



UNIVERSIDAD DE CHILE  
FACULTAD DE MEDICINA  
ESCUELA DE FONOAUDIOLOGIA

## FUNDAMENTOS DE LA IDENTIFICACIÓN VOCAL DE HABLANTES DEL ESPAÑOL DE CHILE: UNA MIRADA FONOAUDIOLÓGICA

INTEGRANTES:

Karin Castillo Frankenberg  
Natalia Cruz Vargas  
Mauricio Escobar Castañeda  
Estela Medina Bustamante

TUTOR PRINCIPAL:

Luis Romero Romero

TUTORES ASOCIADOS:

Carla Badani Schoneweg  
Rodrigo Pérez Pérez

Santiago – Chile  
2011





**UNIVERSIDAD DE CHILE  
FACULTAD DE MEDICINA  
ESCUELA DE FONOAUDIOLOGIA**

## **FUNDAMENTOS DE LA IDENTIFICACIÓN VOCAL DE HABLANTES DEL ESPAÑOL DE CHILE: UNA MIRADA FONOAUDIOLÓGICA**

INTEGRANTES:

Karin Castillo Frankenberg  
Natalia Cruz Vargas  
Mauricio Escobar Castañeda  
Estela Medina Bustamante

TUTOR PRINCIPAL:

Luis Romero Romero

TUTORES ASOCIADOS:

Carla Badani Schoneweg  
Rodrigo Pérez Pérez

Santiago – Chile  
2011

## DEDICATORIA

*A nuestras familias por representar nuestro origen y destino. A nuestras madres por su eterna entrega, a nuestros padres por su incondicionalidad y a nuestros hermanos por apoyarnos siempre.*

## **AGRADECIMIENTOS**

*A nuestros queridos docentes Luis y Rodrigo por su gran disposición y guía en este proceso. Por ofrecernos su sincera amistad.*

*A Carla Badani por su ayuda, preocupación permanente y por compartir con nosotros su sabiduría.*

## ÍNDICE

### RESUMEN

### ABSTRACT

<b>I. INTRODUCCIÓN .....</b>	<b>1</b>
<b>II. PROBLEMA DE INVESTIGACIÓN .....</b>	<b>3</b>
<b>III. OBJETIVOS .....</b>	<b>5</b>
1. OBJETIVO GENERAL.....	5
2. OBJETIVOS ESPECÍFICOS .....	5
<b>IV. MARCO METODOLÓGICO.....</b>	<b>6</b>
1. PROCEDIMIENTOS DE OBTENCIÓN DE INFORMACIÓN .....	7
2. INSTRUMENTO DE RECOLECCIÓN DE INFORMACIÓN .....	7
<b>V. MARCO TEÓRICO .....</b>	<b>8</b>
1. ASPECTOS ANÁTOMO-FISIOLÓGICOS.....	8
1.1. La Voz .....	8
1.2. Aparato Fonador .....	9
1.2.1. Cavidades Infraglóticas .....	9
1.2.2. Cavidad Glótica .....	9
1.2.3. Cavidades Supraglóticas .....	10
2. ASPECTOS ACÚSTICOS .....	11
2.1. Características Acústicas de los Sonidos .....	11
2.2. Resonancia .....	13
3. ASPECTOS FONÉTICO-FONOLÓGICOS .....	16
3.1. Elementos Segmentales y Suprasegmentales .....	16
<b>VI. DESARROLLO DEL TEMA.....</b>	<b>19</b>
1. CONCEPTOS FUDAMENTALES .....	19
2. SOBRE LA ACÚSTICA DEL CRIMEN Y LA FONÉTICA FORENSE .....	23
2.1. Corrientes Principales Presentes en el Análisis de la Voz.....	25
2.1.1. Reconocimiento .....	25
2.1.2. Verificación .....	25
2.1.3. Identificación .....	28

2.2.	Tipos de Análisis Utilizados en el Ámbito Forense .....	29
2.2.1.	Identificación de Hablantes por Testigos de Audio .....	30
2.2.2.	Identificación de Hablantes por Expertos.....	30
2.2.3.	Audioanálisis.....	31
2.2.4.	Estudios Lingüísticos de Autoría .....	31
<b>3.</b>	<b>MÉTODOS PARA LA IDENTIFICACIÓN VOCAL DE HABLANTES .....</b>	<b>32</b>
3.1.	Métodos Objetivos .....	33
3.1.1.	Automático .....	33
3.1.2.	Paramétrico .....	34
3.2.	Métodos Subjetivos.....	35
3.2.1.	Sociolingüístico .....	35
3.2.2.	Espectrográfico.....	35
3.2.3.	Perceptivo-Auditivo .....	36
3.2.4.	Fonético-Acústico .....	36
3.3.	Métodos Combinados .....	37
<b>4.</b>	<b>VARIABLES PARA LA IDENTIFICACIÓN VOCAL DE HABLANTES .....</b>	<b>39</b>
4.1.	Requisitos para la Selección de Variables Acústicas en el Análisis Forense .....	39
4.2.	Clasificación de las Variables Relacionadas con la Identificación Vocal de Hablantes .....	40
4.2.1.	Alta Información .....	40
4.2.2.	Baja Información.....	41
<b>5.</b>	<b>FACTORES QUE INTERVIENEN EN LA IDENTIFICACIÓN VOCAL DE HABLANTES.....</b>	<b>50</b>
5.1.	Circunstancias Dependientes de la Naturaleza del Habla y del Sujeto Emisor .....	50
5.2.	Circunstancias No Dependientes de la Naturaleza del Habla y del Sujeto Emisor .....	51
<b>6.</b>	<b>APROXIMACIÓN A LA SITUACIÓN NACIONAL EN LA IDENTIFICACIÓN VOCAL DE HABLANTES .....</b>	<b>53</b>
6.1.	Policía de Investigaciones de Chile (PDI) .....	53
6.2.	Tesis: “Descripción de los Métodos Utilizados en Reconocimiento Forense de Locutores y su Implementación en Chile” Autor: Pablo Celis, Año 2009. ....	56
6.3.	Trabajos de Investigación. Autores: Claudia Rosas y Jorge Sommerhoff .....	57
6.3.1.	“La Variación Acústica del Hablante: Aplicaciones Forenses” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2008. ....	57
6.3.2.	“Efectos Acústicos de las Variaciones Fonopragmáticas y Ambientales” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2009.....	57
6.3.3.	“Implicancias y Proyecciones Forenses en el Análisis de la Voz” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2010.....	59
<b>VII.</b>	<b>DISCUSIÓN.....</b>	<b>61</b>
<b>VIII.</b>	<b>CONCLUSIONES.....</b>	<b>66</b>
<b>IX.</b>	<b>REFERENCIAS .....</b>	<b>68</b>

## RESUMEN

El presente análisis documental se enmarca dentro de la disciplina forense y contempla una revisión teórica acerca de los aspectos relacionados con la identificación vocal de hablantes. En esta materia, la disciplina fonoaudiológica aún no presenta una participación formal, sin embargo la necesidad imperante de generar un análisis global de los componentes relacionados con la tarea de identificación, ha llevado a que hoy en día esta labor pueda ser realizada por el fonoaudiólogo.

Frente a la insuficiente sistematización de la información en torno a la identificación vocal de hablantes, se propone este documento como respuesta a dicha problemática. En primer lugar, se realiza una descripción referente a los aspectos anátomo-fisiológicos, acústicos y fonético-fonológicos que conforman el sustrato del estudio de la voz. En segunda instancia, se introduce al ámbito de la acústica del crimen y la fonética forense, apartado en el cual se hace referencia a las principales corrientes y tipos de análisis utilizados en el proceso de identificación. A continuación se mencionan los distintos métodos, variables y factores asociados al análisis forense. Por último, se entrega una aproximación a la situación nacional en este ámbito investigativo.

En cuanto a los resultados, se constató un escaso consenso en cuanto a los términos conceptuales y métodos para la identificación vocal. Por su parte, respecto a las variables que intervienen en el análisis forense, se evidenció que no se dispone de un único criterio de clasificación, lo que en conjunto, dificulta su utilización en la práctica y la entrega de resultados certeros sobre la identidad de una persona. A nivel nacional, se observó una acentuación de la carencia investigativa que genera una práctica de insuficiente validez y consenso.



## **ABSTRACT**

This documentary analysis provides a theoretical review about the aspects of vocal identification of speakers within the forensic discipline. Although the speech and language therapist has not yet formal participation in these activities, the current need to generate a comprehensive analysis of the components related to the vocal has opened the possibility that the speech therapist can perform in tasks of this area.

Concerning the lack of systematization of information in reference to the vocal identification of speakers, it is suggested this document as an answer to the problems. Firstly, it has made a description regarding the anatomical- physiological aspects, acoustics and phonetic-phonological that forms the basics of the study of the voice. Secondly, it is introduced the field of the crime acoustic and forensic phonetic, the matter in which it makes reference to the main analysis used in the identification process. In the following, it is mentioned the different methods, variables and factors associated to the forensic analysis. Finally, it is given an estimate to the national situation with regard to the research field.

According to the results, it is stated a limited consensus about conceptual terms and methods of vocal identification. Therefore, respecting the variables that take place in the forensic analysis it was clearly seen that it does not stipulate one main criterion about classification that makes difficult its usefulness in the practical matter and its concretes results about the identification of a person. In national level, it was observed a major lack of research that creates an inadequate validity to this matter.



## I. INTRODUCCIÓN

La voz es el emblema mismo del hablante de forma indeleble en la trama del discurso. Cada una de nuestras expresiones del lenguaje no llevan solo mera información, sino que a través del acento, el tono y la calidad de la voz formamos una declaración sonora de nuestra pertenencia social o nuestra identidad física y psicológica, así como también nuestro propio estado de ánimo (Laver en Rose, 2002). Corresponde además, a uno de los medios más eficientes en el establecimiento de un sistema comunicacional, constituyendo el sustento físico y fisiológico en la transmisión significativa de señales acústicas que permiten la interacción entre sujetos. Sin embargo, la voz no solo cumple la función de vehículo de la interacción social. Para algunos incluso puede transformarse en la herramienta fundamental para desempeñar su función laboral. Dentro de este marco la voz puede entenderse como un medio a través del cual se da vida a una profesión u oficio, como una forma de expresión artística en el caso del canto o el teatro, como soporte económico para quienes constituye un requisito indispensable para realizar su labor, o bien, como fuente de investigación, análisis y estudio para quienes deban emitir juicio sobre la voz de otros.

Una de las cualidades más relevantes del instrumento vocal es su singularidad. Es única, por lo que es capaz de otorgarle una identidad exclusiva a cada ser humano a lo largo de su vida, pese a experimentar ciertas variaciones determinadas con el pasar de los años (o simplemente de las horas), de las circunstancias que nos rodean, del estado de salud o de las emociones que se sientan en un momento dado. Este rasgo hace que por medio de la voz se pueda reconocer a una persona que nos es familiar, sin la necesidad de verla.

Bajo dicho cimiento surge la fonética forense que como disciplina, tiene por fin principal la identificación de individuos mediante análisis acústicos que determinen parámetros vocales característicos de un sujeto en particular. Cabe destacar que dicha ciencia se enmarca en el ámbito legal, funcionando como estrategia pericial en los casos que sea necesario. Desde esta mirada es posible constatar a lo menos dos procesos relevantes: la verificación de identidad y la identificación propiamente tal. El primero se entiende como la comparación de una muestra de habla de un determinado locutor con otro registro de él mismo (Broeders, A., 2001). El segundo, hace referencia al proceso en el cual se comparan distintas muestras de voz y se determina si la identidad de la persona se corresponde con los datos almacenados, es decir, compara el registro de un hablante desconocido con una base de datos conocida.

La verificación de identidad de un sujeto o su identificación propiamente tal, se realiza a través de ciertos parámetros relacionados con el habla, sean estos: fonéticos, acústicos, lingüísticos o extralingüísticos. No obstante, aún no se dispone de un sistema metodológico plenamente validado y aplicado universalmente, lo que sumado al número acotado de investigaciones nacionales que aporten datos concretos, convierte la situación en un problema que exige pronta solución. La dificultad radica en consensuar y establecer un único esquema analítico, que contemple procedimientos y análisis que faciliten la identificación de un individuo mediante el habla.

La voz puede ser vista como el resultado de una sumatoria de distintos elementos entre los que destacan la frecuencia fundamental y los formantes producidos en cada porción del tracto fonatorio, el que a su vez posee una conformación individual y particular para cada ser humano. Desde esta perspectiva, representa un obstáculo al momento de definir cuáles serán los patrones acústicos a analizar en el caso de identificar a un sujeto, ya que varían tanto a nivel de interlocutor, como de intralocutor.

Frente a las escasas investigaciones nacionales y a las múltiples demandas de la disciplina forense actual, se vuelve significativo por un lado, determinar los elementos acústicos variables de la voz, y por otro, la creación de una base de datos fonético-acústica que considere a hablantes nativos del español de Chile. Esto último, debido a que los parámetros vocales en estudio varían entre comunidades de hablantes de distintas locaciones geográficas.

Situados en este contexto, el propósito del presente proyecto consiste en contribuir a la organización de datos nacionales e internacionales, relacionados con parámetros fonético-acústicos pertinentes en la identificación vocal de hablantes. Desde luego, comenzar a satisfacer la falta de base teórica que actúe como soporte en el incipiente desarrollo de la fonética forense en nuestro país, y apoyar de esta forma, el establecimiento de un sistema de reconocimiento e identificación de sujetos por medio del análisis de parámetros acústicos de la voz.

## II. PROBLEMA DE INVESTIGACIÓN

La necesidad de identificar personas mediante su habla tendría sus orígenes en el siglo XVII. Para estos años, la disciplina forense se basaba en la percepción auditiva realizada por personas con habilidades de identificación de voces (Celis, 2009). Desde entonces, se ha gestado un realce paulatino en esta materia.

La identificación es un proceso que pone en evidencia aquellos rasgos que diferencian a un individuo del resto de su comunidad lingüística, con la cual comparte muchos de los rasgos que la caracterizan. Hollien (2002) define identificación de hablantes como un método o conjunto de procedimientos en los que se busca dar con la identidad de un sujeto desconocido.

La problemática radica en que no se dispone de una sistematización suficiente que permita fundamentar la elección de un método válido, estandarizado y consensuado para la identificación vocal de hablantes del español de Chile. En metodología investigativa, la validez hace referencia al grado en que un instrumento realmente mide lo que se ha propuesto (López, 2006) y por tanto requiere de la concordancia entre las variables escogidas y las pruebas o modelos matemáticos-estadísticos a utilizar.

Este conflicto es el resultado, por un lado, de la existencia de múltiples variables personales y ambientales que determinan el perfil vocal de un sujeto; por otro, del escaso consenso en torno a la terminología y conceptualización de los parámetros y procesos que intervienen en la disciplina forense. Este inherente carácter multifactorial de la identificación es el que transforma dicha tarea en una responsabilidad transdisciplinaria, en la que cada una de las especialidades aporta su conocimiento a favor de un proceso de simbiosis compartida y análisis global.

A nivel nacional no se dispone a la fecha, de artículos acreditados que contengan una sistematización de la información disponible para determinar la existencia de marcadores fonético-acústicos. Estos son necesarios en la identificación vocal de hablantes dentro de una comunidad lingüística determinada. La urgencia de suplir este vacío se torna fundamental al considerar que las nuevas políticas jurídicas chilenas contemplan el uso de grabaciones de audio como pruebas complementarias para orientar la resolución de casos, más aun cuando el Artículo 323 del Código Procesal Penal (2003) señala: "Podrán admitirse como pruebas:

películas cinematográficas, fotografías, fonografías, videograbaciones y otros sistemas de reproducción de la imagen o del sonido, versiones taquigráficas y, en general, cualquier medio apto para producir fe. El tribunal determinará la forma de su incorporación al procedimiento, adecuándola, en lo posible, al medio de prueba más análogo”.

En este contexto, el presente trabajo aborda la elaboración de un documento que sintetice la información más relevante del ámbito nacional e internacional, dando cuenta de la existencia de parámetros confiables y científicamente validados en la identificación vocal de individuos. Además, contempla el análisis reflexivo acerca de la competencia fonoaudiológica, y de esta forma responder preguntas referentes a la implicancia y relevancia del rol profesional en el abordaje de esta materia.

### **III. OBJETIVOS**

#### **1. OBJETIVO GENERAL**

1. Sistematizar<sup>1</sup> la información disponible respecto de la identificación vocal de hablantes.

#### **2. OBJETIVOS ESPECÍFICOS**

- 1.1 Realizar una revisión exhaustiva, crítica y actualizada de la evidencia científica disponible acerca de la identificación vocal de hablantes.
- 1.2 Clasificar los métodos más significativos en la identificación vocal de hablantes.
- 1.3 Clasificar las variables más significativas en la identificación vocal de hablantes.
- 1.4 Describir los factores que intervienen en la identificación vocal de hablantes.
- 1.5 Discutir la inserción del profesional fonoaudiólogo en la identificación vocal de hablantes.

---

<sup>1</sup> Según Carvajal (2005), este término es usado comúnmente para hacer referencia al ordenamiento y clasificación de información, de tal manera que estructure de forma precisa categorías y relaciones, posibilitando la constitución de bases de datos. Así, este proceso contempla teorizar y metodologizar, es decir, a partir del ordenamiento, evaluación, análisis, interpretación y reflexión crítica, construir conocimiento y cambio de las prácticas.

#### **IV. MARCO METODOLÓGICO**

La presente investigación se sustenta en un diseño metodológico cualitativo no experimental, de tipo análisis documental, en el que serán revisados un vasto número de artículos, tanto del ámbito nacional como internacional que estén relacionados con el análisis acústico de la voz aplicados a la identificación de hablantes.

El grupo de estudio estará compuesto por unidades de análisis definidas como todo texto o documento publicado en libro, revista, acta de congreso u otras, ya sea en formato impreso o digital. Los artículos seleccionados deberán cumplir con ciertos criterios de exclusión: origen nacional o extranjero, idioma español, inglés o portugués, fecha de publicación igual o posterior al año 2001, publicados en revistas sujetas a índice de impacto o bajo comité editorial.



### ***1. PROCEDIMIENTOS DE OBTENCIÓN DE INFORMACIÓN***

En primera instancia se realizará el proceso de búsqueda y selección de artículos según los criterios de exclusión instaurados. Luego, la información obtenida será organizada de acuerdo a una ficha resumen preestablecida que contemplará bibliografía sujeta a formato APA y la descripción de elementos relevantes del artículo. Posterior a la revisión de cada ficha, se clasificarán los documentos de acuerdo a sus áreas o líneas investigativas. A continuación se procederá a comparar la información y/o hallazgos encontrados por los distintos campos de investigación para finalmente generar un análisis integral de la información obtenida que permita sugerir una propuesta teórica que oriente futuras investigaciones experimentales.

### ***2. INSTRUMENTO DE RECOLECCIÓN DE INFORMACIÓN***

Para organizar y sintetizar la información recabada se utilizará una ficha de resumen con el fin de generar una base de datos bibliográfica referente al ámbito del análisis fonético-acústico.

## V. MARCO TEÓRICO

### 1. ASPECTOS ANÁTOMO-FISIOLÓGICOS

#### 1.1. La Voz

“La voz es el soporte acústico de la palabra. Ella vehiculiza nuestros pensamientos, ideas, emociones... Emerge, se proyecta, se modifica en nosotros mismos y a través de todo nuestro ser.” (Bustos, 2003). Sin embargo, la voz no es un simple traspaso de señales acústicas, sino que en sí misma constituye un complejo fenómeno físico-acústico influenciado por diversos factores como el estado físico y psicológico de un individuo. Posee función comunicativa y varía de acuerdo a los diferentes contextos situacionales (Bonastre, 2001).

Fisiológicamente, la voz corresponde a la vibración de las cuerdas vocales, contenidas en la laringe, que a su vez es el órgano fonador por excelencia. Estos pliegues, por medio de un proceso de aducción, ofrecen resistencia al flujo continuo de aire, obteniendo como resultado el sonido (Romero, 2007). En otras palabras, autores como García-Tapia y Cobeta (1996) plantean que esta producción de energía acústica, a partir del aire obtenido del sistema respiratorio, es la utilización inteligente de los mecanismos de protección de la vía respiratoria que caracterizan a la laringe. No obstante, el sonido producido por las cuerdas vocales no es el definitivo, y la onda sonora resultante debe recorrer otras estructuras pertenecientes a otros sistemas como el neuromuscular o el respiratorio (González, 1981). Estas estructuras son los llamados resonadores, cuya función es enfatizar determinadas bandas de frecuencias del espectro generado por las cuerdas vocales, conformado por un conjunto de frecuencias armónicas de la fundamental. Las señales acústicas resultantes se denominan formantes, correspondientes con cada una de las porciones resonadoras: la faringe, la cavidad nasal, la cavidad oral y el vestíbulo oral (Miyara, 1999).

De esta forma, la voz es uno de los medios de interacción comunicativa más eficientes que existe. La laringe convierte el flujo aéreo en señal acústica, la cual es modificada de manera particular en cada sujeto por los resonadores de la voz. De esta manera la expresión

del lenguaje se concreta en la voz, que es figura fundamental en la articulación de los mensajes ejecutados por el aparato bucofonatorio (Adrian y Casado, 2002).

## **1.2. Aparato Fonador**

Como se ha visto, el conjunto de órganos anatómicos que intervienen en la producción de la voz se denomina aparato fonador. Dichas estructuras no poseen una función exclusivamente lingüística (Guitart, 2004). Con fines didácticos se puede dividir en tres compartimentos principales: cavidades infraglólicas, cavidad glótica y cavidades supraglólicas (Adrian, et al, 2002). Dichas estructuras se describen a continuación.

### **1.2.1. Cavidades Infraglólicas**

Corresponden a estructuras ubicadas en zonas inferiores a las cuerdas vocales, donde despliegan sus principales funciones a nivel del sistema respiratorio. Estas son: diafragma, pulmones, tráquea y bronquios, las cuales proporcionan el aire necesario para la producción de la voz, modulando de esta manera la intensidad y duración de las emisiones (Villayandre, 2009). Es así como la voz es considerada una espiración sonora provocada por el paso de una corriente aérea a través del espacio entre las cuerdas vocales o glotis (Adrian, *et al.*, 2001).

El modo de operación de estas estructuras en la producción del aire para la fonación puede representarse como una secuencia. De esta manera, el diafragma actúa movilizandolas estructuras intraabdominales para favorecer una correcta dilatación de los pulmones, que a su vez actúan como almacenadores y propulsores del aire a utilizar. Esta columna aérea asciende a través de los bronquios, que corresponden a anillos cartilagosos originados de una bifurcación de la tráquea. Esta última es un tubo semirrígido que contacta directamente con la laringe, por lo tanto es la última estructura del esquema infraglólico encargada de proporcionar el aire necesario para la fonación (Romero, 2007).

### **1.2.2. Cavidad Glótica**

Compartimento constituido por la laringe, correspondiente al principal órgano fonador del ser humano. Esta es una estructura cartilaginosa y muscular encargada de proteger la vía aérea durante la deglución, regular la respiración y emitir sonidos. Gracias a esta misma configuración física, incluyendo tanto su musculatura extrínseca e intrínseca, es capaz de

tensar, aducir o abducir las cuerdas vocales, así como también desplazarse verticalmente al interior del cuello, de acuerdo a los requerimientos respiratorios, deglutorios y/o fonatorios (Jackson-Menaldi, 1992).

La laringe está formada por tres unidades anatómicas básicas: esqueleto cartilaginoso, musculatura extrínseca y musculatura intrínseca. (Casado, 2002). En primer lugar, la conformación cartilaginosa del órgano laríngeo está dada por cartílagos mayores (tiroideos, epiglotis y cricoides) y cartílagos menores (aritenoides, corniculados y cuneiformes). En segundo lugar, se encuentran los músculos extrínsecos de la laringe, tanto los infrahióideos como los suprahióideos, ambos grupos insertos en el hueso supensorio de la laringe o hioides. En tercer lugar, se observa musculatura con inserciones en la laringe propiamente tal, por lo que toman el nombre de músculos intrínsecos de la laringe, los cuales participan de manera sinérgica en el funcionamiento propio del órgano (Romero, 2007).

En cuanto a la producción de la voz destacan los músculos intrínsecos de la laringe que conforman las cuerdas o pliegues vocales. Estos son los tiroaritenoides, separados entre sí por un espacio triangular denominado glotis, por donde pasa el aire al subir desde las cavidades infragloticas. Este procedimiento fonatorio se ha tratado de explicar a través de los años y de acuerdo a eso se han generado diversas teorías. Según Jackson-Menaldi (1992), el planteamiento más aceptado corresponde al mioelástico - aerodinámico de Van den Berg, el cual divide la producción de la vibración cordal en tres pasos: a) las cuerdas vocales contactan en la línea media de la glotis, en un proceso de aducción, b) la presión de aire proveniente de las cavidades subglóticas produce un fuerza en los pliegues vocales aducidos, los que se abren momentáneamente, dejando salir la columna de aire y c) posterior a estas etapas se produce el efecto Bernouilli, el cual genera un fenómeno de aspiración de las cuerdas vocales hacia la línea media, con la consecuente producción de sonido.

### **1.2.3. Cavidades Supraglóticas**

Como se dijo anteriormente la producción vocal no basta con el paso de una columna de aire por la laringe, sino que necesita de la acción de diversas estructuras superiores que resalten ciertos parámetros de la onda acústica. Dichas estructuras corresponden a las cavidades supraglóticas, constituidas por la faringe, la cavidad oral y las cavidades nasales, conocidas también como resonadores.

La corriente de aire proveniente de la cavidad glótica es modificada en estas estructuras superiores, gracias al establecimiento de contactos y estrechamientos en diversos puntos de su recorrido. Estas transformaciones de la onda sonora, resultantes de los movimientos conjuntos de los órganos supraglóticos, serán las responsables de la impresión de un timbre característico para cada sujeto. De esta manera, las porciones del tracto superior a la glotis actúan como verdaderos filtros, amplificando solo frecuencias del espectro de la onda laríngea que coincidan con sus propias frecuencias. Este procedimiento tiene por resultado una voz constituida por diferentes formantes o frecuencias amplificadas (Obediente, 2007).

En síntesis, la principal función de las cavidades supraglóticas corresponde al proceso conocido como resonancia, en que el sonido o tono fundamental producido en la laringe es amplificado y enriquecido, para obtener las voces que escuchamos diariamente (Bustos, 2003).

## **2. ASPECTOS ACÚSTICOS**

### **2.1. Características Acústicas de los Sonidos**

La ciencia encargada del estudio del sonido, su producción, propagación, recepción, percepción y sus efectos en objetos y seres vivientes es la *Acústica*. La rama de estudio de la percepción del sonido corresponde a la Psicoacústica. Los parámetros psicoacústicos perceptibles por el ser humano son sonoridad, altura y duración, conceptos que en física se relacionan con amplitud, frecuencia y tiempo (Jackson-Menaldi, 1992). La relevancia de la psicoacústica radica en que permite medir las relaciones cuantitativas entre la sensación perceptiva y el estímulo físico que la produce. Por este motivo, todo experto forense debería considerar esta disciplina al momento de diseñar y ejecutar cualquier procedimiento de análisis de carácter auditivo (Delgado, 2001).

El diccionario de la Real Academia Española define sonido como la “Sensación producida en el órgano del oído por el movimiento vibratorio de los cuerpos, transmitido por un medio elástico como el aire”. Desde la perspectiva física, el sonido es una onda de presión longitudinal producida por una fuente de vibración mecánica, que se propaga a través de un medio elástico (Jackson-Menaldi, 1992).

Desde un punto de vista puramente físico, podemos clasificar los sonidos en dos grandes grupos: simples y complejos (Delgado, 2001). El movimiento recurrente simple o movimiento armónico simple es entendido como aquél movimiento continuo que constituye la base de todos los sonidos del habla, percibiéndose como un tono puro. Los parámetros que determinan este tipo de movimiento son: A o amplitud de la onda, T o período, definido como el tiempo en segundos que demora una partícula en dar una vuelta completa; y F o frecuencia, que corresponde al número de recurrencias (ciclos) efectuados durante un segundo (Jackson-Menaldi, 1992; Llisterri, 2010).

Los sonidos complejos se forman a partir de la suma de armónicos simples de distinta frecuencia (Bernal, 2000). De aquí que el sonido del habla se entienda como un movimiento armónico complejo que se constituye por la suma de movimientos armónicos simples, generados al ritmo de apertura y cierre de las cuerdas vocales. Durante dichos movimientos cordales, las partículas de aire entran en vibración a nivel del tracto vocal dando como resultado el habla (Jackson-Menaldi, 1992).

La onda sonora compleja que en general se produce a partir del sonido del habla, posee ciertos parámetros que permiten caracterizarla, estos corresponden a: frecuencia fundamental o  $F_0$ , definida como el ciclo de apertura y cierre de las cuerdas vocales en cada segundo y su valor se correlaciona con la longitud y masa de dichos pliegues; armónicos o componentes restantes de la onda sonora compleja que son múltiplos enteros de la frecuencia fundamental; resonancia, dada por las modificaciones en amplitud que sufren los armónicos gracias a las cavidades supraglóticas; formantes, que resultan de aquellas frecuencias amplificadas por la resonancia; y finalmente la amplitud o intensidad del sonido, energía acústica cuya medición se realiza en decibeles (dB) (Llisterri, 2010).

Existen diversas representaciones gráficas a las cuales se recurre para realizar estudios acústicos de habla. Estas permiten evidenciar las características o parámetros de la voz hablada, facilitando así su estudio. Dos ejemplos de ello son el oscilograma y el espectrograma. El primero registra y representa de forma gráfica las fluctuaciones de la amplitud a través del tiempo, y el segundo corresponde a una representación visual que permite observar tanto las variaciones de frecuencia (eje vertical) como de intensidad a través de los niveles de colores o grises del sonido graficado. (Llisterri, 2010).

## 2.2. Resonancia

El fenómeno de resonancia es considerado como uno de los componentes indispensables para la producción de un sonido, por lo que puede comprenderse como proceso y resultado acústico. La conjugación de ambos conceptos refiere la participación de diversas estructuras del tracto vocal que otorgan a la onda sonora primitiva la facultad de ser perceptible al oído humano (Guzmán, 2011). En concreto, la vibración de las cuerdas vocales produce la excitación de partículas de aire circundantes a ella, las que transmiten este movimiento a las paredes del sistema resonador y según cómo se aúnen los parámetros acústicos y fisiológicos, será el producto sonoro (Romero *et al.*, 2007). De esta manera la amplificación y/o amortiguación de los armónicos producidos en la laringe será en función del tipo de resonador en el que dicho sonido vibra. En el caso de la voz, las cavidades supraglóticas (en términos generales, cavidades faríngea, oral y nasal) son las que asumen este rol, creando una relación inversa entre forma y frecuencia amplificada: resonadores pequeños refuerzan la amplitud de tonos altos, mientras que resonadores grandes amplifican armónicos de baja frecuencia (Villayandre, 2009). Por su parte, rasgos como la tensión y viscosidad de las paredes (Romero *et al.*, 2007) también son determinantes tanto al definir el perfil acústico de la frecuencia natural o propia de un resonador, como a la manera en que será resonada una determinada onda (Villayandre, 2009).

Bajo este contexto, la resonancia por vibración da lugar a una serie de alteraciones de la onda sonora de acuerdo a cómo esta se asocie con la cavidad supraglótica en cuestión. Así, Villayandre (2009) establece que el resonador enfatiza o amplifica las frecuencias que coinciden con su frecuencia natural, y por otro lado, desestima o filtra aquellas que no concuerdan con sus parámetros tonales propios, siendo responsable del timbre<sup>2</sup>. Como resultado de este proceso de filtración de frecuencias, se obtienen concentraciones de energía acústica en torno a la frecuencia fundamental o natural del resonador conocidas como formantes y sus valores van desde los 100 Hz hasta los 3500 Hz. (Delgado, 2001).

---

<sup>2</sup> Parámetro vocal, resultado del paso de la energía acústica producida por la fuente sonora, a través de una caja de resonancia. Este se caracteriza por su alta capacidad individualizadora (Delgado, 2001). En la práctica, el timbre es aquel rasgo que permite diferenciar dos sonidos que se producen a igual frecuencia e intensidad (Romero *et al.*, 2007).

Como se ha dicho anteriormente, cada cavidad resonadora se corresponde con una gama frecuencial y formante específico. El  $F_1$  o primer formante se relaciona con la faringe, el segundo ( $F_2$ ) representa la cavidad oral y el tercero o  $F_3$  con la región de los labios (Bernal *et al.*, 2000). Marrero *et al.* (2003) postulan que además de ello, el formante tres sugiere información en cuanto a la longitud del tracto vocal. Se postula que  $F_3$  no tiene una función lingüística exclusiva, por lo que podría funcionar como índice de individualidad del hablante. Además se ha comprobado una relación sistemática entre la altura de  $F_3$  y el tamaño de la cavidad oral, en particular de la zona tras los incisivos inferiores, así cuanto menor sea esta dimensión, más alto estará ubicado  $F_3$  (Albalá & cols., 2008).

Desde la perspectiva acústica, las vocales corresponden a ondas complejas periódicas de amplitud más elevada en comparación con las consonantes, lo que implica un mayor nivel de energía asociado a la participación de las cuerdas vocales y su ininterrumpida emisión articulatoria. Para su estudio y descripción se sugieren los formantes  $F_1$  y  $F_2$  (Román, Quezada y Sabaj, 2000). En la práctica se ha podido apreciar que el gráfico del primer formante (ordenada) de una vocal en función de la segunda frecuencia del formante (abscisas), se asimila a la posición de las vocales en el denominado triángulo vocálico, realizado de acuerdo a las variables de apertura y lugar.

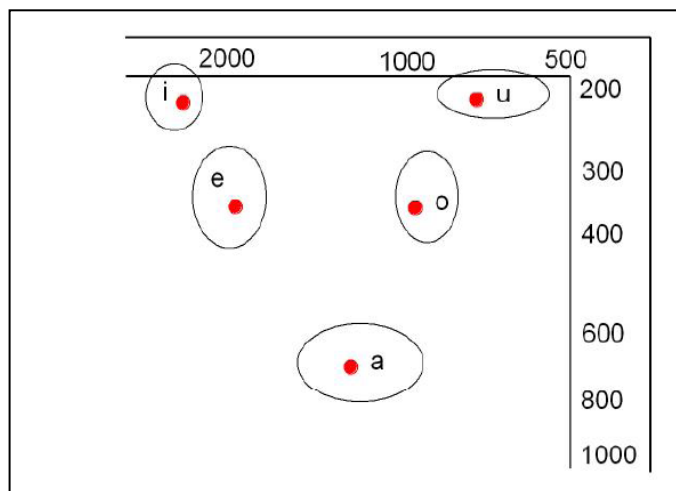


Gráfico 1. Triángulo vocálico del español, formado al graficar  $F_1$  en el eje de las ordenadas y  $F_2$  en el eje de las abscisas (Celis, 2009).

El grado de apertura se relaciona directamente con  $F_1$  (a mayor apertura, mayor será  $F_1$ ) y el lugar de articulación se asocia con  $F_2$  (cuanto más anterior, mayor será  $F_2$ ).



Además de lo señalado, el proyecto VILE reveló que la variabilidad interlocutor superaba los 200 Hz en  $F_2$  y los 100 Hz en  $F_1$  para las vocales. Bajo esta idea se postula que un hablante con elevada variabilidad para  $F_2$  y reducida en  $F_1$ , puede ser identificado con un alto porcentaje de acierto, mientras que un locutor con poca variabilidad de  $F_2$  y alta para  $F_1$  presenta mayores dificultades para su identificación (Celis, 2009).

Por su lado, los sonidos consonánticos se producen con cierta constricción dada por la aproximación de estructuras articulatorias. De acuerdo a su función cordal, estos pueden ser sonoros o áfonos, rasgos que indican presencia o ausencia de vibración de los pliegues vocales, respectivamente (Bernal *et al.*, 2000).

Como fue mencionado, los formantes uno y dos de las vocales permiten una mejor caracterización acústica y descripción del movimiento lingual asociado;  $F_1$  se relaciona con el eje vertical alto-bajo, mientras que  $F_2$  con la dimensión antero-posterior (Ruiz y Soto-Barba, 2005).

A continuación en la tabla 1, se presenta el promedio de valores formánticos ( $F_1 - F_2 - F_3$ ) de las vocales. Luego, en la tabla 2 se muestran las cifras diferenciadas según el género.

	<b>U</b>	<b>O</b>	<b>A</b>	<b>E</b>	<b>I</b>
<b>F<sub>1</sub></b>	350	530	900	440	310
<b>F<sub>2</sub></b>	730	900	1400	2300	2500
<b>F<sub>3</sub></b>	2600	2700	2700	2900	3000

Tabla I. Valores Formánticos de Vocales (Hormazábal 2011).

	Hombres		Mujeres	
Vocal	F <sub>1</sub>	F <sub>2</sub>	F <sub>1</sub>	F <sub>2</sub>
I	265	2318	241	2835
E	454	1995	492	2252
A	657	1215	664	1168
O	475	888	511	981
U	294	669	243	629

Tabla II. Promedio de valores formánticos de las vocales según género (Román, 2010).

### 3. ASPECTOS FONÉTICO-FONOLÓGICOS

#### 3.1. Elementos Segmentales y Suprasegmentales

Se define como segmental, a toda unidad discreta que puede ser limitada en un rasgo diferenciable e individual. Por patrones suprasegmentales se entienden aquellos fenómenos que involucran más de una unidad (Villayandre, 2009). Ahora bien, el estudio fonético del habla se sustenta en el análisis de elementos segmentales de una lengua en particular. Los parámetros que pertenecen a esta categoría, son: el sonido, el fonema, los alófonos o fonos, la grafía y el alfabeto fonético. El primero de ellos corresponde a la unidad del habla que se manifiesta acústicamente y es generado por el órgano fonatorio (Llisterri, 2011). El fonema es la mínima unidad significativa de percepción mental de un sonido (Frías, 2001; Hualde, 2001). Por su lado, Llisterri (2011) define los alófonos como la expresión acústica de un fonema y la grafía como la unidad escrita de una lengua. Finalmente, el mismo autor señala en el año 2003 que el alfabeto fonético es el conjunto de representaciones escritas de los fonemas donde existe una asociación sonido-símbolo.

Los fonemas se dividen en consonánticos y vocálicos. Los primeros “se caracterizan por presentar una energía total reducida. Articulatoriamente, presentan un obstáculo al paso del aire en el canal vocal, que causa una turbulencia o, incluso, una interrupción del flujo de aquél” (Obediente, 2007).

La clasificación de los fonemas consonánticos se realiza a partir de tres parámetros: el punto articulatorio, el modo articulatorio y la sonoridad. El primero hace referencia al lugar

dentro de la cavidad oral en donde se articula un fonema en específico. En la articulación del sonido, el paso de aire de los pulmones se ve total o parcialmente impedido por los órganos articulatorios activos<sup>3</sup> y pasivos<sup>4</sup>. (Frías, 2001; Hualde, 2001). El modo articulatorio hace referencia a la posición que toman los órganos fonoarticulatorios en cuanto al grado y regularidad de apertura en el momento del paso del aire, en el inicio o durante la ejecución del fonema (Frías, 2001; Hualde, 2001). Por último, la sonoridad, es el resultado de la vibración de las cuerdas vocales en el momento en que pasa el aire entre ellas, durante la ejecución de los fonemas sonoros. En los fonemas áfonos, los pliegues vocales no ofrecen resistencia al paso del aire (Hualde, 2001).

Según Miraya (1999) se define a los fonemas vocálicos como sonidos emitidos por la vibración de las cuerdas vocales, sin ningún obstáculo o constricción entre la laringe y las aberturas oral y nasal. Respecto a la articulación vocálica, el mismo autor señala que este acto consiste en la modificación de la acción de filtro que cumplen los distintos resonadores a lo largo del tracto vocal. Asevera que esto, “depende de las posiciones de la lengua (tanto en elevación como en profundidad o avance), de la mandíbula inferior, de los labios y del paladar blando. Estos órganos influyen sobre los formantes, permitiendo su control” (Miraya, 1999, p. 8).

Al igual que para las consonantes, existen tres parámetros que permiten clasificarlos: altura de la porción lingual dorsal, posición de la lengua y posición de los labios. El primer parámetro hace alusión a la cercanía que logra la región dorsal de la lengua con el paladar duro o blando, al momento de la articulación. Las distintas posiciones se pueden dividir en alta (/i/, /u/), media (/o/, /e/) y baja (/a/). El segundo parámetro hace referencia a la posición que ocupa la lengua dentro de la cavidad oral al momento de realizar la articulación, permite hacer la división entre vocales anteriores (/i/, /e/), centrales (/a/) y posteriores (/o/, /u/) (Quilis y Fernández, 1990). Para la mayor parte de los fonetistas, los dos parámetros vocálicos, antes mencionados, son considerados suficientes para establecer rasgos oposicionales que permiten constituir el sistema de fonemas vocálicos, sin embargo, en varias ocasiones, también se considera el tercer parámetro, la posición de los labios. En la descripción de este parámetro, no existe consenso en relación a la división a partir de la posición de los labios al momento de la articulación. Algunos autores afirman que la conformación labial permite separar las vocales en

---

<sup>3</sup> Corresponden a las siguientes estructuras neuromusculares: lengua, velo palatino, mandíbula y labios (Romero *et al.*, 2007).

<sup>4</sup> Corresponden a: paladar duro, dientes, alvéolos y nariz (Romero *et al.*, 2007).

neutras (/a/), estiradas (/i/, /e/) y redondeadas (/o/, /u/). En cambio otros, las dividen solo en redondeadas o no (Quilis et al., *n.d.*).

La voz es también un fenómeno complejo constituido por distintos elementos suprasegmentales. Estos guardan relación con un amplio número de modulaciones naturales de la voz que permiten cambiar el significado de lo que se dice. También son componentes capaces de entregar información sobre el sujeto que las emite como, la nacionalidad, nivel socioeconómico, edad, género, estado emocional, entre otras (Torres, 2005). Estos fenómenos fonético-fonológicos, tienen la capacidad de actuar simultáneamente sobre más de un segmento lingüístico como una sílaba, una palabra, un enunciado, etc. (Iribar, *n.d.*). Tal es la importancia de estos componentes de la voz, que se ha postulado que el estudio de estos segmentos en forma aislada, permitiría identificar a un individuo (Minematsu, Nishimura, Murakami y Hirose, 2006).

Considerando la gran importancia de estos atributos en la voz humana, se hace necesario incluirlos dentro del análisis acústico, prestando gran atención en la capacidad que tienen para modificar características vocales y por ende, alterar distintos patrones en el espectrograma, como por ejemplo, la frecuencia fundamental. Según Llisterri (2003), los elementos que conforman los aspectos suprasegmentales son:

1. Melodía: variación de la altura tonal a lo largo de un enunciado (Llisterri, 2003)
2. Acento: sensación perceptiva que pone de relieve una sílaba sobre el resto de las sílabas de la palabra (Gil, 2007).
3. Tono: grado de tensión y distensión (vibración) que provoca el flujo en los repertorios de patrones vocálicos (Sabido, 2002).
4. Ritmo: sensación perceptiva provocada por la sucesión de determinados elementos en periodos regulares de tiempo (Gil, 2007). De esta forma, "Se puede hablar de ritmo cada vez que sea posible prever qué es lo que va a seguir en función de aquello percibido. Hay por consiguiente, una relación estrecha entre ritmo y previsibilidad" (Jackson-Menaldi, 1992).
5. Entonación: sensación perceptiva que produce, fundamentalmente, las variaciones de tono a lo largo de un enunciado (Gil, 2007).
6. Pausas: silencio o vocalización intercalados en el discurso (Gil, 2007).
7. Velocidad de elocución: número de elementos fónicos (sonidos y pausas) que se pronuncian en una unidad de tiempo determinada (Gil, 2007)

8. Calidad de la voz: características globales debidas a los ajustes laríngeos y de las cavidades supraglóticas de un hablante (Llisterri, 2010).

Los elementos mencionados guardan una estrecha relación con distintos aspectos presentes en el análisis acústico. La frecuencia fundamental, que corresponde a la frecuencia más baja del espectro frecuencial de la voz, determina la melodía, el tono y acento de la emisión. Por su parte, el acento y la calidad de la voz se vinculan fuertemente con la configuración espectral, que varía según las cualidades supraglóticas. La amplitud, que corresponde al aumento o descenso de los máximos de la presión aérea durante el ciclo vibratorio, se correlaciona con el acento y el ritmo. Estos parámetros varían según la actividad de los músculos respiratorios y la presión subglótica. Por último el tiempo, que depende de la organización temporal de la articulación, se vincula con el acento, las pausas, la velocidad de elocución y el ritmo (Llisterri, 2010). Lo anterior se enmarca dentro del estudio de la prosodia, la cual incluye fenómenos como acentuación, entonación, ritmo y fluidez del habla. Las características prosódicas representan todo lo que, en la realización del habla, escapa al análisis de los fonemas (Narbona & Chevie Muller, 2003).

Zañartu (2003) afirma que debe incluirse la información prosódica en caso que se lleve a cabo un análisis acústico, ya que para cumplir el objetivo de las tres áreas básicas de las tecnologías del habla (síntesis, reconocimiento y sistemas de diálogo) se requiere de estos parámetros. Agrega además, que uno de los aspectos más importantes dentro del análisis de los fenómenos prosódicos que permiten la identificación vocal de un sujeto, es consecuencia de la frecuencia fundamental.

## **VI. DESARROLLO DEL TEMA**

### ***1. CONCEPTOS FUDAMENTALES***

El ***lenguaje*** es definido como la capacidad o facultad propia del ser humano para comunicarse. Dicho proceso se lleva a cabo fundamentalmente mediante la modalidad oral, aunque también puede gestarse bajo el modo escrito. Ambas formas constituyen lo que se conoce como lenguaje verbal. Desde otra perspectiva, el lenguaje puede ser entendido como un sistema de signos lingüísticos, convencional y económico que se sustenta en reglas de

combinación tanto paradigmáticas (relaciones intercambiables) como sintagmáticas (relación lineal de cada elemento con su contiguo), base de la relación e interacción con el otro. Se añade a esta idea, que puede ser visto como una representación simbólica de la actividad mental, es decir, como herramienta de transformación de los pensamientos en palabras (planeamiento o procesamiento simbólico), lo que incluye la selección de estas en relación a las ideas, su organización en frases y la coordinación de estas en un discurso. Todo ello bajo una motivación social que regule y adapte su uso de acuerdo al contexto comunicativo (Narbona & Chevrie-Muller, 2003).

Desde la perspectiva lingüística, el lenguaje conjuga tanto una parte social o lengua, como una individual o **habla**. Así, este último es definido como el acto expresivo lingüístico concreto que realiza el individuo para materializar de forma física y fisiológica los datos abstractos del lenguaje y la lengua. Es entonces, la realización motora del lenguaje (Queiroz, 2002). La producción del habla requiere la programación de los actos motores, lo que implica definir qué órganos estarán comprometidos y con qué ritmo, velocidad e intensidad deberán ser realizados los movimientos. Para ello, los impulsos nerviosos (transmisores de esta información) viajan por las vías descendentes hasta llegar a la musculatura orofacial para producir los sonidos del habla.

De forma más particular, el habla responde a una norma determinada en directa correspondencia con quienes la usan. Según plantean Richards, Platt y Weber (1985), a grandes rasgos, el **hablante nativo** es aquella persona que habla la lengua del país donde nació. También se habla de la “natividad” como la encarnación auténtica de un idioma modelo en particular. Sin embargo, años más tarde esta concepción comienza a variar y complejizarse. Algunos especialistas incluso comienzan a cuestionar el carácter de identidad que otorga el concepto al perfil vocal del individuo. Lo anterior responde de alguna u otra forma a las transformaciones vividas en las últimas décadas, que contemplan múltiples migraciones y encuentros transculturales (Byran & Fleming, 1998).

Haughes y Lascaratou (1982) por su parte, definen al hablante nativo como aquel individuo que aprende la lengua en la infancia y además procede de un lugar geográfico en que se hable dicho idioma.

En un intento por aunar criterios para definir al hablante nativo, Davies (2003) propone 5 aspectos a considerar:

- Si la persona adquirió la lengua desde su nacimiento.

- Si la persona es un hablante competente.
- Si la persona adquirió la lengua a través de instrucción formal o mediante el uso cotidiano de esta.
- La variedad de la lengua que presenta el individuo y;
- La raza del hablante.

Para que el habla sea percibida por otras personas y así generar el acto comunicativo, se requiere de un sustrato fisiológico correspondiente a la **voz**. Esta puede considerarse como un sonido, que producido por la laringe y amplificado por las estructuras de la resonancia, nos permite la interacción comunicativa con otras personas. Su proceso se inicia con la voluntad, cuando aparece el deseo de emitir un sonido. Este desencadena en el sistema nervioso central un gran número de órdenes que pondrán en funcionamiento todos los elementos que configuran la voz: mecanismos de la respiración, de la fonación, de la articulación, de la resonancia, de la expresión, etc. (Tulon, 2005). La producción de la voz depende de la relación equilibrada entre las fuerzas ejercidas por los músculos intrínsecos de la laringe que en acción conjunta con otros sistemas, como el auditivo, vestibular, sistema nervioso central y periférico, ponen en movimiento las cuerdas vocales (Bustos, 2003).

Para Tulon (2003) dichos aspectos pueden resumirse de la siguiente forma: el lenguaje es el medio que permite comunicarnos con nuestros semejantes, a través de él expresamos nuestras ideas, emociones y experiencias. Por medio del habla nos mostramos tal cual somos, damos a conocer nuestra identidad tanto en lo físico como en lo psíquico, y finalmente la voz nos representa, es nuestra tarjeta de presentación.

En la actualidad, todas las técnicas o mecanismos que se han propuesto para identificar a un hablante, pretenden el mismo objetivo: caracterizar la voz de una persona a tal punto, que puedan conocerse las variables que comparte con otros individuos y aquellas que la diferencian de los demás. Bajo esta consigna, es importante definir el concepto de **perfil vocal**. Este concepto corresponde al conjunto de características vocales de una persona en particular y que se pueden valorar a través de una evaluación auditivo-perceptiva ( $F_0$ , resonancia, brillo, colocación, TMH, etc.). Este perfil tiene como base el sonido que se produce a nivel cordal y que es modificado por los distintos órganos resonadores. Junto a los ajustes motores de los resonadores, se debe considerar también que las características vocales son afectadas por aspectos socio-económicos, culturales, preferencias individuales, geográficos, emocionales, edad, género, entre otros (Barros y Cavalcanti, 2006).

Toda producción vocal se da en un ambiente o contexto acústico en que el que pueden influir diversas variables externas, como es el **ruido de fondo**. Cuando se lleva a cabo un análisis acústico de la voz o simplemente se atiende a una señal sonora determinada, este concepto se define como cualquier sonido indeseado que se produzca de manera simultánea a la emisión de la señal que se está analizando. En otras palabras, corresponde a señales emitidas por fuentes externas o del entorno que se superponen a una señal en particular que sea de interés para el oyente en ese momento. De este modo, para considerar que una señal sonora corresponde a un ruido de fondo, esta debe superponerse al sonido que esté siendo analizado o percibido por un auditor determinado (Recuero, 1996).

Otro elemento importante, es el **corpus**. Este es considerado como una herramienta útil en fonoaudiología para estudiar y caracterizar el habla de un sujeto. Los relatos que se obtienen son considerados como “habla de laboratorio”, ya que suelen ser reproducciones de frases preparadas y leídas o en algunos casos, frases inducidas por el investigador en conversaciones semi-estructuradas (Cantero y Font, 2009). Desde el punto de vista fonético, un corpus debe representar la frecuencia en que aparecen los alófonos correspondientes a un fonema, tal que sea reflejo de la norma de una lengua específica (Llisterri, 2005). Esto último alude a lo que se conoce como un texto “fonéticamente equilibrado”.

A continuación se presentan los resultados de un estudio realizado en Chile por Pérez en el año 2003, alusivo a la frecuencia de uso de fonemas en el español de Chile.

<b>/i/</b>	7.46	<b>/f/</b>	0.75
<b>/e/</b>	14.13	<b>/s/</b>	9.61
<b>/a/</b>	12.31	<b>/x/</b>	0.74
<b>/o/</b>	9.28	<b>/ch/</b>	0.32
<b>/u/</b>	3.05	<b>/y/</b>	0.69
<b>/p/</b>	2.58	<b>/m/</b>	2.62
<b>/t/</b>	4.92	<b>/n/</b>	7.78
<b>/k/</b>	3.94	<b>/ñ/</b>	0.24
<b>/b/</b>	1.92	<b>/r/</b>	6.19
<b>/d/</b>	4.84	<b>/rr/</b>	0.64
<b>/g/</b>	0.94	<b>/l/</b>	5.05

Tabla III. Distribución porcentual de la frecuencia de los fonemas del español (Pérez, 2003).



## ***2. SOBRE LA ACÚSTICA DEL CRIMEN Y LA FONÉTICA FORENSE***

La fonética es una disciplina lingüística, teórica y experimental. Una de sus aplicaciones más desafiantes y al mismo tiempo más desconocidas es la que se da en el ámbito jurídico. La fonética forense implica el análisis de patrones acústicos a fin de poder identificar al imputado a través de su voz, lo que la vuelve una prueba importante en los procesos legislativos. Pese a ello, en la actualidad se discute hasta qué punto esta definición podría operacionalizarse con exactitud, ya que existe evidencia de factores que alteran la homogeneidad de la muestra, sean estos propios de la persona (estado anímico) o externos a ella, como los ruidos ambientales, distancia hacia el micrófono, entre otros. Es así que se postula que más que una identificación de la voz, se trata de una verificación de ella (Fernández Planas, 1998).

La fonética forense presenta sus inicios a principios de la década de los 90 y su principal función corresponde a la identificación de hablantes. El especialista que se desempeña en esta área se conoce como perito, el cual puede actuar como testigo en corte, entregando conocimientos más allá de los que tiene el juez en un determinado caso. De acuerdo a esto, el profesional en cuestión declara acerca de la procedencia de una determinada muestra de habla a fin de inculpar o exculpar a un sujeto presuntamente involucrado en un crimen. En la práctica, algunas de las actividades realizadas por estos científicos son: caracterización de hablantes (perfil de habla), realce de la inteligibilidad del habla en grabaciones, inspección de la autenticidad e integridad de los registros de audio y análisis e identificación de sonidos no propios del habla o ruidos de fondo en grabaciones correspondientes a evidencia (Ashby en Brown, 2006).

Adicional a las funciones anteriormente mencionadas, el perito debe conocer el código al que se adscribe cada muestra (Romito & Galatà, 2004) e indagar el sistema funcional de pertenencia (el registro, dialecto y sociolecto), es decir manejar las formas que los diferentes valores comunicativos adquieren en los distintos contextos, situaciones y registros.

En fonética forense, el científico recibe dos grabaciones para ser comparadas: una, de un hablante conocido, y la otra, de un hablante cuya identidad está en cuestión. Su objetivo es entonces, determinar qué tan significativas son las semejanzas entre las descubiertas en el análisis. Para lograrlo, se ha propuesto el uso de un método estadístico basado en la aplicación de una fórmula bayesiana, utilizada por los sistemas automáticos de reconocimiento (Robertson

y Vignaux 1995; González-Rodríguez *et al.* 2003). Esta debe calcular el LR (likelihood ratio), variable que entrega información acerca de las diferencias acústicas que existen entre dichas muestras de voz, tanto bajo la hipótesis que estas fueron producidas por el mismo sujeto del hablante conocido, como bajo la hipótesis que fueron producidas por hablantes distintos. Lo anterior se traduce en la siguiente fórmula:

$$LR = \frac{p(\text{diferencia acústica observada/ mismo hablante})}{p(\text{diferencia acústica observada/ distintos hablantes})}$$

De este modo, si LR resulta mayor que uno, se refuerza la hipótesis que ambas muestras fueron producidas por el mismo hablante. Por el contrario, si LR es menor que uno, se apoya la idea de que las voces corresponden a hablantes distintos (Stewart, 2011).

Una de las desventajas de este razonamiento, consiste en la exclusión de los fenómenos lingüísticos asociados al momento de llevar a cabo las interpretaciones pertinentes, abocándose a la caracterización acústica de las muestras analizadas (Robertson y Vignaux 1995; González-Rodríguez *et al.* 2003).

La tarea de identificación en fonética forense se basa en métodos auditivos o perceptivo-acústicos, los cuales son frecuentemente complementados con el análisis acústico de ciertos rasgos como: frecuencia fundamental, pitch, entonación y calidad vocálica. Por su parte, los resultados se deducen a partir de un detallado análisis acerca de: las variaciones del acento y dialecto, calidad de la voz, recurrencia de modismos o de cualquier patrón léxico, sintáctico o paralingüístico, teniendo siempre en consideración los contextos comunicativos en los cuales fueron producidas las muestras de habla, así como también el estado físico y emocional de los hablantes involucrados. El científico que se desempeña en fonética forense, intentará por tanto rescatar aquellos rasgos que se desvían de la norma y los que son relativamente permanentes o difíciles de modificar, ya sea consciente o inconscientemente por el hablante. Cabe destacar sin embargo, que dada la inexistencia de criterios cuantitativos rigurosos, dichas conclusiones deben ser formuladas en términos de probabilidades, y no de certezas definitivas (Ashby en Brown, 2006).

## **2.1. Corrientes Principales Presentes en el Análisis de la Voz**

### **2.1.1. Reconocimiento**

Furui en Delgado (2001) define este concepto como todo proceso automático de reconocimiento de hablantes que utiliza la información individual incluida en la señal de habla. A su vez, indica que este procesamiento se divide en identificación y verificación de hablantes. Concordante a esto, los ingenieros Doddington, Bishnu, Atal y Rosenberg (1972) incluyeron dentro de este término todas aquellas tareas relacionadas con la identificación, verificación, discriminación y autenticación de hablantes.

En el caso de las técnicas de reconocimiento automático, estas deben adaptar las condiciones externas y generar resultados comprensibles por las distintas instituciones. Estos son propuestos en base a una razón de verosimilitud (LR), que como fue dicho anteriormente, mide la fuerza con que la evidencia apoya si el sospechoso ha realizado o no, la voz dubitada (Delgado, 2001). Para estimar el LR se calcula la probabilidad de ambas hipótesis y luego la relación entre ellas (Teorema de Bayes). Con el fin de regular los métodos de reconocimiento forense se han organizado las probabilidades de acuerdo a los llamados "Tippetplots", que representan la mayor probabilidad de razón de verosimilitud. A través de este tipo de análisis se pueden definir niveles indeseables de error (distorsiones por canal de transmisión y/o ruido) y con ello validar o no, la prueba efectuada (Alexander, 2005).

La secuencia de análisis implica una primera etapa de pre-procesamiento del archivo de entrada, extracción de las características a analizar, modelado de funciones, comparación de resultados e interpretación de acuerdo a la razón de probabilidad. Para compensar la variabilidad de condiciones externas se utiliza un sistema de puntuación. Por otra parte, se ha demostrado que la inclusión de parámetros prosódicos en el análisis, mejora el nivel de certeza para la verificación de hablantes (Alexander, 2005).

### **2.1.2. Verificación**

Doddington, Bishnu, Atal y Rosenberg (1972) hablan de verificación, cuando la muestra anónima es aportada por el emisor de forma "cooperativa", es decir, existe un interés en que se verifique una comparación con otro registro de voz para obtener un resultado positivo. Además,

se necesita de una única muestra patrón predeterminada, para llevar a cabo el cotejo. Tosi (en Delgado, 2001) denomina "autenticación" a la verificación precisa de una serie de requisitos para el análisis: buena calidad del material acústico, disposición de una base de datos con un número limitado de muestras indubitadas, canales idénticos para la transmisión y grabación de las muestras de habla. De esta manera se eliminan los efectos derivados del uso de distintas curvas de respuesta.

Hollien (1990) señala que la verificación no requiere interpretar lo que el hablante dice, sino sólo verificar su identidad bajo el supuesto que dicho interlocutor desea ser reconocido. Los potenciales usos de los sistemas de verificación son múltiples, algunos ejemplos son: acceso de determinadas personas a áreas de seguridad, verificación de la identidad de un oficial que da instrucciones vía radio de comunicación, walkie-talkie o cualquier canal en que no pueda ser identificado a través de la vista, verificación de la identidad de sujetos que hablan desde aviones, cápsulas espaciales, cámaras hiperbáricas, tanques o cualquier otra ubicación remota. Existe la idea que la tarea de verificación es relativamente sencilla. En general, el hablante es cooperador, el equipamiento de análisis es de alta calidad, y la naturaleza de la muestra de habla está controlada. No obstante lo anterior, aún persisten ciertos problemas sin solucionar. El principal de ellos es que parece no haber un sistema operacional en línea que permita verificar la identidad de un gran número de hablantes únicamente a través del análisis de sus voces. Los resultados obtenidos en la verificación nos pondrán en dos posibles circunstancias. La primera, en donde el sistema de verificación del hablante arroja los resultados correctos (aceptar un locutor registrado o rechazar un impostor). La segunda, en donde el sistema comete un error y rechaza a un hablante registrado o acepta a un impostor.

Todos los sistemas de VL (Verificación del Locutor) cuentan con una base de datos de usuarios registrados, denominados clientes. Ella está compuesta por modelos que representan las características del habla de cada uno de ellos. Estos modelos se consiguen mediante el análisis de sesiones de entrenamiento en las cuales el usuario del sistema pronunciará varias frases. Existen diversos sistemas VL, entre los cuales destacan los de texto dependiente (el hablante debe elicitar una palabra o frase específica) y los de texto independiente (procesan cualquier palabra o frase). Dentro de estos últimos encontramos los de pronunciación continua (secuencia de palabras sin pausa) y los de pronunciación aislada (secuencia de palabras con pausas entre ellas). Estos sistemas han demostrado tener un grado de error muy bajo (0,3% a 1% a bajos niveles de ruido) (Becerra & Villar, 2000).

Los sistemas VL requieren que los usuarios estén registrados, por esto se deben realizar sesiones de entrenamiento. Los factores que determinarán el funcionamiento del sistema son:

- El número y periodicidad de registros: mientras mayor sean, se obtendrá un modelo más representativo del cliente.
- La duración de las sesiones de verificación.
- Registros separados en el tiempo: permiten considerar las variaciones naturales en la voz, determinadas por el estado emocional o de salud.
- Canal de transmisión: favorecer el uso del mismo canal de transmisión para mejorar el desempeño del sistema.

El siguiente esquema describe de forma general, la manera en que opera el sistema de Verificación de Locutor (VL) (Becerra & Villar, 2000).

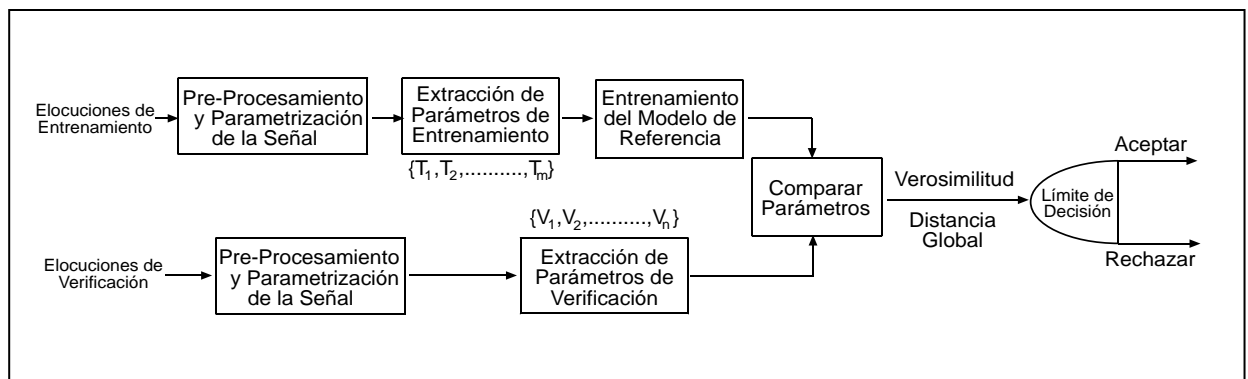


Gráfico 2. Bloques de un sistema de Verificación de Locutores (Becerra & Villar, 2000).

Por otra parte, se deben considerar las características que se espera posean los factores a analizar, los que de acuerdo a Becerra & Villar (2000) son:

- Representar en gran medida las características del interlocutor, y en pequeña medida las del intralocutor.
- Fácil extracción.
- No variar bruscamente en el tiempo.
- Independientes del estado de salud y/o ánimo del locutor.
- No modificables conscientemente.
- Robustos ante ruido ambiental o canales de transmisión.

Finalmente, en el contexto de verificación automática, el problema radica en las variaciones inter e intralocutor, ya que los resultados pueden enmascarse unos con otros, y como resultado, entregar cierta proporción de falsos positivos. Además existen dificultades debido a una alta sensibilidad a los efectos de canal de transmisión (teléfonos, codificadores, etc.), técnicas basadas en la percepción de compresión usada en reproductores de audio y formatos de compresión, así como condiciones del texto y alteraciones debidas a ruido de fondo. Algunos autores consideran que para generar mejoras en el análisis, se deberían contemplar estudios lingüísticos de las variables de la voz, ya que estos parámetros contemplan información específica de cada hablante (Broeders, 2001).

### 2.1.3. Identificación

Para definir esta corriente de análisis, Hollien (2002) propone que corresponde a un método o conjunto de procedimientos en los que se busca dar con la identidad de un sujeto desconocido. Para esta finalidad se dispone de dos tipos de grabaciones de audio: la "evidencia" y los "ejemplares". La evidencia corresponde a la fuente desde la cual se obtiene la voz del desconocido. Los ejemplares por su parte, corresponden a archivos de voces de personas que podrían ser, potencialmente, el individuo desconocido. Por lo tanto, estos últimos pertenecen a personas cuya identidad es conocida. Autores como Doddington, Bishnu, Atal y Rosenberg (1972) definen identificación como el proceso en el que una máquina debe asociar una muestra de habla desconocida con la más similar de entre un número indeterminado de muestras conocidas o indubitadas (Delgado, 2011).

Furui (1994) indica: "(...) la diferencia fundamental entre identificación y verificación es el número de decisiones alternativas. En identificación, el número de decisiones alternativas es igual al número de sujetos de la población que conforma la base de datos, mientras que en verificación sólo existen dos decisiones alternativas, aceptar o rechazar, con independencia de la talla de la población".

En este mismo sentido, Ashby (2006) refiere que la identificación se ocupa de establecer la identidad de un hablante miembro de una población potencialmente ilimitada, mientras que la verificación pretende determinar si un hablante de una comunidad cerrada, es realmente el miembro que él afirma ser.

La identificación forense de hablantes surge como procedimiento formal en el siglo XXI, correspondiente a un método de peritaje que contempla aspectos relacionados tanto con el sonido, como con el lenguaje. En determinados procesos legales se busca establecer si una o más grabaciones de habla, pertenecen al mismo individuo. Una de las primeras instituciones en utilizar este método fue la Policía Criminal Federal de Alemania en 1980, y en 1989 se realiza la primera conferencia referente a la temática en cuestión. Desde entonces, el gran avance en materia de telecomunicaciones y registro de material audiodigital ha permitido manejar de forma más óptima las muestras, reduciendo factores que dificultan el análisis. A pesar del desarrollo actual, aún no existe consenso suficiente por parte de los profesionales involucrados (fonetistas, lingüistas, agencias legales, entre otros), que permita la construcción de una metodología válida (Rose, 2002). Por lo tanto aún quedan muchos desafíos respecto al control de variables para una identificación certera del hablante.

Hollien (1990) afirma que la identificación de hablantes corresponde al desafío más difícil de resolver debido a varias razones. Usualmente el hablante no coopera o derechamente modifica las características de su voz, aspecto al que se suman variaciones en su condición de salud, ingesta de drogas y diversos estados emocionales como estrés o miedo, y por último, la distorsión de la señal acústica y el ruido de fondo. Todos estos factores dificultan el proceso de identificación de un hablante.

## **2.2. Tipos de Análisis Utilizados en el Ámbito Forense**

Bajo la constante actualización y permanente descubrimiento de nuevas técnicas, análisis y aplicaciones, la INTERPOL (2001) realizó un congreso con el fin de difundir aspectos necesarios de aunar. De tal forma, y frente a la necesidad de organizar la información disponible, dicha institución define y describe los siguientes tipos de análisis para lograr la identificación de hablantes. En el presente trabajo se entenderá por tipo de análisis a los procedimientos sugeridos por la INTERPOL para identificar hablantes. Estos contemplan aspectos de voz, habla y lenguaje, por lo que sus variables pueden ser estudiadas por expertos con conocimientos en dichas áreas.

### 2.2.1. Identificación de Hablantes por Testigos de Audio

Por *earwitness* (“*testigo de audio*”) se entiende a toda persona que testifica de acuerdo a lo que escucha (sin posibilidad de ver). Es el método más antiguo de identificación, remontándose incluso a épocas bíblicas. Este procedimiento debe complementarse con la comparación multipersonal de sujetos seleccionados para identificar (line-up), tal que permita mayor certeza en la identificación de una persona, disminuyendo así la probabilidad de obtener falsos positivos (identificación de un inocente como culpable). Este método también ha sido utilizado en el registro de llamados telefónicos para los que se requiere de grabaciones previas del sujeto en cuestión, ya que el análisis se ve fuertemente obstaculizado por la calidad del audio y de los parámetros del habla. En este análisis se distinguen las siguientes variables: naturaleza del discurso, calidad de la voz del hablante, cantidad de expresiones escuchadas por el testigo, diferencia entre el tiempo de exposición de la voz del locutor y el sospechoso, el efecto del teléfono en la calidad del habla, variables intrínsecas del testigo (edad, género, etnia, otras) y el contexto comunicativo.

### 2.2.2. Identificación de Hablantes por Expertos

Sus inicios se remontan a comienzos del siglo XX junto con la aparición de la grabadora y el espectrógrafo de sonido, lo que permitió capturar, reproducir, representar y analizar la voz. Se ha visto que este método puede otorgar resultados altamente variables tanto en el tiempo, como en el hablante (aun así es usado en algunos estados de EE.UU). En la actualidad los expertos pueden agruparse en tres categorías; en la primera se encuentran los fonetistas que utilizan la combinación del análisis fonético-auditivo y mediciones acústicas, forma criticada por la subjetividad inherente y la falta de validez. En el segundo grupo se encuentran aquellos que usan parámetros semi-automáticos como formantes, tasas de articulación, entre otros. El tercero se define como automático global, automático, ya que reduce la subjetividad al mínimo, y global, en el sentido que no se refiere a los parámetros específicos del lenguaje acústico, sino que trata la señal como un fenómeno físico, vibración que varía continuamente. En este caso, se escoge una variable a analizar (por ejemplo, una frecuencia específica de un determinado sonido) y se observan las características que presenta entre los distintos locutores. De esta forma propone determinar un grado de similitud o diferencia, entre el que habla y el sospechoso, en una población de estudio (Broeders, 2001).



### 2.2.3. Audioanálisis

Este procedimiento enfatiza el realce de la inteligibilidad. Para esto utiliza tecnologías de hardware y software que permiten filtrar las palabras a transcribir. Su relevancia en el ámbito forense aún no es de gran impacto y se relaciona más que nada con aspectos estéticos. Este sistema ayuda a eliminar el ruido y a reducir la fatiga del audio obtenido. Por tanto, en la medida que se mejora la calidad de la grabación, se logra una mayor inteligibilidad y se obtiene una calidad más confiable de las transcripciones. Se sabe que mientras más deteriorada está la calidad del registro se requerirán más habilidades del intérprete, asociadas a competencias lingüísticas como: familiaridad con el acento y dialecto del sospechoso (Broeders, 2001).

El mismo autor refiere que para asegurar la integridad y autenticidad de las grabaciones de audio, se debe contemplar: el estudio auditivo de la grabación, el análisis de banda estrecha del espectro de la señal registrada y el análisis de alta resolución de la forma de onda de la señal capturada. Por otra parte, estos factores pueden ayudar en la descripción de cómo fueron tomadas las grabaciones, y así establecer la rigurosidad de su obtención.

Las nuevas técnicas de visualización pueden revelar características que permiten mejorar el poder discriminatorio de estos sistemas, ya que asocia una grabación determinada a una grabadora de audio en particular. La importancia radica en saber la autenticidad de la grabación, es decir, si es original o copia, ya que toda replica no puede ser considerada como prueba que atestigüe contra el sospechoso, debido a que pudo haber sido manipulada y no corrobora la integridad de la cinta (Broeders, 2001).

### 2.2.4. Estudios Lingüísticos de Autoría

En ellos se enmarca la atribución de la autoría, la determinación de la modalidad de producción de texto, la traducción o interpretación jurídica, análisis de transcripción de discursos verbales, así como tipo de lenguaje y discurso del foro. Existe un amplio debate respecto al uso legítimo de esta técnica como método de identificación de hablantes o determinación de la autoría de un texto. Michaelson y Morton llaman Técnica Cusum a la hipótesis de que cada locutor tiene un grupo distintivo de hábitos lingüísticos (longitud de la oración, palabras que comiencen con vocal, cantidad de palabras entre 2 y 3 letras por oración) los que generan

patrones textuales posibles de identificar a través de procedimientos estadísticos (Broeders, 2001).

El mismo autor afirma que Pese a los intentos por plantear variables de identificación universales, los estudios siguen revelando características que pertenecen a un gran grupo de personas que comparten una lengua, por lo que dificulta la distinción entre hablantes. Por otro lado, el análisis cualitativo complica la estandarización estadística de los resultados y sumado a ello se agrega que los textos forenses siguen siendo considerados como breves.

Por autoría de textos se entiende que a partir de un grupo de sujetos sospechosos se busca evidencia lingüística para determinar el posible autor de textos dubitados (aquellos que tienen su autoría en cuestión). El análisis contempla el estudio lingüístico de textos indubitados (textos cuya autoría se conoce) escritos por un único sospechoso, con el fin de atribuirle o no la autoría de un conjunto de textos dubitados (Broeders, 2001).

En la actualidad se sabe que el promedio de longitud de oración caracteriza un estilo de habla asociado a factores como género o tipo de registro, sea este escrito o verbal. En 1993, McMenamin sugiere el hallazgo de marcas idiosincrásicas del estilo comunicativo del autor de un texto dubitado, estas son: diseño de la página, puntuación, ortografía, léxico y gramática. Las investigaciones que siguieron a esta idea se basaron en que el estilo de un autor es la proyección de su carácter y pensamiento, por tanto la observación debe contemplar el análisis de los distintos niveles del lenguaje (fonológico, morfosintáctico, léxico-semántico y pragmático). De esta forma, para construir un perfil estilístico de un sujeto es necesario generar un corpus de referencia, definir la función comunicativa del texto y escoger las variables a analizar (palabras de alta frecuencia, uso de nombres propios, palabras clave, etc.) (Guillén, *n.d.*)

### **3. MÉTODOS PARA LA IDENTIFICACIÓN VOCAL DE HABLANTES**

En general, un método de identificación forense de hablantes se encuentra conformado por distintos sistemas de análisis (Maroto, 2009). TFSR (Technical Forensic Speaker Recognition) es el término utilizado en el reconocimiento de hablantes cuando la decisión está basada en teorías procedentes de disciplinas como la lingüística, fonética, acústica y estadística. En contraste con el método anterior se halla la “Navie Forensic Recognition”

(reconocimiento forense simplista) en referencia a las habilidades corrientes de las personas para reconocer voces (Rose en Maroto, 2009)

Romito postula que los métodos para identificar sujetos a través de su voz pueden ser divididos en objetivos y subjetivos. El primer grupo se encuentra compuesto por los métodos automático y paramétrico. El segundo grupo por su parte, lo integran los métodos sociolingüístico, espectrográfico, fonético-acústico y perceptivo-auditivo. A continuación serán explicadas las principales características de cada uno de ellos.

### **3.1. Métodos Objetivos**

#### **3.1.1. Automático**

Este método realiza un contraste entre la voz del sospechoso y la de un sujeto de identidad conocida, para ello aplica modelos estadísticos y numéricos en los parámetros acústicos de la voz. Su objetivo es cuantificar el grado de similitud entre las características en estudio para determinar la probabilidad relativa de observar esos marcadores de la grabación con las voces de una población potencial. Es decir, compara voces para establecer que el audio analizado es de una sola fuente, en contraste con otras (Alexander, 2005). Este enfoque fue desarrollado por la ingeniería y al igual que la aproximación fonético-acústica, se basa en mediciones cuantitativas de las propiedades acústicas del habla. La principal diferencia con respecto al enfoque anterior radica en que los sistemas automáticos analizan la totalidad de la muestra de habla a partir de la extracción de lapsos de corta duración (20-30 ms), sin considerar la información relativa a las unidades fonéticas presentes en ella (Stewart, 2009). Los sistemas de reconocimiento automático de voz han sido utilizados por ejemplo, para permitir o denegar el acceso de un sujeto a un determinado recinto. A pesar que estos sistemas poseen la ventaja de procesar de forma rápida y a bajo costo grandes cantidades de datos, no resultan de gran utilidad para la fonética forense. Lo anterior se explica principalmente por el hecho que el reconocimiento automático permite una verificación del hablante, y no su identificación como es requerido en esta disciplina, a lo cual se suman las variaciones de la señal acústica producto de las diversas condiciones de grabación, aspecto que añade una dificultad adicional para su reconocimiento (Ashby en Brown, 2006).

En este mismo sentido Stewart (2009) afirma que en la comparación forense de voces, se debe considerar el efecto de la conversión de la señal acústica en eléctrica, y frecuentemente su transmisión vía teléfono, lo cual empobrece la calidad de la señal de voz, así como también los diferentes estilos de habla que puede adoptar un mismo sujeto en diversas situaciones de habla.

Se ha demostrado que el análisis automático del habla genera una alta precisión de resultados bajo condiciones controladas y además reduce el tiempo de investigación en comparación con otros tipos de procesamientos. Durante los últimos años ha aumentado el interés en el uso de esta técnica para el reconocimiento forense del habla. Este hecho ha implicado tener que generalizar los métodos a la comprensión e interpretación de los tribunales judiciales (Alexander, 2005).

### **3.1.2. Paramétrico**

Actualmente, el reconocimiento de hablantes se está afrontando de al menos dos modos significativamente distintos. El primero de ellos corresponde al fonético y lingüístico, centrados en una perspectiva humana, y el segundo, al físico, informático y de ingeniería, focalizados en el procesamiento de la señal de voz. Los parámetros escogidos para la clasificación de la voz, en base a un método acústico-objetivo, son cinco: frecuencia fundamental y los primeros cuatro formantes de las vocales.

Se exponen cinco etapas fundamentales en una comparación fónica utilizando un método acústico objetivo:

- Revisión de la calidad de la grabación mediante la cual se determina la posibilidad de utilizar dicha cinta en un posterior análisis técnico-científico.
- Análisis en el sentido interno en el cual se seleccionan las vocales a analizar (tónicas y átonas) centrándose en la frecuencia fundamental y los cuatro primeros formantes. Para este fin, se consideran los procesos de reducción, centralización y coarticulación.

- Análisis en el sentido externo, correspondiente a la identificación del “Steady state”<sup>5</sup>
- Análisis de la muestra fónica en que se compara el registro anónimo con una grabación realizada en el laboratorio que intenta igualar las condiciones físicas y psicológicas.
- Análisis estadístico que busca la similitud entre los registros de audio mencionados en el punto anterior, a partir de sus mediciones individuales.

### **3.2. Métodos Subjetivos**

#### **3.2.1. Sociolingüístico**

A través de este método es posible acceder a la identificación de una comunidad o grupo de hablantes de un país o de una gran ciudad pero no a la un individuo específico, por tanto, este procedimiento es útil como apoyo complementario a la globalidad de la información sobre un caso particular. Esta técnica se basa en un análisis lingüístico de la variabilidad y morfología de una lengua que identifica con algún grado de aproximación, el lugar de procedencia de un hablante anónimo (Romito, 2000). Además este método contempla la organización y sistematización de rasgos acústicos y articulatorios asociados al estilo de habla de diferentes grupos, los cuales se hallan determinados por el medio geográfico, origen social, nivel sociocultural, profesión, entre otros (Llisterri, 2008). Bajo esta consigna, según Rosas y Sommerhoff (2009) es que existen estudios que enfatizan su investigación en descubrir una relación entre variables físicas (sexo, edad, constitución física) y características acústicas (Van Dommelen 1995), así como hábitos sociales (como los de un fumador, Braun y Rietveld 1995) y otros aspectos psicosomáticos (Van Rie y Bezooijen 1995).

#### **3.2.2. Espectrográfico**

Este método utiliza una tecnología creada en la década de los 40, que permite apreciar las variaciones de la señal acústica en frecuencia e intensidad a lo largo del tiempo, lo cual se representa en un gráfico denominado espectrograma. Generalmente, el eje x corresponde al

---

<sup>5</sup> Suceso que ocurre durante el habla espontánea en el cual el aparato vocal queda detenido en una impostación precisa (20-25 ms.) durante la emisión de una vocal tónica. Esta porción a su vez, está caracterizada por tener una estabilidad de los valores formánticos (Romito, 2000).

tiempo y el eje y, a la frecuencia, mientras que las variaciones de amplitud se observan como la evolución en el tiempo de una degradación del negro presente en el espectro, considerando una escala monocromática (Stewart, 2009).

Romito (2000) aduce que este método parece ser poco conveniente, principalmente, debido a las condiciones que deben cumplir las muestras. Las palabras que se comparan deben ser idénticas, capturadas en iguales condiciones de registro y emitidas con las mismas características fónicas (tonalidad, prosodia, altura, entonación, ajustes articulatorios, estado emocional, entre otras). Se ha comprobado que al comparar la misma palabra en contextos distintos, el sonograma arroja resultados muy diferentes debido a la presencia de diversos factores. Adicional a lo anterior, es necesario que exista un idéntico estado de ánimo, distancia del micrófono, canal de registro, intensidad, condición mental y de salud, ya que cualquiera de estos podría traer como consecuencia un cambio en la "impronta vocal". La situación antes descrita dificulta que en la práctica, el análisis de las muestras sea idéntico.

### **3.2.3. Perceptivo-Auditivo**

Este método se basa en la capacidad del fonetista forense o experto para distinguir aquellos rasgos inusuales, distintivos o sobresalientes a partir de las muestras de habla que escucha, y así poder determinar el grado de similitud de los audios. Algunas de las características de interés en este enfoque son por ejemplo la calidad de la fonación (voz soplada, estridente, opaca, entre otras) o ciertas diferencias que permiten distinguir dialectos. A estas, se pueden agregar según Alexander (2005), rasgos como: características sociales, ritmo de habla y entonación, articulación y dicción. A pesar de la utilidad del enfoque auditivo en ciertos casos, su desventaja radica en la escasa objetividad que lo sustenta, principalmente porque se basa en la experiencia del científico a cargo, quien además requiere varios años de entrenamiento y práctica en la disciplina (Stewart, 2009).

### **3.2.4. Fonético-Acústico**

Este método implica mediciones cuantitativas acerca de las propiedades acústicas de los sonidos del habla. En específico, se aíslan ciertas unidades fonéticas (vocales o fonemas)

para ser medidas, analizadas y comparadas en función de sus propiedades o rasgos fonético-acústicos como son; la resonancia del tracto vocal (formantes), frecuencia fundamental ( $F_0$ ) y segundo formante ( $F_2$ ). Estos dos últimos componentes son de gran utilidad, puesto que se caracterizan por ser relativamente resistentes frente a los efectos producidos por la transmisión de la onda sonora a través del canal (Stewart, 2009). Este enfoque es descrito por Alexander (2005) con el nombre de método auditivo-instrumental y consiste en la medición de parámetros acústicos como  $F_0$ , formantes, tasa de articulación, energía espectral y entonación. Este análisis puede complementarse a través del uso de espectrogramas que grafiquen la muestra de habla.

Antes de continuar con el siguiente método, se procederá a diferenciar el análisis acústico del auditivo.

Rose (2002) propone que al utilizar parámetros acústicos, se realiza una comparación de muestras de habla en base a propiedades de la voz extraídas por equipos especializados. Por el contrario, el análisis auditivo es llevado a cabo por un observador entrenado y la consecuente transcripción de las características auditivas. Respecto a esto se pueden encontrar tres posiciones radicales:

- El análisis auditivo es suficiente por sí solo.
- El análisis auditivo no es necesario en lo absoluto.
- El análisis auditivo se debe combinar con métodos acústicos.

Actualmente, se reconoce que mezclar ambas metodologías es indispensable. Por lo tanto, la posición más aceptada es la número tres, también conocida como “aproximación fonético-acústica” (French, 1994). Un ejemplo de esto es que previo a un análisis acústico de una muestra, se debe determinar auditivamente la calidad de esta y además si es apta para llevar a cabo el proceso de identificación. Adicionalmente, en este análisis auditivo previo se debe determinar la lengua del hablante y los parámetros que serán estudiados en un análisis acústico posterior.

### 3.3. Métodos Combinados

Los métodos combinados son considerados por la comunidad forense de vanguardia, como la alternativa metodológica de mayor fiabilidad. Este carácter les es adjudicado ya que conjugan métodos como el perceptivo-auditivo, acústico y fonético-lingüístico.

Este modelo analiza el mayor número de factores, componentes y contextos que forman parte de una emisión de habla, y los aborda desde distintas perspectivas y opciones de análisis con la finalidad de obtener una estimación comparativa lo más amplia y fiable posible.

La técnica combinada considera ocho elementos fundamentales. Los primeros cuatro corresponden a las características físico-acústicas de la voz que son: tiempo, presión o intensidad acústica, frecuencia y resonancia. Los restantes comprenden la duración, sonoridad, tonalidad y timbre.

A estas variables se agregan componentes idiolectales, dialectales y sociolectales del habla como:

- a) Características fonoarticulatorias: modo y punto articulatorio, producciones o eventos de transición, rasgos distintivos prosódicos y rasgos distintivos intrínsecos, sonoridad y tonalidad, alteraciones de la dicción y alteraciones disfuncionales.
- b) Modulación del flujo articulatorio: ritmo respiratorio, nivel de intensidad relativa del flujo fonatorio.
- c) Rasgos y secuencias de ataque y extinción.
- d) Características de orden lingüístico: nivel léxico-semántico y morfosintáctico.
- e) Características de la expresión comunicativa: rasgos prosódicos suprasegmentales, ratios elocutivos y componentes psicolingüísticos.
- f) Ubicación diastrática: jergas y factor cultural.
- g) Ubicación diatópica: lenguas, dialectos, acentos, códigos alternancia/influencia, entre otros.
- h) Pautas de relación comunicativa: proxémica y dialectos verticales

Para llevar a cabo el método combinado, en primer lugar, el experto realiza un análisis cualitativo y cuantitativo de la muestra dubitada, estableciendo márgenes de admisibilidad que hacen referencia a la propia cualidad de la estructura acústica del registro. Luego corresponde efectuar la obtención de la muestra indubitada. En esta instancia el imputado debe leer un texto de 15 frases en tres ocasiones. Estos enunciados coinciden con el contenido semántico de otros existentes en las grabaciones dubitadas. A continuación, las mismas frases son elicitadas por uno de los especialistas, tratando de reproducir el contexto expresivo utilizado en las expresiones dubitadas. Posteriormente, se repiten los pasos anteriores mediante grabación por



interceptación telefónica. Finalmente, son registrados unos cinco minutos de conversación espontánea entre el/los experto(s) y el sujeto en torno a un tema cotidiano (deportes, lugar de nacimiento, etc.). Para que ambos registros de voz (dubitada e indubitada) sean comparables, se debe contar con muestras de similar contexto a diferentes niveles: léxico-semántico, sintáctico, melódico-expresivo, ambiental, de transmisión, entre otros.

Obtenidas las dos muestras (dubitada e indubitada), se realizan dos análisis perceptivos. El primero, acerca de los componentes físico-perceptivos básicos, idiolectales-dialectales y sociolectales del habla, y el segundo, respecto a los índices acústicos mediante espectrogramas. Esta etapa permite obtener características e índices acústicos para confirmar hallazgos obtenidos a través de otras alternativas de análisis. Delgado (2001) recomienda que, para aumentar la fidelidad de los resultados, se deben complementar los análisis antes mencionados con representaciones gráficas de la señal. Para culminar, se efectúa un análisis de los parámetros mensurables contenidos en la muestra de voz a través de un sistema de reconocimiento automático como complemento a los estudios antes mencionados.

#### ***4. VARIABLES PARA LA IDENTIFICACIÓN VOCAL DE HABLANTES***

Desde una mirada metodológica las variables se entienden como propiedades, atributos o características que pueden ser descritas, enumeradas o medidas. Estas cambian de un sujeto a otro y dentro de un mismo individuo (López, 2010). En el marco de la disciplina forense, son aquellos rasgos que permiten diferenciar a un individuo del resto de su comunidad lingüística con la cual comparte muchos de los rasgos que la caracterizan.

##### **4.1. Requisitos para la Selección de Variables Acústicas en el Análisis Forense**

Battaner et al. (2007) proponen requisitos para la selección de variables acústicas a analizar. De acuerdo a estos, toda variable debe:

- Estar presente de forma natural en el habla frecuente de un sujeto y ser de fácil medición.

- Tener la mayor variabilidad interlocutor, pero la menor intralocutor (onda glotal y sonidos turbulentos).
- Permanecer constante bajo factores como el tiempo o las condiciones psico-fisiológicas del hablante.
- Ser resistentes a los efectos de transmisión y ruidos ambientales.
- Ser indemne a modificaciones voluntarias de la persona.

#### **4.2. Clasificación de las Variables Relacionadas con la Identificación Vocal de Hablantes**

Doddington en Battaner (2003) establece una clasificación para las variables: alto nivel de información (dialecto, estilo, etc.), y bajo nivel de información, como amplitud espectral, frecuencia del tono de voz, frecuencias formánticas, y otros rasgos acústicos. El primer grupo se asocia a las denominadas dimensiones socio-psicológicas (Kuwabara y Sagisaka, 1995), es decir, rasgos dependientes de factores sociales, económicos, geográficos, educativos, psicológicos, físicos transitorios, sexuales o lingüísticos. El segundo grupo, se corresponde con las dimensiones fisiológicas. Destacan, la frecuencia media del tono ( $F_0$ ) y la forma de la onda glotal (frecuencias formánticas), sonidos turbulentos, entre otros.

De acuerdo a este planteamiento las variables serán agrupadas y explicadas a continuación:

##### **4.2.1. Alta Información**

Aun cuando este nivel está conformado por parámetros de carácter social, económico, geográfico, educativo, psicológico, físico transitorio, sexual y lingüísticos, en la clasificación solo será abordado este último.

##### **Lenguaje**

Es escogido de entre los demás parámetros debido a que es capaz de contextualizar el análisis de las variables más específicas, pertenecientes al nivel de baja información. A su vez,

media el análisis de los demás aspectos de alta información que han sido excluidos de la presente clasificación.

#### 4.2.2. Baja Información

Estas variables a su vez, pueden ser clasificadas de acuerdo al objeto de análisis pericial:

##### A. Voz

- Referidos a la fuente

##### Frecuencia Fundamental

Según Battaner (2007), los argumentos para utilizar  $F_0$  como indicador de identidad de un hablante, se sustentan en que es robusta y fácilmente extraíble desde registros de baja calidad, por lo que no es afectada debido a transmisión telefónica. Además está disponible en la mayor parte de los sonidos del habla. Sin embargo, existen múltiples factores que pueden ocasionar variaciones intralocutor. Estos se han agrupado en técnicos, fisiológicos y psicológicos. Los primeros comprenden aspectos como el tamaño de muestra y la velocidad de cinta. En el caso de los fisiológicos, se contemplan rasgos como raza, edad, presencia de humo, intoxicación, entre otros. Finalmente como ejemplo de factor psicológico está el estado emocional.

Los valores de  $F_0$  muestran variaciones significativas entre los distintos estilos de habla. Así, dentro del habla espontánea se producen diferencias significativas, mientras que en la lectura los resultados son más homogéneos. La duración y número de segmentos del grupo fónico<sup>6</sup> son dos variables que predicen la desviación atípica de  $F_0$  (Battaner et al., 2007). Si el análisis se realiza en vocales tónicas, el  $F_0$  varía dentro de ciertos rangos de acuerdo a la edad, lengua, estado emocional o perfil sociocultural del hablante. Las vocales tónicas superan aproximadamente en 6 Hz a las vocales átonas. Esta diferencia es más notoria entre los fonemas /i/ y /o/, de acuerdo a los hallazgos obtenidos en uno de los estudios del Proyecto VILE (Albalá et al., 2008). Cabe destacar que las variaciones interlocutor e intralocutor están

---

<sup>6</sup> Porción de discurso comprendida entre dos pausas (Quilis, 2010).

determinadas por los valores medios de  $F_0$  y la duración total de los silencios. Por el contrario, la desviación estándar de  $F_0$  no es un parámetro que permita distinguir sujetos entre sí (Albalá, 2008).

#### Distribución a Largo Plazo de $F_0$

Se propone esta variable, debido a que permite asegurar que todos los parámetros lingüísticos relacionados con el estilo de habla, se neutralicen y que se pueda analizar el perfil y las propiedades estadísticas que identifican a un determinado hablante. Así, una distribución de largo plazo necesita ser incrementada en una extensión de tiempo de habla suficiente para generar un estudio objetivo de la señal (60 segundos, aunque puede variar de acuerdo al dialecto) en estudio. Corresponde a la forma como los emisores hablaron en esa ocasión, e implica que los resultados son válidos en algunas ocasiones.

- Referidos a los Resonadores

#### Formantes

Celis (2009) refiere que los formantes son frecuentemente utilizados en identificación de voces, ya sea de manera experimental o como aplicación. Éstos presentan tres dimensiones: peak o amplitud (altura de onda en dB), ancho de banda (cantidad de absorción de energía, se expresa en Hz) y la frecuencia central, la cual corresponde a la dimensión más relevante ya que representa el punto de máxima respuesta. Se estima que las primeras cuatro frecuencias centrales de los formantes son: 500 Hz ( $F_1$ ), 1500 Hz ( $F_2$ ), 2500 Hz ( $F_3$ ) y 3500 Hz ( $F_4$ ).

Se sabe que la magnitud de la frecuencia central es inversamente proporcional al largo del tracto vocal, por lo que mientras mayor sea el largo, menor será el valor del formante. Esto se observa en la diferencia existente entre hombres y mujeres. Las mujeres poseen un tracto vocal que es en promedio un 20% más corto que el de los varones, por tanto la frecuencia de resonancia es alrededor de un 20% mayor en las damas. A pesar de lo anterior, es posible encontrar desigualdades entre personas del mismo género (Celis, 2009).

Respecto a los valores de los formantes (valores referenciales), destaca que no son estos en sí mismos los que brindan mayor información, sino el campo de dispersión de una vocal en específico. De acuerdo a ello, uno de los resultados obtenidos por el Proyecto VILE

evidencia una gran variación en los valores formánticos del fonema /o/. La mayor variabilidad de resultados se halla en las emisiones átonas y la  $F_2$  de /o/ (Albalá et als., 2008). Además, los formantes de alta frecuencia y  $F_0$  permiten diferenciar en todos los segmentos vocálicos a un mayor número de sujetos, mientras que  $F_1$  y  $F_2$  distinguen un número más reducido. (Battaner et al., 2007). A pesar de esto, en la práctica forense se destaca el análisis de formantes vocálicos debido a su gran estabilidad frente a la variación intralocutor y resistencia a la distorsión e interferencia (Battaner, 2007).

Celis (2009) plantea que si bien se ha establecido cierta relación entre el tamaño del tracto vocal y valor formántico se debe considerar que esta estructura es flexible y su largo puede cambiar de acuerdo a la movilidad que le confiera cada locutor durante el habla (por ejemplo a través de la postura de los labios). Así, la plasticidad de este tracto y su escasa variabilidad entre individuos, pueden generar bajos porcentajes de variación interlocutor e intralocutor.

#### Sonidos Turbulentos

Respecto al análisis de sonidos turbulentos, se sugiere el uso del alófono /s/ para buscar diferencias interlocutor, ya que dependen de la forma de las cavidades anteriores a la constricción (infraglóticas) y del modo en el que la lengua y el paladar se disponen en la parte inmediatamente posterior a la constricción (supraglóticas) (Battaner et al., 2007).

#### Formantes de Alta Frecuencia

Se definen como todo formante que sea igual o superior a  $F_3$ . Stevens (1971), señala que el  $F_3$  es un buen indicador de la longitud del tracto vocal, mientras que Ladefoged (1993), postula que  $F_4$  y  $F_5$  se asociarían con indicadores del timbre de voz del locutor. Además, Rose y Clermont (2001) mostraron que  $F_4$  se desempeña mejor que  $F_2$  en la discriminación forense intralocutor e interlocutor en pares de palabras.

Se ha postulado que la proporción entre la variabilidad intralocutor e interlocutor tiende a ser mayor para los altos formantes, y por lo tanto la identificación de un hablante tiene mayor éxito al utilizar este marcador. A lo anterior, se agrega que dichos formantes evidencian el comportamiento de las resonancias fijas de las cavidades más pequeñas del tracto vocal, por lo que no se logran constatar cambios apreciables frente a una emisión de otra, lo que disminuye

la variabilidad intralocutor. Por otra parte, los formantes de baja frecuencia no reflejan con exactitud la anatomía individual del hablante. A pesar de su gran potencial identificador, los formantes de alta frecuencia son altamente vulnerables a condiciones de transmisión telefónica o calidad de registro de grabación y en sí son muy difíciles de percibir (Celis, 2009).

#### Espectro a Largo Plazo

Según Celis (2009), el espectro a largo plazo (LTS o LTAS) explora características en amplios intervalos de tiempo, sobre los 20 segundos. El LTAS otorga una distribución promedio de la energía acústica a largo plazo en la voz de un individuo, por lo que se deja de lado el análisis de parámetros lingüísticos y de características acústicas, para perfilar la energía global del emisor.

Para el análisis forense, el LTAS es de gran utilidad ya que es considerado como un parámetro robusto, de gran resistencia a la tensión del emisor, poco cambiante para las diversas configuraciones del tracto vocal supralaríngeo, y estable en ciertos período de tiempo (entre 2 y 21 días). Pese a ello, aún no se determina con certeza el grado de variabilidad que puede tener dentro de un mismo locutor, ya que puede afectarse por las diferencias en la transmisión de canal, revelar cambios significativos en prolongados períodos de tiempo y no presenta resistencia al encubrimiento de la voz (Celis, 2009). Según Battaner et al. (2003) se considera su análisis de esta variable, ya que es muy resistente a los efectos de la acentuación sobre el habla.

- Mixto

#### El Cepstrum o Coeficiente Cepstral

Permite identificar características para valorar la calidad de la voz, este análisis se realiza mediante parámetros como la amplitud y anchura de la onda cepstral. La presencia de un peak con amplitud considerable, señala la existencia de mayor cantidad de energía en dicho armónico, lo que se asocia a voces de buena calidad. En cuanto a la altura, si esta fuera reducida cuando se produce el peak cepstral, nos indicaría una alta estabilidad de la frecuencia de pitch (Alonso et al, 2006). Este indicador ha demostrado ser muy potente en identificación forense de hablantes, sobre todo cuando se realiza con sistemas automáticos, ya que es fácil de extraer y manipular (Celis, 2009).

## B. Habla

- Referidos a Dimensión Temporal

### Parámetros Temporales

Incluyen el tiempo total de habla, la proporción del tiempo de habla, la proporción de los intervalos de silencio, la velocidad de habla y el ratio de la duración consonante/vocal (relación entre el tiempo destinado a la producción de la consonante y el destinado a la vocal en una emisión dada de consonante/vocal) (Battaner et als., 2003).

La presencia de pausas es un parámetro significativo para diferenciar lectura oral de habla espontánea, que puede deberse a la necesidad de respirar, organizar la estructura prosódica o a posibles vacilaciones. Esto porque se ha visto la relación existente entre el uso de pausas y la necesidad de planificar, léxica y sintácticamente el discurso. El análisis temporal es un parámetro significativo para diferenciar los dos estilos de habla, así los grupos de entonación tienen una menor duración en la lectura (1.3 segundos) que en el habla espontánea (1.6 segundos). Ahora bien, si se considera a cada hablante individualmente, el valor total de duración de los silencios es una variable significativa para caracterizar la variación interlocutor, tanto en lectura como en habla espontánea. Además, la duración de los silencios y de las pausas sonoras pueden ser considerados significativos en la variación intralocutor (Albalá et al., 2008).

Otro de los parámetros temporales es el ritmo, el cual es definido por Rose (2002) como la velocidad de articulación. Este se encuentra determinado por aspectos paralingüísticos y extralingüísticos. Su carácter paralingüístico se refleja en el hecho que los diversos ritmos pueden tener interpretaciones culturales particulares. Por ejemplo, un ritmo acelerado puede ser considerado como signo de impaciencia. A su vez, el carácter extralingüístico se advierte en que los diferentes hablantes, habitualmente utilizan diversos ritmos. Ambos aspectos son de utilidad forense, siendo importante destacar que la evaluación del ritmo requiere en primer lugar, que las muestras de habla sean comparables en términos de parámetros paralingüísticos. Distintos estudios han hallado resultados controversiales respecto de su utilidad en fonética forense, principalmente debido a la variabilidad intrasujeto que presenta este parámetro según la estrategia de medición utilizada (habla v/s lectura). En este sentido, se requerirían estudios

que evaluaran a una mayor cantidad de hablantes para determinar el valor funcional del ritmo en esta disciplina.

En la misma publicación, el autor (Rose, 2002) señala que la velocidad de articulación cuantifica la rapidez con la cual es producida una oración por un sujeto determinado, y corresponde al número de sílabas que ella contiene, dividido por el tiempo de duración de dicho enunciado (en segundos), excluyendo las pausas. La velocidad del habla por su parte, cuantifica el tiempo total en que un sujeto realiza una determinada producción lingüística y corresponde al número de sílabas de la serie de oraciones elicitadas por el individuo, dividido por el tiempo de duración de dichas oraciones, con inclusión de las pausas. Considerando que la realización de pausas durante el discurso es frecuente, la velocidad del habla usualmente es menor que la velocidad de articulación.

Otro parámetro descrito por Rose (2002) es la fluidez, la cual se relaciona con la continuidad del habla de un sujeto. La variable más significativa en la fluidez corresponde a la cantidad y tipo de pausas realizadas por el hablante. Existen pausas no silentes, referidas a aquellas que se completan con material fonológico (por ejemplo /e:/ en el español de Chile) o con la prolongación de un segmento (por ejemplo prolongar la /a/ en /kása:/). Por otro lado se identifican las pausas “de unión”, las cuales reflejan aspectos sintácticos, semánticos y entonacionales de las oraciones (por ejemplo, “La casa <pausa> es blanca <pausa>”). También existen pausas de indecisión (titubeo), en las cuales el hablante detiene su producción lingüística mientras realiza algún planeamiento cognitivo de su mensaje. Las pausas de respiración por su parte, corresponden a aquellos momentos en que el hablante inspira un cierto volumen de aire para continuar con su producción lingüística. Por último se identifican las pausas de turno, las cuales indican el término del enunciado de un hablante y dan pie a la respuesta por parte de su interlocutor. La duración de estas últimas está determinada por la cultura a la cual pertenecen los individuos en interacción. A pesar que hace falta mayor investigación en cuanto a la identificación de hablantes basada en la fluidez del habla, algunos parámetros como las pausas no silentes y las pausas de indecisión aparentan ser prometedoras en este ámbito.

Un estudio realizado por Künzel (1997) acerca del grado de titubeo en el habla de determinados sujetos (proporción de pausas no silentes respecto del total de pausas), arrojó amplias diferencias interlocutor, y a su vez, variaciones relativamente estrechas intralocutor. Es



en este sentido por tanto, que dichos parámetros pudieran tener alguna utilidad en fonética forense.

Por último dentro de esta clasificación, el grupo fónico también es considerado un tipo de parámetro temporal. Este presenta diferencias significativas en cuanto al estilo de habla, hecho que no ocurre con la velocidad de habla y la articulación (Battaner et al., 2007). Por otra parte, la duración del grupo fónico y el número de segmentos de este, son parámetros que predicen la desviación atípica de  $F_0$ .

### C. Tipo Mixto

- Grupo de entonación

Corresponde a una porción de discurso comprendida entre dos pausas, entre pausa e inflexión del fundamental, entre inflexión del fundamental y pausa, o entre dos inflexiones del fundamental que configura una unidad sintáctica (sintagma, cláusula, oración). En concreto, corresponde a fragmentos de habla que reciben una entonación determinada sin que haya necesariamente pausa (Quilis, 2010). Las variables relacionadas con el grupo de entonación son: valor medio y desviación estándar de  $F_0$ ; número de sílabas fonéticas en cada grupo de entonación, duración total de cada sesión de habla, duración total de cada grupo de entonación, duración de pausas silenciosa y pausas sonoras (Albalá et al., 2008).

Respecto al valor medio y desviación estándar de la  $F_0$  en el grupo de entonación, estos son más altos en lectura (119 Hz) que en habla espontánea (112 Hz). Por su parte, se obtuvo que la DS de  $F_0$  en los respectivos estilos de habla, no es un valor significativo para diferenciar los estilos de habla analizados. El habla espontánea se caracteriza por: una mayor dispersión respecto al número de sílabas por grupo de entonación (el más frecuente es el de 5 sílabas), mayor duración de los grupos de entonación, y presencia de pausas sonoras. Referente a la lectura, esta tiene una menor duración del grupo de entonación y una ausencia casi total de pausas sonoras (Albalá et al., 2008).

A continuación se presenta el siguiente diagrama de clasificación de las variables recién descritas, como propuesta para sistematizar la información referida a la identificación vocal de hablantes. Para este fin, en primera instancia se agrupan las variables de acuerdo al nivel de información que entregan (alto nivel de información o bajo nivel de información). A continuación las variables son subclasificadas según el objeto de análisis a estudiar, en este nivel se distingue: lenguaje, voz, habla y tipo mixto (habla-voz en esta ocasión). Luego se realiza una nueva división de acuerdo al componente y/o estructura relacionada con la variable en particular. En esta oportunidad se considera lo siguiente: cuando el objeto de análisis es la voz, las variables serán referidas a la fuente, a los resonadores, o a ambos (mixto); si la información es en relación al habla, las variables estarán referidas a la dimensión temporal; y en los casos que se combinen ambos objetos de análisis (voz y habla), el nivel recibe el nombre de mixto e integra al grupo de entonación.

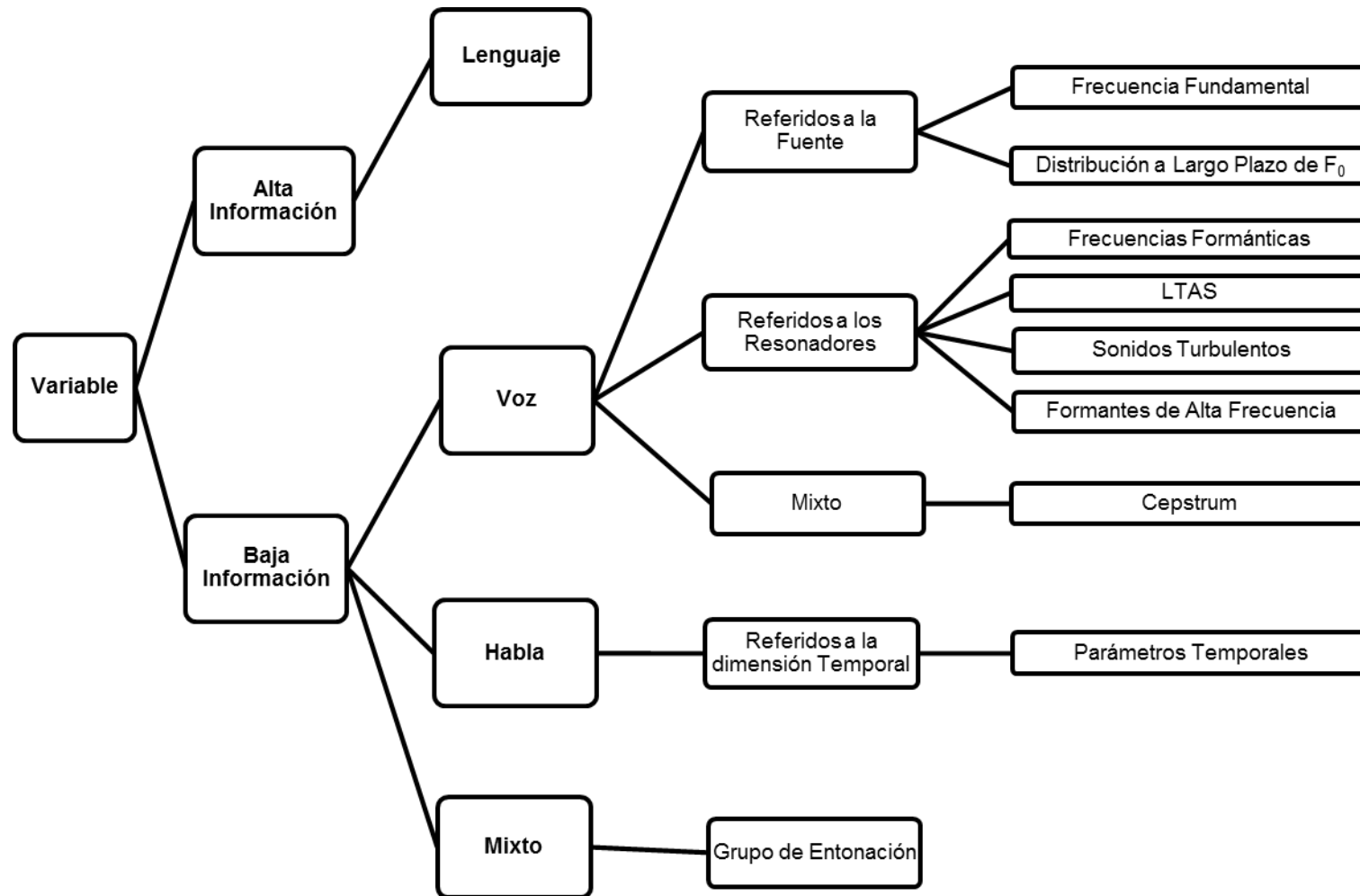


Gráfico 4. Diagrama de Clasificación de Variables para la Identificación Vocal de Hablantes.

## **5. FACTORES QUE INTERVIENEN EN LA IDENTIFICACIÓN VOCAL DE HABLANTES**

La voz de un sujeto se encuentra determinada por factores físicos, fisiológicos, psico-emocionales y extracorpóreos, los que a su vez, se relacionan con las circunstancias en que ella es producida (ruido, transmisión, grabación, etc.). En consecuencia, es difícil presumir que exista la posibilidad de generar dos emisiones iguales (Delgado, 2001).

Delgado (2001) define y organiza los factores involucrados en el proceso de identificación vocal de hablantes en dos agrupaciones. La primera se relaciona con las características internas del emisor, y la segunda, con aquellas no dependientes del locutor. Dicha clasificación será explicada a continuación:

### **5.1. Circunstancias Dependientes de la Naturaleza del Habla y del Sujeto Emisor**

- Contemporaneidad o no contemporaneidad de las muestras: hace referencia a los cambios que se producen en la voz en relación al tiempo, mientras mayor sea el intervalo temporal, mayor será la variabilidad intralocutor. Existen estudios que constatan la existencia de cambios a nivel de  $F_0$  en función de la edad de una persona. Se sabe que el proceso de muda vocal se produce alrededor de los 15 años, resultando en un descenso de la tonalidad de la voz, que se hace más evidente en el género masculino. Pese a ello, la estabilidad se alcanzaría alrededor de los 20 años para luego volver a sufrir variaciones de tipo perceptuales hacia la etapa de la vejez (presbifonía).
- Cambios en los procesos y órganos fonatorios: contemplan modificaciones anatómicas (dentadura, prótesis, tumores, traumas, etc.), fisiológicas (inflamación, secreciones, metabólicas, irritaciones, etc.) y neuropsicológicas (temblor, disfonía psicógena, cambios emocionales, efecto Lombard<sup>7</sup>, entre otras).
- Cambios producidos por agentes químicos exógenos: se asocia al consumo de sustancias que alteran el aparato fonador y los órganos y sistemas relacionados, ya sea a nivel de estructura o función. En este grupo se encuentran: fármacos, drogas, alcohol, cigarrillos, entre otros.

---

<sup>7</sup> Aumentar la intensidad de la voz durante condiciones de ruido en competencia (Boone, 1983).

- Modificaciones de rangos fonatorios y articulatorios estándar, dentro de las cuales se incluyen variaciones relacionadas con: tono, intensidad, grados de tensión/relajación en la articulación, grados de nasalización/oralidad, sonoridad/ensordecimiento, apertura/oclusión, velarización/palatalización, fricativización, bembolización, etc.
  
- Alteraciones elocutivas de elementos fonéticos simples: considera cada variación alofónica de un fonema y los efectos que ello produce a nivel de ataque, extinción, transición y coarticulación.
  
- Alteraciones elocutivas relativas al tiempo y carácter suprasegmental o melódico-expresivo, que agrupa parámetros como: entonación (ascendente, descendente y suspensiva), acentuación, velocidad (ratios silábicos o articulatorios de las emisiones), fluidez, ritmo (normal, bradilálico o taquilálico) y pausas (articulatorias, respiratorias, dubitativas, de selección de unidades léxicas, de configuración sintáctica, de examen o preparación de información comunicativa, de turno conversacional, etc.).
  
- Modificaciones de construcción lingüística y de códigos de relación comunicativa referentes al campo del lenguaje, tales como: variaciones morfosintácticas, léxicas, semánticas, cambios relacionados con el idioma, dialecto, estilo, aspectos paralingüísticos o extralingüísticos y lenguaje no verbal.
  
- Alteraciones transitorias de la cualidad de voz: contempla variaciones involuntarias (voz quebrada, áspera, ronca, cavernosa, fañosa, etc.) y voluntarias (voz cuchicheada, imitada, fingida o disimulada, etc.).
  
- Variaciones de los componentes de construcción emocional o comunicativa del discurso, que conjuga niveles como: excitación, equilibrio, exclamación, tristeza, temor, amenaza, ansiedad, furia, alegría, temor, persuasión, entre otros.

## **5.2. Circunstancias No Dependientes de la Naturaleza del Habla y del Sujeto Emisor**

Este nivel se encuentra compuesto por aquellos elementos que no pertenecen al locutor, en su gran parte provienen del canal telefónico y del ambiente en que se toma la muestra de habla. Dichos factores se refieren a: distorsión de la señal de voz, calidad del dispositivo para la toma de registro de audio, ruido de fondo y todos los factores ambientales asociados al contexto.

- Variabilidad relacionada con los canales, procesos de transmisión y conversión por distorsiones asociadas a presencia de ruido e interferencias como: radiodifusión, inducciones electromagnéticas, saturaciones o distorsiones de transductores microfónicos o de altavoces. Además, la distorsión puede ser derivada de curvas de respuesta telefónica, microfónica, cajas acústicas, alteración en procesos de conversión, codificación y compresión.

- Variaciones relacionadas con emulsiones y soportes magnetofónicos: se deben a diferencias de calidad de los registros magnetofónicos e involucran propiedades físicas (densidad, composición, homogeneidad, etc), magnéticas (magnetización, coercitividad, remanencia), electroacústicas (respuesta en frecuencia, MOL) y mecánicas (resistencia a la rotura o elongación).

- Eventos sonoros simultáneos a la señal de voz analizada: eventos de solapamiento sonoro que deforman la señal a analizar. Mientras más eventos ocurran al mismo tiempo, mayor será el grado de distorsión provocado.

- Variabilidad proporcionada por los diversos diseños y ubicaciones de las fuentes de registro: agrupa parámetros como coeficiente de reverberación y absorción acústica, y ejerce un alto impacto en el timbre e intensidad.

Además de los factores antes mencionados, Delgado (2001) añade otras situaciones que podrían repercutir en el proceso de identificación vocal de un hablante. Ejemplo de ellas son las alteraciones de habla y lenguaje, las cuales pudieran facilitar el trabajo identificativo a aquellos especialistas que poseen algún grado de dominio en dichas disciplinas. A su vez, las patologías de los órganos o mecanismos de producción de voz poseen un carácter intrínsecamente atípico y por lo tanto, identificador o al menos diferenciador. Entre estas encontramos los distintos tipos de disfonía (funcional, orgánica de base funcional y orgánica)

que pueden ser producidas por mal uso o abuso vocal, etiología psicológica, consumo de sustancias (lícitas o ilícitas), etc.

## **6. APROXIMACIÓN A LA SITUACIÓN NACIONAL EN LA IDENTIFICACIÓN VOCAL DE HABLANTES**

### **6.1. Policía de Investigaciones de Chile (PDI)**

En el año 2003 la Policía de Investigaciones de Chile entra en un proceso de especialización de acuerdo a una reforma en las condiciones judiciales vigentes. Bajo este contexto instaura la Sección de Sonido del Laboratorio de Criminalística (LACRIM). Años más tarde se suma a este departamento la especialidad audiovisual, conformándose lo que actualmente se conoce como “Sección Sonido y Audiovisual”, encargada entre otras cosas, del análisis forense de grabaciones de voz. A la fecha, es la única institución de orden público que realiza este tipo de análisis. En específico, existen dos unidades en las que se realizan dichos procedimientos: el Laboratorio Criminalista Regional Temuco (IX Región) y el Laboratorio Criminalista Central (Región Metropolitana). El primero se encarga del análisis forense local, mientras que el segundo, realiza los análisis para todas las regiones restantes. Esta labor es efectuada por nueve especialistas (ocho en la Región Metropolitana y uno en la Araucanía), todos profesionales Ingenieros, ya sea acústicos o en ejecución en sonido (Palma, Silva & Tapia, 2011).

Las funciones que se realizan en la sección de Sonido y Audiovisual son (Palma, Silva & Tapia, 2011).

- Análisis e identificación de voz.
- Autenticación de discos compactos.
- Mejora de la calidad de registros sonoros.
- Digitalización de imágenes de video.
- Fijación y arreglos de imágenes digitales.
- Traspaso de norma y formato.
- Grabación y edición de reconstituciones e inspecciones oculares.
- Autenticación de formato.

De acuerdo a Palma, Silva y Tapia (2011), los pasos a seguir dentro de la labor de identificación son:

1. Solicitud de análisis: se solicita al perito que realice el análisis de una grabación de voz determinada. La solicitud es enviada por el encargado de la investigación y puede provenir desde Tribunales, Ministerio Público, o Unidades o Brigadas Especializadas de la Policía
2. Recepción de la evidencia: las grabaciones de voz pueden llegar al perito en cualquier tipo de formato de audio (WAV, AIFF, MP3, AAC, WMA, MPEG-4, WavPack, Shorten, TTA, ATRAC, entre otros), por lo que la calidad de la grabación es variable según el instrumento utilizado y el formato en que este se grabe.
3. Edición del archivo de voz dubitado: grabación de voz de un sujeto desconocido. En este caso se excluye la voz de otros sujetos que no son de interés, y se eliminan los ruidos sin alterar la voz del sujeto. Además se efectúa una extracción o reducción de la señal en pequeños tramos, lo cual tiene por finalidad reducir la cantidad de información a procesar (inventanado) y expresarla de manera más adecuada para su análisis, proceso denominado parametrización (González, 2008).
4. Citación: previa notificación, el sospechoso acude al laboratorio para entrevistarse con los especialistas. En esta ocasión se graba al sujeto con la finalidad de obtener una muestra de su voz (archivo indubitado).
5. Edición del archivo de voz indubitado: contempla el mismo proceso que la edición del archivo dubitado.
6. Análisis de las muestras: ambas muestras se ingresan al computador para que el Sistema Automático de Reconocimiento Biométrico de Locutor para Aplicaciones Forenses (Batvox) realice el procedimiento. Este programa entrega resultados de forma numérica y gráfica, de acuerdo a las relaciones de verosimilitud (*LR*) halladas. Es importante señalar que la comparación se efectúa en relación uno es a uno, es decir, tanto de la voz que se debe analizar en el registro dubitado, como del sujeto al cual se le imputa esa voz (Palma, Silva & Tapia, 2011).



El sistema Batvox fue implementado en el año 2006, corresponde a una técnica que trabaja con sistemas independientes de texto e idioma, y está diseñado para manejar diversas perspectivas biométricas en distintos ambientes. Dicho sistema está fundado en modelos de mezclas gaussianas, con parametrización en base a coeficientes cepstrales de Mel. Los resultados son establecidos por proporciones de verosimilitud. De forma general, el método se orienta de la siguiente forma (Celis, 2009):

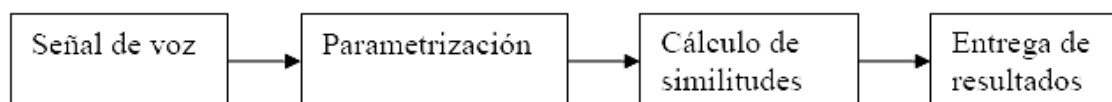


Gráfico 3. Secuencia de Análisis del Sistema Batvox (Celis, 2009).

1. Señal de voz: comprende a la entrada de muestras de voz en formato digital y su edición, de manera que no se escuchen silencios. Como requisito, el sistema estipula que la señal de voz tenga una diferencia de al menos 20 dB con el ruido de fondo.

2. Parametrización: consiste en la reducción de información desde dominios robustos e independientes, a pequeños tramos de la muestra. Este proceso es conocido como “enventanado de la señal”, existiendo para ello varias alternativas como son la ventana rectangular y la ventana Hamming.

3. Cálculo de Similitudes: en términos prácticos, consiste en asignar una identidad al locutor del cual proviene la señal de entrada, cálculo que se sustenta en la determinación de las relaciones de verosimilitud (LR). Para determinar dichas relaciones el sistema estima dos distribuciones de probabilidad: la de intervariabilidad y la de intravariabilidad, cada una con su curva de resultados respectiva. Estos gráficos se comparan con el trozo de audio que no se ha utilizado. Así, la media obtenida representa el máximo en la curva de intravariabilidad. Si la puntuación del test frente al modelo de sospechoso completo cae dentro de la curva de intervariabilidad, implica entonces que lo más probable es que no pertenezca a la misma persona. En caso contrario, si cae dentro de la curva de intravariabilidad, lo más esperable es que corresponda al mismo sujeto.

4. Entrega de Resultados: el sistema Batvox proporciona resultados de forma numérica y gráfica, de acuerdo a las relaciones de verosimilitud (LR) halladas.

## **6.2. Tesis: “Descripción de los Métodos Utilizados en Reconocimiento Forense de Locutores y su Implementación en Chile” Autor: Pablo Celis, Año 2009.**

En el ámbito internacional, la práctica del reconocimiento forense de hablantes ha demostrado ser una poderosa herramienta para aclarar una actividad delictual. Sin embargo, esto no se corresponde con la realidad nacional, ya que su aplicación en tribunales chilenos es todavía extraña y escasa. Prueba de ello es el desconocimiento parcial o total por parte de personas vinculadas estrechamente a la administración de justicia como es el caso de abogados, jueces, fiscales, entre otros. A este nivel, se hacen insuficientes los peritajes en verificación e identificación forense de locutores. Por lo general se ha visto que se utiliza la metodología bayesiana, exceptuando los trabajos efectuados por peritos de la sección sonido del Laboratorio de Criminalística de la Policía de Investigaciones de Chile.

Por otra parte hacen falta datos que describan la realidad fonética-acústica de la población de referencia (español de Chile) y que además, esta sea de uso público. En este campo, tampoco existe consenso en relación a los parámetros estadísticos necesarios para llevar a cabo un peritaje de manera tradicional en base a una metodología bayesiana. Lo anterior se explica porque sólo es factible una investigación utilizando sistemas automáticos.

Una vez realizada la revisión de material en Chile, se evidenció que aún no hay una metodología unificada por parte de quienes conforman el círculo de peritos. Se agrega además, que en ninguno de los casos anteriormente señalados existe una evaluación de las metodologías, con el fin de entregar un grado de certeza o margen de error en sus resultados. A esta carencia se le suma que la difusión de esta disciplina es aún escasa y que el número de expertos que colaboran de forma activa, es mucho menor.

En razón a la situación expuesta anteriormente, se vuelve necesario orientar el trabajo hacia la creación de una base de datos con fines forenses, que contemple la caracterización de hablantes chilenos, considerando aspectos lingüísticos y fonético-acústicos. De acuerdo a los procedimientos utilizados actualmente en la disciplina de reconocimiento forense, se recomienda que la recopilación de datos nacionales responda a la aplicación de la metodología bayesiana, y al desarrollo y evaluación de diversas metodologías empleadas en reconocimiento forense.

### **6.3. Trabajos de Investigación. Autores: Claudia Rosas y Jorge Sommerhoff**

#### **6.3.1. “La Variación Acústica del Hablante: Aplicaciones Forenses” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2008.**

El presente estudio utilizó una base de datos llamada “Grabaciones preliminares del proyecto Fondecyt 2007” que consideró el uso de dos corpus de habla. El corpus uno contempló 20 segmentos vocálicos del fonema /á/ emitidos por una misma mujer en cinco modalidades fonopragmáticas distintas (lectura, evidencia, asombro, atenuación y no conclusión) y reproducidos en cuatro ambientes (anecoico, reverberante, celular y telefónico). Por su lado, el corpus dos consideró la elicitación de segmentos de dos fonemas /á/, dos /é/ y dos /e/, leídos en sala anecoica. Además, se extrajeron segmentos vocálicos de frases con estructura artículo-sustantivo. La investigación consideró como objeto de estudio,  $F_0$  y los tres primeros formantes.

Los resultados mostraron que  $F_0$  se evidencia levemente superior al resto de las variables, logrando los valores más bajos en asombro y no conclusión. Adicionalmente, se constató que  $F_0$  no varía en los diversos ambientes. Respecto a los formantes, se pudo observar que estos variaron en los distintos ambientes de producción de la muestra. Este hallazgo se obtuvo al comparar las cifras en relación a lo obtenido en la sala anecoica, el siguiente orden representa el número de cambio de mayor a menor: reverberante 9 alteraciones en los formantes, telefónica 8 y celular 7. En cuanto al ambiente reverberante, se evidenció que en las modalidades de evidencia y no conclusión, las cifras más altas fueron para los formantes  $F_2$  y  $F_3$ , mayores valores de  $F_1$  y  $F_3$  en lectura y de  $F_2$  en asombro. Por su parte, en el ambiente telefónico, los puntajes de  $F_2$  y  $F_3$  aumentan en evidencia y solo de  $F_2$  en no conclusión. En el caso del contexto celular, hubo mayor  $F_2$  en lectura y evidencia.

#### **6.3.2. “Efectos Acústicos de las Variaciones Fonopragmáticas y Ambientales” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2009.**

La importancia de este trabajo radica en abordaje de factores determinantes al momento de identificar un hablante. El sustento de esta idea se basa en estudios referidos a los efectos de la transmisión telefónica y celular sobre las propiedades acústicas de las vocales

(Künzel 2001; Byrne y Foulkes 2004) y los cambios en la señal que tiene el propio individuo al hablar por teléfono (debido al uso del aparato en sí), como son: pronunciación, volumen y tono (Cf. Wells (1982: 28); Summers *et al.* 1988; Summers *et al.* 1989; French (1998); Hirson, French y Howard (1995)). Bajo este contexto, el objetivo de esta investigación se centra en comparar las señales de voz producidas por una hablante en diferentes situaciones fonopragmáticas (lectura, evidencia, asombro, atenuación y no conclusión) y transmitidas a través de distintos canales de comunicación (anecoico, reverberante, telefónico y celular).

Los resultados indican que en el canal reverberante existe una cierta tendencia a la disminución de los parámetros en las diversas modalidades. En lectura  $F_0$  y  $F_1$  aumentan; en tanto que  $F_2$ ,  $F_3$  y  $F_4$  disminuyen. En el resto de las situaciones, todos los parámetros descienden sistemáticamente, destacando la baja significativa de  $F_3$  en la modalidad de atenuación. Referente al canal telefónico, no existe un patrón evidente que se aplique a todos los parámetros en las distintas modalidades. En la lectura, las frecuencias ascienden de manera estable, siendo una variación significativa en el caso de  $F_1$ . En el resto de las modalidades, los cambios son variables. En el caso del canal celular no se observa una directriz frente al tipo de cambio, pese a ello se distinguen algunas regularidades. En la modalidad de asombro hay un aumento de todas las variables, al igual que en las modalidades de atenuación y no conclusión, exceptuando –en estas dos–  $F_0$  que no registra datos. En cuanto a  $F_1$  y  $F_2$ , estos se elevan en todas las modalidades, y significativamente en la situación de no conclusión. En lectura y evidencia, el comportamiento es más variable.

Por último, se concluye que al comparar los tres canales: reverberante, telefónico y celular, se evidencia que  $F_1$  es la variable más estable en el tipo de cambio producido. Esta muestra un patrón de ascenso en los tres canales de las diferentes modalidades y de forma significativa en la lectura. El resto de parámetros no se comporta de la misma manera, es decir no se constata una variación estable en todos los canales, ni en todas las modalidades. De esta forma, los resultados muestran que las situaciones fonopragmáticas y ambientales causan variaciones no homogéneas en las propiedades acústicas de la señal, siendo  $F_0$  y los formantes afectados bajo distinta manera e intensidad.

### **6.3.3. “Implicancias y Proyecciones Forenses en el Análisis de la Voz” Autores: Claudia Rosas y Jorge Sommerhoff, Año 2010.**

Esta investigación postula que “en el español de Chile existe variación acústica idiolectal producida por factores ambientales y expresivos, posible de reconocer, representar y jerarquizar”. De acuerdo a esta premisa se plantearon objetivos relacionados tanto con la caracterización de la variación acústica idiolectal en el español de Chile, como con la contribución de conocimiento destinado a mejorar las aplicaciones en el ámbito forense y la labor investigativa interdisciplinaria. El estudio utilizó un corpus de 160 vocales extraídas a partir de la base de datos del proyecto de investigación (Ref.: N° 1070210 Fondecyt Regular 2007). En este, las frases fueron producidas y registradas en canal anecoico por cuatro actrices, quienes simularon estar inmersas en distintas situaciones fonopragmáticas (lectura, evidencia, asombro, atenuación y no-conclusión). Terminada esta etapa, se procedió a su reproducción y registro en canal telefónico, celular y reverberante.

Los análisis efectuados permitieron obtener dos tipos de información. El primero se titula central y contempla variables como  $F_0$  y sus formantes. El segundo recibe el nombre de auxiliar y se relaciona con contenido complementario de interés, como la frecuencia fundamental de largo plazo (LTF0).

Respecto de los resultados, se observa que en la información central todas las modalidades presentan variaciones en  $F_0$ . Estas, al ser comparadas con la modalidad de lectura se evidencia un aumento en dicho parámetro (no regular ni significativo). Es por tanto la lectura, la situación en la que se observan las mayores disminuciones, aunque no de forma consistente para todas las sujetos. En cuanto a los datos de los análisis auxiliares, el LTF0 mostró ser el parámetro más rentable para complementar objetivamente los datos del análisis individual de cada locutora; es decir permite determinar aspectos relacionados con lo sistemático-individual, así como rechazar lo casual-individual. En ese sentido, se observó que la curva del LTF0 aumentó sistemáticamente en todos los casos y que en general, ese cambio estaba en relación directamente con el aumento del  $F_0$  temporal relativo.

En conclusión, los resultados relacionados con la información central muestran que las situaciones fonopragmáticas y ambientales (canales de transmisión) causan variaciones en las propiedades acústicas de la señal (Rosas y Sommerhoff 2008b, 2009a, 2009b y 2009c). Estos factores actúan de forma compleja y se asocian a distintos aspectos como: variable

estudiada ( $F_0$ ,  $F_1$ ,  $F_2$ ,  $F_3$  y  $F_4$ ), orientación del cambio (aumento v/s disminución), canales (reverberante, telefónico y celular) y modalidades fonopragmáticas (lectura, evidencia, asombro, atenuación y no conclusión). Ahora bien, los resultados auxiliares con fines complementarios revelaron la importancia de agregar la frecuencia fundamental de largo plazo (LTF0). Este hallazgo apoya la idea de incluir un corpus diferenciado que contenga por un lado parámetros fonéticos y ambientales controlados, y por otro, emisiones espontáneas y de mayor duración. De esta manera observar dichas variables con fines comparativos y que contribuyan en la determinación de lo sistemático-individual v/s lo casual-individual.

## VII. DISCUSIÓN

Es imposible negar la incipiente necesidad de sistematizar la información disponible acerca de la identificación vocal de hablantes. Si bien, esta idea surgió como una hipótesis débil sustentada en la apreciación parcial de un problema, paulatinamente fue tomando mayor fuerza. A medida que más se investigaba acerca de las implicancias de la disciplina forense en el ámbito de la voz, mayor era el número de estudios e investigaciones que manifestaban la dificultad que implicaba no poseer un sustento teórico capaz de conformar el cimiento científico de una metodología válida y consensuada. En concordancia con lo señalado, Morrison (2009) así lo planteaba en una conferencia en China. El autor lo rotuló con el nombre de “cambio de paradigma<sup>8n</sup>” y lo definió como una evolución desde un período en el que un diversas escuelas investigativas buscaban soluciones a diferentes problemas (con algunos solapamientos parciales entre los distintos conjuntos) utilizando marcos conceptuales diversos, hacia una etapa en la que existe un acuerdo en la comunidad científica sobre qué problemas son importantes y cuáles son los procedimientos generales para resolver estas dificultades y la naturaleza de las soluciones adecuadas. La finalidad de promover el nuevo paradigma es enfatizar el aumento de la eficacia y productividad del análisis efectuado, a través de la instauración de una metodología consensuada en la que se considere: la caracterización cuidadosa y precisa del procedimiento científico a fin de que pueda replicarse y validarse de manera indistinta por múltiples peritos; definición y cuantificación fiable de las medidas; aplicación de pruebas normadas, y realización de análisis exactos (Daubert en Morrison, 2009). Es por tanto el problema expuesto en esta revisión, un hallazgo a nivel nacional e internacional. Dentro de este marco, se pudo constatar una amplia cantidad de documentación que en oportunidades se torna dispersa, ya sea por la evidencia de escasa correspondencia de los términos utilizados o la insuficiente información metodológica, entre otras. Como fue descrito por Morrison (2009) y en correspondencia a lo planteado en este trabajo, se considera fundamental ordenar y clasificar la información disponible a fin de delimitar los campos a analizar y dirigir la investigación empírica a objetivos directamente relacionados con la resolución del desafío de la identificación vocal de hablantes.

---

<sup>8</sup> Por una parte, se considera como la constelación de creencias, valores, técnicas, entre otras, compartidas por los miembros de una comunidad dada. Por otra, denota un tipo de elemento de dicha constelación, las soluciones concretas, los métodos, modelos o ejemplos que pueden construir la base de las soluciones Kuhn en Morrison (2009).

Como se ha mencionado hasta ahora, para llevar a cabo un procedimiento de identificación vocal de hablantes, es necesario considerar las múltiples variables que influyen en la producción de la voz. Para poder abarcar y trabajar sobre dichos elementos, se ha hecho necesario involucrar distintas disciplinas que aborden el objeto de estudio desde sus diferentes perspectivas. De este modo, a lo largo del desarrollo de esta tarea forense, se ha evidenciado la participación de profesionales pertenecientes a las disciplinas ingenieril, judicial, fonética, acústica y lingüística. Cada una de estas realiza un aporte específico, concerniente a su propio perfil profesional.

Concordante la situación explicitada, se ha visto desde hace algún tiempo la incipiente introducción de una nueva disciplina en las labores forenses de identificación vocal: la fonoaudiología. Dentro de la realidad nacional, el fonoaudiólogo adquiere competencias concernientes a diferentes áreas del conocimiento, que le permiten describir e identificar características de la voz de acuerdo a sus componentes acústicos (tono, intensidad y timbre) y variaciones determinadas por la influencia de factores culturales, ambientales e individuales (género, edad, estado anímico). Su amplio conocimiento lingüístico, fonético y fonológico, le otorgan la capacidad de describir a un sujeto a partir de los distintos niveles del lenguaje y detectar posibles alteraciones presentes a nivel léxico-semántico, morfológico, sintáctico, pragmático y fonológico. Adicionalmente, tiene la facultad para relacionar los aspectos antes mencionados con distintas patologías de los órganos o mecanismos de producción de voz, que junto con las alteraciones del habla y del lenguaje, poseen un carácter intrínsecamente atípico y por lo tanto, identificador o al menos diferenciador. Entre estas encontramos los tipos de disfonía, como las relacionadas al mal uso o abuso vocal, las de etiología psicológica, las relacionadas al consumo de sustancias, entre otras.

En este sentido, se podría establecer que la fonoaudiología es capaz de enfrentar las variables involucradas en la identificación vocal de hablantes desde una perspectiva holística, representando una suerte de transdisciplina. En otras palabras, el profesional debería estar preparado para emitir juicios integrativos acerca de la diversidad de parámetros que se analizan en la práctica forense. Sin embargo, cabe destacar que para lograr una adecuada participación del fonoaudiólogo en las pericias de identificación vocal de hablantes se vuelve imprescindible el surgimiento de nuevas instancias de especialización concernientes a esta área de trabajo.

Como es sabido, la voz corresponde a un fenómeno de naturaleza variable, esto tanto a nivel interlocutor como intralocutor. Además está compuesta por diversos parámetros e influida



por múltiples factores. Según Delgado (2001), la posibilidad de encontrar dos emisiones iguales es prácticamente imposible, situación que dificulta de manera considerable la identificación de sujetos a través de ella.

En consecuencia a la situación planteada por Delgado y los respectivos obstáculos que se interponen durante investigaciones abocadas a la identificación vocal, se vuelve necesario el despliegue de análisis con una alta tasa de adaptación en relación a la multivariedad de factores que influyen en la voz. En este sentido, se debe organizar el trabajo forense de acuerdo a estas características, que pueden ser dependientes o no de la naturaleza del habla del emisor. Cabe destacar por último, que no bastaría con desarrollar un método de identificación que solo contemple este aspecto. El proceso de peritaje vocal debe contar con la construcción de una base de datos a nivel nacional que sirva de sustento para la elaboración de una caracterización fonético-acústica de la población nacional y desde ahí, avanzar en el territorio de la disciplina forense.

El intento por implantar el cambio de paradigma (mencionado anteriormente), ha acercado el campo forense a áreas relacionadas con la Fonoaudiología. El vínculo se produce a partir de la perspectiva holística que ofrece el análisis y la interpretación plausible de los métodos, variables y factores que intervienen en el estudio del habla y la voz, los que a su vez pertenecen al componente de estudio de dicha disciplina. Según Morrison (2009), para cumplir con los objetivos del nuevo paradigma se requiere de la implementación de ciertas condiciones que aseguren la validez del procedimiento, estas son: disponer de una base de datos representativa respecto a una población relevante (relacionado con: área geográfica determinada, grupo étnico concreto, hablantes de una lengua específica y dialecto particular, etc.), determinar y cuantificar las variables, entre otros aspectos.

Otro de los requisitos necesarios para la ejecución del emergente paradigma de identificación, es la asociada al tipo de metodología a utilizar. La existencia de diversos métodos (sociolingüístico, auditivo, acústico, espectrográfico y automático) ha llevado a algunos investigadores a cuestionarse hasta qué punto los análisis para la identificación vocal han sido practicados con rigurosidad metodológica (Rosas & Sommerhoff, 2010). La respuesta, gestada durante años de práctica forense, ha permitido el uso de los diversos tipos de métodos por parte de los peritos. Sin embargo y hasta el momento, no se ha podido generalizar y consensuar un método exclusivo para la identificación forense de hablantes. Por el contrario, la tendencia actual sugiere la integración de ellos. Delgado (2001) propone el uso del “método combinado”

como procedimiento facultado de generar una mayor confianza y certeza de los resultados. Este carácter es adjudicado, ya que considera elementos de la lingüística, fonología y fonética, indispensables en la construcción de la identidad vocal de hablantes. Por otra parte, el mismo autor indica que sería un error aplicar técnicas automáticas de forma exclusiva. El hecho se explica debido a la presencia de múltiples factores que intervienen en la obtención de un registro de audio ideal, lo que genera que el cálculo de la relación de verosimilitud no indique en sí, el resultado del análisis. En concordancia con esta afirmación, Morrison (2009) advierte que si bien, la relación de verosimilitud ha sido propuesta como un cálculo estadístico que busca objetivar los resultados obtenidos (aceptar o rechazar la hipótesis de que dos muestras tienen igual o distinto origen), no existe una plena aceptación de este elemento como factor resolutivo en la identificación e individualización de hablantes. El debate propone que esta medida debería estar enfocada en detectar la fuerza de la evidencia y no la probabilidad de obtener una de las hipótesis. Las razones lógicas y legales de dicha discusión, se argumentan en la idea de que determinar la probabilidad de culpabilidad carece de sentido si es que antes no se ha estudiado de forma acabada la calidad de la evidencia<sup>9</sup>. Bajo este contexto, es que se propone efectuar un análisis de todas las evidencias relacionadas con la identificación, a fin de conducir a una decisión. La relevancia de contemplar este hecho radica en que si no es considerado, podría condicionar la declaración errada de un juicio (Morrison, 2009).

Es así, como la creación e implementación de un único método representa un desafío en estado de construcción. Es indiscutible que durante el transcurso evolutivo de la disciplina forense, se han propuesto múltiples soluciones frente a esta temática. En la actualidad, la mayoría de los expertos en el área opta por el uso de métodos objetivos exentos de todo tipo de interpretación, dentro de lo posible. Pese a ello, el paradigma todavía no es considerado resuelto.

Las variables corresponden a todos aquellos rasgos que permiten caracterizar las voces de los sujetos, y desde una mirada forense, estas representan el sustento que permitiría identificar a un individuo particular dentro de su comunidad de hablantes. A pesar de las numerosas investigaciones que se han dado en este ámbito, la principal dificultad radica en establecer cuáles son las variables que con mayor exactitud, permiten rescatar la particularidad

---

<sup>9</sup> Este proceso consiste en la autenticación de las grabaciones de audio y el análisis de las palabras contenidas en la muestra. Además considera: el efecto de la conversión que sufre la señal acústica producto de la transmisión telefónica, las potenciales diferencias entre la calidad de los registros de audio de la voz conocida y la cuestionada, así como las variaciones asociadas al estilo de habla

de la voz de un hablante, y que por tanto, son potenciales marcadores de su identidad vocal. En este sentido, nos parece de suma trascendencia el poder contar con una clasificación que permita agrupar las variables a fin de establecer una suerte de modelo que determine el valor funcional de cada variable en el marco de la identificación vocal de hablantes, o dicho en otras palabras, que determine cuáles son las variables que intervienen con mayor fuerza en la voz, habla y lenguaje de los sujetos.

En contraste con la idea antes expuesta, Battaner et al. (2007) argumenta que no sería posible establecer prioridades, ya que cada parámetro puede diferir entre los hablantes. Dicho autor postula que la falta de resultados confiables y la poca concordancia a la hora de definir un parámetro como marcador de la individualidad, puede deberse al tipo de enfoque metodológico utilizado, y que por lo tanto, existiría una interdependencia entre los distintos factores, cuya prioridad relativa estaría dada por cada hablante.

Otra dificultad que se hace evidente a la hora de sistematizar las variables, se relaciona con la pertenencia de ellas a diversas áreas disciplinares. En el curso del presente análisis documental se ha hecho énfasis en la idea de la voz como un fenómeno determinado por múltiples factores, los cuales se enmarcan dentro de dominios como el habla, el lenguaje, la psico-acústica, entre otros. En dicho sentido, la intervención de tan variadas disciplinas complejiza la elaboración de un método válido y consensuado que abarque e integre la totalidad de variables que juegan un rol en la identificación vocal de hablantes.

Por otra parte, la multiplicidad de variables existentes ejerce un impacto sobre el grado de certeza con el cual los expertos informan sus conclusiones periciales a los jueces. En la práctica, la diversa naturaleza de los materiales con los que se trabaja en el laboratorio, las circunstancias en las que se obtienen los datos, el tamaño de la muestra y el tipo de registro, constituyen factores que alejan el análisis de lo ideal. Sumado a ello, se trabaja con muestras breves, bajo condiciones ambientales desfavorables, con canal telefónico o celular, y con ruido de fondo. Dichas circunstancias, en su conjunto, impiden la comparación directa con la voz del sujeto conocido, y añaden mayor dificultad a la labor que deben realizar los peritos forenses (Rosas & Sommerhoff, 2009).

## VIII. CONCLUSIONES

Para Chile se vuelve relevante investigar en torno a la fonética forense a partir de la implementación de la Reforma Procesal Penal el año 2001, la cual contempla el uso de grabaciones de audio como pruebas periciales que permiten inculpar o exculpar a un sospechoso. A este respecto, la presente revisión bibliográfica evidencia una franca desproporción entre los escasos trabajos chilenos disponibles y la bibliografía internacional publicada en múltiples fuentes, apreciándose con claridad que el escenario investigativo nacional no responde las interrogantes ni satisface las necesidades surgidas desde la disciplina forense.

Resolver este vacío se hace más urgente al considerar que estos son saberes fundamentales en la formación de peritos que realicen la identificación vocal de hablantes, capaces de asegurar un desempeño eficaz, eficiente y certero en el procesamiento de las citadas grabaciones de audio, resguardando la ejecución de un análisis más objetivo y menos impreciso.

Desde el punto de vista experimental, se vuelve imprescindible la construcción de una base de datos nacional representativa de la población relevante, que determine los rasgos compartidos y, a su vez, particulares que impactan con mayor énfasis el perfil vocal de un determinado grupo de hablantes. En consideración a lo expuesto, el presente estudio propone una organización teórica al sistematizar la información disponible, orientando a quienes se involucren en el ámbito de la identificación vocal de hablantes.

Las tendencias actuales destacan los beneficios de aunar los distintos enfoques que se emplean en la tarea forense, en desmedro de la aplicación exclusiva y excluyente de métodos aislados. Bajo este nuevo paradigma, se pretende generar resultados más fidedignos capaces de entregar soporte a decisiones más certeras. De esta forma el uso de métodos combinados se levanta como una de las propuestas más sólidas frente a la problemática de la identificación vocal.

En cuanto a la identidad vocal de un hablante, tras la revisión bibliográfica se podría argumentar que la mayor dificultad de su determinación radica en que existe una multiplicidad

de factores, dependientes e independientes del emisor, que obstaculizan la precisión del análisis forense. Esta situación implica un potencial cambio de la expresión de las variables que facultan la identificación, por lo cual sería erróneo considerar este análisis como un proceso lineal y de bajo dinamismo. De esta forma, se considera relevante agrupar dichos parámetros a fin de relacionar los resultados que se obtengan a partir del análisis de estas variables, para posteriormente integrar y contextualizar el análisis forense.

Respecto de la aparición de datos definitorios en el área de la fonética forense, varias investigaciones reportan la existencia de determinadas variables que permiten la identificación de hablantes, tales como  $F_0$  y formantes de alta frecuencia, mientras la gran mayoría solo muestra tendencias significativamente insuficientes para ser replicadas a un conocimiento generalizado (LTAS, Cepstrum). Del mismo modo, algunas publicaciones han establecido cierto grupo de variables que definitivamente no prestan una utilidad significativa al análisis forense, entre ellas la distribución a largo plazo de  $F_0$  y  $F_2$  en segmentos vocálicos.

La incorporación del profesional fonoaudiólogo al escenario actual de la investigación y análisis fonético forense surge como una alternativa legítima frente a las nuevas tendencias en el área, las cuales buscan un quehacer integral y multifactorial del estudio de la identificación de hablantes. A diferencia de la mirada hiperespecializada de otras disciplinas, el profesional fonoaudiólogo entrega una perspectiva capaz de integrar aspectos de la fisiología vocal, parámetros lingüísticos y del habla, además de poseer las competencias para asociar un patrón de perfil sanitario con el sustrato fonético-acústico de la voz que lo emite. Es imperativo entonces que la disciplina fonoaudiológica aborde responsable y comprometidamente este nuevo campo de desempeño profesional.

## IX. REFERENCIAS

- (1) Adrian, J.A., Casado, J.C. (2002). *La evaluación clínica de la voz. Fundamentos médicos y logopédicos*. Málaga: Aljibe.
- (2) Albalá, M., Battaner, E., Carranza, M., Gil, J., Llisterri, J., Machuca, M., et al. (2008). VILE: Nuevos datos acústicos sobre vocales del español. *Language Desing. Journal of Experimental and Theoretical Linguistics. Special Issue 2: New Trends in Experimental Phonetics: Selected Papers From the IV International Conference on Experimental Phonetics (Granada, 11-14 Feb. 2008)*, 1, 1-14. Extraído el día 28 de Marzo de 2011, desde: <http://liceu.uab.es/~joaquim/phonetics/VILE/VILE.html>.
- (3) Alonso, J., Ferrer, M., León, J., Travieso, C. (2006). *Cuantificación de la calidad de la voz para su evaluación clínica por medio del análisis acústico*. IV Jornadas de Tecnología del habla, Zaragoza.
- (4) Ashby, P. (2006). *Phonetics Pedagogy*. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics (2nd Edition, Vol. 9, pp 372-378)*. Elsevier: Oxford.
- (5) Becerra, N. & Villar, M. (2000). Verificación de Identidad de Individuos mediante la Voz. *Revista Chilena de Ciencia, Tecnología y Sociedad: Estudios, documentos y debates*. Recuperado el 25 de Abril de 2011, desde: <http://www.cec.uchile.cl/~cabierta/revista/19/articulos/pdf/edu6.pdf>
- (6) Bernal, J., Bobadilla, J. & Gómez, P. (2000). *Reconocimiento de voz y fonética acústica*. México D.F.: Alfaomega.
- (7) Bonastre, J., Bimbot, F., Boë, L., Campbell, J., Reynolds, D. & Magrin-Chagnolleau, I. (2003, 1-4 septiembre). *Person Authentication by Voice: A Need for Caution*. Ponenciapresentada en EUROSPEECH 2003 - INTERSPEECH 2003. Proceedings of the 8h European Conference on Speech Communication and Technology. Ginebra, Suiza.
- (8) Boone, D. (1983). *La Voz y el Tratamiento de sus Alteraciones*. Buenos Aires: Editorial Médica Panamericana.
- (9) Bustos, S. (2003). *La voz. La técnica y la expresión*. Barcelona: Editorial Paidotribo.
- (10) Byram, M. & Fleming, M. (1998). *Perspectivas interculturales en el aprendizaje de idiomas: enfoques a través del teatro y la etnografía*. Inglaterra: Cambridge UniversityPress.
- (11) Cantero, F., Font, D. (2009) Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental*, XIII, pp. 79-110.

- (12) Celis, P. (2009). Descripción de los métodos utilizados en reconocimiento forense de locutores y su implementación en Chile. Tesis para optar a Ingeniero Civil Acústico. Facultad de Ciencias de la Ingeniería, Universidad Austral de Chile.
- (13) Davies, A. (2003). *The native speaker: myth and reality*. Clevedon: Multilingual Matters.
- (14) Delgado, C. (2001). La identificación de locutores en el ámbito forense. Memoria para optar al grado de Doctor. Facultad de Ciencias de la Información, Universidad Complutense de Madrid.
- (15) Fernández, Ana. (2007). ¿Para qué sirve la fonética?. *Revista de Lingüística, Filología y Traducción*. (1). 39-51
- (16) Furui, S. (1994). An overview of Speaker Recognition Technology. *ESCA Workshop on Automatic Speaker Recognition*, pp. 1-10.
- (17) Frias, X. Introducción a la fonética y fonología del español. *Ianua. Revista Philologica Romanica*. Suplemento 04. p. 3-4, 2001. Recuperado desde: <http://www.romaniaminor.net/ianua/sup/sup04>.
- (18) García Tapia, R., y Cobeta, I. (1996). *Diagnóstico y tratamiento de los trastornos de la voz*. Barcelona: Lebón
- (19) Gil, J. (2007) *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco/Libros (Manuales de formación de profesores de español 2/L).
- (20) González, B., (2008) Batvox: Sistema automático de reconocimiento locutor. *Estudios de fonética experimental*, XVII, 301-316
- (21) González, J., (1981). *Fonación y alteraciones de la laringe*. Cap. 1. Buenos Aires. Ed. Médica panamericana. Citado en Romero, L. & Villanueva, P. (2007). *Eufonía*.
- (22) Guitart, J., (2004). *Sonido y sentido. Teoría y práctica de la pronunciación del español*. Ediciones Universidad de Georgetown. Estados Unidos.
- (23) Guzmán, M. (2011). Producción de la voz. Extraído el día 21 de Abril de 2011, desde [http://www.vozprofesional.cl/index.php?option=com\\_content&task=view&id=13&Itemid=29](http://www.vozprofesional.cl/index.php?option=com_content&task=view&id=13&Itemid=29)
- (24) Hernando, F., (1993). Técnicas de Procesado y Representación de la Señal de Voz para el Reconocimiento del Habla en Ambientes Ruidosos. Extraído el día 28 de Diciembre de 2011, desde [http://www.tdx.cat/bitstream/handle/10803/6911/05\\_hernandoPericas\\_capitol\\_4.pdf?sequence=5](http://www.tdx.cat/bitstream/handle/10803/6911/05_hernandoPericas_capitol_4.pdf?sequence=5).

- (25) Hollien, H. (1990). *The Acoustics of Crime. The New Science of Forensic Phonetics*. New York: Plenum Press.
- (26) Hollien, H. (2002). *Forensic Voice Identification*. Londres: Academic Press.
- (27) Hormazábal, X. (2011). *Rehabilitación. Aspectos Suprasegmentales*. Presentado en Audiología III, Santiago, Chile.
- (28) Hughes, A. y Lascaratou, C. (1982). Competing criteria for error gravity. *ELTJournal*. 36 (3), p. 175-182.
- (29) Iribar, A. (n.d). *Caracterización de los suprasegmentos: La prosodia*. Recuperado 27 abril 2010, desde <http://paginaspersonales.deusto.es/airibar/Fonetica/Apuntes/07.html>
- (30) Jackson-Menaldi, M. (1992). *La voz normal*. Buenos Aires: Ed. Panamericana.
- (31) Llisterri, J. (n.d). Coarticulación. Recuperada el 25 de abril de 2011, desde: [http://liceu.uab.es/~joaquim/phonetics/fon\\_produccio/Coarticulacin.pdf](http://liceu.uab.es/~joaquim/phonetics/fon_produccio/Coarticulacin.pdf).
- (32) Llisterri, J. (2008). El Enfoque Lingüístico en un Informe Pericial de Identificación de un Locutor por su Voz. Trabajo presentado en el I Congreso Internacional de Acústica Forense, Octubre, Madrid.
- (33) Llisterri, J. (2011). Conceptos previos para la enseñanza de la pronunciación y la corrección fonética. Obtenido el día 19 de Abril de 2011, desde: [http://liceu.uab.es/~joaquim/applied\\_linguistics/L2\\_phonetics/Concept\\_basic\\_corr\\_fon/Conceptos\\_basicos.pdf](http://liceu.uab.es/~joaquim/applied_linguistics/L2_phonetics/Concept_basic_corr_fon/Conceptos_basicos.pdf)
- (34) López, I. (2006). Texto Guía de Metodología de la Investigación. Santiago: Universidad de Chile.
- (35) López, I. (2010). Variables y Escalas de Medición. Presentación en Metodología de la Investigación, Junio, Stgo.
- (36) Maroto, S. (2009). Estimación del Peso de la Evidencia en Acústica Forense para la Inferencia Estadística de la Identidad del Hablante mediante la Aplicación de Redes Bayesianas a Rasgos Acústicos. Trabajo fin de Máster. Departamento de Ingeniería Informática, Universidad Autónoma de Madrid.
- (37) Marrero, V., Gil, J., Battaner, E. (2003, 3-9 Agosto) *Inter-Speaker Variation in Spanish. An Experimental and Acoustic Preliminary Approach*, in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, España.



- (38) Martínez, M. (2010). *Manual de Ortografía*. Madrid: Akal.
- (39) Minematsu, N., Nishimura, T., Murakami, T., Keikichi H. (2006). Speech Recognition Only with Supra-segmental Features – Hearing Speech as Music -. Obtenido el día 17 de Abril de 2011, desde: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136> .
- (40) Miyara, F., (1999). *La voz humana*. Extraído desde <http://www.fceia.unr.edu.ar/acustica/biblio/fonatori.pdf> el día 20 de abril de 2011.
- (41) Narbona, J., Chevie-Muller, C. (2003). El lenguaje del niño. Desarrollo normal, evaluación y trastornos. Barcelona: Editorial Masson.
- (42) Obediente, E., (2007). *Fonética y fonología*. Ediciones del Consejo de Publicaciones de la Universidad de Los Andes. Venezuela
- (43) Pérez, H. (2003) . Frecuencia de fonemas. “Fonética segmental del español público de Chile: un estudio de variación estilística de la pronunciación en los noticieros de la televisión chilena”, tesis doctoral . Universidad de Concepción
- (44) Pikabea, I. (2008). *Glosario del Lenguaje*. La Coruña: Ed. Gesbiblio.
- (45) Policía de Investigaciones de Chile (PDI)., (2011, Diciembre). [Entrevista con: Alex Palma, Marlene Silva y Rodrigo Tapia, Ingenieros de la PDI: Labor de la PDI en la Identificación Vocal de Hablantes].
- (46) Queiroz, I. (2002). Fundamentos de Fonoaudiología. Aspectos Clínicos de la Motricidad Oral. Madrid: Editorial Médica Panamericana.
- (47) Quilis, A., Fernandez, J. (1990), Curso de fonética y fonología españolas, Madrid, Consejo Superior de Investigaciones Científica.
- (48) Quilis, A. (2010). Principios de Fonología y Fonética Españolas. Madrid: Arco Libros.
- (49) Real Academia Española (2001). Diccionario de la Lengua Española (22ª ed.). Consultado en <http://www.rae.es/rae.html>
- (50) Recuero, M. (1996). Acústica. Descripción, medición y evaluación del ruido ambiental. Madrid: Paraninfo.
- (51) Richards, J., etals. (1985). Longman dictionary of applied linguistics. Inglaterra: Longman.
- (52) Rodríguez, S. & Smith-Agreda, J. (2003). Anatomía de los órganos del Lenguaje, Visión y Audición. Madrid: Editorial Médica Panamericana.

- (53) Román, D. Estudios de fonética experimental 7(17), 2008. Recuperado de [http://dialnet.unirioja.es/servlet/listaarticulostipo\\_busqueda=ANUALIDAD&revista\\_busqueda=8634&clave\\_busqueda=2008](http://dialnet.unirioja.es/servlet/listaarticulostipo_busqueda=ANUALIDAD&revista_busqueda=8634&clave_busqueda=2008)
- (54) Román, D., Quezada, C. y Sabaj, O. (2000). Manual de introducción al estudio fonético y fonológico. Con especial énfasis en el análisis acústico del habla. Extraído el día 30 de Marzo de 2011, desde: [http://www.domingo-roman.net/Documentos/Manual\\_de\\_introduccion.pdf](http://www.domingo-roman.net/Documentos/Manual_de_introduccion.pdf)
- (55) Romero, L. & Villanueva, P. (2007). *Eufonía*. Santiago de Chile: Ediciones Escuela de Fonoaudiología.
- (56) Rosas, C. y Sommerhoff, J. (2008). La variación acústica del hablante: aplicaciones forenses. Trabajo presentado en el VI Congreso Iberoamericano de Acústica, Noviembre, Buenos Aires.
- (57) Rosas, C. & Sommerhoff, J. (2009). Efectos acústicos de las variaciones fonopragmáticas y ambientales. *Redaylc. Red de Revistas Científicas de América Latina, el Caribe, España y Portugal.*, 44. 195-210.
- (58) Rosas, C. & Sommerhoff, J. (2010). Implicancias y proyecciones forenses en el análisis de la voz. *Redaylc. Red de Revistas Científicas de América Latina, el Caribe, España y Portugal.*, (46). 101-118.
- (59) Rose, P. (2002). *Forensic Speaker Identification*. 1a edición. Londres: Taylor and Francis.
- (60) Ruiz, M. y J. Soto-Barba. 2005. Timbre vocálico en hablantes de español como segunda lengua, *Onomázein*, XI(1), 57-65.
- (61) Sabido, M. (2002). *El tono: andanzas teóricas, aventuras prácticas, el entretenimiento con beneficio social*. Mexico: UNAM.
- (62) Stewart, G. (2009). Forensic voice comparison and the paradigm shift. *Science and Justice*, 49, 298-308.
- (63) Stewart, G., Zhang, C., & Rose, P. (2011). An empirical estimate of the precision of likelihood ratios from a forensic-voice-comparison system. *Forensic Science International*, 208, 59-65.
- (64) Torres, M. (2005). Manejo fonoaudiológico de los trastornos de ATM. *Revista Estomatología de la Universidad del valle*, 13 (1), 36-41.
- (65) Tulon, C. (2005). *Cantar y hablar*. Barcelona: Editorial Paidotribo

- (66) Villayandre, M. (2009). *Fonética y Fonología, tema 3. Fonética Acústica*. España, obtenido el día 20 de Abril de 2011, desde: <http://www3.unileon.es/dp/dfh/Milka/FyF/36.pdf>
- (67) Zañartu, M. (2003) *Aplicaciones del análisis acústico en los estudios de voz humana*. Universidad Perez rosales, Chile. Recuperado 27 de Abril 2010 de <http://web.ics.purdue.edu/~mzanartu/Documents/Paper%20-%20Analisis%20voz%20humana%202003.pdf>