vowel pause [e:] were analysed, since it is the segment with the highest frequency of occurrence found. The data from the acoustic measurements were processed using a parametric statistical test (ANOVA) for formants and a non-parametric one (Kruskal-Wallis) for duration. According to the results, midpoints of F2 and F3 can mainly be considered as potential, as according to the F1 midpoints and duration measurements, the intra-speaker variation is high, preventing to identify differences between speakers.

**Keywords:** Forensic phonetics, forensic voice comparison, filled pauses, instant messaging.

# Contenido

	Pág.
<b>1. Introducción.....</b>	<b>9</b>
<b>2. Marco Teórico.....</b>	<b>11</b>
2.1 Lingüística Forense .....	11
2.2 Fonética Forense.....	13
2.2.1 El papel del lingüista-fonetista forense .....	14
2.2.2 Identificación forense de hablantes .....	16
2.2.3 Comparación forense de voces (CFV).....	17
2.3 Pausas Llenas Como Material Fonético Para el Análisis Acústico Forense.....	22
2.4 Dificultades en el Análisis Acústico Forense.....	26
2.5 Potencial de las Aplicaciones de Mensajería Instantánea (MI) Como Fuente de Datos Para el Análisis Acústico Forense .....	27
2.6 Pausas Llenas y Mensajería Instantánea: Antecedentes de Investigación.....	30
2.7 Preguntas de investigación.....	35
2.8 Objetivo general.....	36
2.9 Objetivos específicos.....	36
<b>3. Metodología.....</b>	<b>37</b>
3.1 Tipo de Estudio.....	37
3.2 Participantes .....	37
3.3 Procedimiento para la Recolección y Selección de Audios .....	39
3.4 Descripción de los Datos Seleccionados .....	40
3.5 Procedimientos.....	41
3.5.1 Segmentación de pausas llenas.....	41
3.5.2 Mediciones.....	42
3.5.2 Conjuntos de datos seleccionados por hablante (muestras).....	43
3.5.3 Análisis estadístico .....	44
<b>4. Resultados.....</b>	<b>45</b>
4.1 Tipos de Pausas y Frecuencias de Aparición.....	45
4.2 Variación Intra e Interhablante.....	50
4.2.1 Primer formante (F1).....	51
4.2.2 Segundo y tercer formante (F2 y F3) .....	53
4.2.3 Duración.....	59
<b>5. Discusión.....</b>	<b>63</b>
5.1 Uso de Aplicaciones de MI como Fuente de Datos Lingüísticos.....	63
5.2 Pausas Llenas: Tipos y Frecuencias de Aparición .....	64
5.2.1 Segmento vocálico [e:] .....	65
5.2.2 Segmento compuesto [e:m:].....	66
5.3 Pausas Llenas Como Parámetro en la CFV .....	66
<b>6. Conclusiones.....</b>	<b>69</b>

## Lista de figuras

	<b>Pág.</b>
<i>Figura 1.</i> Espectrograma de un audio proveniente Whatsapp.....	29
<i>Figura 2.</i> Espectrograma de un audio proveniente de una llamada telefónica. ....	29
<i>Figura 3.</i> Segmentación de una realización de la pausa llena vocálica [e:]. ....	42
<i>Figura 4.</i> Apariciones totales registradas por cada tipo de pausa en cada hablante.....	48
<i>Figura 5</i> Diagrama de dispersión de los valores de F1 para todas las hablantes. ....	52
<i>Figura 6.</i> Diagrama de dispersión de los valores de F2 para todas las hablantes. ....	56
<i>Figura 7.</i> Diagrama de dispersión de los valores de F3 para todas las hablantes. ....	56
<i>Figura 8.</i> Gráfico comparativo de la diferencia de medias de la frecuencia central: F1 F2 y F3. ....	59
<i>Figura 9.</i> Diagrama de dispersión de los valores de la duración para todas las hablantes. ....	61

## Lista de tablas

	<b>Pág.</b>
Tabla 1 Ilustración del funcionamiento de la razón de verosimilitud. Adaptado de Rose (2002, p. 56).....	21
Tabla 2 Rótulos de las participantes y datos preliminares.....	38
Tabla 3 Tipos de pausas observadas. ....	46
Tabla 4 Síntesis del total de apariciones registradas por tipo de pausa. ....	47
Tabla 5 Relación del tiempo de grabaciones analizado por hablante y el total de pausas registradas.....	49
Tabla 6 Estadísticos descriptivos de los valores del punto medio de F1.....	51
Tabla 7 Estadísticos descriptivos de los valores del punto medio de F2.....	54
Tabla 8 Estadísticos descriptivos de los valores del punto medio de F3.....	54
Tabla 9 Diferencias significativas entre hablantes, según valores de F2. ....	57
Tabla 10 Diferencias significativas entre hablantes, según valores de F3.....	58
Tabla 11 Estadísticos descriptivos de los valores de la duración en milisegundos (ms). ....	60

## 1. Introducción

Dentro del campo de la lingüística forense se aplican la teoría y los métodos de la lingüística en el ámbito legal o judicial. Una de sus áreas principales es el estudio del lenguaje como evidencia, en donde la fonética forense y, dentro de esta, la comparación forense de voces (CFV) desempeñan un papel importante. En la CFV se emplea material fonético proveniente de un hablante involucrado en un crimen (voz dubitada) y muestras de uno o más sospechosos (voz indubitada) con el fin de determinar en qué medida pueden haber sido o no producidas por el mismo hablante. Investigaciones previas se han ocupado en determinar qué tipo de material fonético es el más indicado para este fin y, también, qué métodos ofrecen el mejor desempeño en este tipo de comparaciones (Nolan, 1997, Rose, 2002, como se cita en Hughes, Wood, & Foulkes, 2016). En relación con lo anterior, se ha profundizado en la investigación de las pausas llenas como parámetro en la CFV en diferentes lenguas como, por ejemplo, el catalán (Cicres, 2014) o el inglés (Hughes, Wood, & Foulkes, 2016); los resultados indican que tienen alto potencial. Por lo tanto, en la presente investigación el objetivo es explorar el potencial de las pausas llenas del español hablado en Bogotá (Colombia) como parámetro en la CFV, bajo la hipótesis de que en esta lengua las pausas llenas también funcionan para este propósito.

Por otra parte, a través de este estudio se busca, también, ampliar la información existente acerca las pausas llenas en español, especialmente sobre sus características acústicas. Estableciendo un punto de partida para la investigación de estos fenómenos en esta y en otras variedades del español, ya que, en general, tanto el desarrollo teórico de esta disciplina (lingüística/fonética forense) como el estudio de los fenómenos de hesitación son limitados en la



región. Así, el presente documento está estructurado como se describe a continuación.

Inicialmente, en el capítulo 2, se ofrece una contextualización detallada de la teoría que enmarca la investigación, así como la socialización de estudios previos relevantes. En el capítulo 3 se explica a profundidad el método empleado. El capítulo 4 corresponde a la presentación de los hallazgos. En el capítulo 5 se hace una discusión de los resultados y sus implicaciones a la luz de las investigaciones previas. Por último, el capítulo 6 comprende las conclusiones.

## **2. Marco Teórico**

En este capítulo se provee inicialmente una contextualización sobre la lingüística y su aplicación en el contexto forense o judicial. Primero, se destaca que, aunque su aplicación se puede considerar reciente, sus dominios no dejan de ser extensos. Por esta razón, a medida que avanza el capítulo, el lector encontrará aspectos cada vez más específicos, como la descripción de la fonética forense, lo que incluye considerar el papel del lingüista-fonetista forense. Dentro de la fonética forense, se hará énfasis en la comparación forense de voces; así, se describen algunos de sus enfoques, métodos y parámetros que han sido empleados en el análisis fonético-forense, así como algunas consideraciones sobre las problemáticas que pueden estar presentes en su desarrollo. Por último, se considera la relevancia de las pausas llenas como material fonético útil en la comparación forense de voces y, también, el potencial de la mensajería instantánea como fuente para la obtención de muestras de habla en el contexto forense. Complementariamente, se hace una revisión de diferentes estudios previos relacionados. Exponiendo, a su vez, el porqué es importante la investigación en esta área en el contexto específico de Colombia, pues al indagar acerca del estado de esta disciplina en el país, no se encontró información disponible sobre investigaciones previas o casos documentados.

### **2.1 Lingüística Forense**

La lingüística forense se puede entender como una rama de la lingüística aplicada, dado que se trata de la aplicación de teoría y métodos de la lingüística, en un contexto o con un propósito específico. Esto es, la práctica lingüística en el ámbito legal. De lo anterior surge la definición más extendida y general que se ha aceptado hasta ahora: la lingüística forense es la interfaz entre lengua

y derecho, en donde *derecho* incluye aplicación de la ley, cuestiones judiciales, legislación, disputas y procedimientos en derecho, entre otros (Olsson, 2008).

La lingüística forense como disciplina, aunque ha tenido un desarrollo importante en los últimos años, aún se puede considerar reciente. Como expone Olsson (2008), se apunta a establecer su origen en 1986 con la publicación de *The Evans Statements: A Case for Forensic Linguistics* (Jan Startvik), en donde se someten a un análisis lingüístico cuatro confesiones supuestamente hechas a la policía por parte de un joven llamado Timothy Evans. Estas confesiones lo incriminaron en el asesinato de su esposa e hija, siendo condenado a muerte. El resultado de un análisis realizado posteriormente demostró que dentro de las confesiones existían dos estilos diferenciados cualitativa y cuantitativamente y que favorecían la inocencia del acusado. Lo anterior dejó en evidencia la importancia de tener en cuenta el estudio de la lengua y su incidencia en los contextos judiciales. Desde entonces, la disciplina ha tenido un amplio desarrollo; actualmente se encuentra documentación especializada, como libros de texto: *An introduction to forensic linguistics: Language in evidence* (Coulthard, Johnson & Wright, 2017), *The Routledge handbook of forensic linguistics* (Coulthard & Johnson, 2010), *Forensic Linguistics* (Olsson, 2008), entre otros. También, instituciones como la Asociación Internacional de Lingüistas Forenses (IAFL, por su sigla en inglés), con su revista *International Journal of Speech, Language and the Law*; además, diferentes laboratorios de lingüística forense. Por ejemplo, el Laboratorio SQ-Lingüistas Forenses y el ForensicLab en la Universidad Pompeu Fabra, ambos en España; laboratorios enfocados en Fonética Forense (Ver sección 1.2) como el laboratorio *J P French Associates* en Reino Unido, entre otros. Asimismo, se imparten algunos posgrados y cursos en lingüística forense o relacionados, principalmente en universidades del Reino Unido como Aston University, Cardiff University o University of York.

Retomando la contextualización, la lingüística forense se puede dividir en dos grandes áreas de acuerdo con Coulthard, Johnson, & Wright (2017). Por una parte, el lenguaje del procedimiento

legal, incluyendo el lenguaje jurídico, que en ocasiones se ha tenido en cuenta como un área aparte; y por otro, la lengua como evidencia. De manera general, el lenguaje del procedimiento legal comprende el análisis de la lengua en relación con el registro, estilo y géneros empleados en el contexto jurídico, así como el análisis de los discursos empleados en interrogatorios, juicios, llamadas de emergencia, etc. La segunda área, el análisis de la lengua como evidencia, comprende el peritaje de evidencia de orden lingüístico en asuntos como la detección de plagio, atribución de autoría y la fonética forense. Esta última, siendo el eje del presente estudio, merece una descripción detallada que se hará en la siguiente sección. Para una contextualización más amplia sobre las aplicaciones de la lingüística forense se sugiere consultar, entre otros: Coulthard & Johnson (2010), Coulthard, Johnson & Wright (2017) y Olsson (2008).

## 2.2 Fonética Forense

La fonética forense es una rama bastante extensa dentro de la lingüística forense. Se puede definir como la aplicación de la teoría y metodología de la fonética general en el ámbito judicial en donde exista uso o registro (grabaciones) de la lengua oral. Olsson (2008) la define como el uso de técnicas fonéticas, principalmente en el análisis de la voz, aplicadas a las investigaciones criminales y Jessen (2008) añade que también se trata del desarrollo de nuevo conocimiento específicamente forense-fonético, esto es, teorías y métodos específicos. Una de sus aplicaciones principales es el reconocimiento o identificación de hablantes, sin embargo, también abarca otros campos. Por ejemplo, el análisis de audios cuya inteligibilidad pueda estar reducida por razones técnicas, del entorno o conductuales. A este campo se le puede denominar análisis de los *enunciados discutidos*,

según se conoce en la literatura inglesa como *disputed utterances analysis*<sup>1</sup> (Olsson, 2008; Jessen, 2008; Coulthard, Johnson, & Wright, 2017). En ocasiones, dentro de la evidencia en un determinado caso, demasiada importancia puede reposar en una sola palabra o en un solo sonido. Por lo anterior, la interpretación que se haga trae consigo fuertes consecuencias, que representan una serie de dificultades, ya que a diferencia de la lengua escrita, los sonidos producidos en el habla son continuos, sus límites no están definidos y es difícil distinguirlos cuando se producen con diferentes características como la velocidad, acento, ritmo, incluso en grabaciones de calidad (Coulthard, Johnson, & Wright, 2017). Además, los mismos autores resaltan el papel que juega el oyente en la interpretación de una grabación, pues sostienen que a través de diferentes estudios se ha podido determinar que los oyentes no hacen una interpretación neutra sino que en realidad se realiza una construcción de lo que él o la oyente piensa que escucha, dicha interpretación está mediada por el contexto y por el conocimiento del habla que posea el oyente previamente (Coulthard, Johnson, & Wright, 2017). Dada la importancia de un procedimiento de análisis fonético en el ámbito forense, es importante resaltar el papel del profesional o experto en el área, tanto de la lingüística como la fonética forense como se hará a continuación.

### **2.2.1 El papel del lingüista-fonetista forense**

Como se ha mencionado, lingüística y por lo tanto la fonética forense alcanzan hasta el día de hoy un periodo relativamente corto desde que fueron establecidas formalmente como disciplinas. También resalta que su desarrollo ha sido específico de algunos países. Sin embargo, la lengua siempre ha estado involucrada de una u otra manera en el ámbito forense, así que antes, e

---

<sup>1</sup> Este concepto hace referencia a los casos en que no se puede determinar con claridad lo que está dicho en una grabación, puede darse a nivel de palabras o inclusive un solo fonema. Para información más detallada ver Coulthard, Johnson, & Wright (2017, p 130)

incluso aún, se realizan diferentes procedimientos relacionados con el lenguaje en la investigación criminal sin contar con profesionales instruidos y experimentados en el área, restándole importancia y relegando esta labor a profesionales de otros campos. Esto es evidentemente una problemática pues como se mencionó anteriormente, la interpretación errónea de algún elemento lingüístico que haga parte de la evidencia puede traer consecuencias graves. Por otra parte, tampoco es correcto afirmar que la intervención de un lingüista o fonetista forense llevará a resultados absolutamente precisos e infalibles, pero sí, dado un mayor grado de conocimiento e instrucción en el área, se pueden mejorar considerablemente los procedimientos, y se pueden ampliar las opciones que se tienen frente a un caso, pues el conocimiento lingüístico es necesario por ejemplo en la identificación de dialectos y variedades lingüísticas; además, cuando hay escaso conocimiento en algún caso particular, el experto debe saber a dónde acudir, en dónde consultar y, por lo tanto, estar capacitado para entender la información necesaria, como señala Jessen (2010).

Coulthard, Johnson, & Wright (2017) presentan un resumen del quehacer del fonetista forense. Esto incluye: el perfilamiento de hablantes, obtener información del contexto social y regional con base en muestras de habla; comparación de hablantes, es decir, contribuir a determinar si dos o más muestras fueron producidas por el mismo o no. También puede aportar, como ya se mencionó, en la interpretación y transcripción de grabaciones de difícil inteligibilidad; y asesorar a las autoridades en el diseño de procedimientos como los *voice line-ups*<sup>2</sup> (descrito más adelante). Jessen (2008) menciona que también pueden ser considerados en labores de mejora de audio y en análisis de audios de carácter no lingüístico.

---

<sup>2</sup> Se conserva el término en inglés al no encontrar una traducción adecuada.

### 2.2.2 *Identificación forense de hablantes*

Siguiendo el enfoque de Jessen (2008, 2010), la identificación de hablantes es el campo principal de aplicación dentro de la fonética forense, que es también en el cual se enmarca este trabajo, razón por la cual es imprescindible definirlo. Para este fin, se describirán de forma general las categorías que el mismo autor distingue, con el fin de ofrecer una contextualización y se abordará con mayor detalle la categoría específica que sirve como sustento teórico requerido en la presente investigación. Estas categorías son: i) perfilamiento de voz o perfilamiento de hablante, ii) el análisis de la identificación del hablante por víctimas y testigos y iii) comparación de voces.

En el perfilamiento de voz o perfilamiento de hablantes, se parte de la situación en que se tiene una grabación proveniente de un hablante involucrado en un crimen, pero no se tiene un sospechoso, este es el caso en el que un fonetista forense puede contribuir a extraer la mayor cantidad de información posible, únicamente basándose en la muestra de habla y con el objetivo final de reducir el número de posibles sospechosos. Dependiendo de la muestra, la información que se puede obtener varía: aspectos regionales, el lugar de procedencia del hablante, edad, sexo, nivel educativo, contexto social, entre otros, son algunos ejemplos (Jessen, 2008). El perfilamiento de voz/hablante en algunos casos también se puede desarrollar en trabajo conjunto con la fonética clínica, con el fin de determinar características específicas del hablante que puedan ser producto de condiciones médicas o desordenes del habla (Coulthard, Johnson, & Wright, 2017). En el análisis de la identificación del hablante por víctimas y testigos se parte de la situación en donde no se tiene muestra de habla grabada, por lo tanto, la capacidad de descripción y de reconocimiento por parte de un testigo, o en algunos casos la misma víctima, es lo que cobra relevancia como medio para la identificación. Esta categoría está dividida en dos tipos de identificación. En la primera, el testigo conoce previamente al acusado y lo puede identificar al menos por el nombre (Jessen

2008); en la segunda, si el perpetrador es totalmente desconocido y dada la situación de que se encuentre un sospechoso, se realizarán procedimientos de reconocimiento en los que se despliega una sucesión secuencial de voces, de forma similar a una «rueda de reconocimiento», pero de carácter auditivo, con todas las dificultades que esto pueda representar. En esta categoría el papel del fonetista es imprescindible en el diseño y realización de este procedimiento, conocido como *voice line-up*. En la comparación de voces, se empleará el término «voz» siguiendo el enfoque de Jessen (2008), sin embargo, en la literatura el concepto puede ser referido como comparación de hablantes o comparación del habla. Esta categoría, por su relevancia en esta investigación, será abordada con mayor detalle en la siguiente sección.

### **2.2.3 Comparación forense de voces (CFV)**

Como se mencionó previamente, el concepto de comparación forense de voces (CFV) es un eje principal, ya que este trabajo al ser una comparación de características acústicas, se realiza y sustenta a través de la teoría y métodos propios de esta categoría.

En la práctica de la fonética forense, existe una situación común en la que se cuenta con una o varias muestra de habla grabadas, provenientes de un hablante que esté involucrado en un crimen. Estas muestras se denominan voz dubitada puesto que no se puede asegurar quien la produjo; además, se posee o se está en la posibilidad de obtener una grabación de uno o más sospechosos, estas muestras se denominan voz o voces indubitadas puesto que hay certeza sobre a quién pertenece (Machuca, Ríos, & Llisterri, 2014, p. 97). Contando con estos dos tipos de muestra, se puede llevar a cabo una comparación forense de voces (CFV). Esta comparación se puede realizar teniendo en cuenta diversos segmentos y rasgos del habla. Aunque a través de la investigación se ha ido descubriendo el potencial que tienen algunos en particular, este potencial está dado en la medida en que se cumpla que las características de las muestras de un hablante sean diferentes en



un alto grado respecto a las de otro y que las diferencias entre muestras del mismo hablante sean menores. Esta distinción se denomina variación interhablante e intrahablante (Rose, 2002) y se retomará con más detalle más adelante. Además, tales características deben ser evidenciables en muestras cortas de habla, resistentes a la imitación y deformación del habla y ser considerablemente independientes, de otras variables (Nolan, 1997, Rose, 2002, como se cita en Hughes, Wood, & Foulkes, 2016). Bajo estas consideraciones, las pausas llenas (fenómenos de la fluidez del discurso. Ver sección 1.3) se han tenido en cuenta en diferentes investigaciones, estableciéndose como material fonético confiable para la CFV en varias lenguas. Existen además múltiples métodos, como se verá más adelante, pero el objetivo final es determinar en qué grado dos muestras de habla pueden haber sido, o no, producidas por el mismo hablante. Los resultados producto de estas comparaciones son tenidos en cuenta como evidencia en las cortes y el responsable del reporte debe estar en capacidad de sustentarlo (Jessen , 2008).

Para continuar con el abordaje de métodos en la CFV es fundamental reconocer la importancia de la distinción entre la variación interhablante y la variación intrahablante, por lo tanto, estos y otros conceptos relacionados serán descritos a continuación.

Coulthard, Johnson, & Wright (2017) señalan un problema básico en la comparación de voces, esto es, siempre existen diferencias entre dos muestras de habla aún cuando se obtengan del mismo hablante, en las mismas condiciones, incluso en un rango de tiempo reducido. Así que el reto del experto en la comparación de voces, en la práctica de la fonética forense, es determinar si en la CFV es más probable que las diferencias entre dos muestras de habla sean VARIACIÓN INTERHABLANTE O VARIACIÓN INTRAHABLANTE (Rose, 2002, p. 9). La variación interhablante se entiende como el conjunto de diferencias entre las características del habla de dos o más hablantes, mientras que la variación intrahablante hace referencia a las diferencias que se puedan presentar dentro de las características del habla del mismo hablante (Rose, 2002, p. 9). Estos conceptos son relevantes

para introducir el marco de comparación más extendido y aceptado en la CFV; esto es, el ENFOQUE BAYESIANO (Rose, 2002, p. 63). Los aspectos preliminares a tener en cuenta son: el grado de SIMILITUD (en inglés: *similarity*): qué tan similares o diferentes pueden ser dos voces según las características o rasgos fonéticos desde los que se están comparando (Jessen, 2008), sin embargo, hay que tener presente que no se puede concluir inmediatamente que dos muestras provengan de la misma fuente (el mismo hablante) si la comparación de características da como resultado un alto grado de similitud. Esto se debe a que existe otro factor denominado TIPICIDAD (en inglés: *typicality*): las características fonéticas que son relativamente típicas o atípicas dentro de la población en general, es decir, hablantes de la misma variedad de la lengua (Jessen, 2008). El abordaje matemático de esta relación entre similitud y tipicidad está basado en lo que se conoce como el enfoque bayesiano, previamente mencionado, el cual sirve para determinar la probabilidad condicionada, es decir la probabilidad de que suceda un evento A, dada la ocurrencia de otro evento B, cuya notación es  $P(A|B)$ <sup>3</sup> (Kaplan & Kaplan, 2006). Dentro de la CFV, este teorema ha sido adaptado para los intereses específicos del área. Desde este enfoque, el concepto fundamental a tener en cuenta es la RAZÓN DE VEROSIMILITUD (en inglés: *likelihood ratio*), siguiendo la conceptualización detallada que ofrece Rose (2002):

Para explicarlo podemos partir de la siguiente fórmula:

Ecuación 1

$$LR = \frac{p(E|H_p)}{p(E|H_d)}$$

---

<sup>3</sup> Para ampliar la información sobre el teorema bayesiano, ver (Kaplan & Kaplan, 2006, p. 184)

En la Ecuación 1. se observa el cálculo de la razón de verosimilitud ( $LR$ ) a partir de los valores de dos hipótesis, la hipótesis de acusación  $H_p$  (en inglés: *prosecution hypothesis*), que sostiene que las dos muestras de habla son provenientes del mismo hablante y la hipótesis de defensa (en inglés: *defence hypothesis*) la cual sostiene que las muestras provienen de diferentes hablantes. En términos más sencillos, y en relación con los conceptos previamente abordados, el numerador representa el grado de similitud entre las dos muestras en cuestión y el denominador representa el grado de tipicidad. Finalmente, el valor de  $LR$  expresa el grado en que la evidencia soporta la hipótesis de acusación, es decir que entre mayor sea la similitud entre las muestras (siempre respecto a uno o varios medibles específicos), mayor será el numerador y viceversa y, por otra parte, entre más probable sea encontrar esos mismos valores en la población en general (tipicidad), mayor será el denominador y menor en el caso contrario. Si el resultado de la ecuación es menor a uno (1) se soportará la hipótesis de defensa; si el resultado es mayor a uno (1) se soportará la hipótesis de acusación, siempre de forma proporcional al resultado.

En la parte superior de la Tabla 1 se observa que los valores para la similitud se mantienen en 80% y esto es un porcentaje elevado, sin embargo, vemos como el  $LR$  desciende a medida que se aumenta el valor de la tipicidad, hasta el caso en que similitud y tipicidad tienen el mismo valor y esto significa que es igual de probable que las muestras se le puedan atribuir al hablante en cuestión como a cualquier hablante elegido al azar dentro de la población en general. Un valor de  $LR=1$  descarta la evidencia totalmente (Rose, 2002). Por otra parte, cuando los valores de  $LR$  son inferiores a 1, es más sencillo evaluarlos términos del recíproco, por ejemplo, en el primer caso de la parte inferior (10%/80%), el resultado es 0,125, esto es, una probabilidad de 8 a 1 a favor de la defensa ya que  $1/125 = 8$ . Siendo ocho veces más probable que se trate de diferentes hablantes (Rose, 2002).

Tabla 1

Ilustración del funcionamiento de la razón de verosimilitud. Adaptado de Rose (2002, p. 56).

$p(E \mid \text{mismo hablante})^4$	80%	80%	80%	80%	80%
$p(E \mid \text{diferentes hablantes})^5$	10%	20%	40%	60%	80%
LR	8	6	4	2	1
$p(E \mid \text{mismo hablante})$	10%	20%	40%	60%	80%
$p(E \mid \text{diferentes hablantes})$	80%	80%	80%	80%	80%
LR	0.125	0.167	0.25	0.5	1

Para finalizar, son fundamentales varias consideraciones, Jessen (2008) anota que los valores obtenidos bajo este método de análisis son *relativos* y están en función de la demás evidencia y hechos relevantes que se tengan sobre el caso puntual que no tienen relación con el análisis del habla. Por otra parte, este es uno de los diferentes métodos que se pueden emplear en la CFV y en la literatura se insiste en que en ningún caso el fonetista forense se puede basar estrictamente en un solo método. De hecho, lo más usual es que se aborden dos o más en un trabajo conjunto. Finalmente, para una revisión detallada de los métodos empleados en la CFV y en general en la fonética forense ver Jessen (2008); Coulthard, Johnson & Wright (2017).

---

<sup>4</sup> Similitud

<sup>5</sup> Tipicidad

Hasta este punto se ha abordado la aplicación de la lingüística en el ámbito forense, de manera más puntual, la fonética forense y sus categorías en las que se destaca la comparación forense de voces, sus planteamientos teóricos y métodos; ahora, según los objetivos de este trabajo, se hace necesario definir y contextualizar el material fonético que nos ocupa. También, teniendo en cuenta la relevancia del análisis de información acústica en el ámbito forense, en este trabajo se adoptará esta perspectiva. A continuación se describirán las pausas llenas, sus características y rasgos desde la perspectiva forense y acústica.

### **2.3 Pausas Llenas Como Material Fonético Para el Análisis Acústico Forense**

El objetivo de la CFV, es decir, la distinción entre hablantes, se puede llevar a cabo a través del análisis de diferentes materiales fonéticos según sea el caso, entendidos estos como los segmentos (consonantes y vocales) o conjuntos de estos segmentos. Dentro de las posibilidades, encontramos los fenómenos relacionados con la fluidez del habla, denominados disfluencias. La primera consideración es que por paradójico que suene, hablar de fluidez incluye que estén presentes, porque son necesarias, muchas de estas disfluencias: pausas llenas y vacías, repeticiones, prolongaciones, y cambios de tema (McDougall & Duckworth, 2017). Desde otras perspectivas, estos fenómenos se pueden catalogar con el nombre de pausas y se han realizado diferentes tipologías, aunque bastante similares. Por ejemplo, Gil (2007) define estas pausas como: «la interrupción temporal del proceso de producción de sonidos, es decir de la fonación y de la articulación» (p. 298) y distingue entre pausas intencionadas y pausas no intencionadas, asimismo cada categoría se divide en dos: pausas virtuales: admisibles pero prescindibles y pausas obligatorias: necesarias en términos gramaticales o fisiológicos (respiración) (Gil, 2007). Por otra parte, Rose (2002) distingue cinco tipos de pausas según su forma y su función: pausas llenas y pausas vacías, pausas de unión, pausas de duda (hesitación) y pausas de turno (p. 165).

Dada la naturaleza de este trabajo, el enfoque que se tendrá como referente será la taxonomía de la fluencia para el análisis forense (TOFFA, por su sigla en inglés) propuesto por McDougall & Duckworth (2017). Esta taxonomía incluye cinco grupos: i) Pausas vacías, definidas como un silencio menor o igual a 200 ms (milisegundos), que ocurre dentro de un único turno de un hablante. ii) Pausas llenas, definidas como un sonido vocálico que puede estar seguido por una consonante nasal. iii) Repeticiones, esto es, parte de una palabra, una palabra completa o una frase que es repetida. La repetición puede presentarse más de una vez. iv) Prolongaciones, entendidas como uno o varios segmentos de una palabra que se prolongan en su pronunciación. Las pausas llenas prolongadas no son tenidas en cuenta en esta categoría. v) Interrupciones, definidas como el fenómeno en que un hablante interrumpe su enunciado y lo cambia, generalmente se trata de una corrección. (p. 6 – 10). El tipo de pausas que nos ocupan en esta investigación son las pausas llenas y, por lo tanto, a continuación, se profundizará en su conceptualización y la descripción de sus características.

Una característica que hace evidente la relevancia del estudio de las pausas llenas en el ámbito forense es que tienen carácter fonológico, ya que varían de una lengua a otra. Por ejemplo, en español se reconocen como pausas llenas dos sonidos transcritos como [e:], también denominada *vocal de la duda*, y [m:] (Gil, 2007, p. 299), que no deben entenderse directamente como, en este caso, la vocal semicerrada anterior no redondeada [e] y con la consonante nasal bilabial [m], ya que tienen características distintas (Machuca & Rios, 2016). Aunque estos segmentos sirven de referencia, es decir que las pausas llenas funcionan como un tipo de alófono. Cada lengua tiende a seleccionar una vocal no marcada para llenar este tipo de pausas, en francés [ø:] y en inglés [ə:], [a:m] o [ə:m]; asimismo puede haber variación en las pausas llenas incluso

dentro de una misma lengua, de una variedad dialectal a otra (Gil, 2007). Por otra parte, Hughes, Wood, & Foulkes (2016) señalan que las pausas llenas tienen un gran potencial si se comparan por ejemplo con las vocales o consonantes que son usualmente tenidas en cuenta en los análisis de CFV. En primer lugar, porque son bastante frecuentes en el habla espontánea, lo que eleva la cantidad de casos que se pueden obtener de una muestra, además, por su duración y relativa estabilidad, sus características son más fáciles de medir acústicamente. Asimismo, son sonidos menos afectados por la coarticulación y por último una característica bastante relevante es que son usualmente producidos de forma inconsciente, por lo tanto, son parámetros del habla que se escapan en gran medida del control del hablante (p. 2). McDougall & Duckworth (2017) añaden que sus características son relativamente estables en grabaciones de baja calidad, con ruido de fondo y limitaciones producto de la transmisión telefónica (Ver sección 1.4). Estudios como el que proponemos buscan profundizar en la descripción de estos segmentos bajo la consideración de que a nivel interhablante su variación es elevada. En este caso se tendrá en cuenta específicamente el segmento vocálico [e:], por ser el elemento sonoro que presenta más frecuencia de aparición como relleno de pausas (Machuca & Rios, 2016). Por lo tanto, a continuación, se profundizará en describir los posibles parámetros que se pueden examinar.

Dependiendo del material fonético a analizar, son diversos los rasgos y características, que se pueden tener en cuenta en un análisis acústico forense. De acuerdo con McDougall & Duckworth (2017), generalmente se acude al análisis de rasgos fonéticos que están directamente relacionados con la anatomía del hablante: la frecuencia fundamental ( $f_0$ )/tono (*pitch*) que da cuenta de las características de los pliegues vocálicos y permite, por ejemplo, la distinción entre sexos (Coulthard, Johnson, & Wright, 2017). Las frecuencias de los formantes que reflejan la configuración del tracto vocal, por ejemplo, en las vocales el primer formante (F1) está asociado al grado de abertura (Olsson, 2008). Por otra parte, estos rasgos fonéticos son evaluados en conjunto

con aspectos léxico-gramaticales y aspectos sociolingüísticos que pueden condicionar la producción de los sonidos dadas las particularidades dialectales (Coulthard, Johnson , & Wright, 2017). En el caso de las pausas llenas como material fonético forense, el análisis se centra en aspectos de comportamiento, más que anatómicos (McDougall & Duckworth, 2017). Sin embargo, en general se analizan parámetros similares (e. g. estructura formántica) que, además, como se mencionó anteriormente, son proclives a la variación dialectal.

Dado lo anterior, para la descripción de la pausa vocálica [e:] se tendrán en cuenta como aspectos medibles, los parámetros descritos a continuación. Un concepto preliminar es el de FRECUENCIA, en términos generales es el número de repeticiones por segundo (Hertz), de un fenómeno, en este caso la vibración. Todo objeto que vibre emite un sonido y su tono está determinado por la frecuencia de tales vibraciones, así mismo depende de su composición física y de su longitud, de manera que cualquier modificación altera el tono (Coulthard, Johnson , & Wright, 2017, p. 134). Lo anterior es indispensable para abordar el primer parámetro, denominado FRECUENCIA FUNDAMENTAL ( $F_0$ ), que corresponde a la velocidad de las vibraciones de los pliegues vocálicos, esta frecuencia se amplifica a través de la configuración del tracto vocálico y sus múltiplos se conocen como armónicos (Coulthard, Johnson , & Wright, 2017 p. 134), el conjunto de armónicos constituyen los FORMANTES, segundo parámetro a tener en cuenta, estos se definen también como las resonancias que ocurren en el tracto vocálico (Real Academia Española y Asociación de Academias de la Lengua Española, 2011, p. 38), en otras palabras, son las frecuencias sobre las que la voz centra la mayoría de su energía y están basados en la forma del tracto oral (que incluye tracto vocal y nasal) (Olsson, 2008), su medida se expresa en Hertz y usualmente se obtiene del punto de máxima respuesta, denominado frecuencia central (*centre frequency*) aunque esta no es la única dimensión en la que pueden ser medidos (Rose, 2002). Gráficamente se pueden



observar en un espectrograma como bandas oscuras. Su importancia en el análisis forense se debe a que pueden contener información de las características específicas del hablante (Nolan, 1991, 1994, 1997, como se cita en Duckworth, et al., 2011). La discusión sobre cuales formantes contienen más información específica del hablante está abierta (Cao & Dellwo, 2019), sin embargo, en este caso se tendrán en cuenta los tres primeros formantes (F1, F2, F3) como un punto de partida ya que se trata de una investigación pionera en una variedad del español en la que no se cuenta con estudios previos sobre pausas llenas, como veremos más adelante. Por último, se tendrá en cuenta la DURACIÓN, entendida como el lapso en el cual se extiende la producción del sonido. La unidad de tiempo más común empleada en la medición de la duración acústica son los milisegundos, milésima fracción de un segundo (ms) (Rose, 2002).

## 2.4 Dificultades en el Análisis Acústico Forense

Hasta este punto hemos hecho una contextualización de la fonética en el ámbito forense y hemos profundizado en algunos aspectos puntuales que atañen a nuestro trabajo y objetivos. Ahora se hace necesario hacer una revisión de algunas de las diferentes dificultades que se pueden enfrentar en un análisis forense y que en parte se deben a lo relativamente nueva que es la disciplina; por lo tanto, estas dificultades constituyen a su vez focos y oportunidades de investigación. Coulthard, Johnson, & Wright (2017), basados en diferentes autores, recogen algunos de los factores que pueden complicar la CFV, estos incluyen: la calidad de las grabaciones, generalmente por ser obtenidas con equipo de baja calidad y también por el ruido de fondo, la alteración voluntaria de la voz por parte del hablante (*voice disguise*), los gritos y susurros y la alteración de las características de la voz debido al consumo de sustancias como el alcohol. Un ejemplo claro y bastante relacionado con nuestro objeto de estudio es la disminución de la producción de pausas llenas en la medida en que los hablantes consumen alcohol (Christenfeld

and Creager, 1996 como se cita en Hughes, Wood, & Foulkes, 2016). Otras dificultades incluyen aspectos emocionales como el miedo, la ansiedad y estrés, que alteran las características del habla (Hollien, 2013). Existe también un inconveniente que por la naturaleza del presente trabajo, se hace necesario profundizar, se trata de las limitaciones técnicas que supone la transmisión de voz a través de un canal telefónico, sea celular o teléfono fijo. El ancho de banda está limitado a un rango que comprende desde aproximadamente 200 Hz hasta 3500 Hz, esto significa que no se puede acceder a la información acústica que queda fuera de este límite de frecuencia, por ejemplo los formantes altos de las vocales (F4, F5) y las características de las consonantes fricativas (Jessen, 2008). A este respecto, se puede considerar que con la innovación tecnológica han aparecido sistemas de comunicación como las aplicaciones de mensajería instantánea (WhatsApp, WeChat, QQ, entre otras) que representan una mejora en la calidad del audio principalmente en términos del ancho de banda. Esto facilita que se incluyan en los análisis, rasgos que en estudios previos no se podían tener en cuenta (Cao & Dellwo, 2019). Sin embargo aún con estas ventajas, estos sistemas de comunicación no están exentos de otras dificultades como se verá a continuación.

## **2.5 Potencial de las Aplicaciones de Mensajería Instantánea (MI) Como Fuente de Datos Para el Análisis Acústico Forense**

La mensajería instantánea se basa en aplicaciones, generalmente para dispositivos móviles, que facilitan la comunicación en tiempo real. Sin duda, estas han generado un cambio significativo en la interacción y el intercambio de información, desplazando por ejemplo a la comunicación telefónica y a los mensajes de texto. Existen diversas aplicaciones cuyo enfoque es la mensajería instantánea, que a su vez ofrecen distintas funciones. Sin embargo, dentro de estas aplicaciones resalta WhatsApp, pues actualmente cuenta con un inmenso número de descargas y usuarios.

Según la información de su página oficial<sup>6</sup>: «Más de 2 mil millones de personas en más de 180 países usan WhatsApp». Desde la perspectiva forense, el desarrollo de estas nuevas formas de comunicación despierta el interés porque a través de ellas también surgen nuevos medios y modalidades delictivas. Por lo anterior es esperable que, como indica De Gregorio (2020): «las comunicaciones mantenidas a través de las aplicaciones de mensajería instantánea están adquiriendo una gran transcendencia en la investigación de hechos delictivos» (p. 42) sin embargo la obtención de información proveniente de estas aplicaciones no se puede realizar como se acostumbra en la comunicación telefónica, es decir por medio de interceptaciones legales o solicitudes de copias a las empresas prestadoras de servicios de comunicación (De Gregorio, 2020). Dado lo anterior, se han llevado a cabo investigaciones para determinar qué información, producto de estas aplicaciones, se puede encontrar en la memoria interna de los dispositivos e identificar las rutas de acceso (e. g. Mahajan, Dahiya, & Sanghvi, 2013). Esto se puede considerar, a su vez, como una primera limitación porque se hace necesario contar con el dispositivo que contenga la información. Por otra parte, Soto (2020) pone en discusión la pertinencia y validez de la información obtenida de WhatsApp en el ámbito judicial, puesto que representa una serie de dificultades, especialmente en términos de la fácil manipulación del material contenido en las bases de datos locales (en el dispositivo) y la fácil simulación de conversaciones, lo que se escapa incluso de los análisis periciales de la informática. Así, el uso de conversaciones de Whatsapp como evidencia forense, tiene que ser tomado con cautela. Aún considerando lo anterior y desde el punto de vista fonético-forense, Whatsapp sigue siendo una fuente de obtención de muestras, con una calidad superior a la obtenida por medio de canales telefónicos, como se ilustra en las Figuras 1 y 2, que la hace potencialmente apropiada para el análisis por ejemplo en la CFV.

---

<sup>6</sup> [www.whatsapp.com/about/](http://www.whatsapp.com/about/)

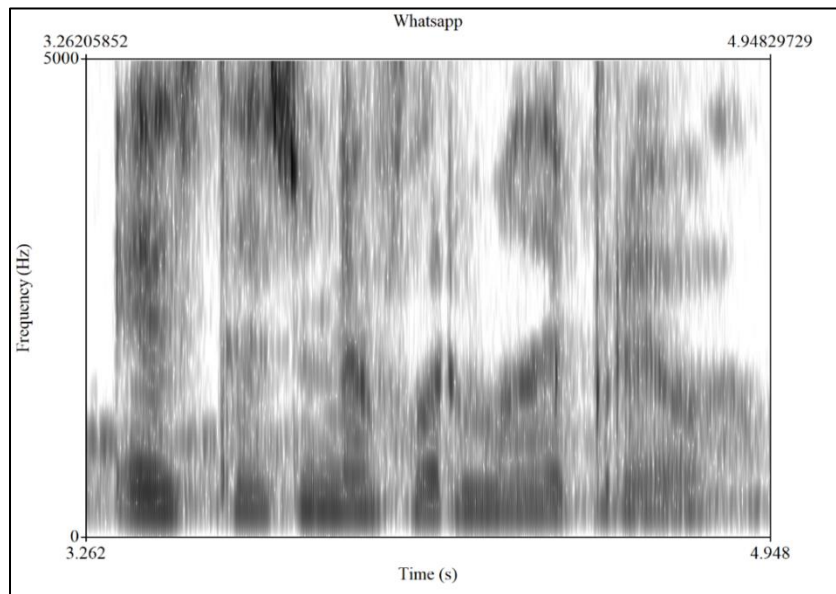


Figura 1. Espectrograma de un audio proveniente Whatsapp.

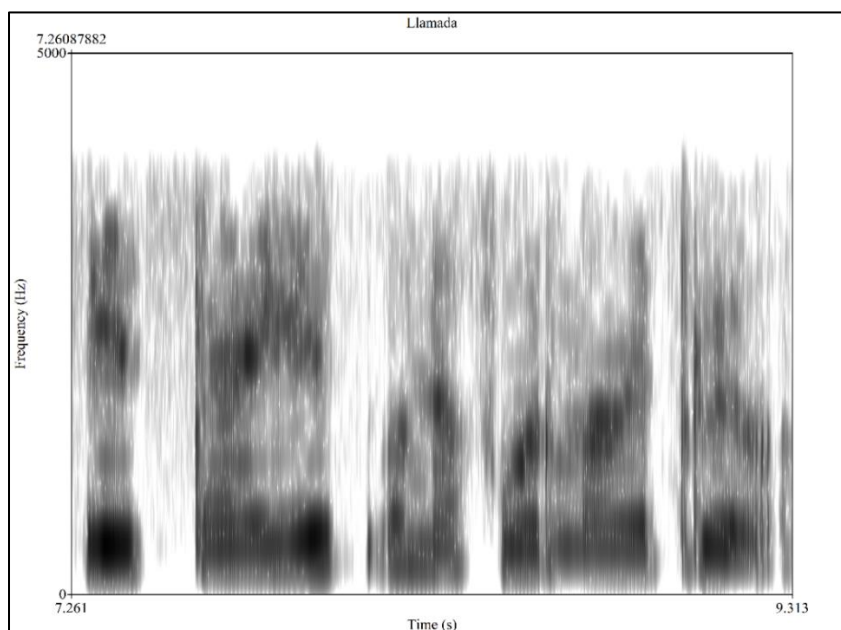


Figura 2. Espectrograma de un audio proveniente de una llamada telefónica.

En la Figura 1 se observa el espectrograma acompañado de la onda sonora de un audio proveniente de Whatsapp en donde claramente se observa que se superan los límites de ancho de banda establecidos en la comunicación telefónica (ver sección 1.4), por lo cual, contiene más

información acústica relevante en el análisis fonético forense. La Figura 2 corresponde al espectograma y onda sonora de un audio proveniente de una llamada telefónica y, como se observa, la información acústica no supera los 4000 Hz. Por esta razón, se hace notorio el potencial de las muestras obtenidas a través de Whatsapp para el análisis en la CFV, lo cual lleva a que este tipo de muestras hayan sido seleccionadas, junto con las pausas llenas, como parte del objeto de esta investigación.

## 2.6 Pausas Llenas y Mensajería Instantánea: Antecedentes de Investigación

En la presente sección se hará una revisión de los conceptos delimitados hasta ahora, así como otros enfoques aplicados en estudios previos que involucran las pausas llenas y otros aspectos relacionados con nuestra investigación, como la atención hacia la mensajería instantánea.

Inicialmente, las pausas llenas se han investigado desde enfoques ajenos a la fonética forense, por ejemplo, Villa, Gil, & Lahoz-Bengoechea (2017) realizan un estudio en el que analizan las muestras de habla de 50 informantes masculinos dentro de lo que ellos denominan la variedad castellana del español, específicamente sobre las características acústicas de las vocales que pueden aparecer como relleno de las pausas llenas en esta lengua, simultáneamente haciendo una comparación con la vocal léxica /e/. Para el análisis son tenidos en cuenta parámetros acústicos como los tres primeros formantes (F1, F2, F3), una duración limitada a 200 ms, la frecuencia fundamental, las distancias interformánticas y, como el trabajo estuvo orientado, también, a identificar el grado de similitud entre las vocales de relleno y la vocal léxica, se consideró la ratio interformántica:  $\frac{F2-F1}{F3-F2}$ , que determina en que medida la vocal es central, anterior o posterior. A partir de los resultados concluyen que la vocal de relleno es similar a la vocal léxica /e/ sin embargo, conserva algunas diferencias significativas. Por ejemplo, en la duración, en el

adelantamiento, y también que se presenta más variación entre las realizaciones de la vocal de relleno en comparación con la vocal lexica. (p. 168) . Resaltan también las diferencias respecto a los valores del segundo formante (F2) que son sistemáticamente más elevados en las vocales de relleno (p. 168).

Desde un enfoque forense, Cicres (2014) evaluó el potencial discriminante de las pausas llenas en la CFV. Su estudio comprendió 15 grabaciones de 10 hablantes, 5 de cada sexo, con edades entre los 25 y los 45 años, hablantes de catalán. El potencial discriminante se evaluó en terminos del porcentaje de de casos clasificados correctamente (muestra indubitada-muestra dubitada), desde un análisis discriminante en un corpus de referencia cerrado. Cicres tuvo en cuenta como parámetros, además de los primeros tres formantes y las distancias interformánticas, aspectos relativos a la cualidad de voz: frecuencia fundamental, sonoridad, jitter, shimmer y ruido. Uno de sus objetivos fue determinar cual conjunto de parámetros ofrecía una mayor capacidad discriminante. De su estudio, Cicres concluye que las pausas llenas sí tienen relevancia en la CFV, además, afirma que los parámetros relativos a la cualidad de voz ofrecen mejores resultados en terminos de casos clasificados correctamente, sin embargo, señala que considerar el conjunto completo de parámetros mejora los resultados (p. 380). Para finalizar señala que en su investigación existen diferentes limitaciones, por un lado, los parámetros sobre la cualidad de voz son altamente susceptibles de ser alterados por factores externos (como los que se describen en la sección 1.4), además el estudio se realizó con muestras que aunque simulaban las que se pueden encontrar en un contexto real (en terminos de calidad del audio), siguen teniendo una calidad que favorece los resultados del análisis empleado.

Hughes, Wood, & Foulkes (2016) también abordan el estudio de las pausas llenas, específicamente *uh* [ə:] y *um* [ə:m] en un conjunto homogéneo de hablantes de una variedad de

inglés británico.. El estudio está enmarcado en el campo de la CFV. Los autores argumentan que la relevancia de este estudio está dada por la implementación del marco de razón de verosimilitud, que hasta el momento no se había aplicado específicamente a la evaluación de las pausas llenas. Tienen en cuenta las propiedades acústicas de los primeros tres formantes y otro aspecto destacable es la consideración de las trayectorias de los formantes (*dynamic features*) (McDougall, 2006), tomadas a partir de nueve puntos de referencia a lo largo de la duración de la vocal, también analizan la duración de la vocal y la duración del segmento nasal en el caso de *um* [ə:m].

Estos autores realizan un estado del arte bastante conciso y pertinente que tiene en cuenta los estudios que se han hecho sobre las pausas llenas desde aproximaciones lingüísticas o fonéticas que tienen en cuenta las pausas llenas así como otro tipo de ‘marcas dubitativas’ que varían según el tipo de discurso, aspectos sociolingüísticos, pragmáticos y también la incidencia de sustancias como el alcohol en la frecuencia de producción de estos segmentos. Sostienen que los estudios sobre las pausas llenas han sido pocos, pero que exhiben el potencial de las pausas llenas como variables para la discriminación entre hablantes. Se describe la importancia de las pausas llenas, estas se acercan al tipo de muestras que muestran poca variación intra-hablante respecto de la variación inter-hablante. Señalan que su análisis es más productivo (en el ámbito forense) que, por ejemplo, el de segmentos vocálicos léxicos (Hughes, Wood, & Foulkes, 2016). En la investigación se buscó, por un lado, describir los patrones de variación acústica grupal e individual, y por otro, evaluar el potencial discriminatorio de las pausas llenas mediante una prueba basada en la razón de verosimilitud. Se trabajó con el corpus DyViS (*Dynamic Variability in Speech corpus*), que fue recolectado para fines forenses. Las condiciones técnicas de grabación simulan un contexto forense, pero siguen siendo alejadas de la realidad, lo que constituye una limitación en la investigación. El conjunto final de datos fue obtenido de 60 hablantes, con 16 muestras de cada pausa por hablante.

Dentro de los resultados del estudio, los autores señalan que las medidas dinámicas (trayectorias de los formantes) fueron más significativas en el caso de *um* [ə:m], lo que pudo deberse a un mayor grado de dinamismo por la transición del sonido vocálico al nasal, con lo que no se cuenta en *uh* [ə:]. Lo anterior sugiere que en medidas dinámicas puede encontrarse información que se ignora con procedimientos que emplean medidas estáticas, como la frecuencia central. Para el caso puntual de [ə:m], tener en cuenta la duración del segmento nasal no representó alguna mejora significativa en los resultados, pero considerar la duración en el caso del segmento vocálico, sí. Como conclusión, ratifican el potencial del uso de pausas llenas como variables en la comparación forense de voces y también se da cuenta de cuales son algunos de sus patrones de variación.

Sin tener en cuenta las pausas llenas como material fonético, pero sí prestando atención a la MI, (Cao & Dellwo (2019) reconocen su potencial desde un punto de vista fonético-forense y desde la relevancia que tiene el avance en la tecnología en la posibilidad de obtener muestras de audio con una mejor calidad, en comparación, por ejemplo, con muestras provenientes de llamadas telefónicas. En su estudio evalúan el rol de los primeros cinco formantes en tres vocales del idioma mandarín. En este caso el corpus está construido simulando las características que tendría un audio proveniente de WeChat<sup>7</sup>, esto es, una frecuencia de muestreo de 16kHz, por otra parte, la recolección de datos no es contemporánea, teniendo un margen de diferencia de entre una semana y un mes. Sin embargo, la producción de los segmentos se hizo de forma aislada, es decir que no corresponde a muestras de habla espontánea. Los resultados muestran que los formantes altos (F4

---

<sup>7</sup> Aplicación de mensajería instantánea popular en China ([www.wechat.com](http://www.wechat.com))



y F5) tienen mayor potencial de discriminación de hablantes (p. 620). Sin embargo, el quinto formante no se presenta en todos los hablantes. Los formantes que aportan información relevante para cada vocal varían, pero el análisis bajo la combinación de los tres formantes (específica para cada caso), mejora considerablemente los resultados. No queda claro por qué razón los autores emplearon muestras simuladas y no extraídas directamente de las aplicaciones, no obstante lo recomiendan para futuras investigaciones.

Por otra parte, Orebaugh, & Allnutt (2009), elaboran un estudio que se sitúa entre la informática y la lingüística forense. Nuevamente, los investigadores centran su atención en el potencial de las aplicaciones de MI, y desde el marco de la identificación del autor/atribución de autoría, evalúan una serie de características estilométricas presentes en las conversaciones por chat de cuatro usuarios de Gaim (actualmente Pidgin) y Adium, ambas aplicaciones de MI. El objetivo de este experimento fue evaluar si se podía identificar a quien produce un texto, basándose en el «comportamiento de autor» y cuáles características de este comportamiento tienen más fuerza discriminante. Fue predefinida una serie de atributos establecidos en cuatro categorías: caracteres especiales, emoticones, abreviaciones, estructura de oraciones. Se crearon 35 muestras de 2500 caracteres cada una y se sometieron a clasificación mediante 3 diferentes métodos y algoritmos. Los resultados muestran que en general los índices de precisión son elevados, la categoría que obtuvo mayor porcentaje de precisión fue la de abreviaciones, sin embargo, al evaluar los atributos de manera individual, se disminuye la precisión.

De la revisión hecha en esta sección surgen algunas consideraciones. Primero, varios trabajos realizados sobre diferentes lenguas, con diferentes metodologías y parámetros de análisis, señalan el potencial de las pausas llenas como material fonético apropiado para la CFV. Segundo, la discusión sobre qué parámetros analizados ofrecen mejores resultados, sigue abierta, lo que representa un foco de investigación. Tercero, la atención que han despertado las aplicaciones de MI

no es reciente, y como ya se ha mencionado anteriormente, ofrecen nuevas posibilidades en cuanto a la información acústica obtenible de una muestra. También, hay algunos vacíos, o bien, sugerencias que señalan los autores, como la necesidad de emplear muestras más cercanas a la realidad (e. g. Cao & Dellwo, 2019; Hughes, Wood, & Foulkes, 2016), aunque esto no deja de ser una dificultad.

Con todo lo anterior, se ha reflexionado que si existen pocos estudios en países en donde la fonética forense tiene un grado de avance significativo, la necesidad de emprender investigaciones similares, o incluso previas, en el español de Colombia es mayor. Lo anterior debido a que el desarrollo, a nivel teórico, de la disciplina en el país no se encontró significativo, no existen casos documentados ni se pueden encontrar investigaciones enmarcadas en el área. Por lo tanto, desde esta investigación pionera surgen los siguientes interrogantes:

## **2.7 Preguntas de investigación**

- ¿Cuál es el comportamiento de las pausas llenas en cuanto a su frecuencia de aparición en el español hablado en Bogotá (Colombia)?
- ¿Cuáles son las características acústicas presentes en la estructura formántica y duración en las pausas llenas?
- ¿Qué tipo variación (inter o intrahablante) es mayor según las características acústicas de las pausas llenas?
- ¿Se pueden considerar las pausas llenas como un referente confiable para la comparación forense de voces en el español de Bogotá (Colombia)?

En consecuencia, los objetivos de investigación son los siguientes:

## **2.8 Objetivo general**

- Explorar el potencial de las pausas llenas del español hablado en Bogotá (Colombia) como parámetro en la comparación forense de voces.

## **2.9 Objetivos específicos**

- Describir el comportamiento de las diferentes pausas llenas según su frecuencia de aparición a nivel intra e interhablante.
- Explicar las características acústicas de la estructura formántica, y la duración de las pausas llenas en el español hablado en Bogotá (Colombia).
- Determinar si existe variación intra e interhablante según las características acústicas de cada parámetro estudiado.
- Identificar qué tipo de variación (intra o interhablante) es mayor en el conjunto de datos analizado.

### **3. Metodología**

En este capítulo se hará la descripción de los diferentes elementos que componen la metodología empleada, organizados de la siguiente forma: i) tipo de estudio (2.1), ii) las características del grupo de hablantes (2.2), iii) la recolección y procesamiento preliminar de los audios (2.3), iv) los segmentos extraídos (2.4) y v) la descripción detallada de los procedimientos y recursos utilizados (2.5).

#### **3.1 Tipo de Estudio**

De acuerdo con los objetivos de este trabajo, se realizó un estudio de carácter cuantitativo, ya que se emplearon métodos estadísticos para el análisis y comparación de los datos fonéticos abordados, es decir, los valores de las mediciones acústicas de la estructura formántica (tres primeros formantes) de las pausas llenas y la duración de los determinados segmentos.

#### **3.2 Participantes**

En la evaluación de la variación intra e interhablante en las características acústicas de un segmento (en este caso las pausas llenas) bajo un enfoque forense, es fundamental contar un grupo de hablantes con características sociolingüísticas homogéneas. Lo anterior con el objetivo de determinar si, aún en condiciones de homogeneidad, existen particularidades en cada hablante. En este caso, el conjunto completo de participantes está compuesto por 12 hablantes, todos femeninos ya que existen diferencias fisiológicas en los pliegues y tracto vocálico entre ambos sexos, que afectan las características acústicas (Rose, 2002); con edades entre 20 y 25 años (edad media: 22 años), con estudios universitarios, y residentes en la ciudad de Bogotá, en donde se llevó

a cabo el estudio. El promedio de edad definido se debe a que la mayoría de los cambios en los patrones del habla se dan en periodos anteriores y posteriores a ese rango (i. e. infancia, pubertad y la transición a la vejez) (Jessen, 2010). En relación con el nivel de estudio y con el fin de controlar esta variable, se delimitó el grupo a personas que contaran con formación profesional. En la Tabla 2 se relacionan las participantes y el rótulo correspondiente con el cual serán referenciadas en lo posterior de este documento.

Tabla 2

*Rótulos de las participantes y datos preliminares.*

Rótulo del hablante	Edad	Audios recolectados	Tiempo total de grabación (minutos)
Hablante 1	23	24	17
Hablante 2	22	15	16
Hablante 3	23	11	16
Hablante 4	25	9	12
Hablante 5	24	19	17
Hablante 6	21	5	16
Hablante 7	22	20	17
Hablante 8	24	22	19
Hablante 9	24	16	14
Hablante 10	21	9	16
Hablante 11	21	16	14
Hablante 12	23	7	12
	<b>Media 22</b>	<b>Total 173</b>	<b>Media 15,5</b>

Las participantes fueron contactadas e informadas acerca de la naturaleza de la investigación. Se les hizo entrega de una copia de la hoja de información del proyecto de investigación, que explica el propósito de la investigación, las condiciones de participación y el compromiso de confidencialidad y anonimato. De igual forma, se hizo entrega del formato de consentimiento informado para su correspondiente diligenciamiento (Ver Anexo A).

### **3.3 Procedimiento para la Recolección y Selección de Audios**

Inicialmente se realizó un sondeo preliminar con la colaboración de dos hablantes. A cada hablante se le solicitó enviar dos audios, con una duración aproximada de un minuto cada uno, esto con el fin de contar con un estimado sobre la frecuencia de aparición de las pausas requeridas, en un tiempo determinado. Una vez surtida esta fase preliminar, las participantes que fueron seleccionadas como sujetos, enviaron una serie de audios por medio de la aplicación WhatsApp.

Emplear una aplicación de mensajería instantánea como medio para la obtención de datos posee ciertas ventajas, por ejemplo, la celeridad en el proceso de recolección. Así mismo, se garantiza que las grabaciones corresponden a muestras de habla espontánea, lo cual es un factor importante en este tipo de investigaciones. Sin embargo, existen otros aspectos a considerar. Por ejemplo, los archivos de audio extraídos de WhatsApp poseen un formato de archivo denominado OGG<sup>8</sup>, que funciona como contenedor de multimedia y comprime la información contenida. Sin embargo, este formato no es soportado en diferentes programas de análisis de audio dado que no es netamente un formato de audio ya que soporta, simultáneamente, diferentes tipos de contenido (audio, video, texto, etc.) . Por lo anterior, fue necesario hacer la conversión de formato, de OGG a

---

<sup>8</sup> Para ampliar la información al respecto se puede visitar la página web de sus desarrolladores: [www.xiph.org](http://www.xiph.org)

WAV, esta tarea se realizó a través de Audacity (Audacity Team, 2020). De igual manera, se empleó este software para seccionar los audios que superaron los 2 minutos de duración, con el fin de facilitar su análisis, manteniendo un promedio de duración de un minuto. Finalmente, cada audio fue etiquetado de manera que fuera posible agruparlos en carpetas separadas para cada hablante.

### 3.4 Descripción de los Datos Seleccionados

De acuerdo con la literatura relacionada con investigaciones sobre pausas en la lengua española, inicialmente se consideraron las pausas [e:] y [m:] como los elementos a identificar y segmentar. Sin embargo, a lo largo del proceso de identificación y segmentación, además de [e:] y [m:], se evidenciaron realizaciones de [e:m:] con una frecuencia de aparición considerable, como se puede observar más adelante en los resultados, por lo que se optó por incluirlo en la segmentación. En el caso de [e:], por su semejanza con la vocal léxica, los datos tenidos en cuenta fueron las mediciones de la frecuencia, tomadas de los tres primeros formantes en su punto medio, al igual que su duración. En el caso de [m:] solo se tuvo en cuenta la duración. Para [e:m:] se midió la duración del elemento vocal, del elemento nasal y también el total. Los audios que no contenían ninguna de las anteriores pausas, fueron separados de los demás audios y rotulados para su posterior identificación.

De acuerdo con lo anterior, el corpus total que se obtuvo para esta investigación corresponde a 173 audios, los cuales en promedio conforman 15,5 minutos de duración por hablante, para un total de 186 minutos de grabación (ver Tabla 2)(Arriba). Es importante resaltar que todos los audios incluidos superan los 30 segundos de duración e, igualmente, se destaca que

fueron producidos y grabados con una antigüedad no mayor a seis meses al momento de su recolección.

### **3.5 Procedimientos**

Una vez recolectado el corpus y organizadas las colecciones correspondientes a los distintos tipos de pausas, se llevó a cabo una serie de procedimientos los cuales se describirán en detalle a continuación.

#### **3.5.1 Segmentación de pausas llenas**

La segmentación de las pausas llenas se hizo de forma manual, procesando cada audio a través del programa PRAAT (Boersma & Weenik, 2021) mediante de la función de anotación en *TextGrids*. Se establecieron intervalos de segmentación, ignorando aproximadamente el 5% del inicio y final de cada pausa. Lo anterior se hizo con el fin de mitigar cualquier posible influencia de la coarticulación, ya que, aunque se trate de elementos independientes (i. e. su producción es aislada de los elementos léxicos y contexto fonológico), en algunos casos son producidos inmediatamente antes o después de una palabra. En la siguiente imagen se ilustra el proceso de segmentación llevado a cabo:



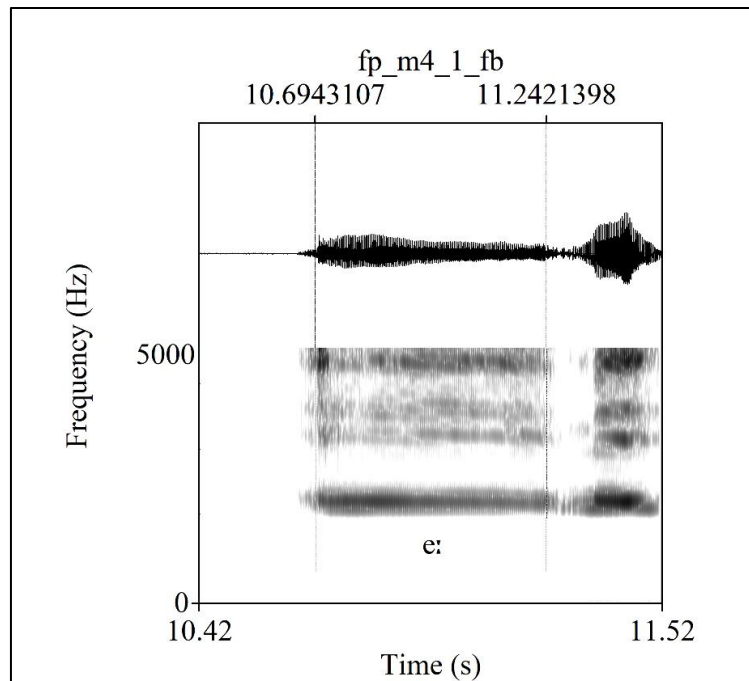


Figura 3. Segmentación de una realización de la pausa llena vocálica [e:].

### 3.5.2 Mediciones

En el caso de [e:], el procedimiento se realizó siguiendo y adaptando la metodología ofrecida en Hughes, Wood, y Foulkes (2016); el valor de los puntos medios de los formantes y la duración del segmento se obtuvieron a través de PRAAT, función *To Formant (burg method)*, mediante un *script* configurado para identificar cinco formantes en un rango de 0 a 6000 kHz. Fue necesario hacer una comprobación manual en algunos casos en donde el *script* arrojó valores demasiado alejados de los esperados. Existieron casos en los que se pudo corregir el valor de forma manual, en los que no, se descartó el segmento para evitar posibles sesgos. La medición de la

duración para [m:] y [e:m:] se tomó de forma manual. Los datos de los tres tipos de pausas fueron tabulados en excel, en una hoja de cálculo por cada pausa<sup>9</sup>.

### **3.5.2 Conjuntos de datos seleccionados por hablante (muestras)**

Según el sondeo previo (ver sección 2.3) y la cantidad de tiempo de grabación recolectada, se esperaba contar con un mínimo de diez pausas de cada tipo ([e:] y [m:]) por hablante. Sin embargo, como se verá con más detalle en los resultados (ver capítulo 3), la frecuencia de realización de pausas varía considerablemente de una hablante a otra; por ejemplo, la hablante 4 alcanza un total de 47 realizaciones de [e:], mientras que para la hablante 8 se registraron solo 9 realizaciones de esta misma pausa. No obstante, se mantuvo el mínimo de diez pausas por hablante y, en consecuencia, la hablante 8 no se tuvo en cuenta en la fase de análisis estadístico del estudio, pero sí en los resultados. Por otra parte, teniendo presente las características que se esperan de un segmento a emplear en la CFV (ver sección 1.2.3), se descartó el análisis de la pausa [m:], dado que la cantidad de realizaciones obtenida no fue suficiente para conformar un conjunto de datos por hablante que se pudiese someter a un análisis estadístico apropiadamente (máximo de apariciones 8, media 3). Por último, se fijaron cuatro conjuntos de datos para la pausa [e:]: F1, F2, F3 y duración. Cada conjunto de datos compuesto por un total de 110 valores y dividido en once subgrupos (i. e. cada una de las hablantes).

---

<sup>9</sup> Esta información se encuentra disponible en: <https://github.com/febaut/linguistics> o por solicitud al correo electrónico lbautista@unal.edu.co

### 3.5.3 *Análisis estadístico*

Los datos anteriormente descritos fueron procesados a través del programa estadístico SPSS, versión 25.0 (IBM, 2020). A cada conjunto de datos se le realizó la prueba de Kolmogorov Smirnov para determinar si contaban con una distribución normal, con el objetivo de poder establecer el tipo de análisis estadístico a realizar (i. e. estadística paramétrica o estadística no paramétrica). A continuación, se aplicó la prueba ANOVA de un factor para los conjuntos de datos correspondientes a la estructura formántica, a saber, F1, F2 y F3. Esta prueba evalúa si hay relación entre una variable independiente y una dependiente, a través de la comparación de las medias de cada grupo; en este caso, los grupos (variable independiente) corresponden a los hablantes, mientras que los valores de los puntos medios de los formantes, corresponden a la variable dependiente. Finalmente, un contraste *post hoc* permitió evidenciar en qué medida difiere cada grupo respecto de los demás. En el caso del conjunto de datos correspondiente a los valores de la duración, se aplicó una prueba no paramétrica para múltiples variables independientes puesto que los datos no presentaron distribución normal. Los resultados obtenidos de los anteriores procedimientos serán descritos detalladamente en el siguiente capítulo.

## 4. Resultados

El objetivo general de este estudio, explorar el potencial de las pausas llenas en la comparación forense de voces, se puede realizar a partir del análisis de diferentes características fonéticas de estos segmentos. En esta caso particular, se analizó su estructura formántica y duración, por lo tanto, su validez como parámetro en la comparación forense de voces está determinada en función de si la variación interhablante es mayor que la variación intra hablante. Teniendo en cuenta lo anterior, en este capítulo se presentan los resultados a través de las siguientes secciones: tipos de pausas y frecuencia de aparición (3.1) y variación intra e interhablante (3.2).

### 4.1 Tipos de Pausas y Frecuencias de Aparición

En la Tabla 3, se observa el inventario final de pausas obtenido, que reúne un total de 389 ítems; de los cuales el mayor porcentaje, es decir 61,1 %, corresponde a la pausas vocálica [e:] con un total de 238 apariciones, seguido de la pausa compuesta por segmento vocálico y nasal [e:m:] con 109 apariciones que corresponden al 28,19 % de las realizaciones; la pausa nasal [m:] se registró 36 veces, lo que equivale al 9,2% y, finalmente, en escasas hablantes (4, 5, 11, 12) se evidenciaron algunos segmentos atípicos, tales como una pausa vocálica [a:] ( 1 observación = 0,25%), cinco segmentos compuestos por vocal y nasal [a:m:] (3 observaciones = 0,77%) y [e:n:] ( 2 observaciones = 0,51%).

Tabla 3

*Tipos de pausas observadas.*

Tipo de pausa	Número de realizaciones	Porcentaje
[e:]	238	61,1
[e:m:]	109	28,19
[m:]	36	9,2
[a:]	1	0,25
[a:m:]	3	0,75
[e:n:]	2	0,51
<b>Total</b>	<b>389</b>	<b>100</b>

Nota: Las filas sombreadas corresponden a los segmentos atípicos evidenciados.

De acuerdo con lo anterior, el número y tipo de pausas producidas por cada una de las hablantes varía considerablemente. El resumen de los registros de las tres principales pausas registradas ([e:], [e:m:] y [m:]), contenido en la Tabla 4, permite observar que existe un rango de aparición interhablante bastante amplio para los tres ítems. Por ejemplo, en el caso de [e:], aunque la media es de aproximadamente 20 pausas, la Hablante 4 supera este valor por más del doble, mientras que la Hablante 8 escasamente alcanza la mitad. Para las pausas [e:m:] y [m:] el comportamiento es similar, incluso, para algunas de las hablantes se registraron cero apariciones.

Tabla 4

*Síntesis del total de apariciones registradas por tipo de pausa.*

Tipo de pausa	Mínimo de realizaciones	Máximo de realizaciones	Media
[e:]	9	47	19,83
[e:m:]	0	27	9,8
[m:]	0	8	3

Complementando, en la Figura 4 se observa que la mayoría de las pausas registradas, especialmente [e:], se concentran en cuatro de los doce hablantes (3, 4, 6, 7). Esta misma pausa es la más frecuente y es empleada por todos los hablantes. Por otra parte, la pausa [e:m:] es la segunda más frecuente. Sin embargo, no se registraron apariciones para la hablante 12 y solamente 2 para la hablante 8. Por último, la frecuencia de aparición de [m:] es bastante reducida en comparación con los otros tipos de pausas y se encuentra concentrada principalmente en las hablantes 1, 2 y 11. Además, las hablantes 3 y 8 tienen apenas un registro de esta pausa, mientras que la hablante 9 no la realiza en lo absoluto.

Según lo anterior, las hablantes muestran preferencias hacia cierto tipo de pausas. Asimismo, algunas hablantes son más proclives a emplear pausas de un solo tipo (hablante 4), mientras que otras presentan un uso con tendencia al balance en los diferentes tipos de pausas (hablantes 1 y 2).

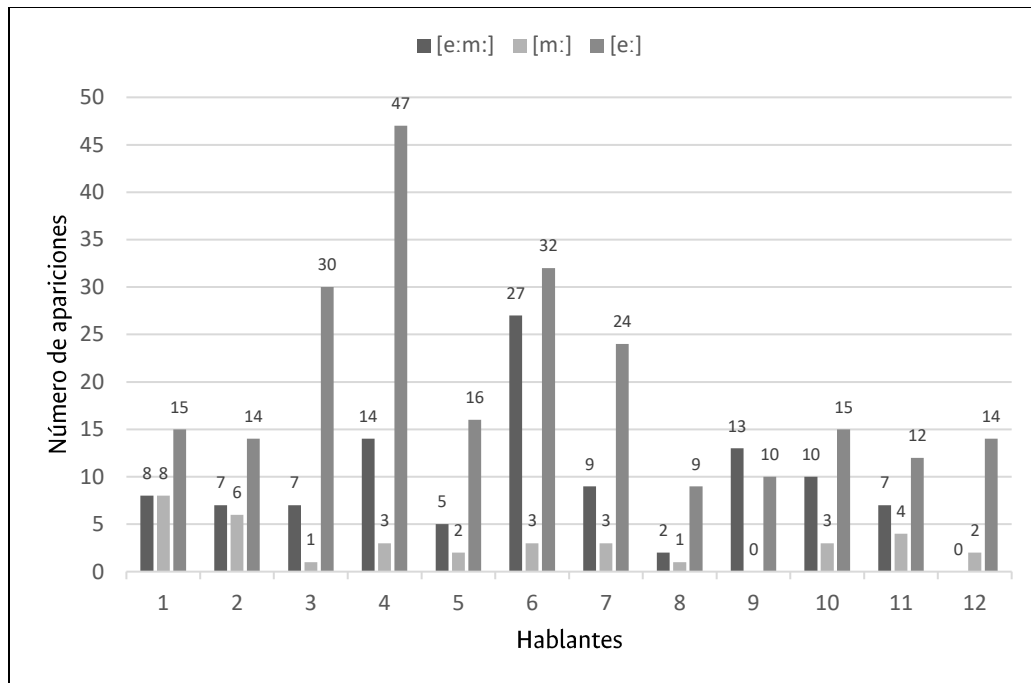


Figura 4. Apariciones totales registradas por cada tipo de pausa en cada hablante.

Finalmente, se evidenció que existen diferencias a nivel intrahablante en la realización de pausas de un audio a otro. Esta información se encuentra resumida en la Tabla 5, la cual relaciona la cantidad de pausas registradas con el tiempo total analizado. Se puede observar también el tiempo total analizado en el que cada hablante no realiza ningún tipo de pausas llenas.

Según el proceso realizado en la etapa de identificación y segmentación (ver 2.5.1), se puede hipotetizar que estas variaciones pueden depender del interlocutor o tema de la conversación. Sin embargo, este fenómeno está fuera del alcance y objetivos de la presente investigación.

Tabla 5

*Relación del tiempo de grabaciones analizado por hablante y el total de pausas registradas.*

Hablantes	Tiempo Total (minutos)	Tiempo total sin pausas (minutos)	Pausas totales	Pausas por minuto
Hablante 1	17	5	31	1,82
Hablante 2	16	4	27	1,68
Hablante 3	16	1	38	2,37
Hablante 4	12	0	64	5,33
Hablante 5	17	6	23	1,35
Hablante 6	16	2	62	3,87
Hablante 7	17	3	36	2,11
Hablante 8	19	10	12	0,63
Hablante 9	14	3	23	1,64
Hablante 10	16	2	28	1,75
Hablante 11	14	6	23	1,64
Hablante 12	12	2	16	1,33

Nota: La fila sombreada corresponde a la hablante con menor número de pausas llenas producidas y fue descartada en el análisis estadístico de la investigación.

De acuerdo con la información presentada en la Tabla 5 (arriba), se destacan las hablantes 4 y 6 con un número elevado de pausas por minuto, además de contar con poco o nulo tiempo sin pausas. Por otra parte, aunque en comparación con las demás participantes, se obtuvo un total de tiempo de grabación mayor para la hablante 8; esta misma hablante cuenta con el menor registro de pausas y el mayor tiempo total sin pausas. Por lo anterior, en lo restante del análisis de los datos obtenidos no se contó con la información proveniente de esta hablante, tal como se mencionó



previamente. En relación con todo anterior, en las siguientes secciones el análisis se centra exclusivamente en la pausa vocálica [e:], ya que tal como se tenía previsto, según la bibliografía consultada, no se contó con el total de apariciones necesario por hablante en el caso de la pausa [m:] y, por otra parte, la pausa [e:m:] no figura en la literatura como una pausa llena del español; razón por la cual no fue incluida en el presente estudio. No obstante, de acuerdo con lo reportado en relación con la frecuencia de aparición, la pausa [e:m:] es la segunda con mayor aparición, lo cual resulta de alto interés para futuros estudios.

## 4.2 Variación Intra e Interhablante

En esta sección se reportan los resultados de la variación observada entre distintas realizaciones de la misma pausa ([e:]) por cada una de las hablantes, en comparación con las demás. Lo anterior con el fin de, por una parte, exponer las características acústicas generales de este tipo de pausa llena, así como determinar qué tipo de variación es mayor (intra o interhablante) según cada parámetro.

Debido a que la pausa llena vocálica analizada es similar a la vocal léxica [e], el análisis se realizó teniendo en cuenta el valor del punto medio de los tres primeros formantes, así como la duración en milisegundos del segmento. Para cada parámetro analizado se presentará un apartado dentro de esta sección en el cual se informará sobre sus resultados.

### 4.2.1 Primer formante (F1)

La Tabla 6 muestra los datos estadísticos descriptivos producto del análisis de once conjuntos de datos, compuestos por los valores de los puntos medios de F1 medidos en diez pausas distintas para cada hablante. Los valores en hercios muestran el valor de la media muestral para cada hablante, la desviación estándar, los valores mínimo y máximo del conjunto de valores medios del formante, y el rango de la media poblacional con un nivel de confianza del 95% (límite inferior y superior). Se observa una desviación estándar con tendencia a la uniformidad entre la mayoría de las hablantes, exceptuando a las hablantes 5, 11 y 12. Lo anterior sugiere que la variación interhablante no es mayor que la variación intrahablante, según en el valor global de la desviación estándar (50,98) comparado con los de cada hablante, lo que representa la poca consistencia en las diferentes realizaciones de la pausa a nivel intrahablante.

Tabla 6

*Estadísticos descriptivos de los valores del punto medio de F1.*

Hablantes	Media	Desviación estándar	Límite Inferior	Límite superior	Valor Mínimo	Valor Máximo
Hablante 1	444,750	35,6628	419,238	470,262	387,8	509,3
Hablante 2	451,940	33,6069	427,899	475,981	412,0	508,1
Hablante 3	421,050	45,3308	388,622	453,478	311,7	466,2
Hablante 4	457,650	36,1038	431,823	483,477	421,4	533,4
Hablante 5	460,640	63,4622	415,242	506,038	353,4	580,0
Hablante 6	448,360	38,3416	420,932	475,788	386,4	509,3
Hablante 7	473,140	49,5070	437,725	508,555	398,5	531,8
Hablante 9	474,260	37,0566	447,751	500,769	426,5	546,3

Hablante 10	497,440	43,2799	466,479	528,401	438,7	555,7
Hablante 11	456,640	82,3867	397,704	515,576	359,5	599,5
Hablante 12	477,680	56,8673	437,000	518,360	409,9	556,0
Global	460,323	50,9856	450,688	469,958	311,7	599,5

Nota: Se encuentran sombreadas las filas de las hablantes con un rango de frecuencia mayor a 200 Hz.

La Figura 5 complementa la información presentada en la tabla anterior. En general se observa que los valores de las diferentes pausas se encuentran dispersos en un rango similar (400 – 500 Hz) para la mayoría de las hablantes. La hablante 3 cuenta con un rango de frecuencias relativamente corto en comparación con las demás; sin embargo, este rango se ve afectado por un valor atípico de 311,7 Hz, que es además el registro más bajo del valor del punto medio de F1 entre todas las participantes. Se observa también una dispersión bastante amplia en las hablantes 5 y 11, lo cual hace más evidente la tendencia a presentar una variación intrahablante elevada en los valores de F1.

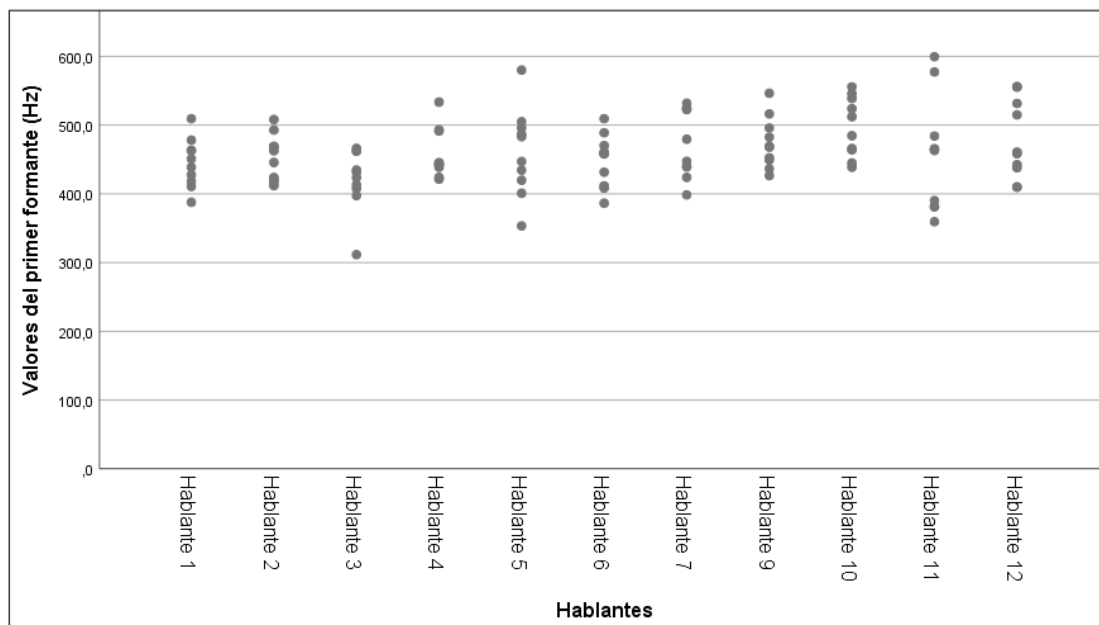


Figura 5 Diagrama de dispersión de los valores de F1 para todas las hablantes.

Finalmente, los resultados del análisis de varianza (ANOVA) muestran que no existen diferencias estadísticamente significativas ( $p > 0,05$ ) entre las medias de los valores de F1 en los diferentes grupos (i. e. conjunto de valores de cada hablante). Según lo anterior se puede establecer que la variación interhablante según los conjuntos de valores de F1 para cada hablante, no es suficiente para determinar que existan diferencias relevantes entre una hablante y otra, lo que sugiere que los datos del punto medio de F1 son un parámetro débil para la CFV.

#### **4.2.2 Segundo y tercer formante (F2 y F3)**

En las Tablas 7 y 8, respectivamente, están contenidos los datos de la media muestral, la desviación estándar, los valores mínimo y máximo del conjunto de valores medios del formante, y el rango de la media poblacional con un nivel de confianza del 95% (límite inferior y superior) para cada formante.

Específicamente para F2, en la Tabla 7 se observa que el valor de la desviación estándar difiere de un hablante a otro y, a excepción de las hablantes 1 y 10, no supera el valor de la desviación global. El rango medio de frecuencias es de 540 Hz, en donde las mismas hablantes (1 y 10) alcanzan un rango de 825,1 Hz y 759,2, respectivamente. Para las demás hablantes, los valores se mantienen en un rango de 484 Hz en promedio.

Para F3, los resultados son similares, la dispersión global es considerablemente mayor que la de cada una de las hablantes como se observa en la Tabla 8. Por otra parte, los rangos de frecuencia más amplios para este formante corresponden a las hablantes 3 y 9 (977,4 Hz y 627,4 Hz respectivamente). Las demás hablantes se mantienen en un rango promedio de 527,3 Hz.

Tabla 7

Estadísticos descriptivos de los valores del punto medio de F2.

Hablantes	Media	Desviación estándar	Límite Inferior	Límite superior	Valor Mínimo	Valor Máximo
Hablante 1	2246,690	277,6859	2048,045	2445,335	1815,3	2640,4
Hablante 2	2539,510	169,3292	2418,379	2660,641	2328,4	2819,5
Hablante 3	2360,980	147,9110	2255,171	2466,789	2153,8	2666,9
Hablante 4	2520,903	106,7916	2444,509	2597,297	2332,3	2735,6
Hablante 5	2641,280	190,7291	2504,841	2777,719	2227,3	2850,8
Hablante 6	2684,730	178,0633	2557,351	2812,109	2468,3	2925,1
Hablante 7	2281,010	167,3717	2161,279	2400,741	2034,3	2608,8
Hablante 9	2379,000	156,7757	2266,849	2491,151	2167,1	2717,8
Hablante 10	2401,350	288,5055	2194,966	2607,734	1938,5	2697,7
Hablante 11	2664,750	121,2759	2577,994	2751,506	2520,3	2862,8
Hablante 12	2564,980	146,4266	2460,233	2669,727	2408,5	2809,0
Global	2480,471	230,7321	2436,869	2524,073	1815,3	2925,1

Tabla 8

Estadísticos descriptivos de los valores del punto medio de F3.

Hablantes	Media	Desviación estándar	Límite Inferior	Límite superior	Valor Mínimo	Valor Máximo
Hablante 1	2926,360	194,9143	2786,927	3065,793	2597,2	3201,5
Hablante 2	3010,770	224,7440	2849,998	3171,542	2766,8	3391,4
Hablante 3	2866,440	299,6634	2652,074	3080,806	2492,9	3470,3
Hablante 4	3126,580	130,8888	3032,948	3220,212	2975,0	3442,1

Hablante 5	3229,850	183,8589	3098,325	3361,375	2993,0	3511,4
Hablante 6	3325,360	151,7457	3216,808	3433,912	3123,9	3596,4
Hablante 7	2915,570	219,7306	2758,384	3072,756	2650,3	3225,4
Hablante 9	3000,350	206,7720	2852,434	3148,266	2502,5	3238,2
Hablante 10	2971,670	141,2834	2870,602	3072,738	2792,2	3148,9
Hablante 11	3294,430	192,5568	3156,683	3432,177	3003,0	3630,6
Hablante 12	3240,500	164,2750	3122,985	3358,015	2956,5	3456,4
Global	3082,535	246,1092	3036,027	3129,043	2492,9	3630,6

Los datos muestran que a diferencia de F1, para F2 y F3 hay tendencia a la uniformidad en las diferentes realizaciones de la pausa para cada hablante, pues la dispersión en cada grupo es menor que la dispersión global, según los valores de F2 y F3. Además, la dispersión de F3 es ligeramente menor para la mayoría de los hablantes en comparación con F2. Lo anterior se puede observar en las Figuras 6 y 7, las cuales muestran los diagramas de dispersión de los valores medios, para cada hablante, de F2 y F3 respectivamente. Según lo anterior se sugiere que para F2 y F3 se presenta una variación interhablante mayor que la intrahablante.

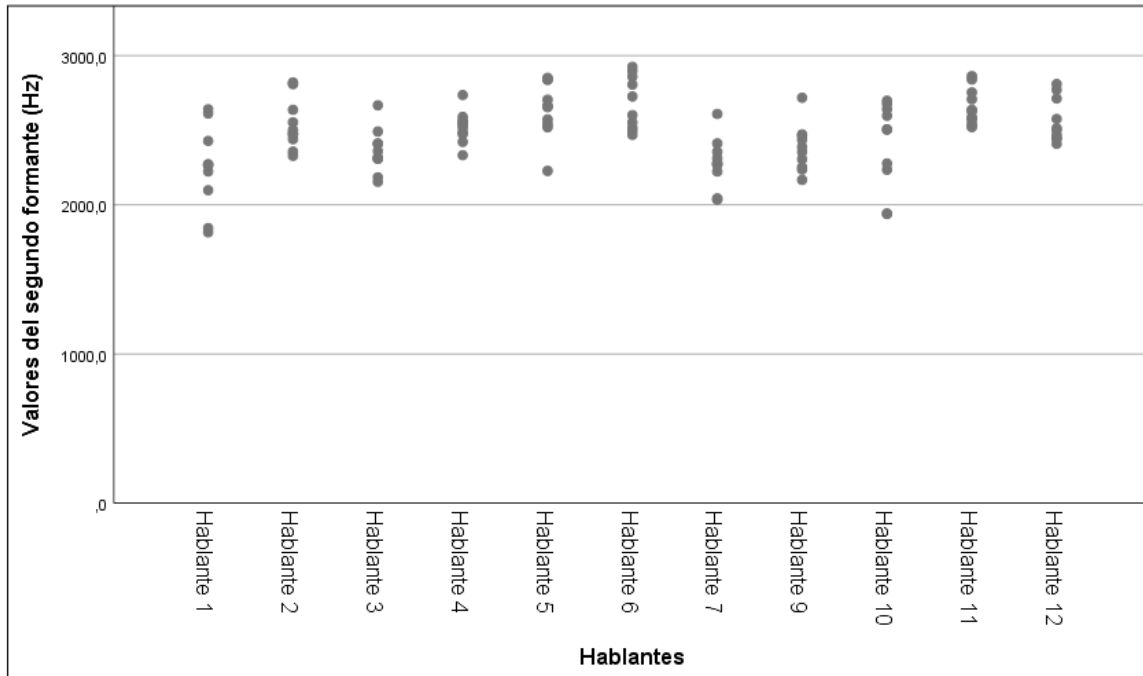


Figura 6. Diagrama de dispersión de los valores de F2 para todas las hablantes.

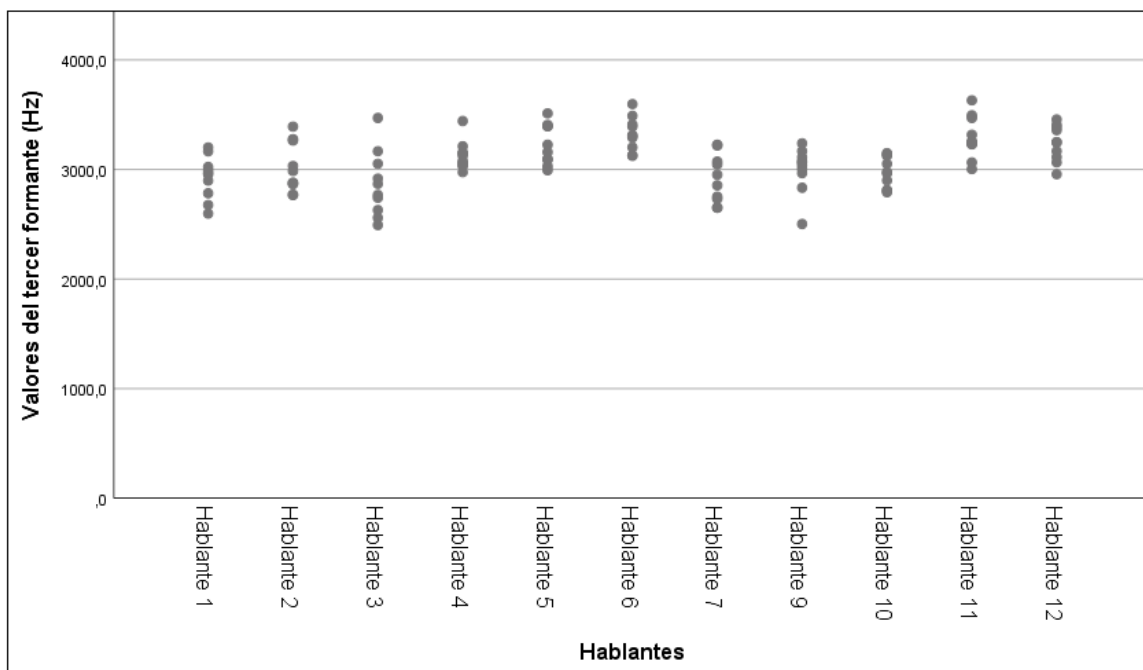


Figura 7. Diagrama de dispersión de los valores de F3 para todas las hablantes.

Como se explicó anteriormente, el análisis de la varianza (ANOVA) se empleó como método para determinar si existía variación interhablante viable para la CFV. Los resultados de este análisis en el conjunto de datos del F2, a diferencia de F1, muestran diferencias estadísticamente significativas entre mínimo dos de las hablantes, con un valor de  $p < 0,05$ . En consecuencia, se realizó un contraste *post hoc*; esto es, una comparación múltiple con el fin de determinar qué hablantes presentaban diferencias significativas entre sí. Los resultados muestran que, bajo F2 como parámetro, no todas las hablantes se diferencian. Por ejemplo, las hablantes 2, 4, 10, no tienen diferencias significativas con ninguna otra hablante. Las hablantes 5 y 12 se diferencian solamente con una de las once hablantes. Las hablantes 6, 7 y 11 son las que mejor se diferencian entre sí como se muestra en la Tabla 9.

Tabla 9

*Diferencias significativas entre hablantes, según valores de F2.*

Hablante	Diferencias significativas
Hablante 1	Hablantes 6 y 11
Hablante 3	Hablantes 6 y 11
Hablante 5	Hablante 7
Hablante 6	Hablantes 1, 3, 7 y 9
Hablante 7	Hablantes 5, 6, 11 y 12
Hablante 9	Hablantes 6 y 11
Hablante 11	Hablantes: 1, 3, 7 y 9
Hablante 12	Hablante 7

Por otra parte, se realizó la misma comparación múltiple para los datos de F3, los resultados se resumen en la Tabla 10. De acuerdo con el análisis, se registran nuevas diferencias



significativas entre hablantes y algunas de las diferencias registradas en los valores de F2 ya no aparecen. El número de hablantes que presentan diferencias se reduce de ocho, según F2, a seis. Las hablantes 2, 4, 5, 9, 12 no presentan diferencias significativas con ninguna otra.

Tabla 10

*Diferencias significativas entre hablantes, según valores de F3.*

Hablante	Diferencias significativas
Hablante 1	Hablantes: 6, 11
Hablante 3	Hablante 6
Hablante 6	Hablantes: 1, 3, 7, 10
Hablante 7	Hablantes: 6, 11
Hablante 10	Hablantes: 6, 11
Hablante 11	Hablantes: 1, 7, 10

Por último, la Figura 8 compara gráficamente las diferencias entre las hablantes según la media de cada conjunto de datos para cada formante. Se puede observar que la trayectoria formada por los valores correspondientes a F1 es casi lineal. F2 y F3 muestran un comportamiento similar y se observa la variación entre las distintas medias de cada una de las hablantes, tal como se relacionó anteriormente. En virtud de lo anterior es posible establecer, de acuerdo con los resultados del análisis de estos dos formantes, que existen particularidades en la producción de esta pausa en diferentes hablantes, con resultados superiores en comparación con F1, en términos de variación interhablante. Por otra parte, F2 y F3 difieren en algunas de las hablantes que se diferencian significativamente, aportando información complementaria, por lo que se sugiere que en conjunto estos dos formantes tienen potencial como parámetro en la CFV.

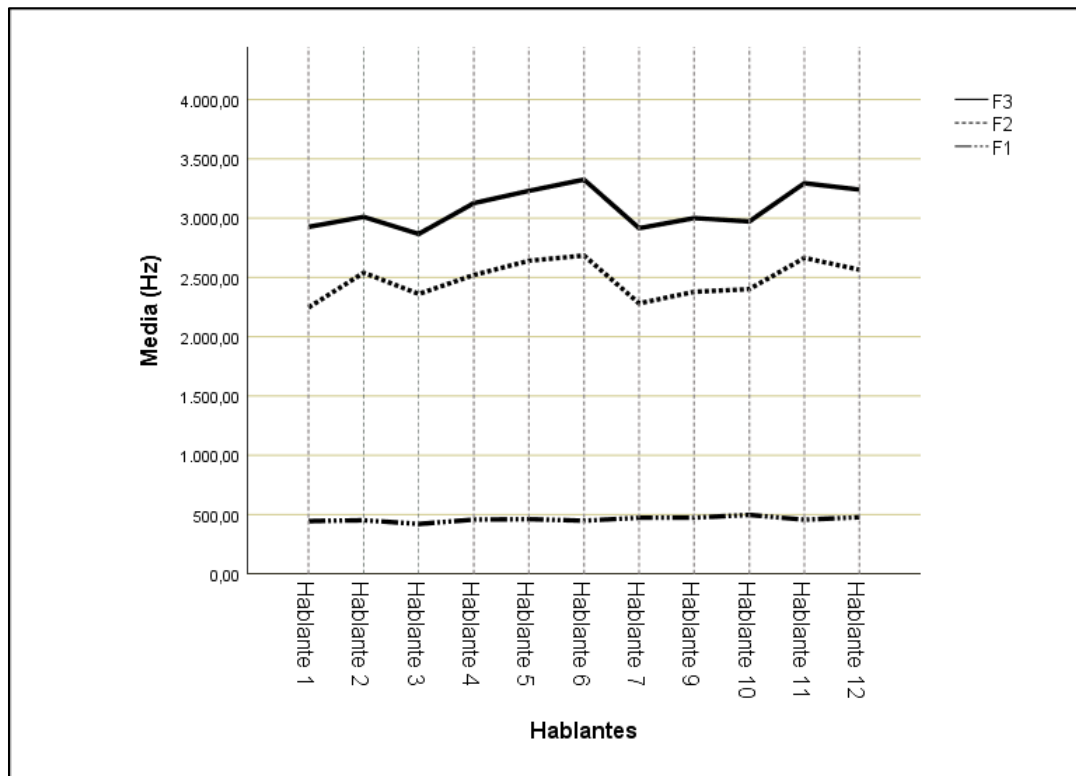


Figura 8. Gráfico comparativo de la diferencia de medias de la frecuencia central: F1 F2 y F3.

### 4.2.3 Duración

La Tabla 11 muestra los datos estadísticos descriptivos resultado del análisis de los valores de la duración medida en milisegundos (ms) en diez pausas distintas para cada hablante. Se observa el valor de la media muestral para cada hablante, la desviación estándar, los valores mínimo y máximo del conjunto de valores medios de la duración, y el rango de la media poblacional con un nivel de confianza del 95% (límite inferior y superior). Se observa una desviación estándar con tendencia elevada para todas las hablantes, lo que hace evidente que la variación interhablante no es mayor que la variación intrahablante, según la comparación del valor global de la desviación estándar (158,628) con los valores de cada hablante.

Tabla 11

*Estadísticos descriptivos de los valores de la duración en milisegundos (ms).*

Hablantes	Media	Desviación estándar	Límite Inferior	Límite superior	Valor Mínimo	Valor Máximo
Hablante 1	403,60	173,284	279,64	527,56	166	699
Hablante 2	441,90	207,106	293,75	590,05	172	855
Hablante 3	471,30	139,716	371,35	571,25	220	682
Hablante 4	436,70	164,393	319,10	554,30	276	759
Hablante 5	375,80	120,828	289,36	462,24	202	551
Hablante 6	334,40	86,650	272,41	396,39	174	509
Hablante 7	395,00	141,347	293,89	496,11	272	677
Hablante 9	390,20	197,849	248,67	531,73	169	718
Hablante 10	296,20	118,468	211,45	380,95	199	545
Hablante 11	284,00	129,072	191,67	376,33	133	476
Hablante 12	447,40	175,504	321,85	572,95	151	777
Global	388,77	158,628	358,80	418,75	133	855

La Figura 9 (abajo) complementa la información presentada en la tabla anterior. Los valores de las diferentes pausas para todas las hablantes se encuentran dispersos en un el rango de 133 ms a 855 ms en donde la mayoría de las hablantes presentan duraciones cercanas a ambos límites (valor mínimo y valor máximo). Sin embargo, algunas hablantes (5, 6, 10 y 11) no superan los 600 ms como límite superior, lo que sugiere que existen diferencias entre hablantes, aunque no significativas como se explica a continuación.

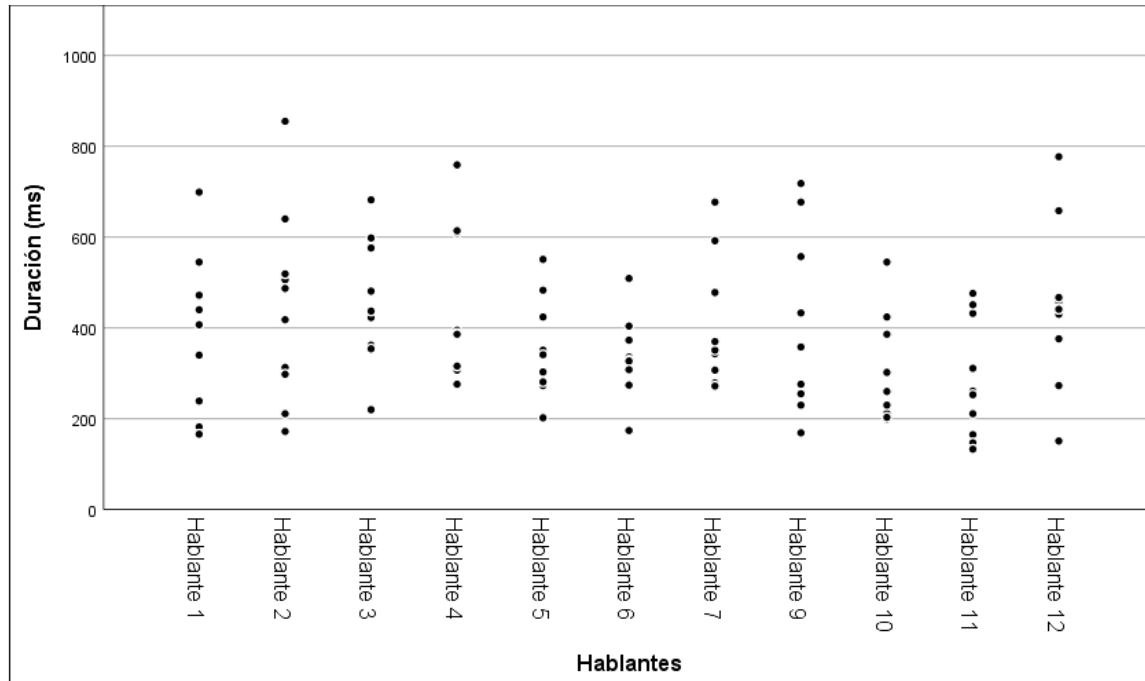


Figura 9. Diagrama de dispersión de los valores de la duración para todas las hablantes.

Debido a que para la duración como parámetro los datos no presentaron una distribución normal, se optó por emplear una prueba estadística no paramétrica para muestras independientes como alternativa a la prueba ANOVA. Los resultados complementan la información expuesta anteriormente, pues retienen la hipótesis de no hay diferencias significativas en la distribución de los valores de la duración entre de las hablantes ya que el valor de significación resultó elevado ( $0,092 > 0,05$ ). En consecuencia, la duración del segmento vocálico puede ser muy débil como parámetro en la CFV.

## 5. Discusión

En esta sección se revisarán los resultados obtenidos y aspectos de la metodología empleada, en relación con la literatura disponible. Primero, se hará un recorrido por el método de recolección de datos empleado y su viabilidad. Segundo, se presentarán y discutirán los hallazgos más importantes sobre las pausas llenas presentes en el español hablado en Bogotá. Por último, se retomará el potencial de estos segmentos en la CFV y se entreverá su proyección en futuras investigaciones.

### 5.1 Uso de Aplicaciones de MI como Fuente de Datos Lingüísticos

Para responder a los objetivos de esta investigación, se requirió contar con el mayor número de pausas llenas posible dentro de un tiempo determinado y un grupo delimitado de hablantes. Para este fin, en las investigaciones relacionadas con el área de la fonética forense usualmente se emplea un conjunto de grabaciones que simulan las características propias de un análisis forense, bien sea, los aspectos técnicos (e. g. Cao & Dellwo, 2019), o el contexto (e. g. Cicres, 2014, Hughes, Wood, & Foulkes, 2016). Sin embargo, los mismos autores sugieren emplear datos más realistas. De acuerdo con lo anterior, para esta investigación se optó por aprovechar la aplicación de mensajería instantánea (MI), WhatsApp, como fuente para la obtención de datos. En principio, bajo la hipótesis de que al tratarse de narraciones libres (habla producida sin seguir un guión), estas ofrecerían muestras más espontáneas en comparación con grabaciones simuladas; además de reducir los posibles efectos que supone grabar directamente a los participantes. Según los resultados obtenidos aquí, se destaca que este método ofrece varias ventajas, resultando efectivo para el fin propuesto. Inicialmente, hubo celeridad en la recolección de muestras puesto que no se

requiere configuración previa ni desarrollo de instrumentos. Además, los participantes no son conscientes de estar siendo grabados para propósitos investigativos, lo cual se garantizó estableciendo dentro del procedimiento la condición para los participantes de aportar únicamente audios que hubiesen sido grabados en un lapso anterior al momento de la solicitud de estos (un mes, aproximadamente). En consecuencia, empleando este método y bajo una instrucción clara, se identificó que se pueden mitigar los efectos adversos anteriormente expuestos, logrando muestras de habla con un grado de naturalidad elevado.

Además, durante la etapa de identificación y segmentación, se encontró que la mayoría de los audios provienen de conversaciones con amigos o personas cercanas, en las cuales resaltaron diversos índices estilísticos de habla informal. Por ejemplo, el uso de groserías, vocativos informales, diminutivos, etc. (e. g. «ush», «pues», «china», «paila», «polas», «marica») que son característicos de un estilo de habla casual y en consecuencia espontánea<sup>10</sup>. Aunque los anteriores elementos no se evaluaron de forma sistemática, ya que el análisis de la espontaneidad de los audios no fue el propósito de este trabajo, valdría la pena a futuro, sobre una muestra de datos más amplia, proveniente de aplicaciones de MI, realizar una evaluación enfocada en este aspecto particular (i. e. Torreira & Ernestus, 2010), de manera que se pueda determinar su potencial en otro tipo de investigaciones.

## 5.2 Pausas Llenas: Tipos y Frecuencias de Aparición

En cuanto a los tipos de pausas que se pueden evidenciar en el español, los hallazgos presentados en la sección anterior concuerdan con estudios previos. Por ejemplo, en su

---

<sup>10</sup> Para un ejemplo del empleo de este léxico en el habla casual, ver (Plazas, 2008)

investigación sobre los fenómenos de hesitación en español, Machuca, Llisterri, & Ríos (2015) anotan la aparición de, en terminos generales, cuatro tipos de pausas: un sonido vocálico [e:], un sonido que denominan como murmullo nasal [m:]<sup>11</sup>, [a:] y una última, considerada como una combinación de [e:] que finaliza con un segmento nasal [m:], [e:m:].<sup>12</sup> Respecto a la frecuencia de aparición, en el estudio antes mencionado, el segmento [e:] posee el mayor porcentaje de aparición, lo que concuerda con los resultados de esta investigación. Por el contrario, existen diferencias en las frecuencias de aparición del segmento [m:], ya que esta es considerada como la segunda pausa más frecuente, constituyendo el 32 % de un total de 156 pausas (Machuca & Rios, 2016), mientras que en el presente trabajo representó un escaso 9,2 % de un total de 389 pausas. Lo anterior puede tomarse como un indicador de la variación que puede existir entre las pausas empleadas por los hablantes entre una variedad del español y otra.

### 5.2.1 Segmento vocálico [e:]

Los estudios sobre las características acústicas de las pausas llenas en español son reducidos. Sin embargo, Machuca & Rios (2016) describen las características acústicas del segmento vocálico [e:], en donde, al comparar los resultados obtenidos en dicha investigación con los obtenidos aquí, observamos coincidencia en los valores de la frecuencia central de F1, los cuales, en ambos, se mantienen en un promedio de 460 Hz. Por otra parte, los valores de F2, en

---

<sup>11</sup> En Machuca, Llisterri, & Ríos (2015), esta pausa se presenta como *murmullo nasal*; sin embargo, en este estudio se optó por la transcripción esta transcripción para conservar la uniformidad.

<sup>12</sup> Lo mismo que el anterior.

esta investigación, son considerablemente más altos a los reportados por Machuca & Ríos (2016). Por último, en cuanto a la duración, hay coincidencia en que la dispersión para este parámetro es elevada para todos los hablantes.

### **5.2.2 Segmento compuesto [e:m:]**

Retomando los tipos de pausas llenas hallados, surgen algunas consideraciones sobre la aparición del segmento compuesto por el sonido vocálico [e:] y el segmento nasal [m:]. Se encuentra en la literatura que en el español «los fonemas /e/ y /m/ son los únicos que pueden constituir verdaderas pausas sonoras, ya que pueden insertarse en el discurso sin que necesariamente deriven del alargamiento de una palabra» (Rebollo, 1997, como se citó en Machuca, Llisterra, & Ríos, 2015). Sin embargo, el segmento [e:m:] cumple con tales características y en este estudio constituyó el 28,19 % de las pausas totales, superando considerablemente a la pausa [m:], lo que sugiere que se debería considerar como una pausa llena del español en sí misma. Además, este segmento abre la posibilidad de realizar investigaciones con nuevos parámetros como las dinámicas de formantes, que muestran excelentes resultados en el ámbito forense, según trabajos como el de McDougall, (2006), Hughes, Wood, & Foulkes (2016), San Segundo & Yang (2019), entre otros.

## **5.3 Pausas Llenas Como Parámetro en la CFV**

Investigaciones previas sustentan que las pausas llenas poseen potencial como variables en la aplicación de la CFV en algunas lenguas como inglés (Hughes, Wood, & Foulkes, 2016) y catalán (Cicres, 2014). Este proyecto apuntó a explorar si en el español se obtendrían los mismos



resultados, analizando las características acústicas (puntos medios de F1, F2, F3 y la duración) de la pausa llena más frecuente [e:]. Los resultados mostraron variación tanto a nivel intrahablante como interhablante, en donde se destacan los valores del punto de medio en F2 y F3 pues la variación interhablante, según estos dos parámetros, resultó significativa, lo cual sugiere que también en español las pausas llenas tienen potencial en la CFV. Sin embargo, se debe tener en cuenta que el presente trabajo es precursor, por lo que se requieren investigaciones complementarias en diferentes aspectos como los expuestos a continuación.

Por una parte, es necesario aumentar el conjunto de datos, replicando la investigación con el fin de evaluar si los resultados observados siguen el mismo patrón en un grupo de hablantes más amplio y en consecuencia un número mayor de pausas llenas. Además, incluir nuevos parámetros acústicos para determinar si también presentan potencial en la CFV, ya que usualmente se evalúan las características de los primeros tres formantes, en los segmentos vocálicos, como señalan Cao & Dellwo (2019). Ahora bien, como se mencionó anteriormente, las tecnologías de comunicación actuales ofrecen alternativas que facilitan emprender investigaciones sobre otros parámetros como los formantes más altos, F4 y F5, lo que complementarían y aportarían a la discusión sobre si estos contienen mayor información específica del hablante (Cao & Dellwo, 2019). Asimismo, retomando la contextualización hecha en la sección 2.2.3 sobre los métodos de CFV, se esperaría poder evaluar el desempeño de datos que poseen potencial en la CFV, a través de pruebas como el cálculo de la razón de verosimilitud (LR), método que ha sido ampliamente usado con fines de investigación forense (Rose, 2002, Hughes, 2017). Esto último no hizo parte de los objetivos de este trabajo dado que se debe evitar aplicar estas pruebas en grupos pequeños de hablantes (Hughes, 2017). Retomando el hecho de que esta investigación fue de carácter exploratorio, el corpus analizado fue limitado, por ende, no se aplicó dicho análisis. En virtud de lo anterior, para emprender un estudio

que incluya este tipo de prueba, se requiere contar con información estadística de la población en general, sobre las características acústicas de cada parámetro a analizar (e. g. Hudson, et al, 2007), que suplan el componente de tipicidad necesario en el cálculo de la razón de verosimilitud.

Finalmente, aunque el interés de esta investigación fue fundamentalmente acústico, Hughes, Wood, & Foulkes (2016) anotan que los factores sociolingüísticos pueden delimitar el rango de variación de las pausas llenas. De acuerdo con lo anterior, se recomienda, para futuras investigaciones, incluir encuestas y/o cuestionarios sociolingüísticos que permitan conocer aspectos acerca del perfil social de los hablantes para evaluar y explicar, si es el caso, la variación relacionada con características sociales, ya que esta fue una de nuestras limitaciones.

## 6. Conclusiones

Este estudio amplía la información existente acerca las pausas llenas en español, especialmente sobre sus características acústicas. Además, establece un punto de partida para la investigación de estos fenómenos en otras variedades de español y específicamente en el español que se habla en Colombia.

Por una parte, los resultados obtenidos permiten concluir que existen similitudes en los tipos de pausas encontradas, en comparación con investigaciones anteriores. Por el contrario, el orden según las frecuencias de aparición para cada tipo de pausa cambia. En este sentido, si bien existen estudios que no tienen en cuenta [e:m:] como una pausa llena, según la presente investigación, [e:m:] sí debería considerarse como tal, ya que cumple con las mismas características y su frecuencia de aparición es suficientemente alta. También, se concluye que, en esta lengua, las pausas llenas demuestran potencial como variable en la CFV, según el análisis de su estructura formántica. No obstante, se considerarían potenciales principalmente los valores provenientes de la frecuencia central de F2 y F3, ya que en el caso de F1 y las medidas de la duración, la variación intrahablante es elevada, impidiendo anotar diferencias entre los hablantes.

Otro aporte que se destaca de este estudio, es la inclusión de un método de recolección de datos que se ha empezado a explorar recientemente y puede considerarse novedoso, esto es, el uso de las aplicaciones de MI como fuente de datos lingüísticos. A través de este recurso se obtuvieron muestras de habla prometedoras en términos de los índices de espontaneidad que estas pueden poseer. Por lo anterior, se recomienda su uso en futuras investigaciones.

Finalmente, una de las principales dificultades en el desarrollo de esta investigación fue la poca, o nula, documentación sobre el estudio de las pausas llenas empleadas en el español hablado

en Colombia. Por lo anterior, no es posible hacer una comparación o análisis de los resultados más allá de la información obtenida aquí. Sin embargo, lo anterior representa, a su vez, un foco de investigación amplio para los interesados en el desarrollo de la lingüística forense en Colombia, especialmente, en la fonética, así como también en el estudio de los fenómenos de hesitación.

# Anexo A. Consentimiento Informado y Hoja de Información

## HOJA DE INFORMACIÓN

### PROYECTO DE INVESTIGACIÓN

*Análisis acústico de las pausas llenas como parámetro en la comparación forense de voces*

#### Información general

Formalmente lo invitamos a participar en un proyecto de investigación de la Universidad Nacional de Colombia. Antes de decidir si desea participar, es importante que usted entienda por qué se está realizando esta investigación y lo que involucra la misma. Por favor, lea cuidadosamente la siguiente información y no dude en consultar si hay algún punto que no esté totalmente claro o si desea más información. De antemano le agradecemos por tomarse el tiempo para leer este documento.

#### ¿Cuál es el propósito de este proyecto?

Investigar el potencial de las pausas llenas como parámetro en la comparación forense de voces en una variedad de español hablada en Bogotá.

#### ¿Por qué he sido escogido?

Para desarrollar el proyecto se debe contar con un grupo de participantes que posea determinadas características (edad, sexo, residir en Bogotá, entre otras) con las cuales usted cumple. Depende de usted decidir si desea o no participar. Si decide participar, se le entregará esta hoja de información para que la conserve y se le pedirá que firme un formulario de consentimiento. Si decide participar, usted está en libertad de retirarse en cualquier momento y no tendrá que dar ninguna justificación.

#### ¿Qué pasará si decido participar?

Se le solicitará compartir con el investigador una serie de mensajes de audio grabados en aplicaciones de mensajería instantánea (WhatsApp). Dichos audios corresponden a conversaciones que haya sostenido recientemente (hasta seis meses de antigüedad) y tengan una duración considerable (30 segundos en adelante).

#### ¿Qué pasa cuando la investigación finaliza?

Si el estudio termina antes de haber sido completado en su totalidad, usted será notificado del porqué. Una vez el proceso de recolección de datos haya finalizado, se procederá al análisis e interpretación de los datos. La investigación será presentada formalmente a la comunidad académica y otros profesionales. Usted podrá ponerse en contacto con nosotros en cualquier momento una vez haya finalizado la recolección de datos. Lo anterior con el fin de resolver cualquier inquietud relacionada con su participación.

#### ¿Mi participación en este proyecto se mantendrá confidencial?

Toda la información que se recoja sobre usted durante el curso de la investigación se mantendrá estrictamente confidencial. Cualquier información sobre usted que se difunda estará en total anonimato de manera que usted no podrá ser identificado a través de ella.

#### ¿Qué pasará con los resultados del proyecto de investigación?

Los resultados podrán ser compartidos con la comunidad académica en general y profesionales pertinentes a través de artículos publicados en revistas, ponencias o comunicaciones, todas estas de índole académica.

#### ¿Quién está organizando y financiando la investigación?

Esta investigación es llevada a cabo por el estudiante de Lingüística Luis Felipe Bautista Peña bajo la dirección de John Freddy Chaparro Rojas, adscrito al Departamento de Lingüística de la Universidad Nacional de Colombia.

#### Contacto para más información:

**John Freddy Chaparro Rojas**

**Luis Felipe Bautista Peña**

**FORMULARIO DE CONSENTIMIENTO INFORMADO PARA PARTICIPAR EN EL PROYECTO****Análisis acústico de las pausas llenas como parámetro en la comparación forense de voces**

	Ponga sus iniciales al lado del enunciado si está de acuerdo
Confirmando que el día ___/___/___ he leído y entendido la <b>Hoja de Información</b> explicando el proyecto de investigación y he tenido la oportunidad de hacer preguntas sobre el mismo.	
Entiendo que mi participación es voluntaria y que soy libre de retirarme en cualquier momento sin dar ninguna razón alguna y sin que haya ninguna consecuencia. Si no deseo responder a alguna o varias preguntas, estoy en libertad de rechazarla(s).  <b>NOTA:</b> En caso de decidir retirarse, una vez cumplido este proceso, todos los datos electrónicos que se tengan se eliminarán de forma permanente y las copias impresas serán destruidas.	
Doy permiso para que los miembros del equipo de investigación tengan acceso a mis respuestas anónimas. Entiendo que mi nombre no estará vinculado con los materiales de investigación, y no voy a ser identificado o identificable en los informes que resulten de la investigación. Entiendo que mis respuestas serán estrictamente confidenciales.	
Estoy de acuerdo con que los datos recolectados sobre mí sean utilizados de forma anónima en futuras investigaciones pertinentes.	
Estoy de acuerdo con mi participación en el proyecto de investigación arriba mencionado e informaré al investigador principal, si mis datos de contacto cambian.	
Confirmando que estoy de acuerdo con asignar los derechos de autor* que tengo sobre cualquiera de los materiales relacionados con este proyecto a Luis Felipe Bautista Peña.	

<b>Nombre del participante</b>	
<b>Firma del participante</b>	
<b>Fecha</b>	
<b>Nombre del investigador principal</b>	
<b>Nombre de quien llevó a cabo la recolección de datos</b>	
<b>Firma**</b>	
<b>Fecha**</b>	

\*Con el fin de poder usar de manera legal la información que usted ha entregado.

\*\*Para ser firmado y fechado en presencia del participante.

Por favor, indique su correo electrónico a continuación para que podamos enviarle una copia del presente documento.

## Referencias

- Audacity Team. (2020). Audacity® Cross-Platform Sound Editor [Programa] Versión 2.4.2.
- Boersma, P., & Weenik, D. (2021). PRAAT: doing phonetics by computer [Programa] Verisón 6.1.41.
- Cao, H., & Dellwo, V. (2019). The role of the first five formants in three vowels of mandarin for. *International Congress of Phonetic Sciences*, 617-621.
- Cicres, J. (2014). Comparación forense de voces mediante el análisis multidimensional de las pausas llenas. *Revista Signos. Estudios de Lingüística*, 365-384.
- Coulthard, M., & Johnson, A. (2010). *The Routledge handbook of forensic linguistics*. Abingdon: Routledge.
- Coulthard, M., Johnson, A., & Wright, D. (2017). *An introduction to forensic linguistics: Language in evidence*. Abingdon: Routledge.
- De Gregorio, J. (2020). *Contribuciones al análisis forense de evidencias digitales procedentes de aplicaciones de mensajería instantánea (Tesis doctoral)*. Universidad de Alcalá: Alcalá de Henares.
- Duckworth, M., McDougall, K., de Jong, G., & Shockey, L. (2011). Improving the consistency of formant measurement. *The International Journal of Speech, Language and Law*, 35-51.
- Gil, J. (2007). *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco .
- Hollien, H. (2013). *The Acoustics of Crime: The New Science of Forensic Phonetics*. New York: Springer Science & Business Media.
- Hudson, T., De Jong, G., McDougall, K., Harrison, P., & Nolan, F. (2007). F0 statistics for 100 young male speakers of Standard Southern British English. *Proceedings of the 16th international congress of phonetic sciences*.
- Hughes, V. (2017). Sample size and the multivariate kernel density likelihood ratio: how many speakers are enough? *Speech Communication.*, 15-29.
- Hughes, V., Wood, S., & Foulkes, P. (2016). Strength of forensic voice comparison evidence from the acoustics of filled pauses. *International Journal Of Speech Language And The Law*, 99-132.
- IBM. (2020). SPSS: Statistical Package for the Social Sciences [Programa] Versión 25.0.
- Jessen , M. (2008). Forensic phonetics. *Language and linguistics compass*, 671-711.
- Jessen, M. (2010). The forensic phonetician. In M. Coulthard, & A. Johnson, *The Routledge handbook of forensic linguistics* (pp. 378 - 394). Abingdon: Taylor & Francis .
- Kaplan, E., & Kaplan, M. (2006). *Chances are... Adventures in probability*. Nueva York: Penguin.
- Machuca, M. J., Llisterri, J., & Ríos, A. (2015). Las pausas sonoras y los alargamientos en español: un estudio preliminar. *Normas: revista de estudios lingüísticos hispánicos*, 81-96.

- Machuca, M., & Ríos, A. (2016). Estructura formántica de las pausas sonoras en español. In A. M. Fernández Planas, *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística* (pp. 67-76). Barcelona: Universitat de Barcelona.
- Machuca, M., Ríos, A., & Llisterri, J. (2014). Conocimiento fonético y fonética judicial. *Quaderns de Filologia: Estudis Lingüístics*, 95-111.
- Mahajan, A., Dahiya, M. S., & Sanghvi, H. P. (2013). Forensic analysis of instant messenger applications on android devices. *International Journal of Computer Applications*, 38-44.
- McDougall, K. (2006). Dynamic features of speech and the characterization of speakers: towards a new approach using formant frequencies. *Speech, Language and the Law*, 89-126.
- McDougall, K., & Duckworth, M. (2017). Profiling fluency: An analysis of individual variation in disfluencies in adult males. *Speech Communication*, 95, 16-27.
- Olsson, J. (2008). *Forensic Linguistics*. London: Continuum.
- Orebaugh, A., & Allnut, J. (2009). Classification of instant messaging communications for forensics analysis. *The International Journal of Forensic Computer Science*, 22-28.
- Plazas, A. (2008). El habla de los jóvenes: reflexiones sociolingüísticas.
- Real Academia Española y Asociación de Academias de la Lengua Española. (2011). *Nueva gramática de la lengua española: fonética y fonología*. Espasa Libros.
- Rose, P. (2002). *Forensic Speaker Identification*. London: Tylor & Francis.
- San Segundo, E., & Yang, J. (2019). Formant dynamics of Spanish vocalic sequences in related speakers: A forensic-voice-comparison investigation. *Journal of Phonetics*, 1-26.
- Soto, J. (2020). *Whatsapp: como medio de prueba en el procedimiento penal (Tesis de pregrado)*. Universidad Miguel Hernández: Elche.
- Torreira, F., & Ernestus, M. (2010). The Nijmegen Corpus of Casual Spanish. *Seventh conference on International Language Resources and Evaluation*, 2981-2985.
- Villa, J., Gil, J., & Lahoz-Bengoechea, J. (2017). Las vocales de relleno en español: nuevos datos y algunas reflexiones. In E. L. Mirayes, M. R. Silva, & A. M. Alvarado., *Nuevos estudios sobre comunicación social*. (pp. 165-169). La Habana: Centro de Lingüística Aplicada.