



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
COLEGIO DE LETRAS HISPÁNICAS

**ANÁLISIS DE LA FRECUENCIA FUNDAMENTAL DE LA VOCAL MEDIA
ANTERIOR, ANTES Y DESPUÉS DE AUMENTAR EL RITMO CARDIACO.
UN ESTUDIO PARA SU APLICACIÓN EN EL ÁMBITO FORENSE**

TESIS QUE, PARA OBTENER EL TÍTULO DE
LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS,
PRESENTA:

MARÍA DEL ROSARIO MARTÍNEZ ROMERO

Asesora: Dra. Fernanda López Escobedo

Ciudad de México, 2014





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis papás,
por apoyarme incondicionalmente en cada paso que doy.

Agradecimientos

Antes que nada, quiero agradecer a la Dra. Fernanda López Escobedo por su guía en la realización de este trabajo, porque sin su apoyo, paciencia y constancia esta tesis no hubiera llegado a su fin. Además de ser mi asesora, fue ella, quien me ayudó a descubrir y a definir mis intereses académicos. Fernanda: gracias por haber confiado en mí.

Agradezco a la Mtra. Gloria Estela Báez Pinal, al Mtro. Javier Cuétara Priede, a la Mtra. Teresita Adriana Reyes Careaga y al Mtro. Octavio Augusto Sánchez Velázquez por haber leído este trabajo y por sus valiosas aportaciones.

En especial quiero agradecer a la Mtra. Gloria no sólo por haberme dado la oportunidad de aprender más en el campo de la docencia, sino también por brindarme su confianza y amistad.

A mis padres debo agradecerles porque siempre me motivaron a seguir adelante y me ayudaron a crecer. Mamá, gracias por tu apoyo en cada momento de mi vida, por cuidarme, por tus sabios consejos y acertados regaños, por enseñarme que a pesar de todo, lo importante en la vida es ser feliz. Papá, te agradezco por alentarme a seguir con mis sueños cuando más lo necesité, por tus abrazos, por ser mi padre.

A mis hermanas, porque en todo momento me han brindado su apoyo, con una palabra cálida o con un abrazo. Porque sé que siempre estarán ahí.

A Arturo, porque eres un combo maravilloso, eres: mi mejor amigo, mi apoyo, mi confidente, mi todo. Tú me inspiras a ser una mejor persona y a dar lo mejor de mí. Gracias por el inmenso amor que me demuestras día tras día.

A mis amigos, gracias por ayudarme a crecer, a madurar y por darme todo su cariño. Siempre están en mi corazón.

Asimismo, quiero agradecer a cada una de las personas que amablemente prestaron su voz, sin importar el esfuerzo físico al que tuvieron que someterse. A todos ellos: muchas gracias.

Agradezco al Grupo de Ingeniería Lingüística, que me brindó el espacio idóneo para realizar este trabajo y me permitió conocer otra cara de la lingüística.

Por supuesto, debo agradecer a la Universidad Nacional Autónoma de México, mi segunda casa, porque me ha brindado la oportunidad de tener una excelente educación.

Agradezco al Consejo Nacional de Ciencia y Tecnología por la beca que me otorgó para la elaboración de esta tesis al ser becaria en el proyecto de investigación con número de referencia 178248.

De igual forma agradezco a la Secretaria de Educación Pública por proporcionarme una Beca de Titulación dentro del Programa de Becas para Educación Superior.

ÍNDICE

| | |
|---|----|
| 1. Introducción | 1 |
| 1.1 Áreas de la fonética forense | 1 |
| 1.2 Desarrollo de la fonética forense en el mundo | 2 |
| 1.3 La fonética forense en México | 3 |
| 1.4 Objetivos | 4 |
| 1.5 Hipótesis..... | 5 |
| 2. Métodos utilizados en fonética forense..... | 7 |
| 2.1 El método Voiceprint..... | 7 |
| 2.2 El método Auditory-phonetic | 9 |
| 2.3 El método Ear-witness Line-ups | 9 |
| 2.4 Métodos semiautomáticos y automáticos | 11 |
| 2.5 Métodos utilizados actualmente en fonética forense..... | 13 |
| 2.6 Método utilizado en esta tesis | 14 |
| 3. La respiración y su relación con el habla | 16 |
| 3.1 La respiración | 16 |
| 3.2 La producción del habla y su relación con la respiración..... | 17 |
| 3.3 Frecuencia fundamental..... | 18 |
| 3.4 Factores que afectan la frecuencia fundamental (F0) | 20 |
| 3.5 Factores que afectan la producción del habla..... | 21 |
| 3.6 El estrés y su relación con la producción del habla | 23 |
| 3.7 Estudios de F0 en fonética forense | 25 |
| 4. Metodología | 29 |
| 4.1 Corpus | 29 |
| 4.1.1 Composición del corpus | 31 |
| 4.1.2 Recopilación del corpus | 35 |
| 4.2 Análisis acústico..... | 37 |
| 4.3 Análisis estadístico..... | 39 |
| 4.3.1 Vaciado de datos | 39 |
| 4.3.2 Contraste de hipótesis..... | 40 |
| 4.3.3 Supuestos de la prueba <i>T.test</i> | 41 |
| 4.3.4 Prueba <i>T.test</i> | 42 |
| 4.3.5 Análisis descriptivo | 43 |

| | |
|--|----|
| 4.3.6 Diagrama de caja o <i>boxplot</i> | 43 |
| 5. Resultados | 46 |
| 5.1 Resultados por género..... | 47 |
| 5.1.1 Resultados de los hablantes masculinos | 47 |
| 5.1.2 Resultados de los hablantes femeninos | 48 |
| 5.2 Resultados por tonicidad..... | 49 |
| 5.2.1 Resultados por tonicidad. Hablantes masculinos | 49 |
| 5.2.2 Resultados por tonicidad. Hablantes femeninos..... | 50 |
| 5.3 Resultados por contexto | 52 |
| 5.3.1 Resultados por contexto. Hablantes masculinos | 52 |
| 5.3.2 Resultados por contexto. Hablantes femeninos..... | 54 |
| 5.4 Resultados individuales..... | 56 |
| 5.4.1 Resultados individuales. Hablantes masculinos | 56 |
| 5.4.2 Resultados individuales. Hablantes femeninos | 59 |
| 6. Conclusiones | 61 |
| Bibliografía | 64 |

ÍNDICE DE TABLAS

4. METODOLOGÍA

| | |
|---|----|
| Tabla 4.1 Condiciones para la realización prototípica de los fonemas alveolares | 33 |
| Tabla 4.2 Contextos analizados para la vocal átona..... | 34 |
| Tabla 4.3 Contextos analizados para la vocal tónica..... | 34 |
| Tabla 4.4 Datos de los hablantes femeninos | 35 |
| Tabla 4.5 Datos de los hablantes masculinos | 36 |
| Tabla 4.6 Ejemplo de organización de datos con realización átona..... | 39 |
| Tabla 4.7 Ejemplo de organización de datos con realización tónica..... | 40 |

5. RESULTADOS

| | |
|--|----|
| Tabla 5.1 Resultados generales: hablantes masculinos | 47 |
| Tabla 5.2 Resultados generales: hablantes femeninos | 48 |
| Tabla 5.3 Resultados por realización: hablantes masculinos | 49 |
| Tabla 5.4 Resultados por realización: hablantes femeninos..... | 51 |
| Tabla 5.5 Resultados por contexto: hablantes masculinos | 53 |
| Tabla 5.6 Resultados por contexto: hablantes femeninos | 54 |
| Tabla 5.7 Resultados por hablante masculino..... | 57 |
| Tabla 5.8 Resultados por hablante femenino | 59 |

ÍNDICE DE FIGURAS

3. LA RESPIRACIÓN Y SU RELACIÓN CON EL HABLA

| | |
|---|----|
| Figura 3.1 Representación de la onda sonora simple y compleja..... | 19 |
| Figura 3.2 Variaciones de la frecuencia fundamental en función de la edad | 20 |

4. METODOLOGÍA

| | |
|---|----|
| Figura 4.1 Etiquetado de /ne/ Tónica | 38 |
| Figura 4.2 Etiquetado que muestra el punto central de la vocal..... | 38 |
| Figura 4.3 Diagrama de caja o <i>boxplot</i> | 44 |

5. RESULTADOS

| | |
|---|----|
| Figura 5.1 <i>Boxplot</i> de los hablantes femeninos y masculinos de ambas realizaciones | 48 |
| Figura 5.2 <i>Boxplot</i> de la realización átona de tres hombres con diferencia significativa en F0..... | 50 |
| Figura 5.3 <i>Boxplot</i> de mujeres que sí presentaron variación en la F0 en la realización átona | 51 |
| Figura 5.4 <i>Boxplot</i> de mujeres que no presentaron variación en la F0, en la realización tónica..... | 52 |
| Figura 5.5 <i>Boxplot</i> de los contextos tónicos que presentaron variación en la F0..... | 53 |
| Figura 5.6 <i>Boxplot</i> de contextos átonos de hablantes femeninos que tuvieron variación en la F0 .. | 55 |
| Figura 5.7 <i>Boxplot</i> de /le/ átono y tónico que no presentó variación | 56 |
| Figura 5.8 <i>Boxplot</i> de H2 en contextos con variación en la F0..... | 58 |
| Figura 5.9 <i>Boxplot</i> de M1 con variación en la F0 | 60 |

1. INTRODUCCIÓN

Existen varias aplicaciones de la fonética, un ejemplo es el ámbito de las tecnologías del habla, en donde se ha trabajado con dos vertientes: la síntesis del habla y el reconocimiento automático de ésta. El objetivo de la primera es generar mensajes orales a partir de un texto escrito. Por otra parte, en el reconocimiento automático del habla se busca “transformar la señal sonora –el habla- en su correspondiente representación simbólica, que en general, será un texto escrito” (Llisterri, 2003: 12).

Entre las aplicaciones de la fonética se encuentra la forense, ésta es relativamente nueva en México, no obstante, ha tenido un desarrollo constante en países como Estados Unidos, Inglaterra, Australia y España.

En general, la lingüística forense se puede definir como la interfaz entre el lenguaje y el derecho¹. Helen Fraser² en su sitio *web* nos dice que la fonética forense “*is the application of expertise in phonetics to assist in legal and law enforcement contexts, and the extension of research on phonetics to topics relevant to the legal system*”.

Por otra parte Joaquim Llisterri hace una descripción más detallada sobre el área:

La fonética judicial o fonética “forense”. Comparación entre los rasgos fonéticos de un locutor “sospechoso” y un locutor indubitado: estudio de la variación fonética interlocutor y de la variación fonética intralocutor en el nivel segmental y en el nivel suprasegmental. Comparaciones auditivas, comparaciones acústicas y comparaciones automáticas mediante técnicas estadísticas. Caracterización de un locutor desconocido a partir de sus rasgos fonéticos (2004: 2).

1.1 Áreas de la fonética forense

Existen diversas áreas de estudio dentro de la fonética forense, entre las que se encuentran la identificación de hablantes, construcción de perfiles lingüísticos, autenticación de grabaciones y evaluación de reconocimiento de hablantes por parte de testigos auditivos (French, 1994: 169). También, dentro del área se inserta el *Linguistic origin analysis* en donde el investigador tiene que recurrir a los recursos lingüísticos que posee para poder determinar si una persona pertenece o no a una comunidad específica. Generalmente este

¹ <http://www.iula.upf.edu/forensiclab/fpreses.htm>

² <http://helenfraser.com.au/forensic/>

tipo de casos se dan cuando alguien solicita asilo político y el país que ampara al sujeto, necesita constatar que el hablante realmente pertenece a determinada comunidad lingüística o si está mintiendo para ingresar al país con otros fines.

La dialectología es otra de las áreas que se trabajan en fonética forense, aquí el fonetista tiene que determinar a qué dialecto pertenece un hablante que haya sido acusado por algún delito y la grabación que le imputan pudiera pertenecer a un dialecto distinto al suyo.

Como vemos, en cada una de las áreas el papel del fonetista es fundamental porque “se convierte en el primordial y último eslabón frente a la sustancia fonética al tener que aplicar en ella todo lo que sabe de acústica, de percepción auditiva, de fonopragmática, etc.” (Rosas y Sommerhoff, 2010: 107)

1.2 Desarrollo de la fonética forense en el mundo

En este punto hablaremos de las asociaciones, grupos y laboratorios que se dedican a la investigación sobre fonética forense en el nivel internacional. Más adelante trataremos el desarrollo de esta área en México.

Actualmente la fonética forense ha tenido un desarrollo considerable en diversos países, muestra de ello es el *Expert Working Group Forensic Speech and Audio Analysis (FSAAWG)*³ en la *European Network of Forensic Science Institutes (ENFSI)*. En este grupo colaboran laboratorios de distintas universidades de más de 20 países entre los que se encuentran Estados Unidos, Reino Unido, Francia y España. Dentro de los objetivos que tiene el *ENFSI* se encuentran el intercambio de información y apoyo en cuestiones técnicas, también promueven ejercicios colaborativos y talleres de capacitación.

La *International Association of Forensic Linguists (IAFL)*, es otra de las asociaciones dedicadas al estudio de la lingüística forense, ésta cuenta con la revista *The International Journal of Speech, Language and the Law (IJSLL)*, en donde dos veces al año se publican artículos sobre aspectos relacionados con el lenguaje forense, así como el discurso y el análisis de audio. La edición de esta revista es compartida con la *International*

³ <http://www.enfsi.eu/about-enfsi/structure/working-groups/speech-and-audio>

Association for Forensic Phonetics and Acoustics (IAFPA), que fue fundada en 1991 en York, Inglaterra.

En Australia se encuentra la *Australasian Speech and Technology Association (ASSTA)*, dedicada al estudio del habla, de la cual se desprende el *Forensic Speech Science Committee (FSSC)*,⁴ grupo especialmente interesado en la comparación forense de voz, que surgió en 1996.

En la universidad *Pompeu Fabra* se encuentra el *Forensic Lab*,⁵ en este laboratorio se dedican a desarrollar actividades formativas y de investigación en lingüística forense, asimismo, cuentan con un servicio de peritaje lingüístico.

En el nivel internacional estas son las asociaciones, que pudimos localizar, dedicadas a la investigación en fonética forense. Muchos de los investigadores que trabajan en esta área se encuentran integrados a estos grupos. Por supuesto, existen varios organismos que desarrollan estudios fonéticos o fonológicos pero como hemos visto son pocos los que se dedican específicamente a esta área.

1.3 La fonética forense en México

Desafortunadamente el desarrollo de la fonética forense en México ha sido muy lento. Una de las instituciones que ha sido partícipe de dicho desarrollo es el Instituto Politécnico Nacional (IPN) de donde surgió la Sociedad Mexicana de Acústica Forense (SMAF)⁶

Otra institución que cuenta con servicios de peritaje en acústica forense es la Procuraduría General de la República (PGR), entre sus especialidades periciales mencionan el Análisis de Voz: “Identificación de una víctima o de presuntos responsables de delitos federales a través de la voz... utilizando técnicas sofisticadas que permiten registrar y cotejar características acústicas y biométricas de la voz como frecuencia, intensidad, timbre, tonalidad y fisiológicas del locutor”.⁷

El Instituto Nacional de Ciencias Penales (INACIPE)⁸ cuenta con un posgrado en criminalística en donde se imparte la materia Audio, Video y Acústica Forense.

⁴ <http://www.assta.org/?q=assta-forensic-speech-science-committee>

⁵ <http://www.iula.upf.edu/forensiclab/fpreses.htm>

⁶ <http://www.somexacusforen.org/>

⁷ <http://www.pgr.gob.mx/>

⁸ <http://www.inacipe.gob.mx/>

Otro lugar en donde se estudia la fonética/ acústica forense es en el Instituto Jalisciense de Ciencias Forenses,⁹ que cuenta con un laboratorio de acústica forense.

En la UNAM, el Grupo de Ingeniería Lingüística ha impulsado el área de lingüística forense a través de la organización de foros académicos y el desarrollo de proyectos de investigación. Asimismo, resulta importante mencionar la asignatura de lingüística forense que se ofrece en la Facultad de Filosofía y Letras de la UNAM ya que de ahí surgió la motivación para realizar este trabajo de tesis.

Asimismo, en la licenciatura en Ciencia Forense de la UNAM se impartirá la asignatura optativa Acústica forense.

Una vez que hemos visto la situación actual de la fonética forense en México, podemos afirmar que existe la necesidad de que se dediquen estudios al área para de esta manera contribuir a acelerar su desarrollo en nuestro país.

1.4 Objetivos

El objetivo general de esta investigación es saber si la frecuencia fundamental (F0) de la vocal media anterior, varía o no, después de aumentar el ritmo cardiaco mediante la actividad física. Este objetivo surge para contribuir a los análisis que se realizan en comparación forense de voz, donde la grabación indubitada (en la que sí conocemos al emisor) no se hace bajo las mismas condiciones en las que se encuentra la grabación dubitada (en la que no conocemos al emisor). En particular, nos interesa saber si el estrés, que en este caso simularemos mediante la actividad física, es un factor que altera la frecuencia fundamental y que por lo tanto, puede afectar la comparación forense de voz.

Para poder cumplir con el objetivo general hemos planteado una serie de objetivos particulares, mismos que enunciaremos a continuación.

- Analizar la frecuencia fundamental de la vocal media anterior en sílaba abierta precedida por consonante alveolar.¹⁰
- Simular una situación de estrés mediante la alteración del ritmo cardiaco de los hablantes, para ello se les pedirá que salten la cuerda.

⁹ <http://cienciasforenses.jalisco.gob.mx/laboratorioacustica.php>

¹⁰ La razón por la que elegimos esta vocal en este contexto es porque tanto la vocal media anterior como las

- Grabar a los hablantes en un estado de reposo y también, después de que salten la cuerda, esto con el objetivo de medir la frecuencia fundamental y ver si existen diferencias significativas.
- Analizar los datos con una técnica estadística para obtener los resultados que nos permitan dilucidar si las hipótesis, que a continuación presentaremos, se cumplen o no.

1.5 Hipótesis

En este trabajo manejaremos una serie de hipótesis, que deberán ser aceptadas o rechazadas a partir del análisis estadístico de nuestros datos. A continuación presentamos cada una de ellas, se espera:

- Que la frecuencia fundamental de la vocal media anterior presente cambios después de aumentar el ritmo cardiaco.
- Que la tonicidad de la vocal sea un factor que afecte de manera distinta a la frecuencia fundamental después de someter a los hablantes a un estado de estrés, mismo que simularemos mediante la alteración del ritmo cardiaco. En particular, se espera que las vocales átonas presenten menos diferencias que las tónicas.
- Que el género sea un factor que afecte de manera distinta a la frecuencia fundamental después de aumentar el ritmo cardiaco. Es decir que existan diferencias entre los resultados obtenidos en los hombres y en las mujeres.
- Encontrar un contexto alveolar en el que la frecuencia fundamental presente menos diferencias entre el antes y el después de saltar la cuerda.

Este trabajo se divide en seis capítulos, en los que se incluye la presente introducción. En el segundo, hablaremos sobre los distintos métodos que se utilizan en fonética forense para la identificación de hablantes. Específicamente se abordará, en primer lugar, el *voiceprint*, la ronda de reconocimiento de hablantes o *Ear-witness Line-ups* y el *Auditory-phonetic*; en segundo lugar, revisaremos los métodos semiautomáticos y automáticos; y en tercer lugar, veremos las técnicas que se usan actualmente para el reconocimiento de hablantes.

Finalmente, de manera breve comentaremos el tipo de metodología que decidimos seguir en este trabajo.

En el tercer capítulo abordaremos el tema de la respiración y veremos cuál es su relación con la producción del habla, después explicaremos cómo se genera la frecuencia fundamental y cuáles son los factores que podrían afectarla. Posteriormente, mencionaremos algunas patologías del lenguaje, cuyo origen puede ser psicológico o fisiológico. En seguida, revisaremos de manera detallada cuáles son los efectos del estrés sobre el cuerpo y haremos énfasis en la relación entre este y el aparato fonador. Por último, destacaremos algunos trabajos realizados en el ámbito forense, en los que la frecuencia fundamental ha sido el parámetro analizado.

En el cuarto capítulo presentaremos la metodología que hemos empleado en esta investigación. En primer lugar hablaremos sobre los tipos de corpus orales que existen, después de ello, explicaremos el protocolo que seguimos para componer el corpus y cómo fue su recopilación, es decir, la forma en la que realizamos las grabaciones. En segundo lugar, explicaremos cómo se llevó a cabo el análisis acústico. En tercer lugar, veremos de qué manera se hizo el vaciado de datos. Para finalizar, describiremos el procedimiento para llevar a cabo el análisis estadístico.

En el capítulo quinto presentaremos los resultados que obtuvimos. Explicaremos minuciosamente cómo se ordenaron los datos con los cuales se realizó el análisis estadístico, que a grandes rasgos consiste en la comparación de la F0 en dos momentos — antes y después de realizar ejercicio— y de acuerdo a tres criterios —género, tonicidad y contexto alveolar— analizados tanto de manera grupal como individual.

Finalmente en el capítulo sexto comentaremos las conclusiones a las que llegamos. Específicamente, veremos en qué medida se cumplen nuestras hipótesis de acuerdo a los resultados que obtuvimos. Asimismo, plantearemos algunos estudios futuros que resultarían pertinentes a partir de esta investigación.

2. MÉTODOS UTILIZADOS EN FONÉTICA FORENSE

Es normal que la mayoría de las personas sean capaces de reconocer la voz de alguien cuando contestan una llamada telefónica, o que puedan distinguir entre la voz de un artista y la de un imitador. Sin embargo, se necesita ser especialista en el área para poder estudiar los rasgos característicos de una voz, no sólo porque se requiere de un buen oído sino también de conocimientos lingüísticos. El fonetista forense, además, tiene que conocer la metodología que se debe seguir en un caso real. Se pueden citar algunas situaciones en las que este especialista puede ser requerido, p. ej., una víctima de secuestro que nunca vio a su raptor, pero que sí lo escuchó, una llamada de extorsión que haya sido grabada, amenazas que se hayan hecho por teléfono, etc.

En este apartado veremos cuáles son los antecedentes de la fonética forense. En primer lugar hablaremos del método *voiceprint*. En segundo lugar, describiremos el método denominado *auditory-phonetic*. En tercer lugar, trataremos el tema de los testigos auditivos y de las rondas de reconocimiento de hablantes, en inglés, *ear-witness line-ups*. En cuarto lugar, ahondaremos en los métodos automáticos y semiautomáticos que se han desarrollado y empleado en la fonética forense. También hablaremos de las técnicas con las que se trabaja en la actualidad. Finalmente, explicaremos el tipo de procedimiento que hemos utilizado en esta tesis de licenciatura.

2.1 El método *Voiceprint*

La primera noticia que se tiene sobre estudios en fonética forense se encuentra en 1945, después de la II Guerra Mundial, cuando Ralph Potter publicó un artículo llamado “*Visible Patterns of Speech*”, en el que describió la manera en la que se pueden observar gráficamente las formas del sonido: sus componentes frecuenciales, la amplitud y el tiempo (Hollien, 2002: 121). El artículo es producto de la investigación dirigida por el doctor Potter en los laboratorios Bell. El proyecto consistió en crear una máquina para identificar a los operadores de radio alemanes, en donde el reconocimiento se hacía mediante el uso del sonógrafo. Los doctores Gray y Koop acuñaron el término “*voiceprint*”, en un intento por

equiparar la representación gráfica de la voz con la de la huella digital o *fingerprint* (Delgado, 2001: 189). Quince años después Lawrence Kersta, un científico de los laboratorios Bell, publicó un artículo titulado "*Voiceprint identification*". Jonas Lindh refiere que la aplicación de este método comenzó, como ya se mencionó, después de la II Guerra Mundial cuando la policía de Nueva York recibió amenazas de bomba hacia diferentes compañías aéreas, esta llamó a los laboratorios Bell para saber si los espectrogramas se podían utilizar para identificar a quienes realizaban las llamadas. Kersta fue asignado para dicho caso (McDermott *apud* Lindh, 2004), después de dos años de estudio, estaba convencido de que los espectrogramas podían utilizarse para identificar hablantes e incluso llegó a comparar literalmente el método *voiceprint* con el *fingerprint*.

Afirmaba que las cavidades vocales y los articuladores son aquellas partes del tracto vocal que determinan la individualidad. Las primeras son resonadores y se caracterizan por sus dimensiones y forma de acoplarse, resultando despreciable la probabilidad de que dos personas tengan idénticas propiedades (Lucena, 2005: 1).

A pesar de que es cierto que las cavidades y los articuladores proporcionan un tipo de fonación característica, varios estudiosos del tema han rechazado el método *voiceprint* como único e infalible. Bolt afirma que "*the most fundamental issue is that it has never been shown that the inter-individual variation of spectra and/or temporal features of a spectrogram is greater than intra-individual variation*" (Braun, 1998: 11). Incluso se afirmó que "*the proponents of the method tended to function pretty much as would any businessman who has a product to promote and, thus, tried to fend off people who would denounce it*" (Hollien, 2002: 122)

Hollien llegó a la conclusión de que el método *voiceprint* recibió tantas críticas porque:

The procedure are that it uses archaic equipment and procedures, its validity has not been established and, indeed, is in serious question, it appears to permit decisions to be made only about one- third of the time and the training are competencies of its operators are largely a mystery. Its effectiveness in the field is so uneven that the only conclusion which can be drawn is that it lacks merit (2002: 133)

Como hemos visto, este método resulta poco científico porque se basa más en la opinión del experto que en pruebas tangibles. Si esa opinión se llevara a un tribunal resultaría poco

fiable porque el estudio no está sustentado en un método científico como sí sucede con la genética forense o la comparación de huellas digitales, por ejemplo.

Debido a las contradicciones que existieron en los estudios realizados por diversos expertos, en los tribunales de esa época, en EUA, se rechazó el método como prueba pericial, en cambio en otros juzgados, se decidió aceptarla con ciertas reservas. (López, 2010: 25)

2.2 El método *Auditory-phonetic*

El método *auditory-phonetic* se desarrolló especialmente en Gran Bretaña entre 1960 y 1980, a la par del *voiceprint* en Estados Unidos (French, 1994). Este método se lleva a cabo por fonetistas y se fundamenta en el estudio auditivo de las características de la voz. El especialista que analiza una grabación anota el sexo, edad, acento de la región y el estrato social al que pertenece la voz (Braun, 1998: 5). Además se debe poner especial atención al estudio de peculiaridades en la realización de consonantes y vocales, también se hace una evaluación del tono y la calidad de la voz.

Al igual que el método anterior, el *auditory-phonetic* tiene graves desventajas, French menciona que acústicamente diferentes datos de una voz pueden generar impresiones auditivas similares, además afirma que es un principio científico el hecho de que la conclusión de algún estudio, incluyendo el de la voz, sea el resultado de dos análisis diferentes e independientes (1994: 173). Los argumentos de French en su artículo son compartidos por la mayoría de los especialistas y hoy se recomienda que se utilicen dos metodologías, en este caso la auditiva y la acústica. Sólo de esa manera se podrá afirmar que la comparación de las muestras resulta confiable. “Actualmente, el sistema británico ha incorporado el análisis acústico y sólo admite evidencia fonética que esté basada en ambos análisis, ya no permite el uso exclusivo del análisis auditivo” (López, 2010: 26)

2.3 El método *Ear-witness Line-ups*

Una de las primeras noticias que se tienen documentadas sobre este método data de 1935, con el juicio de Richard Hauptmann por el secuestro y asesinato del hijo de Charles Lindbergh, quien identificó la voz del criminal dos años después de haber escuchado la voz

vía telefónica. Algunos años después McGehee analizó el decaimiento de la memoria a largo plazo para el reconocimiento de voz; con base en los resultados que obtuvo, afirmó que es poco justificable que alguien pueda reconocer una voz después de tanto tiempo, más aún que este reconocimiento sea sólo producto de un breve encuentro con el hablante. Las conclusiones de McGehee dieron pie a que este tipo de prueba no se aceptara en un tribunal en esa época (Braun, 1998: 11).

Las rondas de reconocimiento de voz pueden tener lugar cuando un individuo, ya sea víctima o testigo de algún crimen, haya escuchado al autor de éste, pero que no lo haya visto. La tarea de dicho sujeto es hacer un juicio acerca de la identidad del victimario, para ello se le pedirá que escuche las muestras de voz del sospechoso. El procedimiento debe ser llevado a cabo de una manera bien estructurada para garantizar que las respuestas del testigo sean verdaderas y no inducidas, y de esta manera garantizar la veracidad de la prueba para ambas partes (Hollien, 2012). Si bien este método puede ser útil, los testigos no tienen la preparación lingüística para reconocer o caracterizar una voz; además, existe el problema de la memoria auditiva. Por si fuera poco, es evidente que no se podrá reproducir el mismo estado de estrés en el que se encontraba la víctima al momento de ser agredida, hecho que puede afectar en la ronda de reconocimiento.

Hollien habla de algunos problemas que se pueden presentar al momento de hacer el reconocimiento de hablantes como lo son la latencia, es decir, el momento que ha transcurrido desde el crimen hasta el momento en que se hace el reconocimiento. La memoria se ve degradada con el tiempo por lo que sería deseable que entre estos procesos transcurriera el menor tiempo posible; otro factor es el orden en el que las grabaciones se presentan al testigo, es decir, la formación de las voces no debe influir en el reconocimiento que se realizará. Por otro lado, refiere que existen problemas cuando los testigos no tienen buena memoria auditiva. Respecto a esto último De Jong estudió los efectos de la memoria, la capacidad auditiva y la habilidad musical sobre la precisión en las rondas de reconocimiento de voz, *“She reported that intelligence (as measured by cognitive processing) was a better predictor of a subject’s ability to identify speakers than were basic auditory and memory skills”* (Hollien, 2012: 4), asimismo concluye que los hablantes que tienen un buen oído musical, obtuvieron mejores resultados en el reconocimiento de hablantes.

Otro de los problemas que se puede presentar es la edad de los testigos, Hollien afirma que las personas son buenas reconociendo el habla de individuos que están en su mismo rango de edad. Un factor más es si la voz que se tiene que reconocer pertenece a una lengua extranjera o a algún dialecto, hecho que presenta dificultades para el testigo auditivo.

Como podemos ver, hay muchos inconvenientes con este tipo de procedimiento, sin embargo, es preciso mejorar la metodología que se lleva a cabo en las rondas de reconocimiento, porque lamentablemente el perito en fonética forense no siempre va a contar con una muestra del habla del sospechoso.

2.4 Métodos semiautomáticos y automáticos

Los métodos semiautomáticos se llevan a cabo con ayuda de programas computacionales. En este tipo de procedimiento es el investigador quien elige los parámetros que se van a estudiar. “Una vez introducidos los datos para su muestreo y procesado en la aplicación del cálculo y análisis, el ordenador proporciona ciertos valores representativos de cada una de las referencias estudiadas” (Delgado, 2010: 216). El cotejo se realiza con los datos dubitados e indubitados. Una vez que estos han sido analizados por la máquina, el investigador tiene que interpretar los resultados obtenidos y llegar a una conclusión.

Por su parte, los métodos automáticos son aquellos en los que la intervención del investigador no es indispensable, Delgado refiere que se ha intentado utilizar este tipo de procedimiento, pero las condiciones en las que las grabaciones de casos reales fueron obtenidas no son nada favorables para un análisis automático (Delgado, 2001: 218).

A pesar de las dificultades a las que los investigadores se han enfrentado en la comparación forense del habla, ellos han persistido en la idea de crear un sistema que sea capaz de hacer la comparación del habla de forma automática. El Departamento de Acústica e Imagen, que pertenece al Servicio de Criminalística en Madrid cuenta con un sistema para el reconocimiento automático de hablantes llamado *IdentiVox*. Este sistema está basado en “tecnología de esquemas de modelado acústico del contenido fonético del locutor, concretamente en modelos de mezclas gaussianas” (González Rodríguez; Lucena,

2005: 120), que es una prueba estadística utilizada en otras ciencias forenses como en la genética y en la dactiloscopia, en donde ha dado buenos resultados.

IdentiVox es un sistema basado en puntuaciones, lo que implica que necesita una base de datos para modelar las distribuciones de las medidas. El sistema trabaja con muestras de habla dubitadas e indubitadas, lo que hace es compararlas y representar el resultado por medio de curvas, entre más alineadas están las curvas, se puede afirmar que se trata del mismo autor. *IdentiVox* permite trabajar con habla espontánea, además no presenta dificultades para trabajar con distintas lenguas o dialectos (González Rodríguez; Lucena, 2005: 131).

Un avance importante que ha permitido *IdentiVox* es la creación de un proyecto llamado *SAIVOX*, que empezó a desarrollarse en 2004. Este proyecto es ambicioso porque pretende recolectar muestras de habla indubitada, provenientes de toda España. Se espera que este sistema permita a los peritos disponer de enormes bancos de voces, lo cual no tiene precedentes (González Rodríguez; Lucena, 2005: 132).

Otro sistema automático es *Batvox*, éste funciona de manera similar a *IdentiVox*. *Batvox* es un sistema comercial que la empresa *Agnitio* patentó y comercializó alrededor del mundo, el uso de este sistema ha sido presentado como prueba pericial en Europa, Asia y América Latina (González Sigüenza, 2008: 315).

Es importante señalar que en Estados Unidos se encuentra el *Speech Group* del Instituto Nacional de Estándares y Tecnología, que realiza evaluaciones constantes sobre los sistemas de reconocimiento automáticos existentes. Se ha reportado que estos resultan perturbados debido a la influencia de la utilización de distinto canal telefónico, la duración temporal de los segmentos, etc.

La importancia del canal telefónico estriba en que éste tiene la capacidad de transmitir frecuencias de entre 300 y 3400 Hz, pero, a pesar de que fue diseñado para la transmisión de la voz humana, un canal telefónico capta toda clase de señales que tengan un rango de frecuencia igual al mencionado anteriormente. La gran mayoría de las voces dubitadas, en el ámbito de la fonética forense, provienen de interceptaciones de telefonía móvil o de línea terrestre, éstas a su vez se registran en diferentes equipos y soportes de grabación, lo anterior aunado a otros factores como ruidos, distorsiones o solapamientos de voz representan un obstáculo en el rendimiento de los sistemas automáticos de

reconocimiento. En cambio el rendimiento de un sistema automático es mayor cuando los segmentos duran más o si se usa el mismo canal telefónico (Delgado, 2005: 118).

Es una gran ventaja que se tengan bien identificados los inconvenientes de la utilización de los sistemas automáticos, para así intentar mejorar el nivel de fiabilidad, no obstante “en la actualidad la tecnología de reconocimiento automático no proporciona los resultados deseables, especialmente cuando se enfrenta a las denominadas condiciones forenses” (Nakasone y Beck *apud* Delgado, 2005: 122).

2.5 Métodos utilizados actualmente en fonética forense

En la actualidad la mayoría de los métodos mencionado se siguen empleando, p. ej., se han hecho diversos trabajos sobre la ronda de reconocimiento de hablantes (*Ear-witness Line-ups*), uno de los autores más citados es Harry Hollien. Este método es indispensable en el área de la fonética forense porque en ocasiones no se tiene una muestra de voz para analizar, pero generalmente habrá un testigo auditivo en el que se pueda apoyar el perito.

En Estados Unidos, en el área de Sección de Tecnología y vigilancia Electrónica del FBI, manejan el método auditivo-espectrográfico y sólo lo usan como apoyo a la investigación de sus propios casos o de organismos policiales que soliciten el servicio. (Delgado, 2001: 201)

Gold y French realizaron un conteo sobre los métodos que más se utilizan en los países, en donde se trabaja con fonética forense. Reportan que el *auditory phonetic* es utilizado en los Países Bajos y en Estados Unidos. En Australia, Austria, China, España, Reino Unido y Estados Unidos, además de trabajar con *auditory phonetic* utilizan el análisis acústico- fonético. Los lugares en donde se utilizan métodos automáticos son pocos, entre éstos se encuentran Alemania, Corea del sur, España, Suecia y Estados Unidos (2011: 296).

Delgado refiere que actualmente las metodologías forenses más practicadas por los laboratorios policiales son las combinadas, éstas “sustentan sus fundamentos de estudio en tres perspectivas: acústica, fonético-lingüística y auditivo-perceptiva. Dichas técnicas, pueden complementarse o desarrollarse a través de sistemas semiautomáticos de cálculo o análisis” (2005: 117). Precisamente, el autor realizó una propuesta de análisis combinado

de las muestras de voz, en su tesis doctoral. La propuesta de Delgado es un examen general de la voz, mediante un cotejo auditivo con apoyo en espectrogramas, uno acústico y además uno estadístico, éste último está basado en modelos de mezclas gaussianas, como el que es utilizado en los sistemas automáticos de reconocimiento de hablantes. Delgado comenzó el estudio de las muestras de voz con el análisis perceptivo-auditivo, hizo una clasificación del sociolecto al que pertenece la muestra de voz, también realizó una lista de los recursos retóricos que más usó el sospechoso, como muletillas, ciertas expresiones y utilización de términos; por otro lado, hizo una anotación sobre la percepción del timbre del hablante. En cuanto a la exploración acústica, el autor comparó minuciosamente las realizaciones de ciertos grupos silábicos, examinó los primeros dos formantes, así como la frecuencia fundamental. Los resultados indican que las muestras de voz podían pertenecer a la misma persona. La conclusión a la que llegó es que la utilización de las técnicas combinadas en una comparación forense de voz lleva a los resultados deseados, es decir, la identificación de hablantes (2001).

Otra de las técnicas que se usa con frecuencia es el análisis multidimensional, éste puede clasificarse dentro de los métodos semiautomáticos. Un ejemplo de su empleo es el experimento realizado en la tesis doctoral de Fernanda López. En su investigación se analizó la frecuencia fundamental (F0) y los tres primeros formantes vocálicos de la vocal /a/ en sílaba abierta. En el análisis acústico la autora no sólo tomó un punto de la vocal, sino nueve a lo largo del segmento /a/. Las conclusiones a las que llegó resultan esclarecedoras porque habla de la utilidad de tomar nueve puntos del segmento o sólo el punto central. En cuanto a la F0, la autora afirma que no hay una diferencia significativa entre tomar nueve puntos a lo largo de la vocal, o analizar sólo el punto central; en cambio en el análisis de los formantes vocálicos, concluye que lo más viable es hacer un análisis de los nueve puntos porque sí aporta más información del hablante (2010: 253).

2.6 Método utilizado en esta tesis

En esta tesis hemos decidido utilizar un método semiautomático. Nos hemos apoyado en dos programas: *Praat* (Boersma y Weenink, 2013) para el análisis acústico, y *R-project*¹¹

¹¹ *R-project* está disponible para su descarga en <http://www.r-project.org/>

para el estadístico. El primero es una herramienta que permite observar el espectrograma de la grabación insertada y obtener diferentes parámetros de la voz como la frecuencia fundamental, formantes vocálicos, entonación, intensidad, etc. La ventaja que tiene *Praat* es que se pueden programar órdenes específicas mediante *scripts*, gracias a éstos se pueden extraer automáticamente los valores numéricos de un parámetro. En este trabajo, esta función nos ha servido para obtener la frecuencia fundamental del punto medio en la vocal /e/. Los *scripts* también facilitan el manejo de grandes cantidades de datos. La función del investigador en este caso es etiquetar los segmentos de los que se desea extraer la información (ver 4.2), porque el programa no realiza un etiquetado automático. Por otro lado, *R-project* se utilizó para realizar el análisis estadístico (ver 4.3).

El papel del lingüista en esta tesis, razón por la cual se clasifica dentro de los métodos semiautomáticos, fue la de elegir el sonido que se iba a analizar, etiquetar la vocal para que *Praat* realizara la extracción automática de los valores numéricos de F0, vaciar los datos, organizarlos, realizar el análisis estadístico en la consola de *R-project* e interpretar los resultados.

Hemos descrito de manera general el método que utilizamos y podemos concluir que en casos como éste, el papel del investigador es fundamental porque tiene que hacer un uso adecuado de las herramientas computacionales que están a su alcance para realizar un manejo apropiado del corpus e interpretar los resultados que arrojen ambos programas.

3. LA RESPIRACIÓN Y SU RELACIÓN CON EL HABLA

En este apartado describiremos el proceso respiratorio, la forma en la que se lleva a cabo la producción del habla y cuál es su relación con la respiración. Hablaremos del objeto de estudio de esta tesis, la frecuencia fundamental; también veremos cuáles son los factores que influyen sobre ella. Asimismo, mencionaremos algunos elementos que afectan la producción del habla, como son algunas disfonías de carácter fisiológico y psicológico, de dichas causas se desprende el estrés, del cual haremos una descripción minuciosa.

Por otra parte, hablaremos sobre algunos trabajos de fonética forense en los que el interés principal es el estudio de la frecuencia fundamental (F0).

3.1 La respiración

Uno de los procesos fundamentales para que se pueda producir la fonación es la respiración, es por ello que hemos decidido comenzar con la explicación de dicho proceso. El aparato respiratorio está compuesto principalmente por los pulmones, bronquios y tráquea. Los pulmones proveen el aire suficiente para que se pueda producir la respiración. El aire de los pulmones pasa por los bronquios y de éstos a la tráquea. La renovación del aire se hace mediante la inspiración y la espiración, es en este último momento en el que se puede producir la fonación.

La respiración se puede clasificar en tres tipos, de acuerdo a la manera en la que se realice el acto.

1. Respiración clavicular, que conlleva elevación y descenso del tórax con hundimiento del pulmón en su parte superior. Esta posición es utilizada en cualquier emoción repentina, provocándonos elevamiento del tórax.
2. La respiración torácica, está asociada a una expansión y retracción del tórax adecuada para el habla conversacional.
3. La respiración abdominal o diafragmática, esta forma es la más sencilla porque permite realizar un mayor intercambio de aire (Delgado, 2001: 28).

Por un lado, si la respiración torácica se usa frecuentemente puede provocar problemas en la voz porque se utilizan los músculos del cuello, lo que provoca insuficiencia

del aire; por otro lado, la respiración abdominal se utiliza si el cuerpo necesita respiración de emergencia, y con ésta se pone en funcionamiento la masa muscular de la columna vertebral para dar mayor apoyo y fuerza a la acción de respirar.

Como observamos hay diversas maneras en las que nuestro sistema respiratorio se adapta a las circunstancias en las que se debe llevar a cabo el proceso de la respiración.

3.2 La producción del habla y su relación con la respiración

Seguida de la respiración, la fonación es el segundo proceso más importante del aparato respiratorio. “La fonación representa una función adquirida tardíamente en la filogénesis por órganos como la laringe, faringe, boca, que ya ejercen funciones vitales y primitivas. [...] Pero la fonación requiere una coordinación muscular y psicosensoresial mucho más compleja” (Segre, 1955: 51). Como anteriormente se vio, los pulmones cumplen una función primordial al momento de producir un sonido, la respiración es el elemento motor de la voz porque produce la presión necesaria para mantener la ondulación de los pliegues vocales (cuerdas vocales), en el habla se producen variaciones manifiestas, que afectan la capacidad respiratoria, el ritmo y sobre todo, la duración de la espiración (Dinville, 1996: 5).

Cuando se inicia la fonación, se cierra la glotis, entonces hay una presión de aire infraglotico contra la tráquea y contra la glotis, los bordes de ésta se separan y dejan salir determinada cantidad de aire, mismo que pasa entre las cuerdas vocales, éstas se aproximan nuevamente para volver a cerrar la glotis (Quilis, 1999: 60).

Cada vez que este proceso se repite, el aire que pasa a través de la glotis provoca que se produzca un movimiento vibratorio. Este movimiento genera energía que se manifiesta en forma de ondas sonoras, éstas son amplificadas en los resonadores, es decir, la cavidad nasal, bucal y faringe, después las ondas son modeladas por los órganos articulatorios: la lengua, el paladar, entre otros, para generar finalmente un sonido del habla.

El proceso de la producción del habla no es sencillo, al contrario, la complejidad que conlleva consiste en que tiene que existir un equilibrio total entre la respiración y la fonación, cuando no lo hay se producen alteraciones en el habla. Perello apunta que en un análisis de la voz debe tenerse en cuenta “la insuficiencia respiratoria nasal [...], así como

la mala coordinación entre la respiración y la fonación, la insuficiencia torácica, debilidad de la musculatura abdominal. Todo ello proporciona a la laringe un trabajo suplementario” (1973:102). Este trabajo suplementario ocasiona que al momento de hablar se produzcan problemas, momentáneos o crónicos, al producir la voz. Las razones por las que se puede perder el equilibrio entre el proceso de la respiración y la fonación son diversas, entre ellas se encuentran algunas patologías del habla de origen fisiológico o psicológico.

3.3 Frecuencia fundamental

Antes de explicar qué es la frecuencia fundamental debemos hablar de la calidad vocal. Rose (2002) denomina *voice quality* o calidad vocal al conjunto de efectos que, tanto los órganos articulatorios como las cavidades supraglóticas, tienen sobre el sonido. Es decir, la calidad vocal engloba todos los aspectos que se mantienen durante la cadena fónica y que están determinados por las propiedades fisiológicas de los hablantes y por el estado de órganos que intervienen en la producción del habla. Rose menciona que la calidad vocal tiene dos componentes, uno orgánico y otro articulatorio. El primero se refiere a “*aspects of the sounds that are determined by the particular speaker’s vocal tract anatomy and physiology, such as their vocal tract length or the volume of their nasal cavity, and which they have no control over*” (2002: 279). El componente articulatorio “*refers to habitual muscular settings that an individual adopts when they speak*” (2002: 279), p. ej.: si el hablante redondea los labios o realiza alguna nasalización.

La frecuencia fundamental tiene una estrecha relación con la calidad vocal, porque se produce en el aparato fonador, cuando se inicia el proceso de producción de la voz. Ya vimos que el aire que pasa a través de la glotis genera vibraciones, éstas se van amplificando por efecto de los resonadores. Una vez que las masas de aire pasan a través de la glotis y provocan que se produzca un movimiento vibratorio, la energía que genera dicho movimiento se manifiesta en forma de onda sonora, a esta onda se le conoce como frecuencia fundamental. La F0 adquiere un carácter complejo porque, debido al efecto de la resonancia en las cavidades supraglóticas, se generan armónicos que conforman la estructura acústica de los sonidos vocálicos. Cada uno de estos armónicos es un múltiplo de la onda sonora más baja: F0. Al conjunto formado por el tono de la F0 y los armónicos formados en las cavidades supraglóticas se le conoce como timbre.

El número de vibraciones de las cuerdas vocales en relación al tiempo depende de una serie de factores a su vez interdependientes: masa de la parte vibratoria de las cuerdas vocales, tensión de las mismas, área de la glotis durante el ciclo, valor de la presión infraglotica y amortiguación de las cuerdas vocales (Delgado, 2001:41).

En la imagen 3.1 observamos las ondas simples que componen la onda compleja.

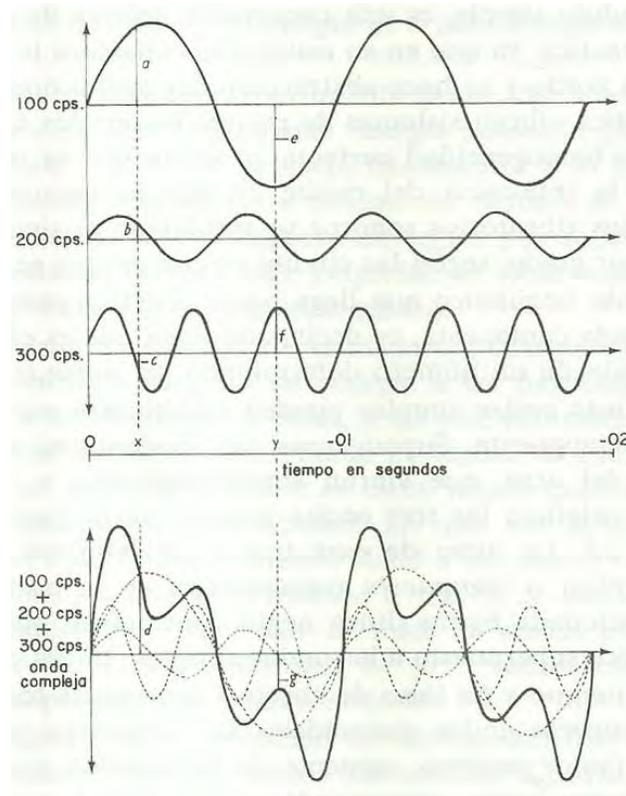


Figura 3.1 Representación de la onda sonora simple y compleja
(Tomado de Quilis, 1999: 92)

La primer onda que aparece en la figura, cuyo valor es de 100 ciclos por segundo (cpc) o Hertz (Hz), es la frecuencia fundamental. La segunda, cuyo valor es de 200 cps, corresponde al segundo armónico y la tercera, que tiene un valor de 300 cps, al tercer armónico. En este trabajo sólo analizaremos la primer onda generada por la vibración de las cuerdas vocales.

La frecuencia fundamental nos puede aportar características fisiológicas de un hablante, p. ej., podemos diferenciar entre la voz de un niño y la de una mujer “uno de los promedios dados para los valores más frecuentes de la frecuencia fundamental es de 350 Hz

para los niños, 250 Hz para las mujeres y 125 Hz para los hombres” (Menaldi *apud* Delgado, 2001: 41).

La diferencia de la F0 entre un niño, un hombre y una mujer se puede apreciar en la siguiente figura:

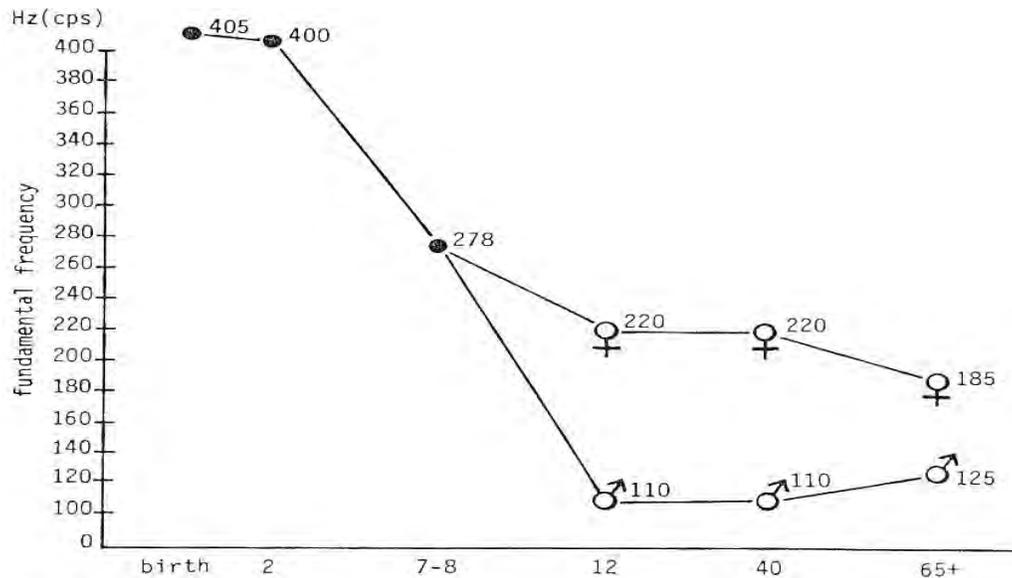


Figura 3.2 Variaciones de la frecuencia fundamental en función de la edad
(Tomado de Llisterri, 2014)

3.4 Factores que afectan la frecuencia fundamental (F0)

La frecuencia fundamental es un parámetro que se ha estudiado mucho en fonética forense porque tiene la capacidad de mantenerse estable aún en ambientes con ruido o en conversaciones telefónicas.

A pesar de ser uno de los parámetros que resultan menos afectados por las condiciones ambientales, se ha comprobado que la F0 puede resultar alterada. Braun (1995 *apud* López, 2010: 37) propone tres grupos de factores:

1. Factores técnicos que se consideran externos al proceso de la producción de la voz.
2. Factores fisiológicos.
3. Factores psicológicos.

En el primer grupo se encuentran aspectos como el tipo de cinta utilizada durante la grabación, la velocidad con la que se hizo, el uso de distintos equipos, etc. Es importante

que cuando se compare la voz de un sospechoso se controlen estos aspectos para no alterar los parámetros acústicos de la grabación.

En el grupo de los factores fisiológicos se encuentran elementos “como la edad del hablante, su procedencia, si es fumador, si durante la grabación se encontraba bajo los efectos del alcohol o de alguna droga o si intentó disfrazar su voz” (López, 2010: 37).

El tercer grupo, el de los factores psicológicos, engloba situaciones anímicas como el miedo, la fatiga vocal, la ansiedad, el estrés, la ira, etc. “Como bien señala Braun uno de los problemas que se presenta con este tipo de factores es que es imposible reproducir aspectos emocionales a la hora de realizar la grabación indubitada” (López, 2010: 38).

En algunos estudios se han tratado de reproducir aspectos emocionales, una de las soluciones ha sido utilizar actores que finjan estar en un determinado estado anímico, sin embargo, no podemos aceptar que el actor que simula un estado emocional, puede reproducir una voz fidedigna como la que aportaría un hablante en una situación real. Jeseen, M. *et al* afirman que la variación intrahablante (entre el mismo hablante) puede complicar la comparación forense de voz, cuando p. ej., si la voz de la grabación fue producida espontáneamente o bajo estrés y, en el momento de comparar, el sospechoso lee tranquilamente y con una voz neutral (2005:175).

Como vemos, existen diversos factores que pueden alterar la frecuencia fundamental, a pesar de ello, diversos autores han tomado este parámetro para la comparación forense de voz. En esta tesis decidimos trabajar con cuestiones que atañen al grupo de los factores psicológicos. Para explicar de una mejor manera los objetivos de esta investigación vamos a comentar otros factores que afectan la producción del habla.

3.5 Factores que afectan la producción del habla

Dentro de los factores que menciona Braun (1995 *apud* López, 2010), se encuentran los fisiológicos y psicológicos. El habla puede ser afectada por diversas causas como la salud del hablante, si fuma o consume alcohol. Entre los factores fisiológicos, se encuentran problemas que van desde alteraciones fisiológicas, como disarmonías entre los órganos de la fonación y la respiración hasta malformaciones en el aparato fonador. Cuando estas patologías son muy acentuadas se pueden presentar graves dificultades, porque obligan a

esfuerzos de compensación, que son seguidos de fatiga vocal o sobreesfuerzo (Dinville, 1996: 46).

La fonación también puede verse afectada por enfermedades pulmonares y cardíacas, como el asma, enfisema, bronquitis crónica e insuficiencia cardíaca, dichas afecciones dificultan la función respiratoria y por ende la fonación (Dinville, 1996: 47).

La voz de un individuo es un indicador muy sensible de sus emociones y sus actitudes, así como podemos saber que una persona está enojada o triste por medio de sus expresiones faciales. Es posible que, también, a través de la voz podamos percibir el humor de alguien, por lo que no es extraño pensar que los estados psicológicos afectan la fonación de un individuo. “La lista de formas en que la mente puede alterar la voz parece interminable: ansiedad, con sus alteraciones respiratorias y posturales aberrantes asociadas, alteraciones de rasgos de personalidad, depresión, etc.” (Morrison *et al*, 1996: 126)

Delgado Romero (2001: 58-60), realizó una síntesis de otros factores que afectan la producción del habla. Los conceptos que ahí se citan son un aporte de la *American Speech Correction Association*. En las siguientes líneas mencionaremos sólo los que se relacionan con el trabajo que vamos a desarrollar, incluiremos las que están relacionadas con factores fisiológicos como psicológicos.

Entre las patologías referidas se encuentran:

- a. Las disfemias. Son alteraciones del lenguaje, que se caracterizan por un desorden del ritmo, espasmos y tics debido a psiconeurosis.
 - La agilofemia. Se refiere al habla agitada o nerviosa.
- b. Las disfonías. Son trastornos de la voz debido a perturbaciones orgánicas o funcionales de las cuerdas vocales o respiración defectuosa, que suele ser torácica y con espiraciones débiles.
 - Entre las disfonías se encuentran las parafonías, se puede decir que son alteraciones mórbidas de la voz.
 - La parafonía copiacca es el cambio de la voz por fatiga.
 - La fonastenia, una disfonía de origen psicológico, se presenta a causa de fatiga, la voz que se produce es débil.
- c. La neumofonía. Son llamados así, los defectos de la voz debidos a falta de coordinación neumofónica, por ejemplo: la voz aspirada.

- d. La disritmia es una alteración más de la voz y consiste en los defectos del ritmo en los que no se incluye la tartamudez, éstos pueden ser por defectos respiratorios o alteraciones endocrinas.
- La disritmia neumafrasia se presenta debido a defectos respiratorios.
 - La disritmia prosódica. Defectos de la acentuación durante la lectura o en el habla espontánea.

Observamos que son muchos los factores que pueden afectar la fonación de un individuo y es de gran ayuda tener estudios que nos permitan enlistar las distintas causas que provocan anomalías en la producción del habla para delimitar los posibles contextos en los que puede estar enmarcada la voz que se va a comparar.

3.6 El estrés y su relación con la producción del habla

Ya hemos señalado que entre los factores de origen psicológico se encuentran causas como la ansiedad, la depresión o alteraciones en la personalidad. Rasgos como la ansiedad, el enojo, el miedo o la ira responden a lo que comúnmente conocemos como estrés. De acuerdo con el DRAE el estrés es la “Tensión provocada por situaciones agobiantes que originan reacciones psicósomáticas o trastornos psicológicos a veces graves” (DRAE, 2001). En la psicología “el estrés es considerado como un proceso básico, adaptativo, que implica diversas respuestas, cuyo objetivo es la supervivencia y la homeostasis de los organismos” (Breva, A. *et al*, 2000: 3). Por otro lado, González de Rivera nos dice que el estrés es “cualquier estímulo o situación que altere o interfiera con el equilibrio fisiológico normal de un organismo, así como los estados de ese organismo que implican sobreesfuerzo o tensión física, mental o emocional” (2010: 15), además del esfuerzo que conlleva, una situación de estrés puede poner al sujeto en riesgo de enfermedad (Dietrich *et al*, 2008: 473).

Existen diversos efectos en el cuerpo mientras se encuentra en estado de estrés, el cuerpo segrega adrenalina y noradrenalina, con la finalidad de movilizar al cuerpo para una posible respuesta, ataque o huida, según convenga. Hay dilatación pupilar, cuya finalidad es mejorar la visión, también aumenta de la coagulabilidad de la sangre, quizá para cerrar posibles heridas. “Además hay un aumento de la frecuencia y fuerza del latido cardiaco, que facilita un bombeo más rápido de oxígeno [...], liberación hepática de azúcar

almacenado para uso de la musculatura” (Gray, 1993: 64-65). Todo se produce en cuestión de segundos o minutos, p. ej., cuando alguien está a punto de ser atropellado sufre en su cuerpo todos los cambios que hemos descrito en unos cuantos segundos. Generalmente la reacción de alarma llega momentos después de que el peligro ha pasado, esto es porque ya no vivimos en un ambiente en el que tengamos que luchar con animales salvajes o cuidar nuestra vida a cada instante.

A pesar de que la reacción de alarma llega un momento después de que ha pasado el peligro, seguimos experimentando el mismo proceso cuando nos estresamos, y no necesariamente tenemos que estar en peligro de muerte para pasar por dichos procesos. “Lo que ordinariamente se conoce como responsabilidad es una de tales circunstancias. Así, el mero requerimiento para la preparación de un informe eleva el ritmo cardíaco [...] en ejecutivos, de manera análoga a lo que ha sido observado en pilotos de guerra antes de entrar en combate” (González de Rivera, 2010: 33).

Existen tres estados anímicos que podemos relacionar directamente con el estrés, la ansiedad, el miedo y la angustia. La ansiedad es una sensación difusa de anticipación de un daño o peligro que no se puede concretar, frecuentemente inhibe las funciones intelectuales y fisiológicas. En cambio, el miedo es un sentimiento claro y surge de la confrontación con un peligro real, es en este estado en el que se presentan con mayor claridad los efectos del estrés sobre el cuerpo, que explicamos más arriba. Esto es porque el peligro es visto como algo concreto a lo que debemos enfrentarnos. Por otro lado, la angustia es una variante de la ansiedad, pero en lugar de anticipar el peligro, éste se vive en el presente (González de Rivera, 2010: 98).

Si pensamos en una situación en la que un hablante se encuentre estresado podremos inferir, conforme a lo que hemos visto arriba sobre el estrés, la fonación y la respiración, que su cuerpo sufre ciertas reacciones como que su ritmo cardíaco se eleve y por lo tanto que su voz se vea afectada. En esta tesis hemos tratado de simular un estado de estrés, mediante la elevación del ritmo cardíaco por medio de la actividad física para observar los efectos que sufre la frecuencia fundamental en cada una de las voces de los hablantes. Es por ello que insistimos en la relación entre la respiración, el ritmo cardíaco y la fonación.

Como hemos visto, el estrés se puede generar por diversas causas, algunas pueden ser más específicas que otras, pero lo que nuestro cuerpo sufre es, en cierta medida, similar. Ahora veremos cuál es la relación entre el estrés y la producción del habla.

Es conocido que existen diversas causas que provocan estrés. Morrison nos dice que “la emoción es la variable intermedia entre los factores psicológicos y la alteración física del tono muscular de la laringe que produce realmente la disfonía” (1996:129). Afirma también, que la ansiedad es la emoción que con más frecuencia causa alguna disfonía, existen distintas causas de la ansiedad, como la percepción de alguna amenaza o peligro externo, los sentimientos de culpa o la intensidad de las propias emociones. “De la misma forma que el rostro aporta pistas sobre la actitud y la emoción de una persona, la voz tiene un efecto similar” (1996: 129)

Ya vimos que la reacción de alarma provoca efectos diferentes en el cuerpo de una persona. Existen otras reacciones que el estrés puede provocar, además del aumento del ritmo cardiaco y la frecuencia respiratoria, como “*bradycardia, miosis, and tracheal constriction through the dorsal vagal motor nucleus and vagal efferent nerves*” (Cheryl L.G., et al, 2013: 290.e27)

Entre los estudios que se han hecho respecto a este tema, Cheryl refiere que se ha reportado un aumento en la F0 cuando los hablantes se encuentran en un estado de estrés además nos dice que “*heart rate contributed 0.5–20% to the total vocal F0 perturbation measure or jitter, whereas Orlikoff demonstrated that the heart rate contributed 11.8% to the total vocal intensity perturbation measure or shimmer*” (Cheryl L.G., et al, 2013: 290.e22). Incluso se han realizado estudios en medicina aeroespacial, en esta área tienen una preocupación por identificar los niveles de estrés y uno de los medios para hacerlo es el análisis de voz, p.ej., “*Kuroda et al [...] reported increases in F0 and associated increases in speech formants with increasing levels of workload (stress) during their attempts to document stress in the voices of male aviators, paratroopers, and air traffic controllers*” (Cheryl L.G., et al, 2013: 290.e26)

3.7 Estudios de F0 en fonética forense

En esta tesis decidimos analizar la frecuencia fundamental porque es un parámetro que se puede extraer fácilmente de las grabaciones, lo cual facilitaría el trabajo del perito en

fonética forense. Además en diversos estudios se ha hablado de la confiabilidad de este parámetro al momento de realizar comparaciones forenses de voz. Kinoshita menciona que además de la facilidad con la que se extrae, “*F0, however, is relatively robust against poor recording quality and differences in transmission channel. Furthermore, in non-tone languages at least, F0 is not affected by the lexical content of the speech samples, so there is no need to locate comparable words or phonemes*” (2009: 92). En este trabajo hemos utilizado el programa computacional *Praat* (Boersma y Weenink, 2013) para extraer los valores de la F0.

En el ámbito forense se han hecho varios estudios sobre la frecuencia fundamental, para el inglés, por ejemplo: Boss, D. en «*The problem of F0 and real-life speaker identification: a case study*». “reporta un caso real donde el promedio de la frecuencia fundamental de un individuo resultó de 228 Hz durante la grabación criminal y de 140 Hz en las grabaciones hechas durante el interrogatorio. La diferencia de 88 Hz la atribuye al estrés que el criminal experimentó durante el robo.” (Boss *apud* López, 2010: 38)

Gold y French encontraron que en todos los trabajos presentados en “*the first international survey on forensic speaker identification on forensic comparison*” se mide rutinariamente la frecuencia fundamental mediante un análisis acústico, el componente mayormente analizado es la media, le sigue la mediana, la moda y la desviación estándar. A pesar de que la frecuencia fundamental es uno de los parámetros más utilizados “*many respondents reported analysing the fundamental frequency, a large proportion point out that it is usually of little help. One respondent stated that usually used as an elimination tool rather an identification tool*” (2011: 301)

Hansen y Patil realizaron la medición de diversos parámetros de la voz bajo distintos tipos de estrés. Analizaron la frecuencia fundamental, intensidad, duración, formantes vocálicos, entre otros más. En este trabajo consideraron la utilización de varios corpus, entre los que se encuentran algunas entrevistas, en donde el interés principal fue analizar el discurso bajo estrés, así como la percepción humana del mismo. También tomaron en cuenta el corpus CU- Move, el cual es un diálogo con una máquina de navegación en automóvil, además utilizaron un corpus que contiene grabaciones bajo estrés físico y cognitivo (Hansen y Patil, 2007: 114). En los corpus analizados se registraron

voces con varios tipos de estrés, como voz rápida, lenta, suave, enojada, que denota miedo, estos registros se compararon con una voz normal.

El análisis de la frecuencia fundamental los lleva a la conclusión de que ésta podría ayudar a clasificar el tipo de estrés que se registra en la voz, nos dicen que “*Mean fundamental frequency is a good indicator over a wide variety of stress conditions. Loud, angry, and Lombard mean fundamental frequency are all significantly different from neutral as well as all other styles considered*” (2007: 115)

Rosas y Sommerhoff realizaron un experimento en el que investigaron si en el español de Chile existe una variación acústica idiolectal producida por factores ambientales como el ruido, y expresivos, es decir, distintos estados de ánimo. Analizaron la vocal /a/ precedida por /s/ y seguida de /p/, /k/, precedida y seguida por /n/ y precedida y seguida por /p/, /t/ (2010: 102). “Las frases fueron producidas y registradas en canal anecoica por cuatro actrices, quienes simulaban estar inmersas en distintas situaciones fonopragmáticas, de lectura, evidencia, asombro, atenuación y no conclusión” (2010: 102). El registro se hizo en canal telefónico y celular. Ellos concluyen que

El comportamiento del F0 y de los formantes permitió caracterizar la forma en que esos parámetros cambiaban frente a diversas condiciones fonopragmáticas y ambientales, mediante la configuración de tendencias- aumentos o disminuciones generalizadas-; sin embargo, ninguna de las tendencias observadas llegó a ser significativa para la totalidad de los locutores (2010: 106).

Jessen *et al* realizaron un trabajo en el que vieron cuál es la influencia del esfuerzo vocal en la frecuencia fundamental. En su trabajo utilizaron el corpus ‘Pool 2010’, en donde están recogidas las voces de 107 hombres alemanes de entre 21 y 63 años de edad. Los investigadores sólo tomaron en cuenta los datos de los hombres que tuvieran una edad aproximada de 39 años. El corpus contiene un registro espontáneo de la voz en condiciones normales y otro en donde los hablantes leyeron en voz alta expuestos a ruido. En este estudio notaron que la F0 aumenta en condiciones de ruido (2005: 202).

Gfroerer and Wagner (*apud* Jessen *et al*, 2005) afirman que en los estudios forenses se presenta un incremento en la F0, debido, quizá, a que los hablantes en los experimentos tienen niveles de estrés más elevados o están en un estado emocional alterado, es decir, las condiciones del laboratorio pueden resultar muy distintas a las que se presentan en un caso real en el ámbito forense (2005:175).

Hemos presentado algunos de los trabajos que se han realizado sobre la frecuencia fundamental en la identificación forense de voz. Observamos que la mayoría de las investigaciones que mencionamos utilizan un corpus ya desarrollado por los investigadores. Como ya habíamos mencionado en el capítulo 1, los trabajos sobre fonética forense en México son pocos, no se cuenta con un grupo que trabaje de manera interdisciplinaria en el área, por esta razón tampoco contamos con corpus que estén dirigidos al estudio de la voz con fines forenses.

4. METODOLOGÍA

En este apartado se presentará el proceso de construcción, recopilación y análisis del corpus. En primer lugar, se hablará de la importancia de un corpus en las investigaciones sobre fonética; en segundo lugar, hablaremos de la forma en la que fue construido el corpus, es decir, mencionaremos qué aspectos de la lengua consideramos para elegir la vocal, los contextos de ésta y las palabras que utilizamos para armar el texto que lo compone; posteriormente veremos de qué manera recopilamos las grabaciones. Finalmente explicaremos la forma en la que se realizó el análisis acústico y estadístico.

4.1 Corpus

Un corpus es la base de la mayoría de los estudios lingüísticos ya que en éste se basan las hipótesis sobre las que se trabaja. Los corpus pueden ser de dos tipos: orales y textuales. En fonética los corpus que necesitamos son orales porque en ellos se encuentran muestras de habla de las que podemos extraer diferentes parámetros, como la frecuencia fundamental, los formantes vocálicos, unidades entonativas, etc. Sinclair (1996) define un corpus como una colección de elementos lingüísticos que son seleccionados y ordenados de acuerdo con criterios explícitos, con la finalidad de ser usada como muestra de la lengua. Según Torruela *et al*, los corpus fonéticos se pueden clasificar en tres tipos: los que están orientados a la descripción fonética de la lengua, los que se utilizan para el desarrollo de sistemas en el ámbito de las tecnologías del habla y los que consisten en transcripciones ortográficas de la lengua hablada (1999:13). En esta tesis de licenciatura, debido al tipo de estudio, nos inclinamos más por los corpus orientados a la descripción de la lengua, éstos “consisten tradicionalmente en materiales grabados en condiciones acústicas óptimas que permitan su posterior análisis experimental en el laboratorio” (1999:14). Una de las características de estos corpus es que tienen un diseño cuidadoso, que se basa en el inventario de elementos segmentales y suprasegmentales de la lengua, tienen un tamaño pequeño debido a que las grabaciones no se hacen con muchos hablantes (1999:14).

El principal objetivo al construir el corpus de esta tesis, fue que éste mostrara la realidad fonética del español de la ciudad de México, es por ello que se tomó como base el estudio de la frecuencia de los fonemas del español de la ciudad de México que realizó Cuétara (2004). Como es bien sabido “la función principal de un corpus, tanto textual como oral, es establecer la relación entre la teoría y los datos; el corpus tiene que mostrar a pequeña escala cómo funciona una lengua natural” (Llisterri y Torruela, 1999: 2). El corpus que recopilamos evidencia la manera en la que se comportan tanto la vocal como los contextos en los que se ha analizado.

Otro aspecto importante en la constitución de un corpus oral es el tipo de habla que se quiere obtener, ya sea controlada o espontánea. A la primera se le conoce generalmente como habla de laboratorio y “engloba cualquier tipo de habla obtenida en un entorno controlado [...] Este término se ha aplicado generalmente a la lectura de listas de palabras frases o textos” (Aguilar *et al*, 1994: 23), a pesar de ser habla que se obtuvo de manera controlada, se intenta que sea lo más semejante al habla que se desarrolla de manera natural. Por su parte, el habla espontánea se compone de los tipos de habla que no sean lectura, sin embargo, es difícil decir que sólo el habla que no sea lectura es espontánea porque dentro de ésta existen distintos estilos. Además resulta complicado que alguien hable de forma natural frente a una grabadora, la espontaneidad se puede ver afectada por la personalidad del hablante y por el conocimiento o interés de un tema durante la entrevista. Uno de los recursos que se ha utilizado últimamente es la grabación de programas radiofónicos o debates políticos, debido a la facilidad para obtener las muestras de habla.

Un punto importante en la constitución de un corpus es la elección de los hablantes. Según Llisterri (2004), se deben considerar cuatro características:

1. Sexo
2. Edad
3. Variedad dialectal
4. Variedad sociolectal

La elección de los hablantes depende en gran medida del objetivo de la investigación, se puede hablar sólo de un sexo, p. ej. El habla de los hombres adolescentes en la ciudad de México, es aquí cuando entran en juego los demás puntos. La edad es un factor fundamental, si queremos analizar parámetros acústicos debemos tomar en cuenta la edad

de los hablantes, es bien sabido que la voz de las personas cambia conforme su crecimiento. En la adolescencia tanto hombres como mujeres sufren cambios en la voz, de la misma forma sucede en la menopausia y la andropausia (Dinville, 1996).

Otro punto importante es la variedad dialectal y sociolectal de los hablantes, en el ejemplo que planteamos se habla de la variedad dialectal de la ciudad de México, bien pudiera ser la variedad de Oaxaca o Chihuahua; la variedad sociolectal corresponde al estrato social al que pertenece un hablante, generalmente este aspecto se mide según la educación del mismo, puede ser habla popular (educación básica) o culta (a partir de educación media superior). En un corpus siempre se tienen que controlar estos aspectos porque de ello dependerá la uniformidad de los datos.

Después de tener definido el objetivo del corpus, el contenido lingüístico y los hablantes, se debe pensar en la forma en la que se va a recopilar dicho corpus. Como ya se ha mencionado, cada aspecto depende del objetivo de la investigación, el corpus puede ser grabado en cabinas anecoicas, una habitación con ruido o bien en un ambiente al aire libre. En este punto se debe elegir el material con el que se va a realizar la grabación, que puede ser una grabadora digital, un micrófono o un programa computacional que permita la grabación del habla.

Cuando el corpus se tiene recopilado, se procede al etiquetado o transcripción ya sea ortográfica, fonológica o fonética, según sea el caso. Finalmente se organizan los datos obtenidos de acuerdo al manejo que se le vaya a dar a éstos.

Una vez que se ha hablado sobre la forma en la que se construye y recopila un corpus, ahondaremos más acerca de la constitución del que hemos compuesto en esta tesis.

4.1.1 Composición del corpus

Como ya se mencionó, para la elección del fonema que se analizará en esta tesis, se tomó como fundamento el estudio que realizó Cuétara (2004) en su tesis de Maestría *Fonética de la ciudad de México. Aportaciones desde las tecnologías del habla*. En dicho trabajo el autor hace un recuento de los fonemas (con sus correspondientes alófonos) más frecuentes en el español de la ciudad de México. Uno de los lineamientos para realizar un corpus según Llisterri *et al* es que “para que el corpus siga siendo representativo de la lengua debe

considerarse [...] la frecuencia de aparición de cada uno de los sonidos y de cada posibilidad de combinación de los mismos” (1994), esta es otra de las razones por la que elegimos el estudio de Cuétara, el cual es minucioso y muestra el comportamiento fonético del español de la ciudad de México. Vemos que en los datos que presenta se muestra que entre las vocales, la de mayor frecuencia es la media anterior /e/ que tiene una aparición de 35.79% del total de las vocales (2004: 78). Por otra parte las consonantes alveolares representan un porcentaje considerable: la fricativa sorda /s/ (20.99%), la nasal /n/ (10.21%), la lateral /l/ (10.39%), por su parte la vibrante simple /r/ representa el 10.81 %, mientras que la múltiple /r/ tiene menor presencia con 0.74% (2004:79).

En el ámbito forense es importante conocer la frecuencia de aparición de cada uno de los fonemas porque de esta manera podemos enfocar el análisis forense en los sonidos que más aparecen en el inventario de los estudios de frecuencia. Es importante hacer el análisis de esta forma porque en el área de fonética forense no es posible definir el contexto ni las condiciones en las que se encuentra la grabación que se va a analizar. En casos así, la frecuencia ayuda a prever cuáles son los fonemas que más encontraremos en una grabación real. Dadas las razones anteriores, el contexto con el que se trabajará es la vocal media anterior en contexto alveolar, salvo la vibrante múltiple porque, como se vio en las estadísticas, es la consonante que tiene menos frecuencia en comparación con los demás fonemas alveolares estudiados.

Debido al carácter experimental de esta investigación se decidió analizar la vocal /e/ y los contextos alveolares /n, l, r, s/ en su forma prototípica, excepto en el contexto alveolar lateral, en donde hemos incluido cuatro palabras que contienen la vocal ante /x/. La teoría menciona que en este contexto la vocal media anterior es semiabierta, sin embargo, debido a la dificultad que presentó encontrar palabras que coincidieran con la coherencia del texto, decidimos hacer esta excepción. La forma prototípica tanto de la vocal como de las consonantes es la más frecuente en el español de la ciudad de México.

En la figura 4.1 se muestran las condiciones en las que la vocal y las consonantes se producen en su forma prototípica (Cuétara, 2004).

| Fonema | Realización prototípica |
|--------|--|
| / e / | Libre en sílaba abierta, en sílaba cerrada trabada por /m,n,s,d/ |
| / s / | Inicio de sílaba o palabra, sin condición. |
| / n / | Libre en sílaba abierta sin condición silábica. |
| / l / | Inicio de sílaba o palabra. |
| / r / | Inicio de sílaba o palabra. |

Tabla 4.1 Condiciones para la realización prototípica de los fonemas alveolares

El parámetro que se analizó es la frecuencia fundamental. Decidimos analizar la vocal tanto en su realización tónica como átona porque, como ya se mencionamos, la frecuencia fundamental cambia si la vocal está acentuada o no. Desde el primer estudio de fonética acústica, realizado por Bolinger y Hodapp se sabe que el tono es el elemento principal en el acento español, mientras que la intensidad y la duración son factores secundarios (Martínez Celdrán, 1988: 136). Para Quilis (1999) el índice principal del acento es la frecuencia fundamental, ya sea sola o acompañada de la duración, “cuando no actúan ni la frecuencia fundamental, ni la duración, es la intensidad la que pone de relieve el prosodema acentual” (1999: 400). La importancia del tono no significa que necesariamente haya una elevación de éste por encima del cuerpo tonal de la frase (1988: 136). La frecuencia fundamental aumenta:

Quando lo hace la corriente de aire espiratorio, la tensión de las cuerdas vocales o ambas simultáneamente [...] Cuando se eleva la lengua para la articulación de una vocal media o alta, la laringe también sube y los músculos se tensan. De este modo aumenta la tensión de las cuerdas vocales y se genera un aumento en el número de las vibraciones (Quilis, 1999: 416)

Con lo anterior se justifica el hecho de que analicemos ambas realizaciones de la vocal. En otras palabras, nuestro objetivo es observar si hay una diferencia en la F0 independientemente de la tonicidad de la vocal, en el ámbito forense. Para ello se seleccionaron veinte palabras que contuvieran la vocal átona y veinte en las que la vocal fuera tónica.

En la tabla 4.2 se pueden ver las palabras que se eligieron con la vocal átona.

| /ne/ | /le/ | /re/ | /se/ |
|------------------|--------------------|------------------|-------------------|
| /ne.bu.ˈlo.so/ | /le.ˈga.do/ | /be. re.ˈni.se/ | /ne.se.ˈsi.dad/ |
| /ne.baɪ / | /e.le.ˈxiɪ / | /se. re.ˈmo.nia/ | /se.le.bra.ˈsion/ |
| / ne.ˈru.da / | /a.le.ˈxo/ | /pa. re.ˈsio/ | /se.ˈk re.to/ |
| /ne.se.ˈsi.dad/ | /se.le. bra.ˈsion/ | /me. re.ˈsi.do/ | /se.ˈma.na/ |
| /te.ne.ˈ bro.so/ | /le.ˈxa.no/ | /ˈai. re/ | /se.le. ri.ˈdad/ |

Tabla 4.2 Contextos analizados para la vocal átona

En la tabla 4.3 se encuentran las palabras, que hemos elegido, con la vocal tónica:

| /ˈne/ | /ˈle/ | /ˈ re/ | /ˈse/ |
|----------------|--------------|---------------|--------------|
| /ˈne.gra/ | /a.ˈle.gre/ | /i.ˈ re.ne/ | /ˈse.na/ |
| /ka.ˈne.la/ | /tsa.ˈle.ko/ | /a.ˈ re.tes/ | /ˈse.ko/ |
| /a.ˈne.mo.nas/ | /ˈle.xos/ | /de.ˈ re.cho/ | /ˈse.saɪ/ |
| /e.ˈne. ro/ | /bio.ˈle.ta/ | /ma.ˈ re.a/ | /ˈse.da/ |
| /a.ˈne.lo/ | /ko.ˈle.ta/ | /a.ˈ re.na/ | /sin.ˈse.ro/ |

Tabla 4.3 Contextos analizados para la vocal tónica

En esta investigación decidimos trabajar con habla controlada, es decir, leída. Nos parece más natural el hecho de leer un texto que palabras sueltas. Es por ello que incluimos la lista de palabras en un texto. Fue fundamental que los hablantes no supieran qué es lo que se buscaba en el corpus porque de otra forma controlarían cada una de las palabras que iban diciendo; en la lectura se preocuparon más por la manera en la que enunciaron de manera general y no por palabras sueltas, de esta forma, se podría decir que el habla leída se acerca un poco a la realidad de la pronunciación espontánea del español.

El texto quedó conformado de la siguiente manera:

En enero Irene sintió la necesidad de tomar un merecido descanso, lejos de todo, por lo que su madre organizó una cena de despedida. La celebración fue muy alegre. Berenice quiso elegir a los invitados, entre éstos había un hombre, que estudiaba

Derecho; él, en secreto, sentía un cariño sincero por Irene. Ella vio la lista y le pareció apropiada. A la semana siguiente, compró un vestido de seda negra, unos aretes y un prendedor para su coleta.

El ambiente de la cena era nebuloso, el jardín estaba cerca del mar, la marea, que comenzaba a subir, cubría la arena, a ratos el aire se respiraba frío y seco, todo tenía un aspecto tenebroso. Las mesas estaban decoradas con anémonas color violeta, el mar se veía lejano desde la terraza. Entonces de un momento a otro, César comenzó a recitar el legado poético de Neruda, mientras comenzaba a nevar. Ella se alejó de la fiesta y él, con un anhelo en el corazón, la siguió con celeridad, la tomó de las manos y en seguida, con una ceremonia la invitó a bailar, Irene pudo percibir el olor a canela de su chaleco.

4.1.2 Recopilación del corpus

Para recopilar el corpus se eligieron diez hablantes, cinco hombres y cinco mujeres, el rango de edad está comprendido entre los 19 y 24 años de edad, además se seleccionaron aquéllos cuya complexión fuera regular, esto con el fin de no tener contrastes entre los hablantes. En la tabla 4.4 se muestran los datos de los hablantes femeninos y en la tabla 4.5, los datos de los hablantes masculinos.

| Hablante | Edad | Estatura | Peso | Frec.Act.F. ¹² | IMC ¹³ | Complexión |
|-----------|------|----------|------|---------------------------|-------------------|----------------|
| M1 | 21 | 1.58 | 51 | 3 H/S | 20.42 | Normal |
| M2 | 20 | 1.6 | 67 | 1 H/S | 26.17 | lig. sobrepeso |
| M3 | 20 | 1.5 | 58 | 2 H/S | 25.78 | lig. sobrepeso |
| M4 | 20 | 1.58 | 58 | Nada | 23.23 | Normal |
| M5 | 22 | 1.6 | 49 | Nada | 19.14 | Bajo |

Tabla 4.4 Datos de los hablantes femeninos

¹² Frecuencia con la que realizan alguna actividad física, tomamos en cuenta las horas a la semana.

¹³ IMC: Índice de Masa Muscular, es una medida que se usa para relacionar el peso y la talla de una persona.

| Hablante | Edad | Estatura | Peso | Frec.Act.F. | IMC | Complexión |
|----------|------|----------|------|-------------|-------|----------------|
| H1 | 22 | 1.6 | 61 | 2 H/S | 23.83 | Normal |
| H2 | 24 | 1.73 | 59 | Nada | 19.71 | Bajo |
| H3 | 21 | 1.77 | 80 | Nada | 23.77 | Normal |
| H4 | 19 | 1.7 | 69 | 1 H/S | 23.88 | Normal |
| H5 | 21 | 1.6 | 66 | 1 H/S | 25.78 | lig. sobrepeso |

Tabla 4.5 Datos de los hablantes masculinos

Debido al área en la que está inscrita esta tesis decidimos hacer las grabaciones sin aparatos especializados y en un ambiente natural, pero cuidando que éstas no perdieran información acústica. Es por ello que se realizaron al aire libre, se procuró que fuera en un lugar solitario para evitar ruidos como coches, ambulancias, música, etc., que pudieran afectar el registro de la grabación; el aparato que se utilizó para realizarlas fue un reproductor *Sony walkman*, que graba voz en formato WAV (Waveform Audio File). Como podemos observar tanto el ambiente y el equipo de grabación pueden ser utilizados por cualquier sujeto, tal como suele suceder en el ámbito forense.

Como ya se mencionó el objetivo de este estudio es observar si hay o no cambios en la frecuencia fundamental de la voz antes y después de haber realizado actividad física, es por eso que se grabó a los hablantes en estado de reposo y posteriormente se les pidió que saltaran la cuerda 40 ocasiones e inmediatamente después de haber saltado leyeron el mismo texto. El objetivo de que saltaran la cuerda fue aumentar el ritmo cardiaco de los hablantes y de esta manera simular un estado de estrés.

Es importante mencionar que las consecuencias del estrés sobre el cuerpo son cruciales para entender la salud del sistema cardiovascular. En los experimentos en los que se ha tratado de ver la relación entre el estrés y el sistema cardiovascular, afirman que es muy importante saber si el individuo percibe control sobre la situación (estrés positivo) o no (estrés negativo). Se ha observado que las situaciones que demandan más exigencias disminuyen la percepción de control del sujeto sobre la situación; es menos probable que un individuo con menor control de la situación finalice la ejecución que se le demanda en el laboratorio (Breva, A. *et al*, 2000: 4).

En los experimentos de laboratorio el sujeto generalmente percibe control sobre la situación, sin embargo, en ciertas ocasiones, cuando la actividad demandada se acerca a una

situación de estrés real, el individuo pierde la sensación de control en la actividad que ejecuta.

En el experimento que realizamos observamos que después de que los hablantes saltaron la cuerda, la percepción de control no fue la misma para todos. Como se esperaba, se presentaron cambios en la forma de emisión de los hablantes, en algunos se notó una mayor fatiga que en otros la mayoría concluyeron el ejercicio cansados, con falta de oxígeno y sólo en tres hablantes la fatiga se notó menos. Pero, no basta con ver que si hay afectación en la fonación después de realizar un esfuerzo, lo que realmente nos hablará de lo que pasó es el análisis de datos, que presentaremos en el apartado de resultados.

4.2 Análisis acústico

Una vez recopilado el corpus se hizo un análisis acústico con ayuda del programa computacional *Praat*. Se realizó un etiquetado fonético, es decir “a cada segmento se le asocia una etiqueta que lo define en términos fonéticos o fonológicos (*labelling*) y se lleva a cabo la alineación (*alignment*) entre la señal sonora y las etiquetas, obteniendo una representación que puede compararse a la de una partitura musical con la letra correspondiente” (Torruela *et al*, 1999: 26). La forma en la que se realizó el etiquetado fue de tal manera que permitiera obtener los resultados de manera automática, por lo tanto, cuando en la grabación apareciera la palabra que contiene el contexto /ne/ tónica, p. ej.: /**en**ero/, se seleccionó la vocal después de la /n/ y se delimitó el inicio y el final. El etiquetado quedó de la siguiente manera: cuando la vocal estaba en posición tónica se marcaba con una T y en posición átona con una A; así, en el caso del contexto tónico nasal el etiquetado quedó como /neT/ y /neA/ en el caso del átono.

En la figura 4.1 se muestra el etiquetado de una vocal media anterior precedida por consonante nasal, la realización de esta vocal es tónica.

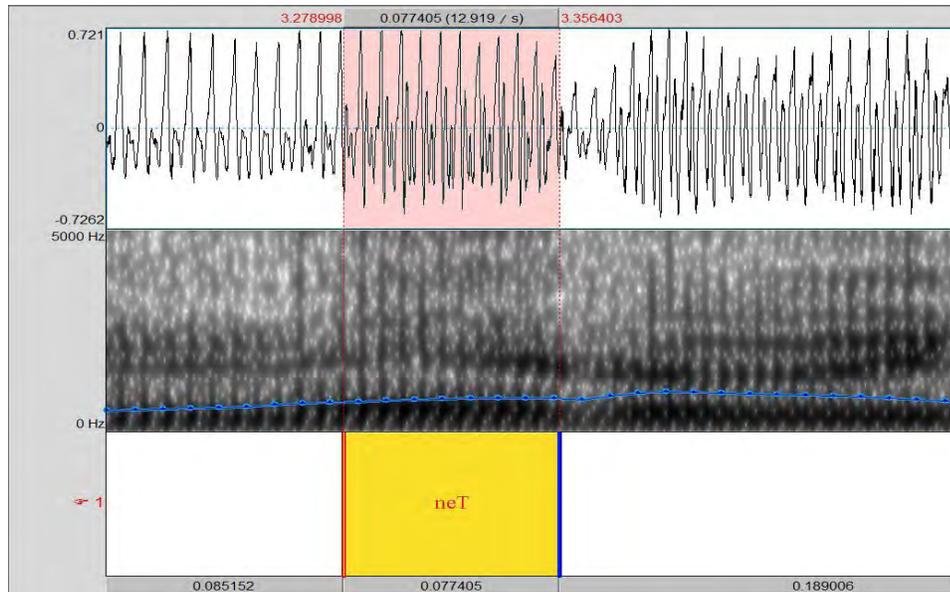


Figura 4.1 Etiquetado de /ne/ Tónica

La frecuencia fundamental es la línea que se marca en puntos azules, en este caso es aproximadamente de 140 Hz. El etiquetado es la base de nuestro análisis porque nos proporciona los datos numéricos, en este caso, la información numérica que se consideró fue el punto central de la vocal como puede observarse en la figura 4.2.

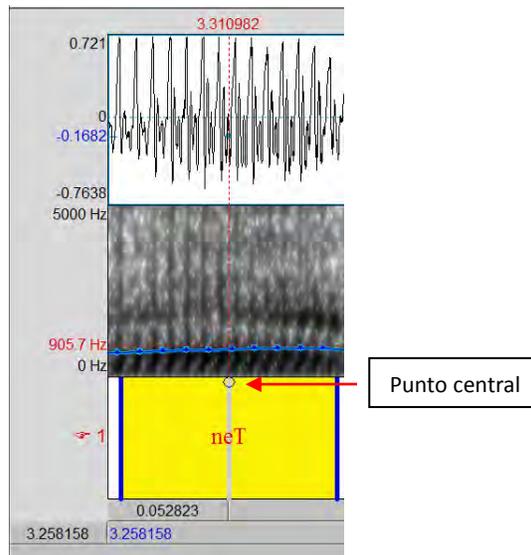


Figura 4.2 Etiquetado que muestra el punto central de la vocal

4.3 Análisis estadístico

Una vez que se ha explicado cómo se realizó el análisis acústico, es fundamental detallar de qué manera se desarrolló el análisis estadístico que utilizamos: la diferencia o contraste de medias. En principio debemos tomar en cuenta que el objetivo de un análisis estadístico es extraer conclusiones a partir de datos particulares, por ello es muy importante seleccionar apropiadamente la información que se va a examinar. (Pardo *et al*, 2002: 169) Como ya lo mencionamos, en esta tesis analizamos la frecuencia fundamental en dos momentos: antes y después de elevar el ritmo cardiaco del hablante, mediante actividad física, para simular un estado de estrés. Antes de pasar a la descripción del análisis estadístico, debemos comentar cómo se realizó el vaciado de datos.

4.3.1 Vaciado de datos

Una vez que se obtuvieron los valores de la frecuencia fundamental proseguimos a vaciar la información en hojas de cálculo, los datos se ordenaron en tablas individuales para cada uno de los hablantes, éstas se organizaron atendiendo dos criterios: la tonicidad de la vocal y cada uno de los contextos en los que está enmarcada. Es decir, para el hablante H1 se hicieron dos tablas, la primera para la realización átona de cada uno de los contextos /ne/ /le/ /re/ /se/ y la segunda para la realización tónica de los mismos contextos.

En la tabla 4.6 se muestra cómo se realizó la organización de los datos de la vocal átona y en la tabla 4.7 se aprecia el ejemplo para la vocal tónica en todos sus contextos. En ambas se observa la realización y el contexto en la columna alineada a la izquierda. En la parte superior de la tabla vemos el valor de la F0 y en la fila que sigue indicamos el *Antes* y el *Después* de realizar ejercicio.

| Realización | F0 | | | F0 | | | F0 | | | F0 | |
|-------------|--------|---------|-----|--------|---------|-----|--------|---------|-----|--------|---------|
| | Antes | Después | | Antes | Después | | Antes | Después | | Antes | Después |
| neA | 136.23 | 167.15 | leA | 131.36 | 134.85 | reA | 137.67 | 175.56 | seA | 135.64 | 127.09 |
| neA | 120.55 | 148.54 | leA | 133.25 | 145.61 | reA | 141.4 | 141.07 | seA | 154.96 | 172.84 |
| neA | 116.85 | 105.22 | leA | 137.4 | 141.45 | reA | 119.53 | 133.83 | seA | 148.61 | 148.98 |
| neA | 132.48 | 137.92 | leA | 145.5 | 156.94 | reA | 112.01 | 114.74 | seA | 143.42 | 187.28 |
| neA | 112.95 | 116.35 | leA | 136.49 | 146.45 | reA | 149.49 | 160.35 | seA | 120.39 | 127.88 |

Tabla 4.6 Ejemplo de organización de datos con realización átona

| Realización | F0 | | | F0 | | | F0 | | | F0 | |
|-------------|--------|---------|-----|--------|---------|-----|--------|---------|-----|--------|---------|
| | Antes | Después | | Antes | Después | | Antes | Después | | Antes | Después |
| neT | 134.07 | 173.91 | leT | 123.27 | 138.21 | reT | 146.54 | 149.78 | seT | 158.35 | 154.01 |
| neT | 133.92 | 140.03 | leT | 110.68 | 115.38 | reT | 155.68 | 162.55 | seT | 125.88 | 135.6 |
| neT | 141.04 | 162.94 | leT | 153.33 | 155.17 | reT | 109.41 | 112.71 | seT | 160.97 | 183.91 |
| neT | 145.56 | 164.67 | leT | 140.78 | 121.63 | reT | 126.92 | 169.51 | seT | 142.14 | 154.57 |
| neT | 150.91 | 184.49 | leT | 117.57 | 132.57 | reT | 154.16 | 166.44 | seT | 149.62 | 176.6 |

Tabla 4.7 Ejemplo de organización de datos con realización tónica

4.3.2 Contraste de hipótesis

Un contraste de hipótesis también puede ser llamado prueba de significación o prueba estadística, ésta permite decidir si una proposición acerca de una población o muestra puede ser mantenida o rechazada (Pardo *et al*, 2002:177). El primer paso para obtener una hipótesis es plantear un problema que nos permita proponer algún tipo de solución, generalmente es una afirmación, a ésta se le denomina hipótesis científica. En nuestro caso, podemos ejemplificar la hipótesis científica más general de este trabajo, que queda como sigue:

- La frecuencia fundamental, tanto en hombres como en mujeres, de la vocal media anterior, en contexto alveolar es igual antes y después de aumentar el ritmo cardiaco.

Una vez planteada la hipótesis científica podemos verificarla y para ello debemos formular una hipótesis estadística, en nuestro caso es la siguiente: $\mu_a = \mu_d$. Es decir, la media de la F0 antes de aumentar el ritmo cardiaco es igual a la media de la F0 después de aumentarlo.

Las hipótesis estadísticas son representadas con la letra H seguida de una afirmación que da contenido a la hipótesis. No debemos confundir la hipótesis científica con la hipótesis estadística, la primera se refiere a algún aspecto de la realidad y la segunda a algún aspecto de la distribución de la probabilidad (Pardo *et al*, 2002: 180). Cabe mencionar que el contraste de hipótesis se da entre:

1. La hipótesis nula (H_0)
2. La hipótesis alternativa (H_1)

La H_0 es la hipótesis que suele someterse a contraste, mientras que la H_1 es la negación de la primera. Se puede afirmar que mientras que la H_0 es exacta (implica igualdad), la H_1 es inexacta (refiere a que tal cosa es distinta, mayor o menor que otra). El contraste de

hipótesis es exclusivo, es decir, si la H_0 es verdadera necesariamente la H_1 es falsa y viceversa (Pardo *et al*, 2002: 181).

En esta tesis la hipótesis principal se ha trabajado de diferentes formas, con el fin de contrastar la F_0 en distintas situaciones, es decir, se ha partido de la hipótesis principal para plantear distintas hipótesis de acuerdo con las variables por las que está compuesto nuestro corpus. P. ej., la siguiente hipótesis está enfocada en un aspecto particular de la investigación.

H_0 : La frecuencia fundamental, en los hombres, de la vocal media anterior átona en contexto alveolar es igual antes y después de aumentar el ritmo cardíaco.

Otra de las hipótesis que hemos planteado es para cada uno de los contextos alveolares, en este caso, nuestra H_0 será:

H_0 : La frecuencia fundamental, en las mujeres, de la vocal media anterior tónica en contexto nasal alveolar es igual antes y después de aumentar el ritmo cardíaco.

Como ya hemos visto la hipótesis nula se puede aceptar o rechazar, entonces, nuestro objetivo es saber si las hipótesis ejemplificadas se aceptan o no.

Para llevar a cabo el análisis estadístico nos apoyamos en *R-project* (*R*), un programa computacional, en donde se pueden realizar diversos análisis estadísticos. *R* fue diseñado por Robert Gentleman y Ross Ihaka, miembros del departamento de la Universidad de Auckland, en Nueva Zelanda. Una de las grandes ventajas de *R* es que hoy en día es fruto del esfuerzo de miles de personas en todo el mundo que colaboran en su desarrollo y el acceso al dicho programa es gratuito (Sáez Castillo, 2010:10).

En *R* se deben insertar instrucciones para poder obtener los resultados que se esperan, la prueba final se denomina *t.test*, pero antes de poder aplicarla se debe de cumplir con dos supuestos, mismos que explicaremos a continuación.

4.3.3 Supuestos de la prueba *T.test*

El análisis estadístico tiene como objetivo arrojar resultados confiables, sin embargo, como ya lo mencionamos, los datos deben de cumplir con ciertos supuestos antes de realizar la prueba estadística denominada *t.test*. Un supuesto es un dato que se asume como cierto en la planificación de una investigación, pero para poder llevar a cabo el contraste de hipótesis

de manera correcta, los supuestos se deben de comprobar. El primer supuesto consiste en verificar la normalidad de los datos que se van a comparar y se indica en *R* con la instrucción *shapiro.test*. Si el *p-value* es mayor a 0.05 se entiende que los datos sí siguen una distribución normal y por lo tanto se puede aplicar la prueba *t.test*. Si el *p-value* es menor a 0.05 no se puede asumir que los datos siguen una distribución normal y por lo tanto deberá utilizarse la prueba *wilcox.test* que es una prueba no paramétrica y que no asume la normalidad los datos.

El segundo supuesto es el de la homogeneidad de varianzas y en *R* se indica con la instrucción *var.test*. Al igual que en el supuesto anterior se revisa el valor del *p-value* y si es mayor a 0.05 asumimos que la varianza de los datos es homogénea. Si el valor de *p-value* es menor a 0.05 no podemos asumir la homogeneidad de varianzas y se tiene que indicar en las instrucciones de la prueba *t.test* con la orden *var.equal=false*.

En los casos en los que ninguno de los supuestos se cumple, es decir normalidad y homogeneidad de varianza de los datos, se debe aplicar la *wilcox.test*.

4.3.4 Prueba *T.test*

La prueba *T* para dos muestras relacionadas permite contrastar hipótesis referidas a la diferencia entre dos medias relacionadas (Pardo *et al*, 2002: 265). “Para que el estadístico *T* se ajuste apropiadamente al modelo de distribución de probabilidad *t* de Student, es necesario que la población muestreada sea normal” (2002: 257). Sin embargo, entre mayor sea la cantidad de datos a analizar, es decir, a partir de 20 o 30 casos “el ajuste de *T* a la distribución *t* [...] es lo suficientemente bueno incluso con poblaciones originales sensiblemente alejadas de la normalidad” (2002: 257).

En *R* esta prueba se introduce con la instrucción *t.test* y, al igual que con los supuestos, si el valor del *p-value* es mayor a 0.05 aceptamos la H_0 , pero si es menor a 0.05 rechazamos H_0 y aceptamos H_1 porque como ya mencionamos arriba, estas dos hipótesis son excluyentes y complementarias a la vez, es decir, si una H_0 se rechaza, inmediatamente aceptamos H_1 . En nuestro caso si en la hipótesis:

H_0 : La frecuencia fundamental, en los hombres, de la vocal media anterior átona en contexto alveolar es igual antes y después de aumentar el ritmo cardiaco.

Obtenemos un *p-value* menor a 0.05 rechazaremos la H_0 , que designa igualdad en el parámetro analizado y aceptaremos H_1 , que indica diferencias significativas en la frecuencia fundamental. Dicho de otra manera, si aceptamos H_0 estamos asumiendo que la frecuencia fundamental no presenta diferencias significativas antes y después de aumentar el ritmo cardiaco; en cambio, si aceptamos H_1 , nuestra conclusión será que sí existe una diferencia significativa en la F_0 después de realizar aumentar el ritmo cardiaco de los hablante.

4.3.5 Análisis descriptivo

Este análisis es importante porque nos permite describir de manera gráfica los resultados que se obtuvieron en el análisis estadístico de diferencia de medias. Existen muchas formas de representar los resultados obtenidos. En esta tesis hemos decidido utilizar el diagrama de caja o *boxplot*.

4.3.6 Diagrama de caja o *boxplot*

El análisis de diferencia de medias se ilustra mejor con el diagrama de caja o *boxplot*. En la siguiente figura se aprecia cómo está compuesto un gráfico de este tipo.

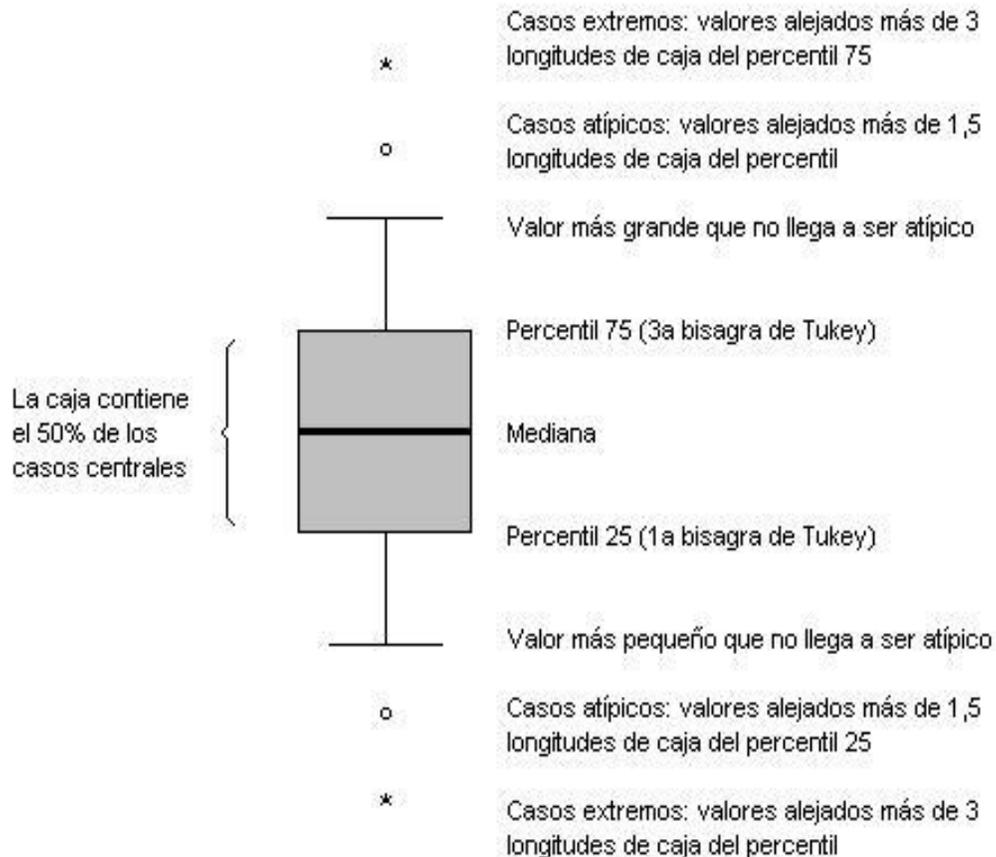


Figura 4.3 Diagrama de caja o *boxplot*. Fuente (Pardo *et al*, 2002 *apud* López, 2010: 106)

Como su nombre lo indica son representaciones en forma de caja que contienen los datos del análisis estadístico. La caja contiene el 50% de los casos centrales, la altura de ésta muestra el grado de dispersión. Dentro de la caja se observa una línea negra que es la mediana. Las líneas que salen de la caja son los valores atípicos, también se conoce a estas líneas como bigotes. Fuera de los bigotes se muestran los datos atípicos en forma de bolitas, y con un asterisco se representan los casos extremos, es decir, los valores más alejados de la caja. “El diagrama incluye [...] los percentiles 25 y 75 (en realidad son las bisagras de Tukey), y una serie de valores (atípicos, extremos) que junto con la mediana y la propia caja proporcionan información bastante completa sobre, entre otras cosas, el grado de dispersión de los datos y el grado de asimetría de la distribución” (Pardo *et al*, 2002: 211).

En el diagrama de caja hay dos aspectos que nos interesan particularmente, el primero se refiere a la mediana, ésta representa el valor medio que hay entre los valores que

se comparan, p. ej., si tenemos los valores 2, 2, 3, **5**, 6, 7, 9, 11. El valor 5 es el que se encuentra en la mitad de los datos. El segundo concepto corresponde a la variabilidad, que se refiere a la variación que hay entre los datos con los que se realiza la comparación, en este caso, entre más pequeña sea la caja, menor es la variabilidad y sucede de modo contrario cuando la caja es más larga. Si comparamos el ejemplo anterior contra el siguiente conjunto de valores 4, 4, 5, 5, 6, 7 podríamos concluir que la variabilidad de éste es menor pues va de 4 a 7, mientras que el anterior va de 2 a 11.

5. RESULTADOS

De la recopilación del corpus se obtuvieron 800 muestras de la frecuencia fundamental de la vocal media anterior: 400 corresponden a la realización tónica y 400 a la realización átona. Como ya se mencionó, la información se agrupó en tablas, una por cada hablante, y se ordenó de cuatro maneras diferentes, es decir, se comparó la F0 antes y después de realizar la actividad física y se organizó de acuerdo al cruce de las siguientes variables: la tonicidad, los contextos consonánticos, el género y los hablantes; todo esto con el fin de hacer un análisis estadístico más completo. A continuación veremos con más detalle cuáles son estas formas.

La primera comparación se realizó de acuerdo a dos criterios: la tonicidad de la vocal y la diferencia de género. Es importante señalar que se consideró a los hablantes como dos conjuntos según su género y no de forma individual, así, tanto en el grupo de hombres como en el de mujeres se analizó la variación de átonas y tónicas.

La segunda comparación es similar a la primera, sólo que aquí la variación entre átonas y tónicas se analizó con respecto a los datos arrojados por cada hablante y no por género como en el primer caso.

En la tercera confrontación se recabó la información de cada uno de los contextos que preceden a la vocal media anterior, específicamente, las consonantes alveolares; también se consideró la diferencia entre la tonicidad, asimismo vimos a los hablantes como un grupo de acuerdo a su género.

El último análisis es el más particular de todos porque se examinaron los siguientes aspectos en cada uno de los hablantes, por un lado la realización de la vocal y por el otro el contexto que la precede; es decir, se comparó cada uno de los contextos en su respectiva realización.

En primer término, decidimos agrupar los datos de esta manera, para tener una visión general de qué es lo que sucede con la F0 en la realización tónica y átona. En segundo lugar, para observar lo que pasa con la F0 en los diferentes contextos alveolares en los que está enmarcada la vocal. Y finalmente, para tener una visión específica de los resultados, tanto por género como por hablante. Es importante señalar que antes de realizar

la prueba final *t.test* comprobamos que el supuesto de normalidad y el de homogeneidad de varianzas se cumplieran. Cuando no fue así proseguimos a realizar la prueba correspondiente (ver 4.3.3).

5.1 Resultados por género

Como ya se mencionamos en la primera comparación sólo hicimos la distinción entre la realización. Se cotejaron las 20 emisiones átonas de la vocal en cada hablante masculino (ver tabla 4.2), que si se multiplican las 20 emisiones de la vocal por los 5 hablantes masculinos obtenemos en total 100 emisiones. La confrontación entre estas 100 emisiones se realizó antes y después de hacer ejercicio. Hicimos lo mismo para el caso de la realización tónica. Hicimos el mismo examen con los datos de las mujeres en ambas realizaciones.

5.1.1 Resultados de los hablantes masculinos

En la tabla 5.1 se presentan los resultados de la comparación entre los hablantes masculinos. En la columna izquierda se indica la realización de la vocal y en la derecha se puede ver el valor que se obtuvo del *p-value*. En la tabla se observa que los datos de los hombres, tanto en la realización átona como en la tónica, muestran un *p-value* menor a 0.05, lo que significa que sí hay una diferencia significativa en la F0 de los hablantes antes y después de haber realizado actividad física.

| Realización | <i>p-value</i> |
|-------------|------------------|
| Átona | 0.01184 |
| Tónica | 0.0002338 |

Tabla 5.1 Resultados generales: hablantes masculinos

5.1.2 Resultados de los hablantes femeninos

En la tabla 5.2 observamos los resultados de la comparación entre el antes y el después de los hablantes femeninos. Vemos que también aquí hay una diferencia significativa en las dos realizaciones.

| Realización | <i>p-value</i> |
|-------------|--------------------|
| Átona | 0.000003237 |
| Tónica | 0.0005308 |

Tabla 5.2 Resultados generales: hablantes femeninos

En la figura 5.1 se muestran los diagramas de caja (*boxplot*) de los hablantes femeninos y masculinos en ambas realizaciones. Presentamos los resultados de manera conjunta porque uno de los objetivos fue hacer la comparación por género.

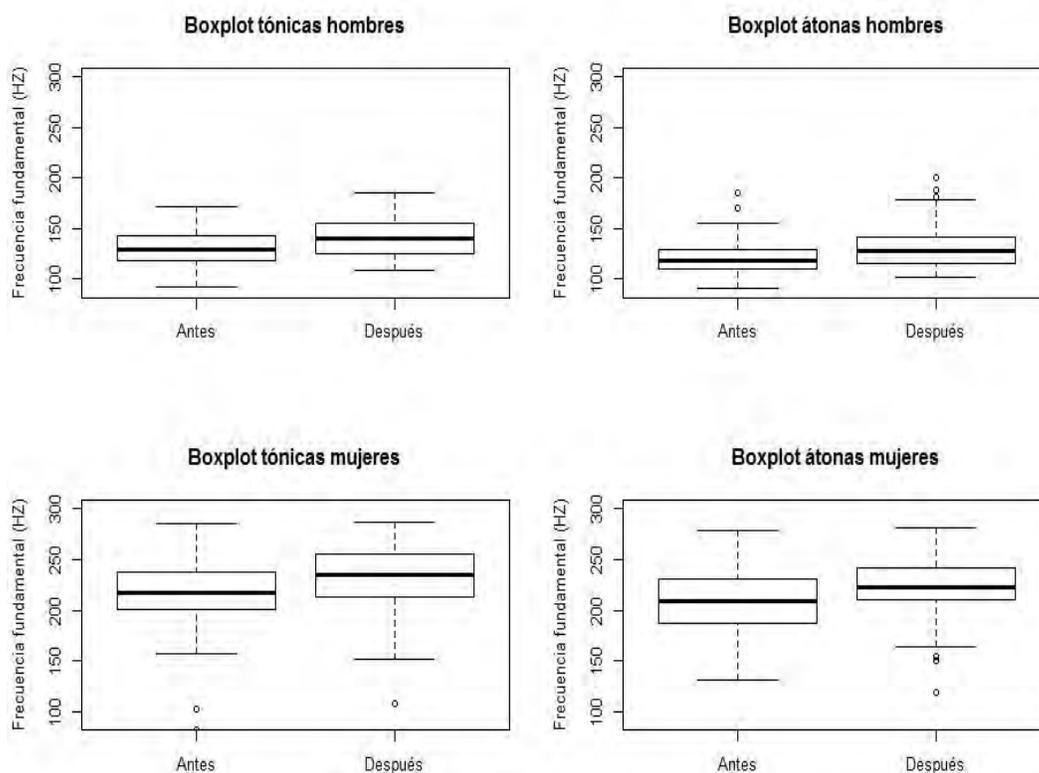


Figura 5.1 *Boxplot* de los hablantes femeninos y masculinos de ambas realizaciones

El valor del *p-value* (ver tabla 5.1) nos indica que hay una diferencia significativa entre las emisiones, antes y después de hacer ejercicio. Por un lado, en el *boxplot* de los hombres, tanto en su realización tónica como átona, se observa un ligero aumento en la caja que corresponde a las emisiones después de realizar la actividad física.

Por otro lado, el valor del *p-value* de las mujeres (ver tabla 5.2) nos indica que la diferencia entre antes y después de hacer ejercicio es significativa. En los *boxplot* de las mujeres, en la realización tónica y átona, se observa, de forma más evidente, un aumento de la mediana (línea negra dentro de la caja) en los datos del *Después*.

5.2 Resultados por tonicidad

5.2.1 Resultados por tonicidad. Hablantes masculinos

Este análisis es similar al primero, sólo que en este caso la comparación se hizo para cada uno de los hablantes.

En la tabla 5.3 se muestran los resultados de los hablantes masculinos.

| Realización | <i>p-value</i> | | | | |
|-------------|------------------|--------------------|-----------------|--------------------|-------------------|
| | H1 | H2 ¹⁴ | H3 | H4 | H5 |
| Átona | 0.002571 | 0.000002066 | 0.004681 | 0.000007884 | 0.00004257 |
| Tónica | 0.0006118 | 3.88E-05 | 0.05799 | 0.00002595 | 0.002762 |

Tabla 5.3 Resultados por realización: hablantes masculinos

Podemos observar que en la realización átona todos los hablantes masculinos presentaron una diferencia significativa; en la tónica, se observa que solamente H3 no presentó una diferencia significativa, mientras que los demás, sí.

¹⁴ En las muestras de H2 esta prueba no se realizó con 20 muestras sino con 15 porque en los datos de /le/ el hablante ensordecía la vocal en tres ocasiones.

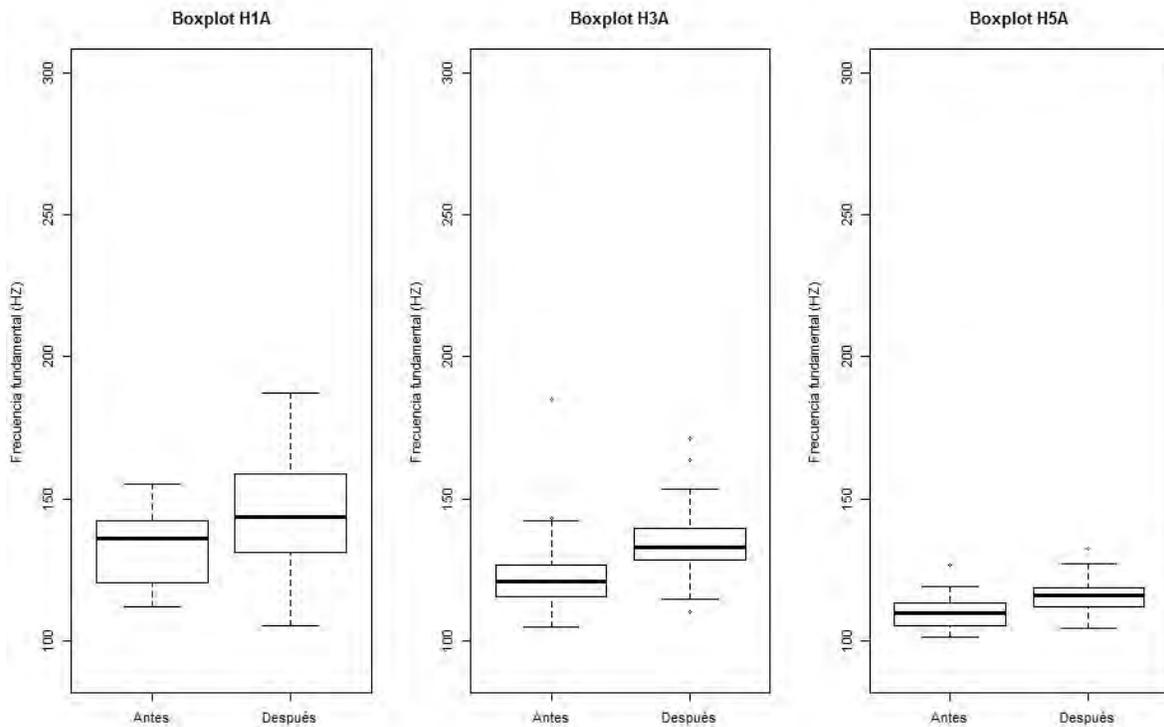


Figura 5.2 *Boxplot* de la realización átona de tres hombres con diferencia significativa en F0

Lo que podemos observar en el *boxplot* de H1 (ver figura 5.2) es que la variabilidad aumenta después de hacer ejercicio, esto se puede observar en el tamaño de los *boxplot* porque mientras que en el *Antes* los bigotes son más cortos, en el *Después* se vuelven más largos. En los otros dos hablantes la variabilidad es menor, es decir, tanto los bigotes como las cajas tienen un menor rango.

Por otra parte vemos que, en los tres hablantes, las cajas que corresponden al *Después* se encuentran más arriba que las del *Antes*, lo que significa que la F0 aumentó después de que saltaran la cuerda.

5.2.2 Resultados por tonicidad. Hablantes femeninos

En los datos de las mujeres se observan resultados muy distintos, mismos que se muestran en la tabla 5.4.

| Realización | <i>p-value</i> | | | | |
|-------------|-------------------|------------------|---------|----------------|--------|
| | M1 | M2 | M3 | M4 | M5 |
| Átona | 1.362E-07 | 6.246E-07 | 0.05077 | 0.01542 | 0.6276 |
| Tónica | 0.00009105 | 0.3983 | 0.1081 | 0.2084 | 0.531 |

Tabla 5.4 Resultados por realización: hablantes femeninos

Se aprecia que en la realización átona hay una diferencia significativa en tres hablantes, éstos son: M1, M2 y M4; en cambio, no la hubo en los hablantes M3 y M5. Por otra parte, en la realización tónica M1 presentó una diferencia significativa, mientras que en las demás no sucede lo mismo. Para ilustrar los resultados, en la figura 5.3 presentamos los diagramas de caja de los tres hablantes femeninos que presentaron variación en la F0 en la realización átona después de hacer ejercicio.

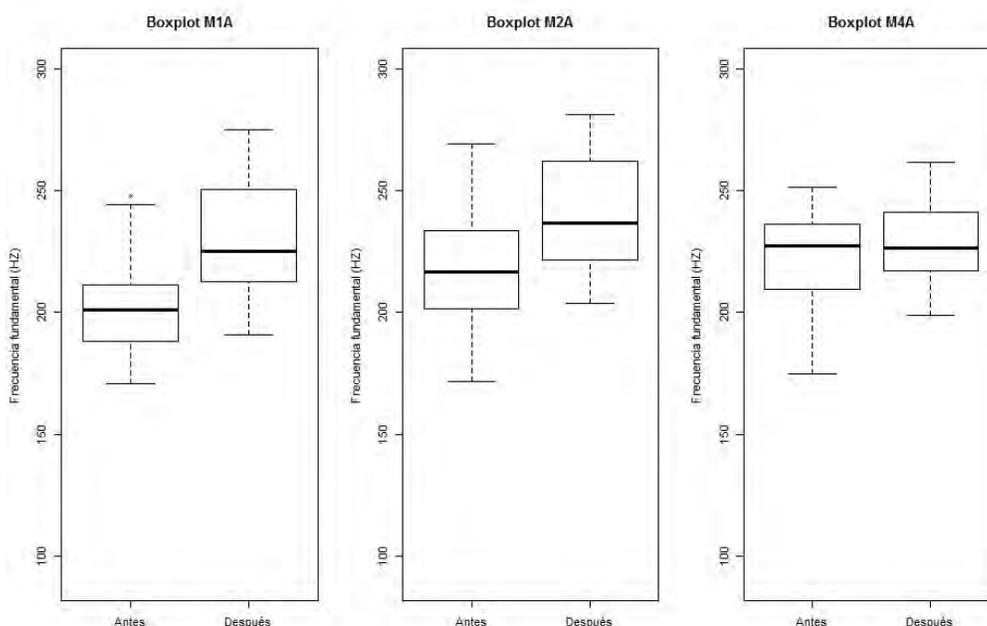


Figura 5.3 *Boxplot* de mujeres que sí presentaron variación en la F0 en la realización átona

Una vez más se observa que la F0 aumentó después de elevar el ritmo cardiaco de los hablantes femeninos. Este aumento en la F0 se observa claramente en el *boxplot* de M1 y M2 (ver figura 5.3), en donde se aprecia que la mediana del *Después* está arriba de la del

Antes. En M4 el aumento de la F0 se observa porque la caja del *Después* está más elevada que la del *Antes*, tal como sucede con M1 y M2.

A continuación se presentan los diagramas de caja de tres mujeres que no presentaron variación en la F0, en la realización tónica (Figura 5.4).

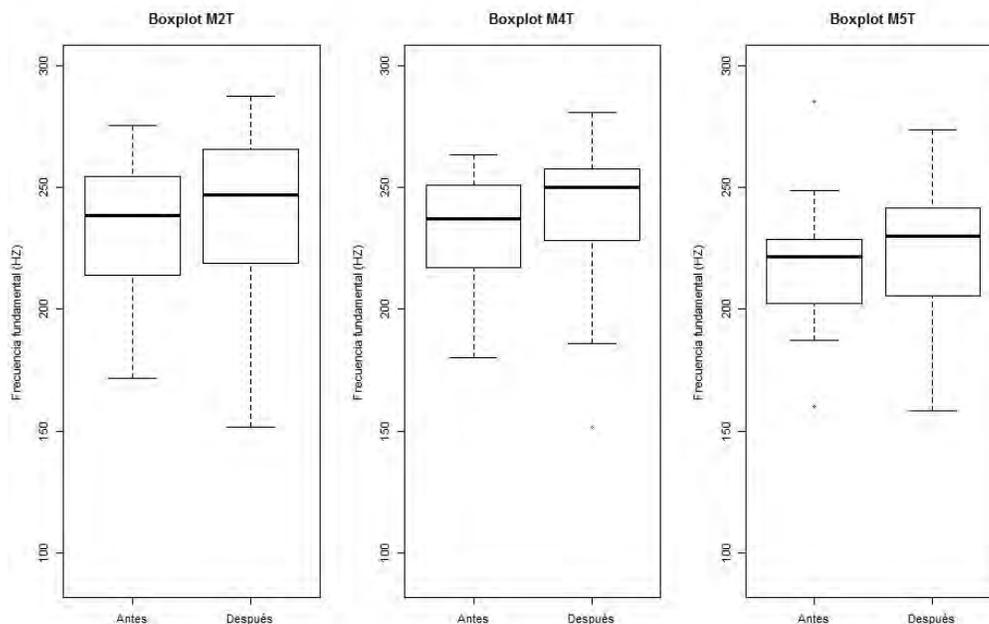


Figura 5.4 *Boxplot* de mujeres que no presentaron variación en la F0, en la realización tónica

Podemos observar que en los *boxplot* de M2, M4 y M5 (ver figura 5.4), las cajas se empalman, no obstante en los tres diagramas vemos que las medianas aumentan ligeramente en el *Después*, aunque este aumento no resultó significativo en el análisis estadístico (ver tabla 5.4), que arrojó un *p-value* mayor a 0.05, lo que indica que la F0 no presenta diferencias significativas antes y después de hacer ejercicio.

5.3 Resultados por contexto

5.3.1 Resultados por contexto. Hablantes masculinos

En la tercera confrontación se consideró cada uno de los contextos y la realización de la vocal. En esta prueba hubo veinticinco emisiones para cada análisis: cinco por cada

contexto alveolar (/n, l, r, s/) de cada uno de los hablantes, p. ej., se analizaron todas las vocales tónicas de los hablantes masculinos, precedidas por consonante nasal. Los resultados se muestran en la tabla 5.5.

| | Contexto | | | |
|--------------------|--------------------|--------------------|------------------|------------------|
| Realización | /ne/ | /le/ | /re/ | /se/ |
| | p-value | | | |
| Átona | 0.00000852 | 0.000000395 | 0.08072 | 0.04139 |
| Tónica | 0.000007969 | 0.3408 | 0.0000305 | 0.0003434 |

Tabla 5.5 Resultados por contexto: hablantes masculinos

Los resultados muestran que en la fonación de los hombres hubo diferencia en la F0, en casi todos los contextos. En la realización átona el único contexto que no presenta una diferencia significativa es el que está precedido por la vibrante simple. En cambio, en la realización tónica se observa que la vocal no presentó diferencias significativas ante la consonante lateral.

Para ejemplificar mejor la variación en los contextos de los hablantes masculinos se presentan los diagramas de caja de los contextos tónicos con diferencia significativa en la F0. Los diagramas se encuentran en la figura 5.5.

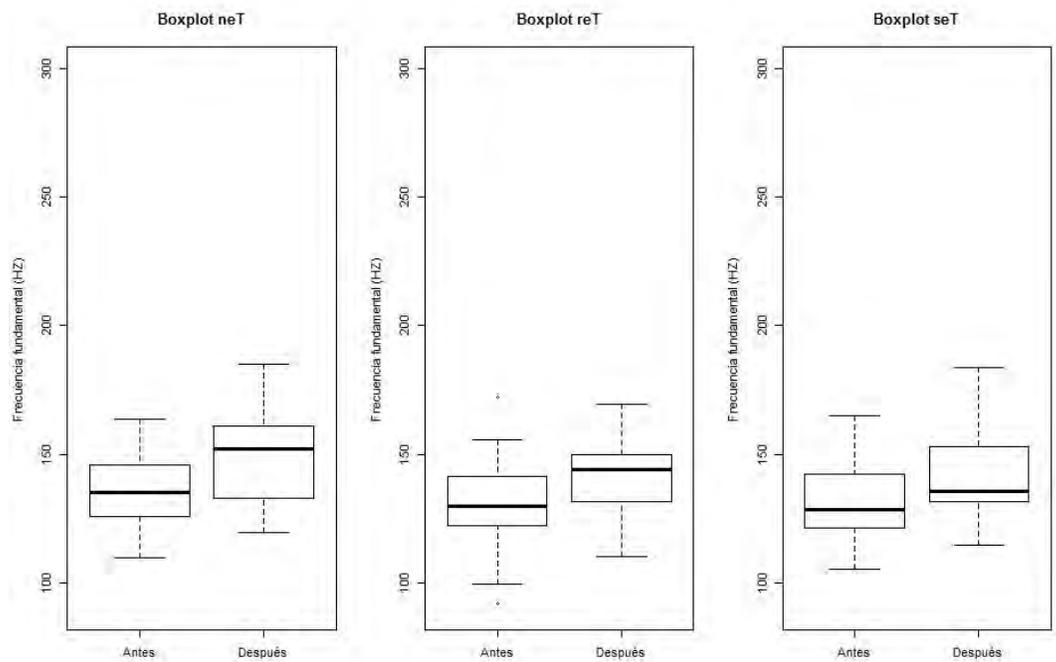


Figura 5.5 *Boxplot* de los contextos tónicos que presentaron variación en la F0

Es interesante observar que la variabilidad (ancho de la caja y altura de los bigotes) no cambio drásticamente antes y después de hacer ejercicio, sino que se mantiene. En los *boxplot* de los tres contextos vemos que, tanto las cajas como las medianas del *Después* están más arriba que las del *Antes*, lo que indica que la F0 aumentó después de aumentar el ritmo cardiaco.

5.3.2 Resultados por contexto. Hablantes femeninos

La prueba que se realizó con los datos de los hombres también se hizo con la información de las mujeres. En la tabla 5.6 se muestran los resultados.

| | Contexto | | | |
|-------------|-----------------|--------|----------------|-----------------|
| Realización | /ne/ | /le/ | /re/ | /se/ |
| | <i>p-value</i> | | | |
| Átona | 0.002102 | 0.2022 | 0.01328 | 0.009222 |
| Tónica | 0.738 | 0.1014 | 0.0336 | 0.001613 |

Tabla 5.6 Resultados por contexto: hablantes femeninos

En la tabla 5.6 se observa que en la realización átona se presenta una diferencia significativa en casi todos los contextos, excepto en la vocal antecedida por consonante lateral; por otra parte, en el contexto tónico se observa que hay diferencias significativas en la vocal precedida por vibrante simple y por fricativa.

En la figura 5.6 se presentan los contextos en los que las mujeres presentaron diferencias significativas en la F0 antes y después de hacer ejercicio en la realización átona.

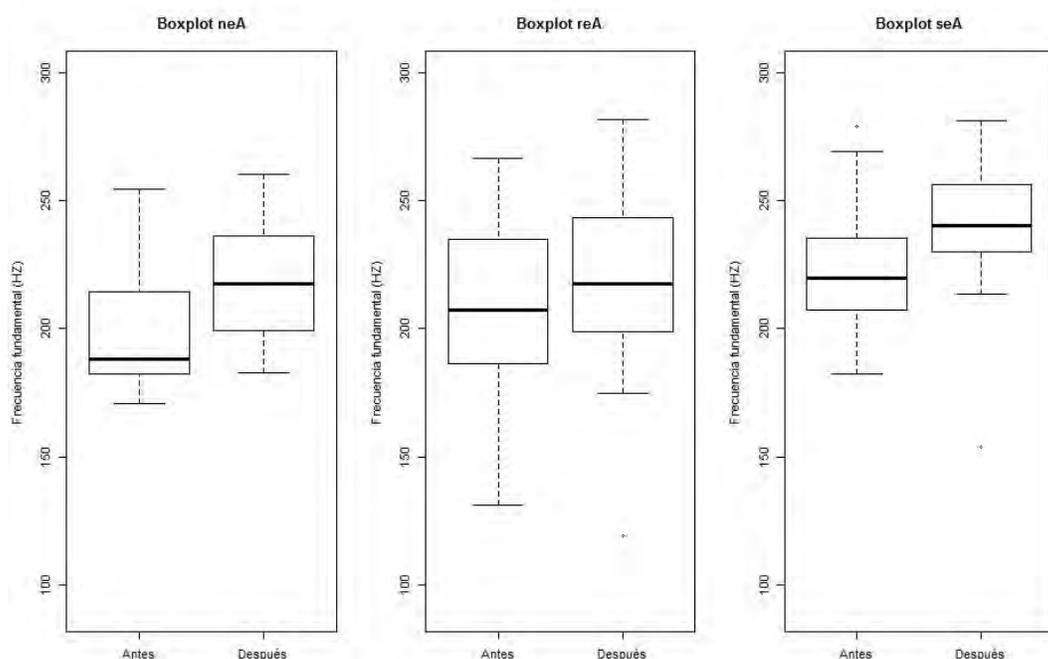


Figura 5.6 *Boxplot* de contextos átonos de hablantes femeninos que tuvieron variación en la F0

Es interesante destacar que en el contexto /re/ y /se/ en realización átona, la variabilidad es menor en la emisión después de realizar ejercicio, esto se observa en la disminución del tamaño de la caja y la longitud de los bigotes. En este análisis también vemos que las cajas, y sus respectivas medianas, están más arriba en el *Después* que en el *Antes*, lo que indica que la F0 aumentó después de elevar el ritmo cardiaco.

En la figura 5.7 presentamos los diagramas del contexto lateral en sus dos realizaciones. Presentamos estos diagramas para ver cómo se observa un *boxplot* en donde no hay una diferencia significativa en la F0.

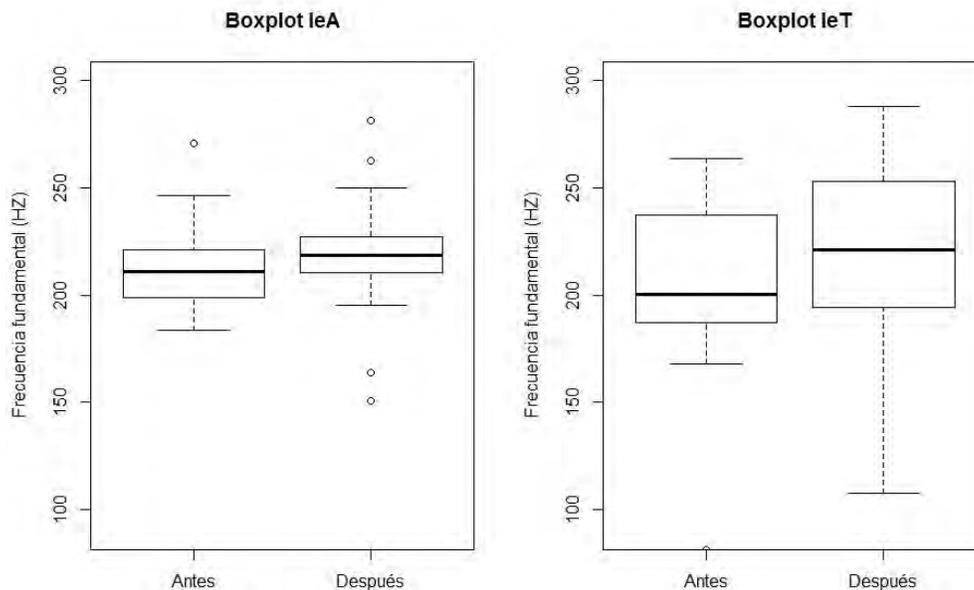


Figura 5.7 *Boxplot* de /e/ átono y tónico que no presentó variación

Lo que podemos observar es que hubo menos variación (ancho de las cajas y altura de los bigotes) en la realización átona que en la tónica. En la tabla 5.6 vimos que este es el único contexto que no presenta diferencias significativas en ninguna de sus realizaciones, pero lo que apreciamos en los diagramas es que las medianas del *Después* están más arriba que las del *Antes*, lo que indica que hay un aumento en la F0 después de hacer ejercicio, sin embargo, lo que destaca es que este aumento es tan mínimo que no es significativo para el resultado estadístico.

5.4 Resultados individuales

5.4.1 Resultados individuales. Hablantes masculinos

El último análisis que se realizó fue por hablante, en éste se utilizó la información de cada uno de los informantes; aquí se tomó en cuenta el contexto y la realización de la vocal. En la tabla 5.7 se presentan los resultados de los hablantes masculinos. En la columna

izquierda se aprecia tanto el contexto (C) como la realización (R), en las columnas de la derecha se anota el número de hablante y el *p-value* que se obtuvo.

Los resultados muestran en negritas los contextos en los que sí hubo una diferencia significativa. A continuación se harán notar dichas diferencias por contexto, es decir, se explicará en qué contexto los hablantes presentaron una variación significativa en la F0.

| C/R | <i>P-value</i> | | | | |
|-----|----------------|------------------|----------------|------------------|----------------|
| | H1 | H2 | H3 | H4 | H5 |
| neA | 0.2342 | 0.02181 | 0.01792 | 0.02483 | 0.06795 |
| neT | 0.01484 | 0.0007807 | 0.01416 | 0.1081 | 0.4884 |
| leA | 0.01164 | 0.01084 | 0.05556 | 0.05556 | 0.02616 |
| leT | 0.6085 | ∅ ¹⁵ | 0.8865 | 0.1091 | 0.02032 |
| reA | 0.1241 | 0.02197 | 0.5724 | 0.06625 | 0.02389 |
| reT | 0.1395 | 0.1038 | 0.0123 | 0.06747 | 0.1076 |
| seA | 0.2473 | 0.0625 | 0.03988 | 0.09836 | 0.1962 |
| seT | 0.0694 | 0.004618 | 1 | 0.0005786 | 0.2846 |

Tabla 5.7 Resultados por hablante masculino

Se observa que en la vocal átona precedida por nasal tres de los cinco hablantes presentan una variación significativa, éstos son H2, H3 y H4. En la realización tónica del mismo contexto H1, H2 y H3 también presentan una diferencia en la F0 antes y después de hacer ejercicio.

El siguiente contexto es la vocal frente a consonante lateral. En la realización átona se observa que H1, H2 y H5 presentaron una diferencia significativa en la F0 mientras que en la átona sólo el hablante H5 presentó diferencias significativas.

En la vocal átona precedida por vibrante simple sólo H2 y H5 presentaron una diferencia en la F0 antes y después de hacer ejercicio, y en la vocal tónica sólo H3

¹⁵ Es importante mencionar que la información de H2 no se pudo analizar debido a que ensordeció en tres ocasiones la consonante lateral y por ello la vocal también resultó afectada. El ensordecimiento no permitió que hubiera datos suficientes para el análisis estadístico en este contexto.

El último contexto es el de la vocal frente a consonante fricativa. En la realización átona sólo H3 presentó un *p-value* menor a 0.05, mientras que en la realización tónica H2 y H4 presentaron una diferencia significativa, los demás no tuvieron una diferencia importante en la F0.

En la figura 5.8 se muestra el *boxplot* de la comparación de la F0 antes y después de hacer ejercicio, en el hablante H2 en diferentes contextos. Decidimos presentar los diagramas de este hablante porque es el que más variación presentó.

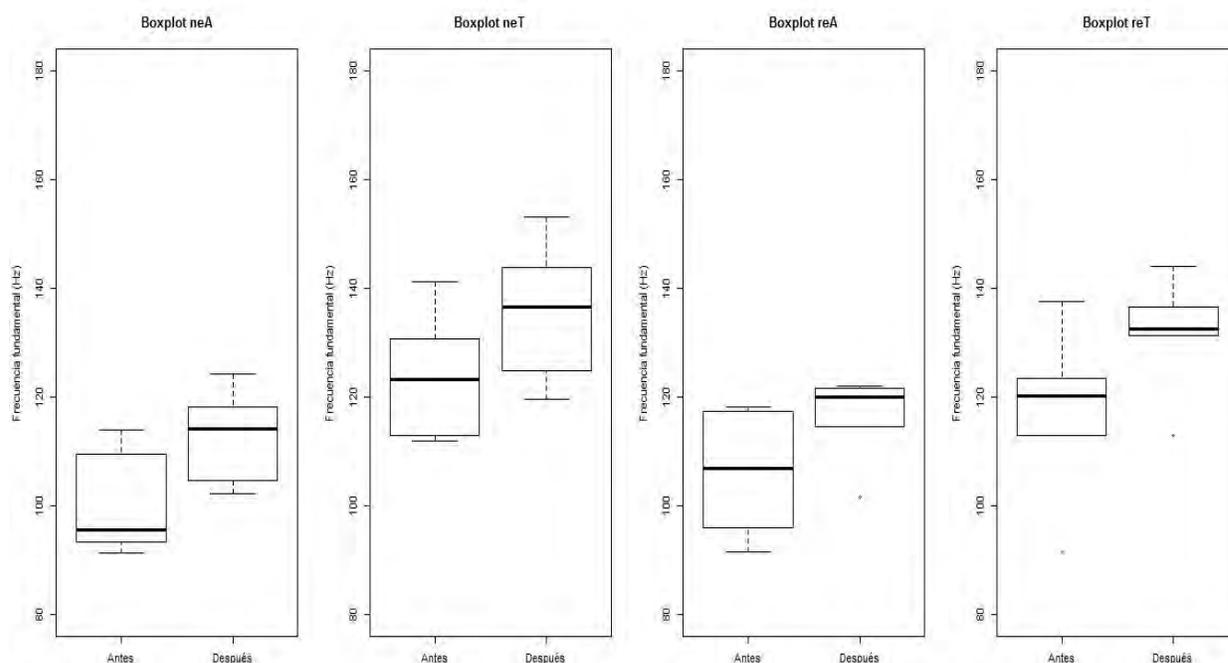


Figura 5.8 *Boxplot* de H2 en contextos con variación en la F0

Lo que podemos apreciar es que la variación no es similar en todos los contextos, lo que hay que resaltar es que la variación es menor en el *Después* del contexto precedido por vibrante simple.

Nuevamente observamos que la mediana de la caja que corresponde a los datos del *Después* está más elevada que la del *Antes*, lo cual indica que la F0 aumentó después de que el hablante realizó actividad física.

5.4.2 Resultados individuales. Hablantes femeninos

En la tabla 5.8 se muestran los resultados de los hablantes femeninos que se obtuvieron de la comparación de la F0 antes y después de aumentar el ritmo cardiaco.

| C/R | <i>p-value</i> | | | | |
|-----|-----------------|-----------------|----------------|----------------|--------|
| | M1 | M2 | M3 | M4 | M5 |
| neA | 0.01648 | 0.04318 | 0.00968 | 0.02503 | 0.3095 |
| neT | 0.1508 | 0.3573 | 0.09421 | 0.1811 | 0.9495 |
| leA | 0.01432 | 0.007814 | 0.5887 | 0.9663 | 0.2222 |
| leT | 0.02942 | 0.003038 | 0.8403 | 0.9197 | 0.1519 |
| reA | 0.03927 | 0.07665 | 0.7033 | 0.25 | 0.8933 |
| reT | 0.009158 | 0.5476 | 0.5063 | 0.6905 | 0.9421 |
| seA | 0.007118 | 0.006208 | 0.02755 | 0.2664 | 0.7926 |
| seT | 0.01587 | 0.008743 | 0.03175 | 0.07935 | 0.4206 |

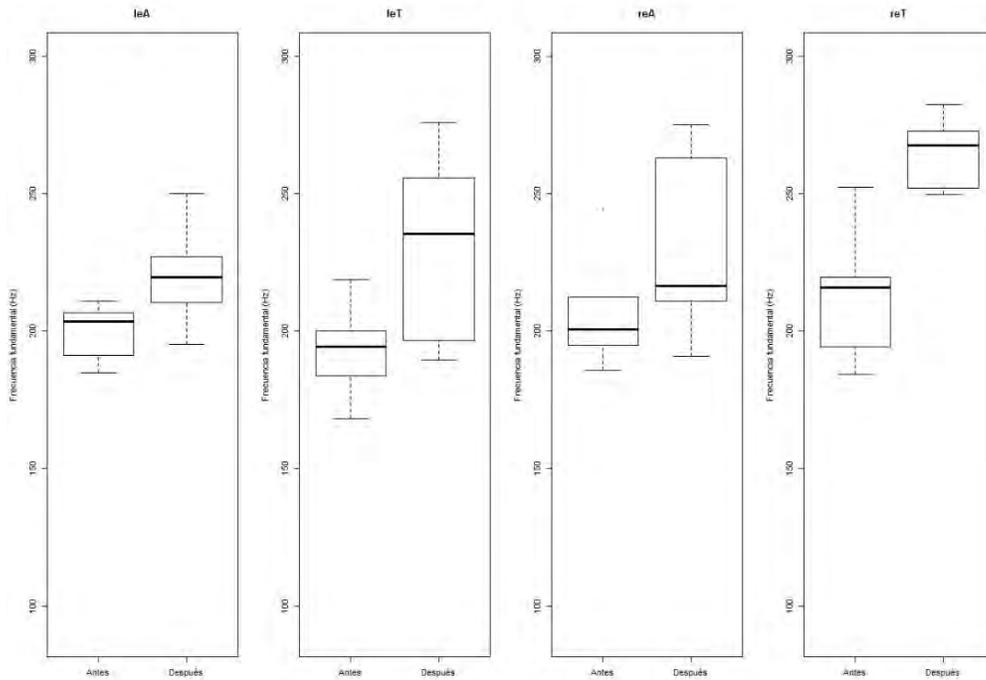
Tabla 5.8 Resultados por hablante femenino

En los resultados de las mujeres se puede ver que M1 presenta una variación significativa en todos los contextos, excepto en la realización tónica de la vocal frente a nasal. Otro de los hablantes que presenta más variación en la comparación de datos es M2, que tiene una diferencia significativa en casi todos los contextos, excepto en la realización tónica del contexto nasal y en ambas realizaciones de la vocal antecedida por vibrante simple.

En lo que respecta a los demás hablantes, se observa que tanto M3 como M4 tuvieron una variación significativa en la vocal átona precedida por nasal. Por su parte M3 tuvo variación en el contexto fricativo en ambas realizaciones.

Finalmente, lo que resalta es que M4 haya presentado diferencias en la F0 sólo en un contexto y que M5 no haya presentado diferencias significativas en ninguno de los resultados obtenidos.

En la figura 5.9 se presentan los diagramas de caja de M1, hablante que presentó más diferencia en la F0 antes y después de elevar el ritmo cardiaco.



5.9 *Boxplot* de M1 con variación en la F0

En los diagramas de la figura 5.9 observamos que la variabilidad es mayor en las emisiones de *Después* de los contextos /le/ tónica y de la /re/ átona. En los demás contextos vemos que la variabilidad es más uniforme. Por otro lado podemos ver que la F0 aumenta después de hacer ejercicio porque una vez más se aprecia que las cajas que corresponden a las emisiones después de realizar actividad física están más arriba.

A lo largo de este apartado se han mostrado los resultados que se obtuvieron en el análisis estadístico. Explicamos de qué manera fueron comparados los datos de la F0 que se obtuvieron del análisis acústico y porqué ordenamos la información de esta forma. También mostramos algunos diagramas de caja o *boxplot* que ilustran mejor la diferencia que se indica en el *p-value* mostrado en cada una de las tablas. En el siguiente apartado se hablará de las conclusiones a las que hemos llegado después de sopesar los resultados aquí descritos.

6. CONCLUSIONES

El objetivo principal de este trabajo fue estudiar si la frecuencia fundamental se ve afectada cuando el hablante se encuentra en una situación de estrés. Para ello fue necesario simular el estrés a partir del aumento del ritmo cardiaco del hablante. Se les pidió a los hablantes que leyeran un texto en estado de reposo y que después saltaran la cuerda y volvieran a leer el mismo texto para observar si existe o no una diferencia en la F0. Se analizó la frecuencia fundamental de la vocal media anterior en sílaba libre antecedida por las consonantes alveolares /n/, /l/, /r/ y /s/.

En general observamos que la F0 sí varía después de que los hablantes son sometidos a una situación de estrés. En concreto, podemos afirmar que coincidimos con las conclusiones a las que se han llegado en otras lenguas, como el inglés, en donde se afirma que la F0 aumenta en el momento en el que el hablante entra en un estado de estrés (Jessen *et al*, 2005). Es decir, los factores psicológicos, en este caso, simulados por los fisiológicos, sí afectan la comparación forense del habla. Por lo anterior, podemos concluir que si en una grabación dubitada el hablante se encuentra estresado es muy probable que la F0 presente una frecuencia más alta que si se analiza al hablante en un estado de reposo.

Ahora bien, siguiendo las hipótesis que planteamos al inicio de este trabajo, podemos dar paso a las siguientes conclusiones específicas.

- Si se analiza la vocal media anterior con base en su tonicidad, observamos que para los hombres la F0 se ve afectada de manera similar en el contexto tónico y en el átono; mientras que en las mujeres la realización átona resulta más afectada que la tónica.
- Si se analiza la vocal media anterior con base en su tonicidad pero de forma individual, observamos que dos de los hablantes femeninos no presentaron diferencias en la F0 antes y después de saltar la cuerda, tanto en la realización tónica como átona. En el caso de los hombres, casi todos presentaron diferencias en la F0, en ambas realizaciones. Excepto el hablante H3 en la realización tónica.
- Si se analiza la vocal media anterior con base en el contexto, observamos que en las mujeres hay un menor número de contextos que presentan diferencias significativas en las medidas de F0 respecto a los hombres. En particular, la F0 en el contexto /le/,

tanto tónico como átono, no se ve afectado por el estrés en el caso de las hablantes femeninas.

- Si se analiza la vocal en cada contexto y realización se observa que en los hombres sólo uno de ellos presentó diferencias significativas en la F0 en los siguientes contextos /le/ y /re/ en realización tónica y en /se/ átona. En los resultados de las mujeres se presentaron casos similares, en la realización átona y tónica de /re/ sólo una de ellas presentó diferencias en la F0.

De manera general podemos decir que en el español de la ciudad de México la frecuencia fundamental de la vocal más común, /e/ en contexto alveolar, se eleva después de aumentar el ritmo cardiaco. Por otro lado, la realización que se ve más afectada es la átona, quizá debido al menor gasto de aire que se le imprime para su articulación.

Ahora bien, gracias al análisis individual podemos afirmar que el contexto que resulta más viable en la comparación forense de voz, es el de la vocal tónica precedida por vibrante simple porque es el que obtuvo menos diferencia en la F0.

Consideramos que para poder hacer conclusiones más precisas en cuanto al nivel de afectación que sufre la F0 debido al estrés, en hombres y en mujeres, y en contexto tónico y átono, es necesario ampliar el corpus. El corpus de este trabajo está compuesto por cinco hablantes masculinos y cinco hablantes femeninos, de los que en total obtuvimos 800 muestras, 400 para la realización átona y 400 para la tónica. Creemos que diez hablantes no son una muestra representativa del español de la Ciudad de México, por lo que resulta necesario ampliar el número de hablantes, y por ende el número de muestras. No obstante, esta tesis representa un primer acercamiento al problema que resulta comparar la F0 de grabaciones que no fueron hechas en las mismas condiciones, particularmente en hablantes sometidos a una situación de estrés.

Otra de las propuestas que surgen a partir de la realización de esta tesis es estudiar otros sonidos que también sean frecuentes en el español de México, con el fin de determinar si la F0 se ve menos afectada por el estrés y ver si se puede utilizar para realizar comparación forense de voz. En particular, se podría analizar la vocal /a/ precedida de una vibrante simple, que como vimos en este trabajo fue el contexto que presentó menos diferencias en la F0 de la vocal media anterior.

Finalmente, consideramos que otro planteamiento interesante es estudiar el efecto del estrés en la F0 pero de acuerdo al índice de masa corporal de los hablantes. Lo que pudimos observar en los resultados de esta tesis es que los hombres con un mismo índice de masa corporal coincidieron en el número de contextos que presentan diferencias significativas de la F0. Por lo tanto, este puede ser un factor que también influye en las mediciones de F0.

BIBLIOGRAFÍA

- Aguilar, L.; Machuca, M. (1994). "Problemas de definición y categorización de los estilos del habla". Departamento de Filología Española. Universidad Autónoma de Barcelona. En http://liceu.uab.es/~maria/Investigacion/aesla_94.pdf
- Braun, A; Künzel, H. (1998). "Is forensic speaker identification unethical- or can it be unethical not to do it?". *Forensic Linguistics*. 5(1): 10-21.
- Breva, A. et al. (2000). "Ira y reactividad cardiaca. Adaptación en una situación de estrés real". *Anales de psicología*. España: Universidad de Murcia, 16 (1): 1-11.
- Boersma, P. ;Weenink, D. (2014). *Praat: doing phonetics by computer* [Programa computacional]. Versión 5.1.25, en <http://www.praat.org/> [09 de febrero de 2012]
- Cuétara, Javier. (2004). *Fonética y fonología del habla espontánea de la ciudad de México. Su aplicación en las tecnologías del habla*. Dirigida por Concepción Company. Tesis de maestría. Universidad Nacional Autónoma de México.
- Delgado Romero, C. (2001). *La identificación de locutores en el ámbito forense*. Dirigida por Francisco García García. Tesis doctoral. Universidad Complutense de Madrid, Facultad de ciencias de la información.
- Delgado Romero, C. (2005). "Comentarios sobre el contexto actual de la identificación forense de locutores" en *Lingüística forense, lengua y derecho: conceptos, métodos y aplicaciones*. Ma. Teresa Turell (ed.) Barcelona: Universitat Pompeu Fabra, Institut Universitari de Linguística Aplicada.
- Dietrich, M., et al. (2008). "The Frequency of Perceived Stress, Anxiety, and Depression in Patients with Common Pathologies Affecting Voice". *Journal of Voice*, 22 (4): 472-488.
- Dinville, C. (1996). *Los trastornos de la voz y su reeducación*. Barcelona: Masson.
- Fraser, H. (2011). Forensic Phonetics. En <http://helenfraser.com.au/forensic/> [20 de noviembre de 2013]
- French, P. (1994). "An overview of forensic phonetics with particular reference to speaker identification". *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 1(2): 169-182.
- Giddens C., et al. (2013). "Vocal Indices of Stress: A Review". *Journal of Voice*, 27 (3): 390.e21- 390.e29.

- Gold, E.; French, P. (2011). "International practices in forensic speaker comparison". *Forensic Linguistics. The International Journal of Speech, Language and the Law*. 18 (2): 293-307.
- González de Rivera, J.L. (2010). *Los síndromes de estrés*. Madrid: Síntesis.
- González Rodríguez, J.; Lucena Molina, J. (2005). "IDENTIVOX: un sistema automático de reconocimiento de locutores por la voz para acústica forense" en *Lingüística forense, lengua y derecho: conceptos, métodos y aplicaciones*. Ma. Teresa Turell (ed.) Barcelona: Universitat Pompeu Fabra, Institut Universitari de Linguística Aplicada.
- González Sigüenza, B. (2008). "BatVox: Sistema automático de reconocimiento de locutor". En <http://www.raco.cat/index.php/EFE/article/viewFile/140074/216441> [30 de diciembre de 2013]
- Gray, J. (1993). *La psicología del miedo y el estrés*. Barcelona: Editorial Labor.
- Hansen, J.; Patil, S. (2007). "Speech Under Stress: Analysis, Modeling and Recognition". *Speaker Classification I. Fundamental, Features, and Methods*. Christian Müller (Ed.). Alemania: Springer.
- Hollien, H. (2002). *Forensic voice identification*. San Diego, California: Academic.
- Hollien, H. (2012). "On Earwitness Lineups". *Investigative Sciences Journal*. 4 (1). Agosto. Versión electrónica en <http://www.investigativesciencesjournal.org/article/view/10868> [25 de noviembre de 2013]
- Jessen, M., et al. (2005). "Influence of vocal effort on average and variability of fundamental frequency". *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 12 (2): 174-213.
- Lindh, J. (2004). "Handling the 'Voiceprint' Issue". Dept. of Linguistics. Göteborg University. Stockolm. Versión electrónica en http://www2.ling.su.se/fon/fonetik_2004/lindh_voiceprint_foneteik2004.pdf [02 de enero de 2014]
- Llisterri, J. (2003) "Lingüística y tecnologías del lenguaje", *Lynx. Panorámica de Estudios Lingüísticos* (Departament de Teoria dels Llenguatges, Universitat de València) 2: 9-71. http://liceu.uab.es/~joaquim/publicacions/TecnoLing_Lynx02.pdf [14 de enero de 2014]

- Llisterri, J. (2004). *El análisis fonético del habla y sus aplicaciones*, en: www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CCwQFjAC&url=http://liceu.uab.es/~joaquim/phonetics/Valencia_04/Guion_bibliografia.pdf/ [15 de noviembre de 2013]
- Llisterri, J. (2014). *La fonación en* http://liceu.uab.es/~joaquim/phonetics/fon_produccio/fonacion.html [12 de enero de 2014]
- López Escobedo, F. (2010). *El análisis de las características dinámicas de la señal de habla como posible marca para la comparación e identificación forense de voz: Un estudio para el español de México*. Dirigida por Ma. Teresa Turell y Luis Alberto Pineda Cortés. Tesis doctoral, Universidad Pompeu Fabra, Instituto Universitario de Lingüística Aplicada.
- Lucena Molina, J. (2005). "La acústica forense". *Instituto Universitario de Investigación sobre Seguridad Interior*. En http://www.uned.es/investigacion/publicaciones/Cuadernillo_octubre200503.pdf [15 de agosto de 2013]
- Martínez Celdrán, E. (1988). *Estudios de fonética experimental*. Universidad de Barcelona: Laboratorio de Fonética
- Olsson, J. (2004). *Forensic linguistics. An introduction to language, crime and the law*. London: Continuum
- Pardo, A.; Ruiz, M. Á. (2002). *SPSS 11. Guía para el análisis de datos*. Madrid: McGraw-Hill.
- Perello, J. (1973). *Alteraciones de la voz*. México: Científico-Médica.
- _____ (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.
- Potter, R. (1945). "Visible patterns of speech." *Science*, Noviembre: 463–470.
- Quilis, A. (1999). *Tratado de Fonología y Fonética Españolas*. Madrid: Gredos.
- Real Academia Española. (2001). *Diccionario de la lengua española* (22.^a ed.). En <http://www.rae.es/rae.html>
- Rosas, C.; Sommerhoff, J. (2010). "Implicancias y proyecciones forenses en el análisis de la voz". *Estudios Filológicos*. Chile: Universidad Austral de Chile, 46: 101-118.
- Segre, R. (1955). *Tratado de foniología: trastornos de la voz y del habla y su corrección*. Buenos Aires: Paidós.

Sinclair, J. (1996). "Preliminary recommendations on Corpus Tipology". En <http://www.ilc.cnr.it/EAGLES96/corpus/typ/node5.html#SECTION00041000000000000000>, [11 de noviembre de 2013]