

CCC
AI
0346

Los estereotipos vocales

Elena Castillo Rivera

INDICE

Introducción	5
Justificación	8
1. MARCO TEÓRICO	
1.1 Los estereotipos vocales	16
1.2 Definición del Objeto de estudio	24
1.3 Metodología General	25
1.4 Objetivos e hipótesis	28
• <u>El estudio de estereotipos vocales para la investigación de los métodos de síntesis y reconocimiento automático del habla.</u>	29
• <u>El estudio de estereotipos vocales para la investigación y producción de mensajes orales en los medios de comunicación audiovisual.</u>	36
• <u>El estudio de estereotipos vocales para las investigaciones sobre las formas de representación social.</u>	38

MARCO METODOLÓGICO

2. LA PERCEPCIÓN SONORA

2.1 El uso de la voz y de estereotipos vocales en los medios audiovisuales de comunicación.....	42
2.2 La percepción	46
2.3 La memoria humana	47
2.4 La memoria auditiva	49

3. LA VOZ Y SU PERCEPCIÓN

3.1 Señales sonoras de la voz que transmiten algún tipo de información en el receptor.	53
3.2 La voz	55
3.3 Información generada por las formas sonoras simples	56
3.4 Información generada por las formas sonoras complejas.....	60
• <u>Información semántica</u>	61
• <u>Información sintomática</u>	62
• <u>Información encuadrativa</u>	63
• <u>Información empática</u>	63
• <u>Información emocional</u>	63
• <u>Actitud comunicativa</u>	66

4. ANÁLISIS DE LOS ELEMENTOS ACÚSTICOS DE LA VOZ Y SU NIVEL EXPRESIVO

4.1 Intensidad	68
4.2 Tono o frecuencia de la voz	71
4.3 Duración	73
4.4 Timbre	74

5. ANÁLISIS DE LAS FORMAS ACÚSTICAS DE LA VOZ EN LA PRODUCCIÓN DEL LENGUAJE ORAL.

5.1 Entonación	79
5.2 Acento	81
5.3 Ritmo o velocidad	82
5.4 Dicción	83
5.5 Recursos paralingüísticos	83
6. PROPUESTA DE UNA TAXONOMÍA DE LA VOZ	
6.1 Formas sonoras simples y complejas	86
6.2 Señales vocales independientes del texto	89
• <u>Direccionalidad</u>	89
• <u>Distancia</u>	89
• <u>Definición acústica de la voz</u>	90
• <u>Duración</u>	90
• <u>Cualidad acústica</u>	90
• <u>Expresión idiográfica</u>	92
6.3 Señales vocales dependientes del texto	93
• <u>Expresión semántica</u>	94
• <u>Expresión encuadrativa</u>	95
• <u>Expresión empática</u>	95
• <u>Expresión emocional</u>	95
• <u>Actitud comunicativa</u>	97
6.4 Propuesta conceptual de estereotipos vocales	97
7. DISEÑO METODOLÓGICO PARA UNA CONTRASTACIÓN EMPÍRICA	
7.1 Diseño de la muestra	103
7.2 Construcción del Corpus	104
7.3 Diseño del instrumento para recopilar la información	107
7.4 Diseño del cuestionario. Primera propuesta	108
7.5 Primera Prueba Piloto para la evaluación del cuestionario y presentación de voces	114
7.6 Segunda Prueba Piloto	117
7.7 Resultados y conclusiones de la segunda prueba piloto	120
7.8 <u>Propuesta final del cuestionario</u>	122
8. PROPUESTA DE UNA HERRAMIENTA DE MODELIZACIÓN Y MEDICIÓN DE LAS VOCES	128
8.1 Presentación general de la Taxonomía de la voz ...	129
9. CONCLUSIONES	130
BIBLIOGRAFÍA	133

Introducción

El presente trabajo es la primera parte de un proyecto más amplio que persigue *localizar y analizar los estereotipos vocales* construidos en la memoria colectiva de un grupo de escuchas, con el fin de enriquecer el estudio de la voz y el lenguaje oral durante el proceso de la comunicación humana.

En esta investigación haremos una revisión conceptual de los principales elementos físicos, psicológicos y sociales que intervienen en la producción de la voz para la elaboración de una propuesta metodológica capaz de indagar en los mecanismos perceptivos que originan la configuración de estereotipos vocales.

El modelo metodológico que desarrollaremos en esta primera etapa servirá como sustento teórico para la segunda parte de la investigación en la cual se contrastarán nuestras hipótesis a través de un estudio empírico. Los resultados que se obtengan permitirán realizar una observación objetiva de la voz a través de diferentes instrumentos técnicos de medición para, finalmente, poder hacer un análisis sobre la relación entre las formas acústicas, los mecanismos internos de su percepción y las interpretaciones estereotipadas que de las voces realiza un determinado grupo social.

Es decir, partiremos de una revisión teórica-conceptual que sustente debidamente a nuestro objeto de estudio; posteriormente, en la segunda etapa, utilizaremos pruebas de recepción con sujetos externos a la investigación que verifiquen las variables y discriminen o clasifiquen un determinado número de voces para concluir con el análisis acústico de las clases que nos permitan establecer relaciones entre la voz y los efectos producidos durante su percepción.

Esta línea de trabajo es desarrollada por el *Laboratorio de Análisis Instrumental de la Comunicación (LAICOM)* del Dep. de Comunicación Audiovisual y de Publicidad de la Universidad Autónoma de Barcelona, cuya experiencia de 12 años en la investigación experimental y aplicada orientan nuestro estudio.

En esta tesina, nos centraremos en describir las principales ideas, teorías y bases conceptuales válidas para acercarnos a los mecanismos mentales que originan la elaboración de estereotipos y de estereotipos vocales.

Revisaremos los procesos psicológicos de percepción y memoria auditiva, relacionándolos con el uso de la voz y de estereotipos vocales en los medios de comunicación audiovisual. Nos detendremos específicamente en el estudio de la voz y su percepción. Analizaremos las señales sonoras de la voz que

transmiten algún tipo de información y que el receptor es capaz de interpretar.

Posteriormente, desglosaremos los elementos acústicos de la voz: intensidad, tono, timbre y duración, resaltando su nivel expresivo, e intentaremos relacionarlos con la producción del lenguaje oral: entonación, acento, ritmo, dicción y recursos paralingüísticos.

A partir de esta revisión proponemos una Taxonomía de la voz la cual deberá ser sometida a una contrastación empírica y de la cual obtendremos la clasificación de voces estereotipadas para su posterior análisis acústico.

Esta primera etapa culmina hasta el diseño metodológico de la investigación aplicada. Explicaremos el diseño de la muestra, la construcción del corpus, el diseño del cuestionario que utilizaremos como instrumento para recopilar la información y presentaremos los resultados de unas pruebas piloto con un grupo de discusión conformado por 6 personas con el fin de validar nuestro modelo antes de someterlo a un grupo mayor de escuchas.

Creemos que este trabajo muestra el camino que deberá seguirse en la segunda etapa con el fin de localizar estereotipos vocales que nos ayuden a interpretar sus mecanismos de elaboración mental.

Justificación

“Oía de vez en cuando el sonido de las palabras,
y notaba la diferencia.
Porque las palabras que había oído hasta entonces,
hasta entonces lo supe, no tenían ningún sonido,
no sonaban; se sentían; pero sin sonido,
como las que se oyen durante los sueños
Juan Rulfo, Pedro Páramo

En el campo de los medios de comunicación audiovisual el término *estereotipo vocal* puede sugerir distintas cosas, unas positivas y otras negativas. En un proceso creativo, por ejemplo, puede hacernos referencia a la adecuada utilización de una voz que facilite y lleve a buen término el proceso de comunicación. Pero también, puede producir un efecto contrario, de que se trata de una simplificación de la voz que coarta muchas de sus potencialidades y variedades expresivas, minimizando la capacidad representativa, evocadora e imaginativa del oyente.

Sin embargo, el estereotipo vocal tiene su origen y funcionamiento en la comunicación oral cotidiana de todo individuo. El estudio de su configuración puede ayudarnos a comprender las representaciones que realiza el ser humano de la realidad y a analizar el proceso de la comunicación en nuestra sociedad. Su naturaleza abarca tres dimensiones, la física, la psíquica y la social, ya que se trata de un proceso cognitivo de clasificación que realiza el individuo al momento de percibir el carácter acústico de una voz, pero condicionada por la asignación de valores y atributos aprendidos por su pertenencia a una colectividad social. Es decir, que es un mecanismo perceptivo individual pero de origen o configuración social. Bajo esta triple perspectiva, la *físico-psico-social*, es que el presente estudio analizará los *estereotipos vocales*, con el fin de comprender mejor los procesos comunicativos que den respuesta a “las repercusiones sociales que son origen, en parte, y fundamento de determinados comportamientos humanos, sociales y nacionales” (CUESTA, 2000: 21).

A través del estudio de los *estereotipos vocales* es posible detectar algunos de los imaginarios colectivos de un grupo social, sus mitos y prejuicios, muchos de los cuales giran en torno a valores tan complejos de abordar como lo *bueno*, lo *malo*, la *belleza* o *fealdad* de las cosas, la *agradabilidad* o *desagradabilidad* de una persona, entre otros, y que los medios masivos reproducen constantemente. Esos valores pueden tener

un acercamiento de estudio si se analizan las representaciones grupales formadas a partir de los estereotipos vocales.

Asimismo, y con la misma importancia, su estudio puede ayudarnos a entrar en ese terreno tan fascinante que es el psiquismo humano, en la medida en que conozcamos cómo y de qué manera se configuran los estereotipos vocales en la mente de los individuos.

Las bases conceptuales en las que nos apoyaremos serán, en primer lugar, las otorgadas por las Ciencias de la Comunicación, principalmente, en los principios desarrollados por un grupo de investigadores conocido con el nombre de escuela de Palo Alto (Bateson 1951) , quienes postulan que "todo comportamiento es una comunicación:

La comunicación es un proceso social permanente que engloba modalidades muy variadas de comportamiento: la palabra, el gesto, la mirada; en esta perspectiva, no existe oposición entre comunicación verbal y comunicación no verbal: la comunicación es un todo integrado. Hay que considerar todo lo que, en el comportamiento, cobra, en un momento u otro, valor de comunicación. Esto significa que no se pueden aislar los mensajes, puesto que sólo tienen sentido en el contexto del conjunto de los modos de comunicación. (FISCHER, 1990: 134).

Es decir, partiendo de que en todo proceso de comunicación hay un emisor, un mensaje, un canal de informaciones, un receptor y un repertorio de signos o elementos referenciales comunes entre emisor y receptor, en la producción e interpretación de la voz humana interviene un amplio sistema comunicativo formado por distintos niveles expresivos.

Para explicar con más detalle estos niveles, nos serán también de utilidad las bases otorgadas por la Física, la Psicología de la Percepción y las de la Psicología Social de la Comunicación:

Cuando nos ocupamos de la Ciencia de la Comunicación, la *percepción* –entendida como el primer proceso mental que se debe producir para que acontezca cualquier otro– viene a ser, sin lugar a ninguna duda, el de mayor relevancia, puesto que se ve comprometido en todos los fenómenos analizados desde esta disciplina y, además de comprometerse como parte constitutiva del fenómeno, lo hace como parte primaria y esencial. La *percepción* es la *unidad psicológica básica del conocimiento sensible*. (CUESTA, 2000: 110).

Por su parte, la Psicología Social de la Comunicación

“adopta el supuesto según el cual la comunicación es un hecho social (...) en el que existen procesos psicológicos (es decir, procesos mentales de percibir, sentir, pensar, recordar, evaluar, etc.) que determinan la forma en que funciona la sociedad y la forma en la que tiene lugar la interacción social. También adopta el supuesto según el cual los *procesos sociales*, a su vez, determinan las características de la psicología humana. Es esta *determinación mutua* de mente y sociedad lo que estudia la Psicología Social” (TURNER, 1995. Citado por CUESTA, 2000: 23).

Como cualquier otro sonido, la voz es perfectamente analizable y la podemos acotar para distinguir sus características y cualidades físicas, pero su uso expresivo y, sobre todo, entendiéndola como el vehículo del lenguaje verbal hacen que el estudio de su percepción e interpretación sea muy complejo.

Además de los aportes de la Psicología de la Percepción y los de la Psicología Social de la Comunicación, para el estudio de la voz nos serán fundamentales otras áreas como la Lingüística, la Fonología y la Fonética.

Si bien en esta investigación no realizaremos análisis de discurso, los principios de la lingüística no pueden pasar desapercibidos cuando hacen referencia al lenguaje oral y, por consiguiente, la utilización de estructuras sonoras de la voz aún cuando sean “formas elegidas arbitrariamente para ser asociadas a un sentido determinado” (RODRÍGUEZ BRAVO, 2002: 161).

La fonología estudia la producción de sonidos y sus reglas de combinación en secuencias adecuadas. Establece una distinción entre un nivel segmental (que abarca la identificación, discriminación y articulación de fonemas, sílabas y cadenas orales) y un nivel suprasegmental (que incluye aspectos como el ritmo, la entonación, la acetnuación, etc.). (LÓPEZ-HIGES, 2003: 19).

“La fonética, (...) estudia los sonidos del lenguaje en su realización concreta, independientemente de su función lingüística.” (CHION, 1998: 91, citando al

Dictionnaire de linguistique). "Estudia experimentalmente los mecanismos de producción y percepción de los sonidos utilizados en el habla a través del análisis acústico, articulatorio y perceptivo". (MIYARA, 2003).

Sin embargo, existen muchas investigaciones sobre la voz que arrojan valiosos datos acerca del potencial de las señales vocales para transmitir información más allá de las que nos transmiten las señales verbales al momento de hablar. Es decir, no todo se centra en la mera transmisión de ideas a través del lenguaje articulado, sino que, la manera en como decimos las cosas determina muchas veces lo que en realidad queremos decir. La modulación en las entonaciones, el volumen de nuestra voz para acentuar o remarcar una palabra o enunciado, las pausas para ceder la palabra, etc. son importantes en la comunicación verbal.

La rama que se encarga de su estudio es la *Paralingüística*. Fernando Poyatos define el *paralenguaje* como "las cualidades verbales de la voz y sus modificadores y las emisiones independientes cuasiléxicas, así como los silencios momentáneos, conscientes o inconscientes, que se utilizan para apoyar o contradecir los signos verbales" (POYATOS, 1994: 28).

La *paralingüística* abre las posibilidades de estudio del lenguaje oral que los lingüistas habían pasado por alto. Abarca "toda la información expresiva que transporta la voz en sus sonidos y que no está codificada dentro de los modelos de la fonética y la fonología clásicas" (RODRÍGUEZ BRAVO, 1988-1989: 16). Sin embargo, no abre el abanico lo suficiente para lo que nosotros intentamos investigar y analizar. Aunque todas esas manipulaciones las realicemos de manera consciente o inconsciente al momento de construir los mensajes, hacen referencia exclusiva a los sonidos que añaden información al texto lingüístico, es decir, que dependen del lenguaje. Sin embargo, existen otras señales vocales que en nuestra comunicación cotidiana transmiten también información. Nos referimos a las señales sonoras producidas por los rasgos físicos propios de cada persona, por ejemplo, la edad y el género del hablante que determinan, muchas veces, que una voz suene más aguda, grave o ronca que otra.

El término Expresión Fonoestésica propuesto por Rodríguez Bravo puede servirnos como punto de arranque al momento de delimitar conceptualmente nuestro objeto de estudio.

Llama Expresión Fonoestésica a:

la expresividad sonora transmitida mediante los rasgos de la voz que comunican acústicamente información sobre el gesto, la actitud, el estado emocional el carácter, el aspecto físico y el contexto de un emisor; o bien sobre la forma, el tamaño, el color, la textura, el tipo de movimiento, etc., de aquello que describe oralmente el emisor (RODRÍGUEZ BRAVO, 2002: 160).

Por último, la Fonética Acústica definida por Martínez Celdrán (1996) como "el estudio de las cualidades de las ondas sonoras y cómo transportan los distintos parámetros que dan soporte al mensaje", nos permitirá realizar el análisis físico objetivo de las formas sonoras con instrumentos de medición que supone:

Encontrar o diseñar una magnitud, una escala de valores numéricos definida respecto a un patrón de referencia estable, de modo que sean posibles desarrollar comparaciones objetivas entre los elementos del objeto de estudio. (RODRÍGUEZ BRAVO, 2004: 38).

Norminanda Montoya agrega que en esta rama, la Fonética Acústica, "se utiliza además la *electroacústica* y aparatos e instrumentos técnicos para objetivar parámetros, útiles a la hora de establecer los estilos de locución, partiendo de conceptos como la sílaba, el acento, la entonación, etc." (MONTROYA, 1998:40).

Asimismo, como nos interesa estudiar los *estereotipos* vocales para entender su aplicación y funcionamiento en los medios audiovisuales de comunicación debemos tener en cuenta que si una voz humana es manipulada a través de instrumentos técnicos que nos permiten grabarla, cortarla, distorsionarla y reproducirla en condiciones temporales y físicas que nada tienen que ver con la persona que la produjo (su fuente original), su análisis debe estar condicionada por esta nueva realidad y olvidarnos de interpretar las características físicas, mentales y emotivas reales del hablante emisor y centrarnos más bien en las características y cualidades físicas, mentales y emotivas que *sugiere* la voz del emisor, así como sus coincidencias interpretativas de manera colectiva.

1. MARCO TEÓRICO

1.1 Los estereotipos vocales

*Allá me oirás mejor (...) Encontrarás más cercana la voz
de mis recuerdos que la de mi muerte, si es que alguna vez
la muerte ha tenido alguna voz".*
Juan Rulfo, *Pedro Páramo*.

Cuando una persona escucha una voz, no sólo la percibe sino que hace una interpretación de ella. Es decir, realiza un proceso inconsciente de clasificación que le permite situarla en un contexto y obtener algún tipo de información. De manera casi inmediata puede saber si se trata de una persona conocida, de un amigo o un familiar. También puede interpretar si se trata de una persona joven, reconocer el estado anímico de quien habla o saber si la voz que escucha es emanada desde un televisor, entre otras muchas cosas.

Pareciera, entonces, casi imposible pensar que con tantos matices se pueda hablar de estereotipos vocales. Pero es precisamente esa enorme cantidad de señales expresivas lo que origina que realicemos un proceso de simplificación.

Para entender lo que son los estereotipos en general, no sólo los vocales, las teorías de la percepción interpersonal dan respuesta a su configuración. Asch destacó que "los sujetos desarrollan percepciones interpersonales rapidísimas a partir de unos pocos rasgos iniciales, categorías que los sujetos manejan con facilidad y que atribuyen a la persona percibida en base a unas pocas claves audiovisuales" (citado por CUESTA, 2000: 149).

Piñuel y Gaitán (1999, 324) resumen que un estereotipo es "una colección de rasgos sobre los que un gran porcentaje de gente concuerda como apropiados para describir a alguna clase de personas". Y citan la definición del investigador John Mackie que para ellos reúne las características esenciales: "el estereotipo alude a aquellas creencias populares sobre los atributos que caracterizan a una categoría social y sobre los que hay un acuerdo sustancial".

El estereotipo, etimológicamente significa carácter sólido, del griego *stereos* (sólido) y *typos* (carácter). El término se utilizó frecuentemente como *estereotipia*: "proceso consistente en atribuir características generalizadas y simplificadas a grupos de gente en forma de etiquetas verbales" (PIÑUEL Y GAITÁN, 1999: 324). Posteriormente, en 1922 fue introducido por Lippmann en Psicología Social para definir los fenómenos de creencias y representaciones. Para Lippmann los estereotipos son aquellas imágenes en nuestra cabeza que se interponen entre la realidad y nuestra percepción, provocando una esquematización para situar a los demás.

Un estereotipo tiene cuatro características fundamentales:

En primer lugar, es el componente cognitivo del *prejuicio*; en segundo lugar, encierra una orientación y una evaluación sobre su objeto; en tercer lugar, constituye un componente conductual; y en cuarto lugar, se da generalmente en un contexto intergrupacional,

de relación mayoría/minoría. (PIÑUEL Y GAITAN, 1999: 324)

Resulta difícil entender la diferencia entre un estereotipo y un prejuicio porque las dos son elaboraciones mentales condicionadas socialmente.

El término prejuicio, se utiliza para entender las actitudes hostiles y de aversión hacia otros grupos sociales:

Si la elaboración de los estereotipos se define como un proceso de racionalización de ciertas características en cuanto a un grupo, en una determinada sociedad, la génesis de los prejuicios está ligada a factores de aprendizaje social (...) Es una actitud del individuo que implica una dimensión evaluativa, con frecuencia negativa, con respecto a personas y grupos, en función de su propia pertenencia social" (FISCHER, 1990:105- 106 y 107).

Por su parte, los estereotipos hacen referencia a la esquematización de los rasgos percibidos sobre los otros individuos, que "simplifican el proceso de comunicación, desarrollando posiciones de discriminación en función de su pertenencia a una categoría" (FISCHER, 1990: 106). En función de la elaboración de dichos estereotipos, "se adoptan determinadas actitudes, positivas o negativas sobre los otros grupos humanos" (PIÑUEL Y GAITÁN, 1999: 324), es decir, se elaboran los prejuicios.

Nos detendremos en las observaciones realizadas por Asch, que vienen a dar respuesta a lo que comúnmente llamamos *la primera impresión* cuando conocemos por primera vez a una persona. "Según Asch, la impresión global que se tiene acerca de alguien no equivale a la suma de impresiones obtenidas para cada rasgo particular, sino que determinados rasgos desempeñan el papel de organizadores centrales" (CUESTA, 2000: 149).

Como parte de ese mismo estudio –que intenta entender cómo nos formamos una idea del otro– se formula el concepto de *teoría implícita* que consiste en un "conjunto de inducciones perceptivas desarrolladas a partir de unos pocos datos y apoyándose en unas correlaciones previas, almacenadas en nuestra memoria, las cuales pueden ser correctas (de hecho en muchas ocasiones lo son) o no". (Op. cit.: 152).

Organizamos nuestra percepción del otro simplificando las informaciones de que disponemos. Conocemos la realidad social por esquematización, es decir, por un proceso de evaluación gracias al cual, con menor esfuerzo, sabemos lo que es el otro y su modo de funcionamiento. Intentamos situar a los otros en categorías sumarias, para formarnos una idea coherente de ellos. (FISCHER, 1990: 96).

Vinculado a todos esos procesos de generalizaciones e inducciones perceptuales se encuentra el llamado *efecto halo*, "que consiste en percibir a los sujetos no ya a partir de unos pocos rasgos considerados centrales, sino de uno o dos rasgos que conduce a inducir el resto de los caracteres del personaje o actor observado" (FISCHER, 1990: 97).

Se identifican cuatro fuentes en la configuración de estereotipos: las fuentes físico-perceptivas con las que interpretamos cualquier forma perceptible, las fuentes psicológicas, las psicosociales y las sociales.

- Fuentes Físico-perceptivas:

Las bases que proporciona el conocimiento de las fuentes físico-perceptivas para la configuración de estereotipos son de gran importancia ya que la manera en como percibimos determina todas las demás. Para explicarlas nos apoyaremos en las palabras de Rodríguez Bravo:

El sistema perceptivo tiende siempre a actuar de manera global y coherente (...) Los sentidos no actúan aisladamente, en consecuencia, el sentido de la audición actúa siempre de forma simultánea con los demás sentidos: la vista, el tacto, el olfato, las sensaciones motoras, etc. (...) Todo proceso perceptivo es múltiple, simultáneo, complementario, de apoyo mutuo y tiende siempre a una coherencia global. La consecuencia de esto es que desarrollamos una etapa fundamental del aprendizaje perceptivo en la que experimentamos que los movimientos de aproximación tienen un sonido determinado y los de alejamiento otro, que el resplandor rojo de la incandescencia se ve, huele y quema... Descubrimos, en definitiva que todo fenómeno físico se siente de muchas maneras a la vez, y que todas estas sensaciones son siempre coherentes con las variaciones del fenómeno que las

produce, y coherentes entre sí. (RODRÍGUEZ BRAVO, 1998: 211)

Es decir, que estas bases fundamentales de la percepción humana nos dan idea sobre cómo y porqué tendemos a organizar la información que vamos adquiriendo a lo largo de nuestra vida. Más adelante mencionaremos específicamente las bases perceptuales concernientes al sentido de la audición.

- Fuentes psicológicas:

En lo que respecta a las fuentes psicológicas, las *actitudes*, los *prejuicios* y la *personalidad* del individuo, condicionan la configuración de un estereotipo

Fischer (1990, 114-115) resume que la función de los estereotipos "corresponde a la reducción del coste psicológico de una representación:

-para los partidarios del psicoanálisis, son sistemas de defensa que tienden, sobre todo, a reducir la angustia;
-para la psicología social son procesos mentales que permiten vivir socialmente a un menor costo psicológico.

Constituyen por tanto un sistema de regulación social, en la medida en la que las esquematizaciones que hacen ayudan a los individuos a formarse una idea de las cosas y a realizar elecciones sin demasiados riesgos".

- Fuentes psicosociales:

El *habla* puede ser considerada una de las fuentes psicosociales más decisivas. Pero también:

Otra de las fuentes psicosociales importantes de estereotipos a nivel de individuos, grupos, colectivos, es la posición social de la *integración -marginación* grupales. A medida que uno está más integrado en un grupo, y si dicho grupo es muy "contestatario", éste será el origen de estereotipos y prejuicios más fuertes". (PIÑUEL Y GAITÁN, 1999: 326).

- Fuentes sociológicas:

Las fuentes sociológicas que configuran los estereotipos son muy variadas: unas son las *religiones*, los *medios de comunicación de masas*, o las diferencias *raciales*, la *educación* y los *nacionalismos*. (Op. cit.: 325, 326 y 327).

Con estas bases podemos definir lo que entenderemos como estereotipos vocales los cuales se conforman a partir de los rasgos sonoros de la voz.

Por **estereotipo vocal** entenderemos que es:

La elaboración cognitiva de clasificación que realizamos al momento de escuchar el conjunto de rasgos sonoros de una voz, asignándoles patrones y atributos.

Por un lado, es a través del sistema físico-perceptivo del sentido auditivo, que se inicia el proceso de interpretación de las señales vocales y categorización de sus rasgos. A través de valores aprendidos, el individuo los agrupa y puede estereotipar la voz escuchada.

Al igual que cualquier otro estereotipo, los estereotipos vocales pueden configurarse a través de las cuatro fuentes antes mencionadas, reflejándose en dos niveles de construcción interpretativa: la perceptiva y la social o histórico-cultural.

- a) Nivel Perceptivo. La configuración de un estereotipo vocal está condicionado por las características perceptivas de nuestro sistema auditivo. De tal manera que puede ser de agrado o rechazo según rebase los patrones de tolerancia perceptiva de nuestro sistema auditivo, por ejemplo, volumen o intensidad demasiado altos o por un tono extremadamente agudo.
- b) Nivel Social. La formación de un estereotipo vocal puede estar condicionado por valores y actitudes aprendidos de nuestro entorno inmediato familiar o aprendidos desde la pertenencia a una sociedad histórica-cultural más amplia. Por ejemplo: los integrantes de una familia pueden compartir un específico grupo de estereotipos vocales que difieran con los de su comunidad. Los miembros de una misma comunidad ideológica o socio-económica pueden compartir muchos estereotipos vocales y discrepar con los de otros grupos. Un grupo parlante de una misma lengua puede compartir muchos estereotipos vocales pero no hacerlo con los de otros grupos parlantes de la misma lengua pero de diferente región o país. Un grupo parlante de una misma lengua puede compartir estereotipos vocales y no compartir ninguno de los que otros grupos que hablen otros idiomas tienen en común.

Al momento de hablar de estereotipos vocales, debemos ser concientes de estos niveles.

Norminanda Montoya (2000: 70) por su parte, utiliza el término de estereotipo vocal de carácter para referirse al "conjunto de atributos de actitud y personalidad que el receptor asigna a un locutor cuando escucha el sonido de la voz". Siempre en términos de agrado o rechazo y que son utilizados comúnmente en la publicidad y en los medios audiovisuales:

La forma de dirigirse (de los medios de comunicación) a la audiencia por parte de un locutor, que centra su poder de convicción en el sonido de los personajes y en lo que estos personajes comunican con su voz. Y lo que comunican son estados emocionales de alegría, euforia, entusiasmo, familiaridad y tipos de personalidad como extroversión, simpatía, firmeza, dureza, madurez, etc.". (MONTROYA; LÁZARO, RODRÍGUEZ, 2000: 70)

Pero, ¿existe algún patrón perceptivo que determine a qué aspectos de la voz reacciona el oyente? ¿cómo y por qué respondemos de determinadas maneras ante estímulos vocales?. ¿Qué señales sonoras de una voz ejercen una influencia en nuestra percepción para que se configure un estereotipo y no otro?.

Estas interrogantes, aún no resueltas, son las que motivan el presente trabajo.

1.2 Definición del objeto de estudio

Nuestro objeto estudio podemos definirlo como:

aquellos rasgos sonoros de la voz humana que al ser escuchados conjuntamente en locutores distintos generan en el receptor la sensación de pertenencia a un mismo estereotipo vocal.

Para poder encontrar esos rasgos sonoros, primero debemos localizar, conceptualmente, todas las señales sonoras de la voz que transmitan algún tipo de información.

De esta manera, nuestra primera tarea será:

La sistematización de los conceptos acústico-descriptivos que permitan clasificar cualquier voz.

Esta sistematización debe dar como resultado, la propuesta de una Taxonomía de la Voz que separe en categorías cada una de sus señales sonoras, para poder observar por separado los diferentes niveles expresivos percibidos en conjunto por nuestro sistema auditivo.

Esta herramienta conceptual servirá para diseñar una metodología capaz de encontrar los *estereotipos vocales* de un determinado grupo social pero que nos oriente a entender el cómo agrupamos las voces de manera inconsciente en nuestra memoria.

1.3 Metodología General

Los estereotipos vocales pueden ser estudiados desde muchas perspectivas metodológicas según lo que se quiera conocer y según la disciplina que los observe, todas válidas e interesantes. En el caso de la comunicación audiovisual, a nosotros nos interesa analizar los mecanismos perceptivos y su configuración mental.

El método que pretendemos utilizar en esta investigación está basado en los principios científicos que pueden clasificarse en tres categorías:

El método *inductivo-deductivo* "acepta la existencia de una realidad externa y postula la capacidad del ser humano de percibirla por medio de sus sentidos y entenderla por medio de su inteligencia" (ÁLVAREZ-GAYOU, 2003: 14).

El método *a priori-deductivo* se basa en que "el conocimiento científico se adquiere por medio de la captura mental de una serie de principios generales, a partir de los cuales se deducen sus factores particulares, que pueden o no ser demostrados objetivamente" (Op. cit.: 14).

El método *hipotético-deductivo* postula "la participación inicial de elementos teóricos o hipótesis que anteceden y determinan a las observaciones" (Op. cit.:14).

Al momento de plantearnos la elaboración de una Taxonomía de la Voz estamos realizando un **método de de-fragmentación** para resaltar cada una de las señales sonoras de cualquier voz y su nivel expresivo y así localizar las

simplificaciones mentales que de ellas realiza el ser humano al momento de percibirlas.

Partimos de que cualquier voz es percibida como un todo unitario y que al ser interpretada se produce un proceso mental que la estereotipa, que la esquematiza y la reduce a algo aún más simple, más general, pero para realizar esa simplificación se debe haber realizado un proceso de selección de entre una amplia gama de rasgos particulares que la conforman. Para localizar los estereotipos primero debemos conocer esos subgrupos y sus correlaciones los cuales producen algún tipo de significación o sentido para poder comprender cómo se realiza ese mecanismo mental.

El pensamiento estructuralista parte de que "la estructura se define como la forma según la cual las partes de un ser concreto se organizan en un todo" (IBAÑEZ, 2001: 106). Sin embargo, aunque decimos que la voz es percibida como un todo unitario constituido por partes, nuestra perspectiva supera algunas de las exclusiones que esta corriente de pensamiento planteaba, como la exclusión del referente, los contenidos, el sujeto y la historia. (MARX, M y HILLIX, W, 1983: 50-149). Es decir, no pretendemos desvincular la voz ni el conjunto de sus partes de lo que el receptor interpreta del sujeto emisor, así como del contenido del mensaje, ni de los referentes sociales y culturales del receptor. Todos los elementos sonoros que intervienen en la interpretación y tipificación de una voz deben aparecer en nuestra Taxonomía, tanto los formados por el lenguaje oral como los que no.

Una vez que tengamos el diseño de la taxonomía propuesta utilizaremos el Método de Análisis Instrumental de la Comunicación (RODRÍGUEZ BRAVO, 2004) el cual se desarrolla en cuatro etapas complementarias:

1. Utilización de Métodos Cualitativos "para localizar y definir las variables". (Op. cit.: 35). En nuestro caso, utilizaremos grupos de discusión para afinar el instrumento de una observación empírica. Posteriormente, someteremos a juicio de un amplio número de escuchas un corpus vocal para la localización de estereotipos. Nos detendremos en etapa pues las siguientes tres se desarrollarán en la Tesis Doctoral.
2. "Análisis físico objetivo de las formas visuales que constituyen los mensajes" (Op. cit.: 37) Una vez que los jueces hayan clasificado nuestro corpus según sus coincidencias estereotípicas, pasaremos a analizar las voces a través de instrumentos de medición.

3. "Estudio de los efectos que ha producido cada uno de esos mensajes concretos sobre un grupo suficientemente amplio y representativo de receptores, utilizando instrumentos de control objetivo de la recepción" (Op. cit.: 38). Cuando las voces estereotípicas hayan sido medidas construiremos un nuevo corpus con otras voces que contengan cualidades sonoras similares a las estudiadas y las someteremos a juicio de otro grupo de escuchas para verificar si existen nuevamente coincidencias.
4. "Análisis de las relaciones entre las etapas 2 y 3 es decir: entre las formas visuales, sonoras, etc. que han sido localizadas y formalizadas numéricamente, y los efectos que éstas han producido durante su recepción" (Op. cit.: 39). Será hasta este momento que podremos sacar conclusiones.

1.4 Objetivos e hipótesis

Los objetivos que perseguimos al estudiar los estereotipos vocales son:

- Conocer de qué manera el ser humano estereotipa las voces.
- Conocer si existen algunas señales sonoras que, de manera sistemática, ejerzan más influencia que otras para la configuración de estereotipos y por qué.

Con esto pretendemos:

1. Ofrecer un aporte conceptual al desarrollo de instrumentos de síntesis y reconocimiento de voces así como para los procedimientos de identificación con fines judiciales.
2. Ofrecer un aporte conceptual al estudio y producción de mensajes orales en los medios de comunicación audiovisual.
3. Ofrecer un aporte conceptual a todas aquellas investigaciones que estudian las formas en que los grupos sociales construyen sus representaciones acerca del otro, a partir de la localización de sus estereotipos vocales.

Definiremos dos hipótesis de trabajo.

Primera hipótesis: Los estereotipos vocales no se construyen de la suma de particularidades

sonoras de cada voz, sino que, algunos cuantos rasgos sonoros fundamentales, son suficientes para que se conformen.

Segunda hipótesis: Existe un número finito de estereotipos con los que un oyente "medio" clasifica en su memoria cualquier voz.

- El estudio de los estereotipos vocales para los métodos de síntesis y reconocimiento automático del habla

Decíamos que estudiar cómo el ser humano clasifica y agrupa en su mente las voces, los *estereotipos vocales*, puede contribuir al desarrollo de instrumentos de síntesis y reconocimiento de voces.

Pues bien, haremos primero una revisión general de los principales conceptos y de los avances que se tienen en esta área con el fin de entender con más exactitud a qué nos referimos.

El desarrollo tecnológico de las telecomunicaciones ha estado orientada en lograr una relación más directa y eficiente entre máquinas y usuarios. Por ejemplo, la telefonía, la Internet, la industria automovilística y algunos servicios bancarios, entre otras, han implementado procesadores automáticos que operan a través de *indicaciones orales* proporcionadas por los usuarios. Con decir o pronunciar algunas palabras la máquina es capaz de procesar la información y ejecutar la orden. Sin embargo, ¿a cuántos nos ha pasado que por más alto, claro y breve que seamos, la máquina nos acaba comunicando con un operador?

Y es que si bien es cierto que día a día se tienen nuevos avances y salen al mercado novedosas máquinas con mejores resultados, aún no se cuenta con un equipo capaz de reconocer cualquier tipo de voz en su forma natural de hablar. Por ejemplo, la empresa *Intel* acaba de crear un nuevo programa capaz de reconocer las órdenes verbales del usuario apoyándose en la grabación de su rostro para la lectura de los labios y así maximizar su efectividad.

Una de las tendencias sobre las que caminan los desarrolladores de hard y soft es la de crear sistemas que interactúen de manera más fluida con los usuarios ampliando las vías de entrada a algo más que el teclado y el mouse e incorporando la vista, el habla y los movimientos de la cabeza. Intel Developer Forum (que se

desarrolla en Alemania) presenta un software que permite a los desarrolladores construir computadoras que ven y "leen los labios" como lo hacen los seres humanos para entender mejor los comandos hablados.

Es sabido que el software que reconoce voz no es infalible ni mucho menos: el ruido ambiente, el volumen y la dicción poco clara, entre otras cosas, dificultan su precisión. Para mejorar este rendimiento la empresa presentó un software de Reconocimiento Audiovisual del Habla (Audio Visual Speech Recognition o AVSR) que hace que las computadoras, mediante una cámara o webcam, puedan detectar el rostro de la persona que habla y rastrear los movimientos de su boca. El sincronizar los datos de video con la identificación del habla permite un reconocimiento más exacto de lo que el usuario quiere expresar, mejorando así una amplia variedad de aplicaciones incluso en ambientes ruidosos. (<http://www.ciudad.com.ar/ar/portales/tecnologia/nota/0,1357,47651,00.asp>. Fecha de consulta: 15 de marzo de 2004.)

Otras investigaciones han encontrado que la idea de eliminar el teclado y el ratón de una computadora para que sólo con las voz del usuario se ejecuten las indicaciones no es del todo posible ya que:

Según estudios realizados en el *Human Computer Interface Lab (HCIL)* en la generación del habla el cerebro utiliza la memoria auditiva, que comparte espacio con la memoria a corto plazo y la memoria de trabajo. En otras palabras *no es fácil trabajar concentrado en una tarea y hablar al mismo tiempo porque el habla utiliza importantes áreas de memoria a corto plazo que necesitamos para concentrarnos en la tarea que estamos haciendo.* Por este motivo, aunque es obvio que la voz encontrará un espacio en la interfaz entre el ser humano y la máquina, no parece que vaya a ser la vía preferente de interacción. (http://www.infovis.net/Revista/2002/num_87.htm. Fecha de consulta: 15 de marzo de 2004)

Sin embargo, son muchas las empresas en todo el mundo que están invirtiendo grandes cantidades de dinero para conseguir una máquina capaz de procesar los sonidos de la voz de un modo muy similar a como los procesa el ser humano.

Desde los años 30 se iniciaron investigaciones sobre la señal de voz pero fue hasta los años 40 que se comenzó a

estudiar el comportamiento del sistema de percepción del ser humano. "Sobre 1940 se desarrollaron los primeros experimentos para estudiar el comportamiento de la membrana basilar." (<http://www.gtc.cps.unizar.es/~eduardo/investigacion/voz/introframe.html>. Fecha de consulta 25 de marzo de 2004).

Pero el verdadero avance se dio con el invento del espectrógrafo de voz en la misma década de 1940. Después de la Segunda Guerra Mundial, en 1947, la *Bell Telephon* comenzó una importante investigación que consistía en convertir las señales acústicas de hablantes en gráficos visuales, con el fin de ayudar a que los sordos pudieran hablar y "escuchar" por teléfono. Sin embargo, el proyecto fracasó debido a la complejidad visual de las gráficas que convertían sencillas palabras en formas difíciles de interpretar por la gente no especializada en acústica. Pero a partir de este proyecto comenzaron los primeros pasos en el conocimiento acústico del habla.

El espectrógrafo de voz permitía obtener un registro gráfico de la evolución de la energía de la señal de voz en diferentes bandas frecuenciales. Esta visión tiempo-frecuencia de la señal de voz conjuntamente con los estudios sobre el sistema auditivo humano abrió el campo a la investigación en una de las áreas más fascinantes de las tecnologías del habla: el reconocimiento automático del habla. La gran revolución en las tecnologías del habla se comenzó a gestar con el advenimiento de las computadoras digitales y la informática. Desde mediados de los años 60 se comenzaron a desarrollar algoritmos de procesamiento de señal, a desarrollar las bases de la inteligencia artificial y comenzaron a aparecer los primeros desarrollos e investigaciones en sistemas de reconocimiento automático del habla utilizando computadoras digitales. Pero es en la década de los 70 cuando se da el impulso definitivo con la financiación de grandes proyectos de reconocimiento del habla por parte de países como EE.UU., Japón, Francia, Alemania, etc. (<http://www.lavanguardia.es:8000/ciencia/portada/p481.html>).

En los años 70 un ambicioso proyecto que duró 5 años realizado por el ARPA (Advanced Research Projects Agency of the USA Department of Defense), encontró que la capacidad computacional necesaria para el funcionamiento de una

máquina era alrededor de 100 MIPS, es decir, 1 millón de instrucciones por segundo, y el costo era insostenible para su comercialización. Actualmente se siguen haciendo investigaciones para obtener mejores resultados.

Debemos distinguir dos conceptos fundamentales. Por un lado, se encuentra el sistema ASR (reconocimiento de voz) que tiene como objetivo que la máquina reconozca la voz del usuario como una orden; y por otro, el sistema TTS (síntesis de voz) para convertir el texto de la respuesta proporcionada por la máquina en una frase hablada.

Como estamos hablando de la creación de máquinas "oyentes" y "parlantes", la ingeniería electrónica ha sido el área responsable de este tipo de investigaciones apoyándose en la teoría matemática que "lo que le importa es saber si la información que circula en estas comunicaciones puede transmitirse o no por determinado canal y cuál es la forma más económica de hacerlo." (RODRÍGUEZ BRAVO, 1988-1989:16).

Un sonido es, al final, una fórmula matemática que se puede "representar en un par de ejes cartesianos. Las curvas esenciales que lo forman son muy simples (senos y cosenos), pero las posibles combinaciones de estas curvas pueden llegar a ser tan complicadas como se quiera". Esta complejidad es la que acompaña al fenómeno del habla.

Si a esto le añadimos a que el lenguaje es ya de por sí un complejísimo sistema de producción e interpretación de comunicación, sea el idioma del país que sea, no resulta difícil comprender que la mayoría de los trabajos encaminados a la síntesis y reconocimiento de la voz sólo hayan conseguido máquinas que entiendan, con una eficiencia similar a la humana, palabras sueltas, con un vocabulario mucho muy reducido y secuencias de frases cortas.

Como hemos mencionado, cada uno de nosotros habla de diferente manera, hasta tal punto que el espectrograma de una voz es considerado un signo de identificación como puede serlo una huella digital. Sin embargo, todos estos experimentos bajo la fórmula matemática dejan de lado un gran número de aspectos importantes o esenciales de la comunicación oral: todos aquellos matices sonoros de la voz derivados de aspectos psicológicos o sociológicos como el estado de ánimo de la persona, del estilo propio de hablar o de los aspectos fisiológicos del locutor, etc.

Por su parte, la Fonética considera otros aspectos importantes:

Los programas de reconocimiento del habla se encuentran ante graves dificultades, una de las más importantes radica en el hecho de que la onda acústica -el sonido- es absolutamente continuo. El hablante no se para a hacer pausas y el problema está en segmentar los sonidos. Tenemos la idea ortográfica de que es muy fácil separar la palabra "mesa" en cuatro sonidos: m/e/s/a, pero fonéticamente es un todo continuo de principio a final. Como consecuencia, en la vocal e hay elementos de la m y de la s. Esto es lo que se denomina 'coarticulación'. Además, para poder hablar deprisa no llegamos a efectuar los movimientos completos, los dejamos a medias que cabalguen unos con otros. Las fronteras entre los diferentes sonidos las pone nuestro cerebro, pero acústicamente no existen". "Los lingüistas hacemos hincapié en el concepto de fonema como unidad lingüística algo que, a diferencia de las unidades físicas como los hertzios o los decibelios, es un concepto mental con el que sólo se puede trabajar si se tienen en cuenta todos los aspectos lingüísticos. Cuando se hace referencia al reconocimiento de voz hay que hacer una diferencia entre dos tipos de programas: los de aproximación estadística y los de aproximación lingüística. El primero es el que, desde el punto de vista práctico, puede tener más éxito y, de hecho, se encuentran ya en el mercado. Para los sistemas de aproximación lingüística no basta con la fonética, también es necesario tener en cuenta aspectos gramaticales como la morfología o la sintaxis.

(MARTÍNEZ CELDRÁN, entrevista publicada en: <http://www.lavanguardia.es:8000/ciencia/portada/p481.html>)

Así como en la identificación del hablante se busca lo que le diferencia, en el reconocimiento del habla interesa lo que de común tienen todos los hablantes de una lengua. Son conocimientos complementarios. Es aquí donde la localización de los estereotipos vocales pueden aportar información que sirva para el reconocimiento y síntesis de la voz.

Lingüistas, comunicólogos y psicólogos no terminan de investigar toda la gama de variables que intervienen en la producción del lenguaje y en la interpretación de las voces. Sin embargo, Rodríguez Bravo lanza una serie de observaciones encaminadas a resolver estos problemas y que sugieren nuevos métodos integrales que merecen atención:

Pero esta variabilidad no es casual, sino que responde a estructuras fisiológicas y psicológicas que están íntimamente relacionadas; la alegría, el malestar, influyen en la tonicidad muscular de todo el cuerpo, sin excluir la boca y la laringe, y por tanto alteran los sonidos, los distorsionan, alejándolos de los modelos que las máquinas son capaces de reconocer; pero estas distorsiones no son aleatorias, sino sistemáticas, susceptibles de ser identificadas e incorporadas por un sistema informático de reconocimiento de formas. (RODRÍGUEZ BRAVO, 1988-1989: 16).

Si tomamos estas consideraciones, aspectos como la emoción y su relación en la producción de la voz, así como los aspectos socio-culturales y psicológicos de la comunicación verbal pueden arrojar información manejable en las pruebas de síntesis y reconocimiento de la voz.

Los alcances que se tienen proyectados si se consigue crear una máquina que pueda reconocer y reproducir fielmente la voz humana, son muchos. Por ejemplo:

La habilidad de transmitir emociones es un área tan importante de desarrollo como lo es tomar muestras de una voz humana para sintetizar voces en esa voz. Así, por ejemplo, sería posible coger una grabación antigua de la voz de Stephen Hawking y usarla como modelo para síntesis de voz, de forma que no sólo hablase con acento británico, sino que además sonase como si realmente fuese Stephen Hawking quien está hablando. En el futuro, aplicaciones como estas podrían ser usadas para escuchar a autores leyendo sus libros. Así, por ejemplo, los fans de Tolkien podrán escuchar a Tolkien - mediante la ayuda de la síntesis de voz - leer "El Señor de los Anillos".

(<http://es.computers.toshiba-europe.com/cgi-bin/ToshibaCSG/>. Fecha de consulta: 26 de marzo de 2004).

Finalmente, estamos convencidos de que si localizamos aquellas señales sonoras fundamentales de la voz que determinan la conformación de un estereotipo y no de otro, y si existe un número finito de estereotipos primarios, esta información puede traducirse en un valioso aporte para el empleo de esta área de investigación.

- El estudio de Estereotipos Vocales para la investigación y producción de mensajes orales en los medios de comunicación audiovisual.

Para los que estudiamos los medios de comunicación masiva no es ningún secreto el enorme vacío de investigaciones sobre la voz humana en la producción audiovisual. La mayoría de la bibliografía existente entiende a la voz como una mera portadora de ideas y sentimientos a través de la palabra, es decir, vinculándola a la producción del lenguaje oral. Encontramos, por ejemplo, que los manuales de producción radiofónica dan una explicación del aparato fonatorio, es decir, los órganos fisiológicos del cuerpo humano que participan en la producción de la voz. Explican qué es el sonido y algunos de sus conceptos físico-acústicos. Podemos encontrar las técnicas de manipulación o, simplemente, la descripción de los valores y potencialidades narrativas del lenguaje oral.

Si nos vamos a los textos sobre producción cinematográfica, pasa lo mismo, y quizá con menos nivel de profundidad. Pocos trabajos (Rodríguez Bravo, Michel Chion, Balsebre) le confieren al sonido y al carácter acústico de la voz la importancia real que tienen en la construcción de los mensajes audiovisuales y dejan claro, los huecos de conocimiento aún existentes que se tienen en esta materia.

El estudio de los *estereotipos vocales*, por ejemplo, no ha sido suficientemente desarrollado y atendido por los que trabajan en los medios, a pesar de su frecuente utilización ya que, la mayoría de las veces, lo hacen de manera intuitiva.

Es decir, que aunque en una producción se pida a un actor que realice su interpretación adoptando una voz TRISTE, no existe un manual que explique cuáles rasgos sonoros de la voz producen la sensación de tristeza.

Asimismo, es muy difícil realizar una descripción de cualquier voz a través de sus rasgos más relevantes:

En muchas ocasiones los oyentes no especializados quieren describir algún rasgo que recuerden con claridad, pero su mayor dificultad radica en no poseer un repertorio de términos, un vocabulario claro, explícito y compartido (RODRÍGUEZ; LÁZARO; MONTOYA; BLANCO; BERNADAS; TENA; LONGHI; OLIVER;, 2003: 27).

Pero esto no sólo se da entre oyentes no especializados, sino también entre quienes cotidianamente trabajan la producción audiovisual. A cuántos no nos ha tocado escuchar que el productor solicita al locutor que su voz la *pinte de rojo*, por ejemplo. ¿A qué se refiere?, ¿cómo suena una voz *roja*?

Este es sólo alguno de los baches de conocimiento que giran en torno de la voz y que falta por resolver. El estudio de la voz a través de la conformación de estereotipos vocales en la mente de los oyentes puede contribuir a disipar algunos de ellos.

- El Estudio de los estereotipos vocales para las investigaciones sobre las formas de representación colectiva.

Existen muchos trabajos de tipo sociológico, antropológico y filosófico que dedican sus esfuerzos al estudio de los mitos, los imaginarios colectivos y las formas en que los grupos sociales construyen las distintas representaciones sobre el *otro*. Estamos hablando de conceptos como identidad cultural, la conformación de prejuicios, racismos, nacionalismos, entre otros.

Por un lado, el idioma empleado por un grupo social condiciona el uso de la voz y su forma acústica, y por otro, determina muchas de las formas de pensamiento y representación de identidad. Volvemos a la dicotomía entre los mecanismos internos de percepción individual que condicionan las formas de representación y cómo inciden las estructuras sociales en la creación y elaboración de nuevas representaciones.

Por supuesto que no sólo el idioma interviene en la construcción de representaciones sociales, pero sí en la conformación de estereotipos vocales, por consiguiente, el estudio de los procedimientos internos de discriminación y simplificación de las formas sonoras vocales, así como la localización específica de estereotipos dentro de una sociedad puede ayudarnos a comprender mejor el grupo social que se estudie bajo estas líneas de investigación.

Cuando hablamos de representaciones sociales nos referimos concretamente a las creencias y valores compartidos colectivamente por un grupo de gente.

Una representación social es un proceso de elaboración perceptiva y mental de la realidad que

transforma los objetos sociales (personas, contextos, situaciones) en categorías simbólicas (valores, creencias, ideologías) y les confiere un estatuto cognitivo que permite captar los aspectos de la vida ordinaria mediante un reenmarque de nuestras propias conductas en el interior de las interacciones sociales. (FISCHER, 1990: 117).

Si entendemos que el lenguaje aparece como la forma más elaborada de la capacidad que posee el ser humano para producir símbolos y signos que representan la realidad, el estudio de la voz como sostén acústico del lenguaje oral necesariamente arrojará valiosos datos acerca de la construcción de esas representaciones.

Las "categorías-tipo" que localiza Fuzellier –aislando las voces de sus propietarios-- están estrechamente ligadas a unas huellas culturales históricas. Existe una relación directa entre la tesitura de la voz con la edad y el sexo. Y estos, a su vez han determinado a lo largo de la historia del hombre qué función social y qué tipo de actividades, hazañas y comportamientos se debían seguir en cada momento" (RODRÍGUEZ BRAVO, 1984: 110).

Como lo mencionábamos en la Justificación, es posible que a través de los estereotipos vocales nos acerquemos a la comprensión de valores construidos socialmente como la belleza o la fealdad de las cosas. Si cuando escuchamos una voz nos produce irritación y logramos localizar los rasgos sonoros que lo provocan puede ayudarnos a comprender cómo se construyen esas representaciones.

Las representaciones sociales se construyen a partir de una serie de materiales de muy diversas procedencias, gran parte de estos materiales provienen del fondo cultural acumulado en la sociedad a lo largo de la historia. Este fondo común circula a través de toda la sociedad bajo la forma de creencias ampliamente compartidas, de valores considerados como básicos y de referencias históricas y culturales que conforman la memoria colectiva y hasta la identidad de la propia sociedad. Todo ello se materializa en las diversas instituciones sociales, por ejemplo la lengua, pero también en una serie de productos sociales que están, a veces, directamente relacionados con el ámbito de la cultura, y otras veces lo están de forma menos directa como cuando se trata por ejemplo de los múltiples

objetos manufacturados que nos rodean o de las propias casas en las que vivimos. (IBÁÑEZ, 1998: 40)

MARCO CONCEPTUAL

2. LA PERCEPCIÓN SONORA

2.1 El uso de la voz y de estereotipos vocales en los medios audiovisuales de comunicación

*El patrón don Pedro le suplica. El niño Miguel ha muerto.
Le suplica su compañía.
Ya lo sé --le dije--. ¿Te pidieron que lloraras?
Sí, don Fulgor me dijo que se lo dijera llorando.
Pedro Páramo. Juan Rulfo*

Se dice que la voz es la expresión sonora del alma, el vehículo más directo para expresar nuestras emociones. Cuando escuchamos una voz podemos sentirnos seducidos, atraídos o hipnotizados por ella; pero también puede ser una señal que nos confunda, irrite y provoque rechazo o miedo. Michel Chion resalta que en la historia de la humanidad, a la voz se le han atribuido poderes específicos: “los mitos nos hablan de voces que atraen o extravían (sirenas, ondinas, voces que se esconden en las aguas o sobre las orillas), que remontan el tiempo, lo eternizan o lo fijan.(1999: 172)”

Si, cuando se trata del sonido, hacemos un mito de ello, es porque vinculamos, en la palabra y en lo que ésta evoca, lo abstracto y lo concreto, y lo espiritual y lo material. (Op. cit.: 172).

En la actualidad, todos estos poderes han sido aprovechados por los medios audiovisuales de comunicación. La voz de un locutor, por ejemplo, no sólo transmite noticias por la radio o la televisión, sino que además transmite seguridad y credibilidad de lo que está informando. En el cine hollywoodense el héroe protagónico no sólo es físicamente bien dotado, sino que su voz nunca es chillona, estridente o gangosa. ¿Por qué? No hay que hacer un largo recorrido conceptual para observar que la utilización de la voz en los medios masivos de comunicación está basada en el empleo y fabricación de personajes-tipo que responden a unos esquemas que los receptores interpretan sin dificultad. Pero ¿cómo lo logran?

En general, se parte del supuesto de que a la voz se la vincula directamente con la producción del habla, es decir, que a través de la palabra oral se pueden expresar ideas y sentimientos y sólo es necesario conocer el idioma para comprender los mensajes; pero no es tan sencillo, sobre todo si tomamos en cuenta que los diferentes sonidos que producimos al momento de hablar, condicionados por la edad o el estado anímico que tengamos, hacen que nuestro discurso adquiera la totalidad de lo que deseamos expresar. El sonido que emitimos también evoca, expresa y comunica; es más, en muchos casos, la palabra sólo nos sirve para reforzar o contradecir lo que estamos sintiendo o pensando.

Pues bien, la voz es uno de los principales elementos que conforman el lenguaje audiovisual, y su uso y manipulación es fundamental para la creación de mensajes a través de los medios de comunicación de masas.

Cuando escuchamos un spot publicitario de radio, vemos una película, o una caricatura de dibujos animados por la televisión, pocas veces nos detenemos a pensar en la cantidad de recursos y técnicas de engaño con que están hechos para que los percibamos como reales. Y la voz, es un recurso más que se emplea en esa manipulación ayudando a que los mensajes convenzan o persuadan según el grado de verosimilitud que ésta alcance. Podemos escuchar hablar a un animal o a una criatura inexistente, o ver a un actor y creer que la voz que se escucha es la suya cuando en realidad se trata de una superposición de otra voz. En fin, los ejemplos del uso de la voz manipulada son muchos y estamos expuestos a ella más de lo que creemos.

(...) lo que suena verdadero para el espectador y el sonido que es verdadero son dos cosas muy diferentes.
(...) el espectador cuando oye en una película un sonido considerado realista, no está en condiciones de compararlo con el sonido real que podría oír junto a él en el lugar de la acción; se refiere más bien, para juzgar su verdad, a su recuerdo de ese tipo de sonido, recuerdo resintetizado a partir de varios datos no únicamente acústicos, e influido por la visión de las películas. (CHION, 1998: 106).

Existen bases perceptivas que nos ayudan a entender el porqué aceptamos algunos engaños como reales y otros no. Los avances tecnológicos hacen que los productos audiovisuales se perfeccionen cada vez más, es decir, la fidelidad en la imagen y en el sonido alcanzan los mejores niveles, los efectos especiales logran un realismo casi perfecto, lo que hace que su efectividad dependa más que nada de nuestros gustos, preferencias y/o necesidades que cubran los mensajes.

Como hemos mencionado, si bien es cierto que cada uno de nosotros tenemos una visión de la realidad según nuestra propia experiencia con nuestro cuerpo y con las condiciones familiares, educativas, sociales, económicas y culturales en las que nos desenvolvemos, también lo es que, como seres humanos, tenemos atributos físicos y cognitivos similares propios de la especie. Nuestro contacto con la realidad se da gracias a nuestros cinco sentidos y, al vivir en sociedad, aprendemos a comunicarnos a través del lenguaje, de las costumbres y hábitos que nos han sido heredados; en fin, aprendemos colectivamente todo aquello necesario para la supervivencia y convivencia en grupo.

Para que el ser humano sea capaz de reconocer e identificar una voz de todas las demás, y configurar una serie de interpretaciones sobre ella, es gracias a que memoriza señales similares que se van reiterando a lo largo de su vida, desde su propia experiencia como hablante, y desde la experiencia de escuchar a los que le rodean.

“Cuando hablamos, todos los rasgos de nuestra expresividad sonora se organizan en primer lugar desde un doble punto de vista que tiene su frontera y su eje de simetría en nuestra propia piel. O bien damos información sobre nosotros mismos (EXPRESIÓN AUTOACÚSTICA), o la damos de todo aquello que nos rodea (EXPRESIÓN ECTOACÚSTICA)”. (RODRÍGUEZ BRAVO, 1989: 49-50).

Los psicólogos sociales hablan de marcos o *esquemas* mentales que nos ayudan a organizar grandes cantidades de información diversa de manera efectiva. “Los esquemas son marcos mentales que contienen información relevante de situaciones o acciones específicas, las cuales, una vez establecidas, nos ayudan a interpretar estas situaciones y lo que pasa en ellas” (BARON y BYRNE, 1998: 86).

Pero la interpretación que podamos hacer de una persona según su voz, está condicionada por los llamados *prototipos* que son también “modelos mentales de las cualidades típicas de los miembros de algún grupo o categoría” (BARON y BYRNE: 1998, 89). Es decir, que si la radio o el cine nos ha mostrado un *prototipo* de voz de *galán*, cuando conocemos a alguien físicamente atractivo, esperamos que su voz concuerde con esa categoría aprendida, de lo contrario suele modificarnos la percepción que teníamos a primera vista.

Si anteriormente citábamos que “la percepción es el primer proceso mental que se debe producir para que acontezca cualquier otro” (CUESTA, 2000: 110), nos detendremos un momento para revisar este concepto fundamental.

2.2 La Percepción

El acto de percibir es un proceso que requiere de un estímulo durante cierto tiempo para la captación del significado de lo que se percibe. Es decir, no se trata de una simple recepción de datos. Cuesta (2000, 110) apunta que “el hombre percibe a través de sensaciones, pero la percepción es, en realidad, un análisis de todas ellas. Este análisis perceptual es un

proceso *sensocognitivo*. Es decir, la percepción hace referencia a un proceso activo, cognitivo, elaborado. No consiste en un proceso fisiológico mecánico, pasivo, automático”.

Es por eso que lo percibido depende de la psicología del receptor, y se deben tener en cuenta, en todo momento, sus procesos cognitivos. Como decíamos en la introducción, existen principios perceptivos resultantes de nuestra condición humana:

La psicología perceptiva ha demostrado de forma experimental, que cuando no se dan unas condiciones especiales como: ambigüedad, confusión, tensión, agotamiento, miedo, etc., se comprueba sistemáticamente que frente a estímulos idénticos los sujetos coinciden en sus percepciones dentro de unos márgenes muy concretos y estrechos” (RODRÍGUEZ BRAVO, 1998: 111).

Sin embargo, no olvidemos el carácter social del percibir:

Actualmente no se discute que éste es un proceso de clara interacción entre lo innato y lo adquirido, una especie de urdimbre donde no cabe hablar de mayor o menor peso del uno sobre el otro, puesto que no se trata de procesos yuxtapuestos, sino de un mismo proceso interactivo. (CUESTA, 2000: 124).

A continuación haremos una revisión conceptual de lo que es la memoria humana ya que gracias a ella es que se puede dar este doble mecanismo y nos permite formar una interpretación de las cosas.

2.3 La memoria humana

La memoria es “un sistema para el almacenamiento y la recuperación de información, información que es obtenida mediante nuestros sentidos. Nuestra memoria es un registro de percepciones (...) y la percepción implica necesariamente alguna forma de almacenamiento del estímulo en el tiempo” (BADDELEY, 1999: 11).

Con más detalle, Dolores Sáiz y Josep Baqués (1996: 69) describen estas tres fases:

(...) para que una información se recuerde, en primer lugar ha de ser registrada, es decir, la fase de

adquisición va a suponer algún tipo de codificación que facilite la entrada de la información. En segundo lugar, la información registrada ha de ser conservada o retenida, se ha de guardar. En tercer lugar, cuando se necesita, ha de ser recuperada, es decir, no sólo ha de estar disponible sino que ha de ser accesible".

A su vez, estas tres fases forman unos modelos *estructurales* de almacenamiento. Los *estructuralistas* postulan que los estímulos percibidos pasan por diferentes elaboraciones que permiten su adecuado procesamiento. El esquema o estructura inicial es cuando el sujeto percibe un estímulo y lo retiene para analizar sus características físicas más básicas y desarrollar una clasificación "muy elemental en términos de variables tales como forma –visual, auditiva, táctil o de otro tipo–. A este proceso de almacenamiento se le denomina *almacén sensorial*". (CUESTA, 2000: 174). Posteriormente, se produce una selección de información según sea del interés del receptor y pasa a otra estructura temporal denominado *almacén o memoria a corto plazo*. Y finalmente, "existiría un mecanismo responsable de integrar esta información con los sistemas propios del sujeto al que se denomina *almacén o memoria a largo plazo*". (Op. cit.: 177).

Cuando hablamos de memoria nos referimos a la forma y modo en como representamos el mundo. Manzanero precisa diciendo que "la mayoría de las personas creen que el cerebro graba percepciones de hechos particulares de forma similar a un video. Sin embargo, la memoria no graba, sino que interpreta y reconstruye la realidad" (MANZANERO, 1996: 274) "La memoria a largo plazo no es un almacén infinito, sino que se producen fenómenos de distorsión y olvido." (CUESTA, 2000: 177).

2.4 La memoria auditiva

La información que percibimos por el sentido del oído ha sido estudiada y clasificada bajo el nombre de *memoria auditiva*. Esto significa, que existe un amplio registro sonoro en nuestra mente que se va ampliando a lo largo de nuestras vidas. Así como memorizamos imágenes, olores y sabores somos capaces de memorizar sonidos y voces. Pero no todo lo que percibimos por nuestros oídos es registrado, ni toda la información registrada es recordada de la misma manera y en todo momento.

Es imposible pensar que toda la información que escuchamos es memorizada. Los psicólogos hablan de que la memoria auditiva también puede dividirse en "*memoria ecoica*, que se extiende por milisegundos; la *memoria auditiva a corto plazo*, que se extiende hasta quizá los cinco o diez segundos, y la *memoria auditiva a largo plazo*" (BADDELEY, 1999: 22-25-28). Y la recuperación de esta información dependerá, entre otras cosas, del nivel de almacenamiento en que fue registrada.

Pero nos vamos a centrar en las investigaciones que se han realizado sobre la memoria auditiva a largo plazo, las cuales han estado interesadas en la capacidad que tenemos para recordar e identificar las voces.

El estudio sobre el reconocimiento de las voces toma especial auge en la identificación de sujetos en los tribunales como testimonio judicial. Intentaremos resumir algunos resultados encontrados. Se demuestra que una voz conocida o familiar puede ser identificada hasta en un 95% de los casos pero con una disminución considerable cuando esas voces cambian su registro utilizando un disfraz, "esto es, hablando como una persona anciana, con voz exageradamente nasal o con voz ronca". (KNAPP, 1982: 288). Y el olvido es alarmante cuando, en condiciones óptimas, se trata de una persona desconocida.

"Las implicaciones generales de estas investigaciones son claras: que aunque el reconocimiento de la voz de una persona conocida puede ser razonablemente bueno, es probable que el de un extraño sea pobre, incluso en condiciones óptimas. Parece lógico que el rendimiento será aún peor cuando el estado emocional del hablante difiera de una ocasión a otra, o cuando se realice algún intento para distorsionar la voz, susurrando o adoptando una voz falsa. Finalmente, es improbable que los sujetos que no prestan particular atención a la voz sean buenos testigos, especialmente tras una demora." (BADDELLEY, 1999: 29- 30 y 31, citando a POLLACK, PICKETT y SUMBY (1954), y a CLIFFORD y FLEMMING).

Sin embargo, otras investigaciones dan cuenta sobre "el probable de que cuando se intenta reconocer la voz del hablante, los oyentes utilizan muy pocas de las muchas características de la voz" (KNAPP, 1982: 33).

Esto tiene su base en la manera en como reconocemos cualquier forma sonora asignándole un valor. Ángel Rodríguez Bravo en su libro *La dimensión sonora del lenguaje audiovisual* (1998, 139) explica que esto lo hacemos:

“ya sea por la repetición reiterada de la misma forma sonora, el impacto emocional de la situación a la que está asociada, el valor que supone identificar esa forma para la supervivencia del oyente, etc. Como consecuencia de esa experiencia, la forma en cuestión se almacenará en nuestra memoria como un patrón sonoro”.

Los patrones sonoros los entiende como:

“las formas sonoras que al ser reiteradamente identificadas y asociadas a otros fenómenos concretos adquieren valor de modelo sonoro de referencia y valor de signo. Así, en la memoria auditiva se van acumulando formas-patrón y (...) su uso reiterado (...) hará que el ser humano tienda con mucha fuerza a la percepción categorial, es decir, a escuchar sólo aquello que busca, o aquello que está acostumbrado a reconocer habitualmente, a menudo, incluso, ignorando todo lo demás. (Op. cit.: 139).

Es así que algunos esfuerzos han sido más bien encaminados a encontrar si las personas somos sensibles a ciertos rasgos de la voz que nos permiten emitir juicios sobre personalidad, emoción, características físicas, sociales o de la actitud que sugiere la voz del emisor.

Una de las investigaciones más completas en este terreno fue la realizada por Addington, quien descubrió que muy frecuentemente se emitían juicios estereotipados sobre señales vocales. (...) Llegó a la conclusión de que la personalidad masculina era percibida generalmente en términos de poder físico y emocional, mientras que la personalidad femenina lo era en términos de facultades sociales. (KNAPP, 1980: 291).

De la misma manera, unos con más éxito que otros, se han encontrado similitudes en juicios sobre características personales como la edad, sexo, lugar de nacimiento, actividad profesional, rasgos físicos o emotivos del hablante, pero sin llegar a resultados contundentes. Y es que son muchas las variables que entran en juego y que entorpecen las investigaciones, como lo son los errores que se producen al interpretar los instrumentos tecnológicos de medición de la voz, en la dificultad para

construir un "corpus" adecuado, basados generalmente por voces que monologan cuando muchas señales vocales surgen más en la forma de diálogo, o en la utilización de actores que fingen o simulan estados de ánimo y que probablemente no alcanzan la verosimilitud correcta; o también, debido a las imprecisiones de los juicios que se han de formular.

Por su parte, y más recientemente, Ángel Rodríguez Bravo, en su tesis doctoral *La construcción de una voz radiofónica* (1989), escribe en sus conclusiones:

"No es cierto que cada oyente de radio se imagine al locutor "X" tal y como a él le gustaría que fuese, ni que se imagine a este locutor de una forma diferente. Muy al contrario, el sonido de cada voz lleva asociada una imagen, imagen que se repite casi fotográficamente en la mente de la gran mayoría de los oyentes. Y esto ocurre a pesar de que la imagen mental que genera la voz del radiofonista "X" no necesariamente tenga que ver con su imagen física real" (:254).

Esas coincidencias perceptivas formadas en los escuchas a partir de voces radiofónicas, demuestran, una vez más, la existencia de patrones sonoros en la mente del ser humano por lo menos a los pertenecientes a un mismo grupo social. Es decir, al momento en que un actor de radio o de doblaje televisivo es capaz de transmitir emociones o de personificar a un hombre como *guapo* o *villano* a través de su voz y, nosotros de reconocerlo como tal, estamos hablando de que existen esas formas patrón en nuestra mente.

3. LA VOZ Y SU PERCEPCIÓN

*"¿No me oyes? –pregunté en voz baja.
Y su voz me respondió (...),
su voz parecía abarcarlo todo. Se perdía más allá de la tierra.
Tú sabes cómo hablan raro allá arriba; pero se les entiende".
Juan Rulfo, Pedro Páramo.*

3.1 Señales sonoras de la voz que transmiten algún tipo de información al receptor

Si por un momento nos detenemos a observar toda la información que podemos obtener al escuchar una voz, no resulta difícil identificar los siguientes niveles de interpretación:

1. Información acerca de la fuente de donde proviene la voz: si pertenece a la de una persona físicamente próxima al oyente, o si es emitida desde un aparato de radio, de TV, etc.
2. La dirección exacta de donde proviene la voz.
3. La distancia aproximada de su emisión.
4. El entorno acústico del espacio en el que se produce, por ejemplo, desde algún lugar cerrado o abierto, desde una cueva o la calle.
5. Si se trata de un hombre o una mujer.
6. Si se trata de una persona joven o anciana
7. La cualidad acústica de la voz, por ejemplo, muy grave, gangosa o potente.

Estos primeros siete niveles que cualquier oyente es capaz de interpretar y dar sentido, se les puede clasificar como *formas sonoras simples de percepción* que se generan cuando la voz ha sido escuchada y ha pasado a un nivel mental de reconocimiento y comprensión del sonido. (RODRÍGUEZ BRAVO, 1998: 209. Véase también SHEFFER, 1988: 61-66).

De manera simultánea, también somos capaces de interpretar:

8. Información semántica, es decir, la identificación de la lengua, las palabras y su correspondiente interpretación sobre lo que nos está diciendo.
9. Información sintomática del hablante, cuando éste presenta algún trastorno de salud o psicológico. (término propuesto por RODRIGUEZ BRAVO, 2002: 163).
10. Información encuadrativa, que proporciona al oyente información sobre "los distintos grupos sociales, culturales, étnicos, geográficos, etc. en los que se puede encuadrar al locutor" (Op. cit: 164).
11. Información "empática, es decir, sobre aquello que describe el locutor" (Op. cit: 164)
12. Información sobre la actitud emocional de quien habla y, por último
13. Información "idiográfica", es decir, sobre el aspecto físico de quien habla. (Op. cit.: 163).

Estos seis últimos niveles de sentido que le asignamos a una "voz cualquiera" son *formas sonoras complejas perceptivas*, que adquieren esa dimensión por nuestra condición social de pertenencia a un grupo. (RODRÍGUEZ BRAVO, 1998: 206)

A continuación, definiremos lo que es la voz humana y nos detendremos a explicar cada uno de los niveles de información que hemos enumerado.

3.2 La voz

Se dice que el hombre desde que nace posee voz pero el lenguaje y su manera de expresarlo es aprendido. "Desde el instante en que nacen los seres humanos comienzan a emitir sonidos; (...) el niño no habla, sino que emite sonidos al igual que cualquier cachorro que posea voz" (KONDRÁTOV, 1973: 112).

Por su parte, Guy Cornut, en su libro *La voz*, dice que "representa en esencia el sostén acústico de la palabra por lo que es energía vital originada en un impulso interior vinculado a las emociones fundamentales".(1985, 56).

Si tenemos en cuenta que ninguno de los órganos fisiológicos del cuerpo humano que intervienen para la producción de la voz tienen "como función principal esa actividad: los pulmones sirven para respirar; la lengua y los dientes intervienen en el proceso de ingestión de alimentos; sin embargo, todos ellos son fundamentales en la articulación del lenguaje" (ALCOBA, 2000: 38), podemos decir que la voz como sonido articulado es privativo del hombre. La voz "es producto, en última instancia del cerebro humano, porque sonido articulado quiere decir, expresión racional por medio del lenguaje" (CASTARLENAS, 1995: 85 y 86).

Al conjunto de órganos que intervienen en la producción de la voz se le conoce como sistema fónico o fonador.

De esta manera y resumiendo, decimos que la voz *es el sonido de la expresión verbal producido por un complejo sistema neuro-fisiológico en el que actúan distintos órganos de nuestro cuerpo y que se proyecta en el espacio con la intención de producir un efecto en el auditor.*

3.3 Información generada por las formas sonoras simples

El término de *forma sonora* es definido por Rodríguez Bravo (1998: 141) como "toda configuración acústica que, aun siendo analizable en dimensiones más simples, tiende a ser percibida como un bloque sonoro unitario y coherente".

Podemos decir entonces, que cualquier voz es una *forma sonora*, conformada a su vez, por formas sonoras simples y complejas.

Una forma sonora simple es "cualquier configuración sonora constituida por las variaciones de una única dimensión acústica, que tiende a ser percibida como un bloque sonoro unitario y coherente". (RODRÍGUEZ BRAVO, 1998, 173).

Las *formas sonoras simples* son el resultado de la combinación de los cuatro elementos que componen el sonido: tono, intensidad, timbre y duración, y pertenecen a un primer nivel de nuestro mecanismo perceptivo antes de que medie la educación. Dependen de las características físicas del hablante y del entorno acústico por donde viaje el sonido.

El ser humano es capaz de asociar los estímulos sonoros que percibe por su sistema de la audición a los distintos objetos que los originan aún cuando se encuentren fuera del alcance de su vista. Puede reconocer, sin ningún problema, si el sonido es emitido por una persona, un animal, un objeto o un altavoz. Si se trata de una sola voz o de varias. Asimismo, a partir del tiempo que tarda en llegar el sonido de un oído o a otro, puede saber la dirección y orientación del sonido que está percibiendo. Otra característica de la percepción auditiva es que a través de ella podemos reconocer con "gran precisión sensaciones espaciales" (RODRÍGUEZ BRAVO, 1989: 222). Por un lado, a través de la intensidad de los sonidos somos capaces de conocer su distancia pero, también, por las "reflexiones del sonido y su envolvente espectral" (Op. cit.:222), de identificar las formas y volúmenes del espacio donde se produce.

Decíamos que los escuchas podemos percibir a través de una voz, la edad, el sexo y la complexión física del hablante. Cabe recordar que esto no significa que necesariamente exista una coincidencia con el aspecto físico real del hablante.

Definimos como edad a través de la voz al:

Registro vocal que sugiere la edad del hablante debido a los cambios físicos ocurridos a lo largo del tiempo de vida del individuo. Según las distintas etapas: infancia, niñez, adolescencia, edad adulta temprana, media y avanzada, la voz sufre cambios en tono, flexibilidad, velocidad, intensidad, cualidad vocal y en el control articulatorio.

Ángel Rodríguez Bravo (1989, 50) encuentra que "cuando un oyente decide qué edad tiene un emisor solamente escuchando el sonido de su voz, lo hace mediante el siguiente doble mecanismo:

- a) vinculando desde la propia experiencia comunicativa la evolución de la calidad del timbre de la voz con el envejecimiento físico del cuerpo humano;
- b) Oponiendo entre sí todas las calidades tímbricas asociadas por la experiencia auditiva a distintas etapas de envejecimiento, que tiene almacenadas en su memoria acústica.

Algunas investigaciones han demostrado que a través de la voz se puede calcular con mucha precisión la edad del hablante. Aunque, como hemos venido mencionando no nos interesa si esas apreciaciones coinciden o no con la realidad, vale la pena citar algunos resultados obtenidos.

A través del tono de la voz, se encontró que "en varones durante la infancia, la niñez, la adolescencia y la edad temprana, media y avanzada, existe un descenso general del tono desde la infancia y a través de la edad madura. Posteriormente se da una tendencia inversa y la voz se hace ligeramente más aguda a medida que la edad avanza" (KNAPP, 1982: 298). Estudios similares sobre las mujeres no han proporcionado datos tan notorios pero si tomamos en cuenta que a través del tiempo nuestro cuerpo cambia, es seguro que la voz sufra cambios por la flexibilidad, velocidad, intensidad y control articulatorio.

Asimismo, una voz nos puede sugerir con gran precisión el sexo de quien la produce. Definimos que es el:

Registro vocal que sugiere el sexo del hablante, es decir, sugiere si se trata de un hombre o una mujer.

El sistema que organiza la transmisión de informaciones útiles para que un oyente radiofónico pueda reconstruir al locutor que está escuchando está basado esencialmente en la interpretación de índices acústicos. Es decir, está basado en el reconocimiento de formas sonoras que se asocian automáticamente a la influencia que la edad, la actitud, la personalidad, la salud, etc. tienen en el aparato fonador, y, por tanto, en la voz humana. (RODRÍGUEZ, 1989: 51).

Aunque esta cita corresponde al estudio que el autor realizó sobre la voz en la radio, nos es de gran utilidad debido a que en la radiodifusión, el único elemento que se tiene para transmitir los mensajes es el sonoro, y la voz humana es sustancial. Es así que sus resultados corresponden a los de cualquier voz, sea ésta manipulada por algún medio de comunicación o no, y a los de cualquier sujeto que la perciba.

Knapp (1982: 294) expone que según un estudio realizado por Pear sobre señales vocales y juicios de características personales, dejó establecido que "se podía estimar con mucha precisión la edad del hablante; con notable precisión el sexo; con poca precisión el lugar de nacimiento, y que, en algunas ocasiones, también se podía establecer con sorprendente precisión la actividad del hablante".

Por último, y siguiendo el mismo esquema, definimos que la complexión física sugerida por la voz es:

Registro vocal que sugiere la complexión física del hablante, es decir, sugiere si se trata de una persona alta o baja, robusta o delgada, atractiva o poco atractiva.

En la Tesis Doctoral de Rodríguez Bravo (1989: 139-143) dice que la "variable "AGRADIBILIDAD" está fuertemente asociada con "ATRACTIVO" (...) o belleza del locutor. (...) Cuanto más apasionada sea la actitud del locutor mayor será el tamaño con que se lo imagine el oyente. Y, contrariamente, que la humildad es una voz producirá el efecto de que el locutor sea imaginado como alguien pequeño. (...) Por su parte, la forma física se relaciona positivamente con la seguridad y la extroversión y ésta a su vez con la agilidad, podemos interpretar, entonces, que una voz con actitud segura y extrovertida generará en el radioescucha la imagen de un locutor o una locutora de aspecto ágil y atlético".

Por otro lado, encontramos que existen muchos sonidos de la voz producidos también por las características físicas de cada individuo (tamaño de las cuerdas vocales, caja torácica, etc) pero que no tienen que ver con las condiciones de género, edad o complexión física del hablante. Por ejemplo, podemos encontrar una voz es mas débil que otra, mas grave o mas aguda, mas ronca o trémula que otra. El término "cualidad de la voz" se refiere "al sonido casi permanente que presenta más o menos todo el tiempo una persona cuando habla" (ABERCROMBIE, 1967: 91. La traducción es mía).

3.4 Información generada por las formas sonoras complejas

A partir del entrenamiento cultural y educativo es que podemos encontrar otros niveles de estructuración perceptiva como una cadena sin fin de formas y super-formas.

Las *formas sonoras complejas* se forman "cuando agrupamos perceptivamente como una única unidad auditiva sonidos que estén compuestos por dos o más formas simples, ocurra esto oyendo las distintas formas simples de manera sucesiva, o bien simultáneamente en el tiempo" (RODRIGUEZ BRAVO, 1998, 192).

Las *formas sonoras complejas* de una voz las encontramos en las formas entonativas, rítmicas y tímbricas.

- Información Semántica

Si nos centramos en la información semántica, en lo que respecta a la producción del habla sus características de producción y escucha también requieren de un mecanismo perceptivo complejo que los lingüistas se han encargado de estudiar:

El habla está basada en la variación. (...) A pesar de que en la palabra *dado* la unidad /d/ se pronuncia de forma distinta en la primera y segunda sílaba, un hablante de español es capaz de hacer abstracción de tal diferencia y clasificar los dos segmentos como pertenecientes a una misma categoría: el fonema /d/. La eficacia de nuestro sistema reside en la capacidad humana de encontrar elementos invariables en lo que aparentemente es variable. Conocer una lengua implica agrupar en clases –fonemas– lo que es diferente en la realidad física. (...) la percepción está condicionada por los principios de categorización de la lengua materna. (ALCOBA, 2000: 36).

Como hemos venido mencionando, la voz humana transmite información sonora más allá de su intención léxica o gramatical y que la paralingüística entiende como fundamental al momento de hablar:

Estos rasgos no verbales de un enunciado son precisamente tan relevantes para determinar un significado como el significado de la palabra y el

significado gramatical ambos codificados en el componente verbal." (LYONS, 1983: 34, citado por RODRIGUEZ BRAVO, 1984: 16).

Asimismo, no olvidemos que el registro acústico de la voz de cada hablante también influye en la transmisión de signos y determina la totalidad de los mensajes:

Están determinados por factores fisiológicos –como el grado de apertura de las bandas vocales–, a su vez afectados por variables personales de índole psicológica y emocional y por razones culturales. Su verdadera importancia reside en las funciones socioculturales que cumplen en la comunicación, ya que producen una serie de efectos de voz percibidos socialmente de acuerdo con las apreciaciones y el sistema de valores estéticos imperante en un momento dado; por ejemplo, la voz trémula de miedo o el susurro de la confidencialidad. (CÁCERES, 2003: 161).

- Información sintomática

Este término es empleado por Rodríguez Bravo y hace referencia a la interpretación de los estados de salud y psicológicos del hablante.

Esta función expresiva está asociada a lo que en medicina se denominan disfonías o disodeas y se caracterizan a partir de la comparación con las capacidades acústicas medias de los hablantes sanos. Se diferencia entre disfonías psíquicas (tartamudeo, contracciones espasmódicas de la glotis, etc) y las disfonías orgánicas (afonía, diplofonía, dicrotismo, etc.) Patologías psíquicas como la depresión o la esquizofrenia, son también reconocibles en el sonido del habla) (2002: 163).

- Información encuadrativa

Cuando escuchamos hablar a una persona podemos reconocer la región o país de procedencia, así como su status cultural y étnico.

Esta función está vinculada a determinadas desviaciones de la forma de emitir los sonidos del habla que son características y comunes entre los miembros de determinados grupos a los que pertenece el orador. Un

ejemplo claro es lo que se suele denominar como "acento regional", y que permite identificar con facilidad el origen geográfico del hablante. (RODRÍGUEZ BRAVO, 2002: 164).

- Información empática

Este término hace referencia a:

Los matices sonoros de la voz que aportan al receptor información sonora de aquello que describe el locutor. Se desencadena en las situaciones en que el orador, al hablar de algo externo a sí mismo, construye el sonido de su voz desarrollando intuitivamente un proceso de identificación o de imitación de lo descrito o nombrado. Actúa, por ejemplo, cuando un locutor narra la velocidad de los movimientos de una carrera aumentando su propia velocidad de locución, o cuando explica la potencia y la energía de un gesto haciendo sonar su voz con mucha más fuerza. (Op. cit.: 164)

- Información emocional

Las emociones humanas se ven indudablemente reflejadas en la voz al momento de hablar. Existe suficiente evidencia de que cuando experimentamos una emoción se producen cambios en la respiración, la articulación y la fonación. Las distintas combinaciones en la intensidad, timbre, tono, entonación, articulación y uso de diferenciadores nos llevan a reconocer en las voces información relevante sobre el estado emocional del emisor. La identificación de las emociones a través de la voz se da gracias a la información de algunos rasgos sonoros y que posiblemente repercuten en la conformación de estereotipos. Si un hombre de 40 años, que tiene una voz grave y profunda, habla con una intensidad fuerte, a una velocidad rápida pero expresándose de manera fluida y segura, probablemente nos dará la sensación de que la persona está anímicamente contenta. Si esa persona varía alguno de los elementos acústicos mencionados, la percepción puede variar de manera drástica. Es decir, si el hombre en vez de 40 tiene 65 años y hablara a una velocidad demasiado lenta, pero respetando todos los atributos anteriores, quizá nos pueda hacer sentir que escuchamos a un hombre deprimido o cansado,

inseguro al expresar sus ideas. Así, el caso de esa voz tendría muchas connotaciones diferentes si en vez de la velocidad se hace alguna variación en la intensidad, entonación, tono, timbre y acento al momento de hablar.

Las mismas palabras se pueden decir de forma completamente distinta, expresando diferentes emociones e incluso diferentes significados (animación, afecto, diversión, ira, aburrimiento, admiración, desesperación, disgusto, antipatía, miedo, impaciencia, alegría, satisfacción y sorpresa). El cambio lo producimos jugando con la calidad de la voz: el volumen, el tono, la velocidad, la calidad (jadeo, lágrimas etc.) y la fluidez. Todos estos aspectos están relacionados con los estados emocionales. Por ejemplo, una persona animosa habla muy deprisa y con un tono elevado. Una persona deprimida habla despacio y en un tono bajo. (GARRIDO, 2000: 179).

En el artículo titulado *Modelización acústica de la expresión emocional en el español* (RODRÍGUEZ; LÁZARO; MONTOYA; BLANCO; 1999: 1135) se analiza la voz emocionada desde una perspectiva híbrida entre la "universalista" con la "culturalista" formalizada por Scherer y que nosotros tomaremos como válida para este trabajo. Esto supone que:

"La voz sufre cambios acústicos que están motivados directamente por las alteraciones fisiológicas que se producen en el cuerpo humano cuando el individuo experimenta alguna emoción; Y los rasgos acústicos de la voz producidos por las emociones resultan parcialmente modificados durante una locución por la influencia específica de la lengua que está utilizando el hablante emocionado".

En muchas investigaciones sobre voz y emoción, uno de los problemas centrales es la confusión al momento de juzgar una emoción. Issac Garrido (2000: 178) hace mención de tres dimensiones de semejanza que generan esa confusión: cualidad, intensidad y valencia.

La primera dimensión, la cualidad de una emoción es cuando las emociones son semejantes pero diferentes en intensidad, por ejemplo *rabia* e *irritación*, *tristeza* y *desesperación*. "Esto se debe a la falta de definición entre la forma intensa y la forma débil de la emoción respectiva" (Op.

cit.:179). La segunda dimensión es la intensidad. Esto es cuando una emoción difiere claramente en cualidad con otra, pero sus intensidades son semejantes, por ejemplo *desesperación* y *rabia*.

Por su parte, la valencia es cuando las emociones positivas son relacionada o confundidas con otras emociones positivas y no con emociones negativas.

Según el artículo *Modelización acústica de la expresión emocional* antes mencionado, tras un análisis de tres aspectos específicos vinculados al estado emocional: frecuencia fundamental, presión sonora o intensidad y duración o ritmo del discurso, se encontraron variaciones muy claras y específicas de la voz de los locutores según la emoción expresada y sus resultados nos serán de gran utilidad al momento de analizar nuestros resultados en la segunda etapa de nuestra investigación.

- Actitud Comunicativa

Por otro lado, hemos resaltado que la voz la utilizamos para comunicar ideas, sentimientos y emociones, de ahí que la actitud o intención que tenga el emisor al momento de hablar influya en la interpretación global del mensaje.

Según Piñuel y Gaitán, (1999: 316 y 317) la actitud es: “la organización relativamente duradera de predisposiciones de comportamiento frente a la interacción con otros, y que induce o anticipa respuestas preferentes hacia el objeto de la interacción, o hacia situaciones de interacción, o hacia el mantenimiento de las predisposiciones mismas”. Estas actitudes o predisposiciones se manifiestan a través del habla (no sólo a través del lenguaje verbal) ya que hacen referencia a normas y valores que los emisores desean expresar.

Es el posicionamiento frente al otro que adopta el emisor al momento de emitir su mensaje oral y que condiciona la forma y manera de hablar (ya sea de manera consciente o no), y que el receptor es capaz de interpretar. A través de una voz podemos reconocer la ironía, la calidez, la hostilidad o la ternura, entre otros.

4. ANÁLISIS DE LOS ELEMENTOS ACÚSTICOS DE LA VOZ Y SU NIVEL EXPRESIVO

(...) la reseña del pregonero se oía más cercana, hasta que, detenido en una esquina, abocinando la voz entre sus manos lanzaba sus gritos agudos y filosos:
"Alazán Tostado... Veinte pesos a quien lo encuentre... Sin averiguatas..."
Esta última frase era larga y destemplada (...)
hasta que el pregón se alejaba de nuevo
y luego se disolvía en los rincones más apartados del pueblo"
Juan Rulfo, Pedro Páramo.

"(...) y pácatelas y pácatelas, oímos los riflonazos".
Juan Rulfo, El llano en llamas.

En este apartado revisaremos los principales conceptos acústicos de la voz y su relación perceptiva, descriptiva y expresiva que el oyente es capaz de identificar.

Como cualquier sonido, la voz se compone de tres elementos: intensidad, tono, timbre y duración los cuales se organizan perceptualmente como un todo influyéndose mutuamente. Cada uno de estos elementos acústicos determinan su dimensión expresiva.

4.1 Intensidad de la voz

Para definir lo que es la intensidad citaremos a Regidor Arribas quien dice que es "la impresión subjetiva que permite

distinguir los sonidos entre débiles y fuertes, y depende principalmente de la amplitud del movimiento vibratorio que origina el sonido" (REGIDOR ARRIBAS, 1977: 88., citado por MONTOYA, 2000: 41).

Para entender su relación en la conformación de estereotipos vocales, haremos una diferenciación entre la intensidad de una voz que transporta contenidos sobre el emisor, y la intensidad de una voz que el emisor utiliza para transmitir elementos externos a él mismo (*expresión ectoacústica*).

En el caso de la primera, cada persona tiene una capacidad distinta de incrementar el sonido de su voz según su fuerza espiratoria. Encontramos que una voz puede ser:

- Fuerte
- Media alta
- Media baja
- Débil

Una voz muy fuerte, medida en decibelios, registra una intensidad de unos 70 dB. En cambio una voz muy baja registra unos 40 dB. Una voz en una conversación normal suele registrar unos 50 dB.

La capacidad intensiva se traduce en una relación de poder: aquel que tiene más voz es el más importante (...) Una forma característica de dominio es gritar más que el adversario, estar acústicamente por encima de él; interrumpirlo mientras habla apagando su voz con la nuestra. (RODRÍGUEZ BRAVO, 1984: 116).

La intensidad de una voz genera también el efecto de distancia entre quien emite y quien escucha, crea la sensación de relación espacial entre ambos. Esto se debe a que hemos aprendido, a través de la experiencia perceptiva, que nuestro sentido de la audición sólo nos permite escuchar sonidos que están próximos a nosotros, y como al mismo tiempo somos capaces de relacionar el sonido con el objeto que lo produce, sabemos qué tan lejos o cerca se encuentra de nosotros. De igual manera, en la comunicación cotidiana utilizamos diferentes niveles de volumen según con quien estemos hablando y el contexto en el que nos encontremos. Podemos distinguir entre la distancia íntima, la personal, la social y la pública (BALSEBRE, 1981).

Por su parte, la intensidad que el emisor utiliza para transmitir elementos externos a él (*expresión ectoacústica*), tiene una función claramente en la producción del habla, en el que se

realizan movimientos sucesivos de aumento y disminución de la energía del impulso espiratorio:

Existen diferencias de intensidad entre los sonidos —/a/ es más intensa que /i/, de ahí la mayor perceptibilidad de las palabras integradas por esa vocal— y entre las sílabas de una palabra (en español, la intensidad es un correlato perceptivo de las sílabas acentuadas). Pero también puede usarse la intensidad en fragmentos largos de discurso con un valor pragmático; por ejemplo, solemos gritar cuando estamos enfadados o usamos menor intensidad al hacer confidencias o incluir una aclaración en el discurso. La intensidad desempeña un papel importante en la distribución melódica de las unidades del habla, al formar grupos de intensidad (o tónicos), definidos como conjunto de sílabas subordinadas a un solo acento. No siempre el grupo de intensidad tiene una línea tonal independiente, pero sus variaciones afectan de manera directa al aspecto dinámico de la elocución: el movimiento creciente o decreciente, los cambios bruscos o suaves configuran una gran parte de los efectos prosódicos del discurso. (ALCOBA, 2000: 107-108).

Podemos utilizar la intensidad de nuestra voz para describir o recrear objetos de mayor o menor magnitud apoyándonos en las relaciones analógicas con la realidad que aproximan al oyente hacia lo que estamos viendo o imaginando.

Podemos resumir diciendo que la intensidad o volumen de una voz tiene como principal función hacer llegar el mensaje hasta un potencial oyente, pero también puede transmitir información extra y generar diferentes reacciones como por ejemplo, dar la sensación de lo que se dice es interesante.

4.2 Tono o frecuencia de la voz

Como todo sonido, la voz también presenta cualidades según la frecuencia de sus vibraciones. “Nos estamos refiriendo al concepto físico de altura del sonido” (MONTROYA, 2000: 43). Y es que de acuerdo con la intensidad con que expulsamos el aire, su velocidad y grado de excitabilidad habrá vibraciones de mayor o menor altura. En el caso de la voz, se entiende por vibración, la velocidad de salida del aire por las cuerdas vocales.

Durante la voz hablada la altura tonal oscila alrededor de una frecuencia media que, en la práctica, es siempre idéntica para un sujeto dado en la conversación espontánea, a la cual Tarneaud ha dado el nombre de *fundamental usual de la palabra* (Fo). Ese sonido producido por la frecuencia fundamental es la base sonora resultante de la superposición de otras vibraciones y es la que marca con más fuerza el sonido tonal de una voz.

La voz de cada persona tiene una altura tonal que puede alterar por arriba o por abajo pero hasta un cierto límite, condicionado por las posibilidades anatómicas. Cada uno podemos hacer que nuestra voz suene más aguda o más grave. A esta posibilidad de alteración se le conoce como la extensión tonal de la voz. Pero cuando no realizamos ningún tipo de alteración y emitimos nuestra voz de manera que nos resulta más cómoda de manejar, a esa forma tonal se le conoce como la tesitura de la voz.

En la conformación de estereotipos, existe una relación estrecha entre la tesitura de una voz y la imagen que sugiere del emisor. En el teatro, en la ópera y en los medios de comunicación podemos encontrar algunos de los estereotipos más básicos del arte dramático: la voz de galán, del villano, del héroe, del traidor, del noble o cómico, entre otros.

Y es que la tesitura de una voz va cambiando según las etapas de vida de los individuos y según su género. "El tono medio de una voz es diferente según se trate de un niño, una mujer o un hombre." (CORNUT, 1985: 57).

A lo largo del tiempo, y gracias a las huellas históricas culturales de la humanidad es que hemos ido asignando esos valores a las voces.

Existen muchas maneras de clasificar las voces según el tono o frecuencia de su voz. La más común es la que se utiliza en la música para la voz cantada. Pero nosotros utilizaremos sólo cuatro categorías para la voz hablada:

- Agudo (500 Hz). Tono de una voz femenina muy aguda.
- Medio Alto (240 Hz). Tono de una voz femenina normal.
- Medio Bajo (120 Hz). Tono de una voz masculina normal.
- Grave (70 Hz). Tono de una voz masculina grave.

Finalmente, debemos hacer la misma distinción que con la intensidad y separar el tono de una voz que comunica información acerca del emisor, de la utilización y manejo del

tono por parte del emisor para evocar, transmitir y representar información externa a él mismo.

Dentro de las variaciones tonales que hacemos de nuestra propia voz al momento de expresarnos, muchas van relacionadas a la expresión de las emociones. Sean aprendidas culturalmente o por las reacciones fisiológicas de nuestro cuerpo cuando experimentamos alguna emoción, el tono de nuestra voz varía: cuando estamos tristes tendemos a bajar el tono de la voz y la aumentamos en actitudes alegres o vitales.

Por otro lado, "los tonos ascendentes se relacionan, a su vez, con objetos que suben en el espacio (...) cuando el tono de voz asciende haciéndose progresivamente más agudo hay una clara percepción de lo que se describe está subiendo. Cuando la construcción tonal es inversa, es decir, si pasamos de emitir tonos agudos a bajar progresivamente hasta los más graves de nuestra tesitura, la percepción es de que el objeto al que hacemos referencia está bajando". (RODRÍGUEZ BRAVO, 1984: 112).

4.3 Duración

Como toda variable del sonido, la *duración* en la voz es simplemente in adherente a ella, y es determinante en el tipo de percepciones que genera en los receptores: el alargamiento de un fonema, de una pausa al hablar, de la velocidad de un enunciado, etc. Su medición aproximada es en segundos y cualquier método técnico que ayude a analizar acústicamente la voz (oscilograma, sonograma y espectograma) toman como base el *tiempo* o *duración* del sonido.

4.4 Timbre

El concepto de timbre ha sido muy difícil de definir por parte de los investigadores. Muchos, se limitan a decir que se trata del sonido particular que emite cada persona con su voz de acuerdo a sus características fisiológicas, haciéndola única e irrepetible.

Y es que "el timbre de la voz, según se analiza al salir de la boca, es el resultado de la transformación y modelación del sonido laríngeo por las cavidades de resonancia" (CORNU, 1985: 59).

El timbre de la voz es una onda sonora compleja:

Formada en la laringe, pasando en su recorrido a las cavidades supraglóticas, las que actúan como filtros que

dejan pasar solamente las frecuencias que coinciden con las cavidades de resonancia. Este conjunto formado por el tono fundamental, más los armónicos, constituyen el timbre del sonido. Este timbre del sonido está directamente relacionado con las condiciones específicas de cada individuo. Sus cuerdas vocales, su aparato resonador, su constitución anatómica, su situación anímica, en determinado momento, va a darle carácter único al sonido de cada persona, y distingue a una de otra. (CASTARLENAS, 1995: 104).

Cuando percibimos el timbre de una voz nos da la sensación de que escuchamos un sólo sonido pero en realidad es el resultado de muchas otras ondas sonoras superpuestas. Este fenómeno perceptivo no es exclusivo de la voz humana sino de cualquier sonido compuesto. Ciertamente, las "vibraciones que tienen mayor amplitud son las que van a predominar sobre las elongaciones características de la onda resultante; pero no por ello las vibraciones que tienen una intensidad menor (menor amplitud) van a dejar de influir sobre el sonido final". (RODRÍGUEZ BRAVO, 1984: 68).

Los sonidos secundarios que acompañan a la frecuencia fundamental se les conoce como armónicos cuando las ondas son múltiplos de la frecuencia fundamental y parciales cuando no lo son.

Sin embargo, con lo mencionado hasta el momento pareciera que el timbre es un sonido complejo inflexible de nuestra voz. Pero es justamente lo contrario, ya que si existe alguna variación de algunos de los resonadores de nuestro cuerpo el timbre sufrirá también cambios perceptibles.

El ejemplo más básico lo encontramos al momento de hablar, cuando el sonido de los fonemas varía por el cambio de posición de los labios, la lengua, etc., aún cuando la intensidad y el tono permanezcan invariables. Ese cambio perceptivo es una variación del timbre. Sin embargo, no debemos confundir el timbre que se produce en un instante temporal determinado (en la producción de un fonema), de la sensación global que se produce en el oído de todo el conjunto tímbrico al momento de hablar. La enorme capacidad perceptiva del sentido humano nos ayuda a identificar un fonema producido por distintas voces diferenciando una voz de otra y también a reconocer los distintos fonemas producidos por una misma persona.

Por otro lado, también en el nivel expresivo, existe una relación entre la manipulación del timbre y los estados de ánimo. Por ejemplo, la tensión de la mandíbula cuando

experimentamos miedo o enojo, repercuten en el sonido tímbrico. Al igual que la intensidad y el tono, esta manipulación se puede hacer de manera consciente cuando impostamos un tipo de emoción.

En la conformación de los estereotipos vocales, el timbre es sin duda una señal sonora de la voz que debe estudiarse.

Pero los resonadores del cuerpo humano, no sólo son los únicos que intervienen en la producción de la sensación tímbrica de una voz. Como cualquier otro sonido, la voz necesita de un medio para propagarse. Al momento en que se emite una voz, las características físicas del espacio donde se produce el sonido también funcionan como resonadores que influyen en su percepción. Las reflexiones acústicas del espacio añaden a la voz vibraciones secundarias que nos permiten distinguir si la voz es emitida desde un lugar cerrado o abierto así como reconstruir e imaginar la forma de ese espacio: si es desde una cueva, una habitación o un campo abierto.

Ángel Rodríguez Bravo (1998: 185, 186 y 187) hace una clasificación de estos niveles adicionales de información: la definición auditiva, la impresión espectral y la armonicidad.

Definición auditiva

Cuando escuchamos una voz que es emitida por el altavoz de un televisor, somos capaces de reconocer las variaciones sonoras provocadas por la manipulación técnica del medio, haciendo que la voz no suene exactamente igual a como si se produjera en el mismo espacio que nosotros.

Estamos hablando de

“La sensación de máximo grado de precisión, exactitud o detalle sonoro que percibe el oyente al escuchar atentamente un sonido”.

Una definición baja, es cuando en la emisión de una voz se produce una reducción o eliminación de muchas de sus frecuencias secundarias. Por ejemplo, la sensación que se produce de una voz cuando es escuchada desde un teléfono. Una definición baja está compuesta por una gama de frecuencias muy reducida, de los 300 Hz a los 3,000 Hz.

Una definición media, es cuando la gama de frecuencias audibles van de 160 a 6,300 Hz. Es la sensación que produce el sonido del cine de los años cuarenta y cincuenta.

Y una definición alta, es cuando todas las frecuencias audibles no sufren ninguna reducción y abarcan todas las que el ser humano es capaz de percibir, de 20 a 20,000 Hz. Da la sensación de gran exactitud.

Impresión espectral

“Es la sensación de diferente matiz auditivo que percibe un receptor, cada vez que escucha un sonido de idéntica composición de frecuencias, pero con distinta envolvente espectral. Entendiendo como envolvente espectral de un sonido la forma obtenida al trazar una línea que une todas las puntas de cresta de cualquiera de sus espectogramas posibles” (Op. cit.: 187).

La impresión espectral es la que nos ayuda a determinar el espacio físico desde donde se propaga el sonido de una voz.

Impresión Oscura: es cuando la intensidad de su zona de frecuencias bajas es proporcionalmente superior a la intensidad del resto de frecuencias. Es el matiz que adquiere una voz. Impresión Brillante: Es cuando la intensidad de su zona de frecuencias altas es proporcionalmente superior a la intensidad del resto de frecuencias. Es el matiz que adquiere una voz cuando se emite en una habitación de paredes lisas y duras. Impresión Mate: es cuando la intensidad de su zona de frecuencias medias es proporcionalmente superior a la intensidad del resto de frecuencias. Es el matiz que adquiere la voz cuando se emite en una pequeña sala de paredes acolchadas o enmoquetadas; por ejemplo, el locutorio de un estudio de radio. (Op. cit.: 188).

Como lo mencionamos anteriormente, la sensación de impresión espectral en la voz también se puede conseguir desde el resonador bucal.

Armonicidad

Por último, la armonicidad del timbre de una voz repercute en la percepción e interpretación que hacemos de ella.

La armonicidad es “el distinto grado de limpieza y agradabilidad que percibimos al escuchar un sonido compuesto, dependiendo de la relación que existe en su espectro entre armónicos y parciales. Cuanto mayor sea la gama de frecuencias organizada armónicamente, mayor será la sensación de limpieza y agradabilidad, es decir: mayor será la armonicidad”. (RODRÍGUEZ BRAVO, 1989: 188-189).

Se puede hablar de voces sucias, bastas o transparentes. En los medios de comunicación, las voces de villanos suelen tener una textura sucia. Por el contrario, los héroes suelen tener texturas transparentes.

El grado de armonicidad de una voz está estrechamente relacionada a la belleza o fealdad de una voz. A nivel perceptivo podemos decir que un sonido compuesto por sólo frecuencias parciales producen molestia en el sentido auditivo. El sonido de una licuadora eléctrica es un ejemplo de sonido sin armónicos.

5. ANÁLISIS DE LAS FORMAS ACÚSTICAS DE LA VOZ EN LA PRODUCCIÓN DEL LENGUAJE ORAL

*Te dejaré en paz, Susana.
Conforme vayas repitiendo las palabras que yo diga,
te irás quedando dormida.
Sentirás como si tú misma te arrullaras.
Y ya que te duermas nadie te despertará (...)
Nunca volverás a despertar".
Juan Rulfo, Pedro Páramo*

*"Y su voz era secreta, casi apagada, como si hablara consigo mismo".
Juan Rulfo, Pedro Páramo.*

En este capítulo, abordaremos los conceptos acústicos mencionados en el capítulo anterior: intensidad, tono, duración y timbre pero a partir de su articulación durante la producción del lenguaje oral.

Nos referimos a los conceptos lingüísticos de entonación, acento y ritmo. Asimismo, revisaremos los recursos paralingüísticos como los diferenciadores, los disfraces y los alternantes que son parte importante de nuestra comunicación oral.

5.1 Entonación

La *entonación* es considerada como "la música de las palabras". Y es el resultado perceptivo de las variaciones de tono, duración, intensidad y timbre.

Cumple diversas funciones:

Distingue preguntas de declaraciones, a la vez que refleja el estado de ánimo del hablante; señala el núcleo sintáctico de la frase pero también puede

realzar los fragmentos de interés en beneficio del hablante o del oyente; indica tanto la procedencia geográfica y social de los hablantes como el tipo de género discursivo. (ALCOBA, 2000: 115).

Norminanda Montoya dice: "Si el locutor dispone de numerosas herramientas léxicas, sintácticas y prosódicas para poner de relieve su argumentación y matizar su pensamiento, las estructuras entonativas y rítmicas vinculan sobradamente las actitudes y las emociones". (1998: 45).

La *entonación* está compuesta por las distintas pausas o inflexiones que hacemos al hablar y, como decíamos, son las variaciones tonales y los cambios de intensidad, en algunos casos de manera arbitraria y en otros no. Algunas disponen de reglas fonológicas, "a las que se agregan el gusto, el arte y la personalidad del hablante". (CASTARLENAS, 1995: 154). Otras responden a la intención o grado de emoción que se esté experimentando. "En el nivel del enunciado es la entonación el medio más importante de la expresión de la afectividad" (MONTROYA, 1998: 45).

Tomás Navarro (1948: 11) dice que "la unidad melódica es la porción más corta de discurso con sentido propio y forma entonativa específica". Pero volviendo a citar a Santiago Alcoba, este autor dice que "hablamos de grupo de entonación para referirnos a aquellos fragmentos del habla que no necesariamente están delimitados por silencios y, sin embargo, reciben una entonación determinada" (ALCOBA, 2000: citando a QUILIS, 1981).

Según la terminología de Tomás Navarro, las unidades melódicas se denominan tonemas y considera cinco tipos: anticadencia ↑, cadencia ↓, suspensión →, semicadencia ↓ y semianticadencia ↑. (NAVARRO, 1948: 12).

Como podemos ver, la entonación empleada en un discurso oral es de mayor relevancia al momento de ser interpretada. A través de la entonación, el hablante puede indicar intenciones a través de las formas de enunciación e interrogación. Expresa emociones que

"determinan y a veces deciden la valoración del sentido e intención de las palabras. En determinados casos el objeto principal de las palabras es la expresión de un deseo. La declaración aparece bajo la modalidad de mandato o ruego, los cuales podrán ir envueltos, además, en sentimientos diferentes. Y por último, la entonación muestra la habitual del habla de cada país y hasta el sello que imprimen en un individuo su actitud o

profesión habitual y su propia manera de ser " (NAVARRO, 1974: 10).

5.2 Acento

En la producción del lenguaje verbal se le denomina *acento* cuando se pone en relieve una sílaba con respecto a las otras sílabas de una palabra. También se puede enfatizar una palabra completa del resto que conforman un enunciado. Y por último, se puede remarcar un fragmento completo del discurso para realzar su interés de lo demás. (ALCOBA, 2000: 93-99).

El acento se puede conseguir utilizando recursos fonéticos: mediante la intensidad o volumen, variando el tono y a través de la duración al pronunciar una sílaba, una palabra o un fragmento. También se da el caso que la combinación o utilización de dos o de los tres parámetros se den de manera simultánea. (MONTROYA, 1998: 46).

5.3 Ritmo o velocidad

Se dice que el ritmo es la "sucesión periódica de elementos marcados y no marcados por algún rasgo contrastivo en intervalos regulares de tiempo (...) Además de los cambios acentuales, las modificaciones de intensidad y de velocidad de elocución crean efectos rítmicos marcados". (ALCOBA, 2000: 109).

Con el ritmo de la voz estamos expresando nuestra relación con el tiempo. Es decir, estamos indicando cómo organizamos nuestro discurso en el tiempo (...) Para Paul Fraisse "un ritmo se caracteriza por la unidad perceptiva que puede ser marcada de diferentes maneras, pausa o repetición de elementos análogos por su duración, su intensidad, su altura ". (Citado por MONTROYA, 1998: 47).

A través del ritmo se puede diferenciar entre una lectura oralizada del de la palabra espontánea.

Las pausas utilizadas se pueden deber a "la necesidad fisiológica de respirar, o bien a razones de tipo lingüístico como señalar el fin de un enunciado completo, separar los miembros de una enumeración, variar el significado de una emisión" (MONTROYA, 2000: 48).

Por su parte, Ángel Rodríguez Bravo (1984: 139) menciona que el ritmo "es una estructura sonora muy próxima al

movimiento, es un reflejo del movimiento mismo en su sentido más amplio”.

Además del movimiento, el ritmo puede transportar información sobre sentimientos y emociones que experimenta el emisor de una voz. En situaciones de gran excitación emocional, como la euforia o el enojo, la velocidad de nuestra elocución puede ser más rápida de la que habitualmente manejamos.

5.4 Dicción

Hace referencia a la articulación de los sonidos vocálicos que “incide de manera determinante en la inteligibilidad y claridad de la cadena de habla resultante”. (ALCOBA, 2000: 73)

Las desviaciones articulatorias respecto a lo que cada uno tiene como estándar nos permiten distinguir el grupo social o geográfico del hablante (*información encuadrativa* de las señales vocales, ver capítulo IV)).

5.5 Recursos Paralingüísticos

Como hemos visto a lo largo del estudio, cuando hablamos producimos una amplia gama de matices comunicativos que pueden reforzar o modificar el discurso verbal. Y es que también producimos una amplia cantidad de sonidos difíciles de clasificar o recopilar y que no son tomados en cuenta por los lingüistas, pero que en la conversación cotidiana utilizamos como elementos expresivos perfectamente reconocibles por los demás.

Podemos distinguir tres tipos de sonidos paralingüísticos y que no han sido tratados en los conceptos anteriores: los diferenciadores, los disfraces y los alternantes.

Diferenciadores

Los diferenciadores son todos aquellos sonidos producidos por los estados fisiológicos o las reacciones emocionales de los hablantes y que tienen un potencial expresivo, muchas veces determinante en la percepción global de lo que se está diciendo.

Los diferenciadores son:

- Reír
- Llorar

- Gritar
- Suspirar
- Jadear
- Bostezar
- Toser
- Carraspear
- Escupir
- Eructar
- Hipar
- estornudar

(CÁCERES, 2003: 162).

Si escuchamos a una persona que habla riendo, seguramente nuestra percepción será distinta de si lo hace bostezando. Al mismo tiempo, si la risa es muy estridente y parecen más gritos que risas, la percepción de persona contenta o alegre puede pasar a ser molesta y negativa.

Disfraces

Son todas aquellas imposturas, fingimientos o engaños que pueden parecer verdaderos o no, según el contexto y el grado de verosimilitud que alcancen o se quieran alcanzar.

Cuando conversamos, recurrimos muchas veces a narrar sucesos personales o de otros (comentar una película, etc). e impostamos nuestra voz para imitar la voz de la persona que estamos recordando en ese momento. Muchas veces exageramos algún rasgo sonoro de aquello que a nuestro parecer se asemeja al producido por la otra persona o que simplemente, percibimos que contenía ese grado de emoción o intención comunicativa.

Una voz susurrada, tipo confidencial es otro ejemplo de disfraz.

Alternantes

“Son emisiones no verbales, de carácter voluntario o involuntario, conscientes o no, más o menos articuladas y para las que no siempre tenemos un término preciso: sonidos como sisear, clics, ronquidos, gruñidos, gemidos, chisteos, bisbiseos, resoplidos, carraspeos, chasquidos, soplos, jadeos, sonidos de vacilación, etc.”. (CÁCERES, 2003: 162).

Los alternantes sirven para “expresar estados emocionales (*¡bah!* de incredulidad, *bfff!* de alivio) y sentimientos (*¡uff!* de cansancio, *¡yupiiii!* de felicidad), para regular el flujo de la conversación o para representar acústicamente algo. (Op. cit.: 164).

6. PROPUESTA DE UNA TAXONOMÍA DE LA VOZ

*Con su boca casi pegada a la oreja de ella
para no hablar fuerte,
encajaba secretamente cada una de sus palabras:
“Tengo la boca llena de tierra”.
Luego se detuvo.
Trató de ver si los labios de ella se movían.
Y los volvió a balbucir, aunque sin dejar salir ningún sonido”.
Juan Rulfo, Pedro Páramo*

Es momento de articular todos los conceptos antes mencionados. Tomando en cuenta que las señales vocales transmiten información, por lo menos en los 11 niveles detallados anteriormente, nuestra *Taxonomía de la voz* debe considerarlos para no perder ninguno de los posibles registros sonoros que los oyentes interpretamos y utilizamos para conformar estereotipos vocales en nuestra mente.

6.1 Formas sonoras simples y complejas

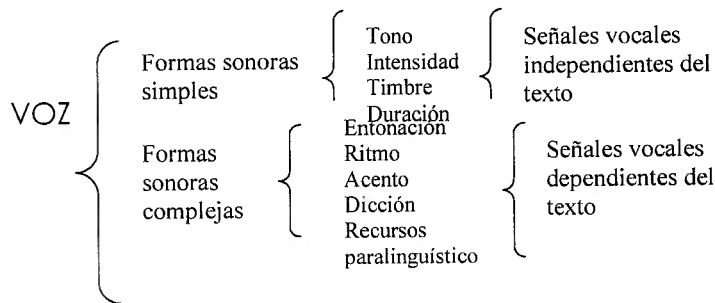
Nuestro clasificador de la voz parte del reconocimiento acústico de las formas sonoras simples y las formas sonoras complejas.

Como habíamos mencionado, las formas sonoras simples se configuran en un primer nivel de percepción e interpretación de las cuatro características que conforman el sonido: de la intensidad, tono, timbre y duración antes de que medie la educación. Su interpretación depende de las características físicas del hablante, del entorno acústico donde se produzca el sonido y del medio por donde viaje. En el caso de la voz humana, partimos de que las formas sonoras simples son señales vocales independientes del texto, es decir, de lenguaje verbal, ya que éste requiere de un aprendizaje especializado más allá del aprendizaje obtenido por la propia experiencia auditiva.

Por su parte, las formas sonoras complejas son interpretaciones de las formas entonativas, rítmicas y tímbricas producidas al momento de hablar, transmitiendo señales vocales

que guardan una dependencia con el texto, es decir, con lo que se dice o se quiere decir.

Como podemos observar, las formas sonoras simples y las complejas están estrechamente relacionadas a la dependencia o independencia del lenguaje verbal al momento en que es emitida y percibida una voz. Es por esto, que proponemos relacionar las formas sonoras simples a las señales vocales independientes del texto, y a las formas sonoras complejas con las señales vocales dependientes del texto, aún cuando estas señales actúen perceptivamente de manera simultánea. Sin embargo, como hemos visto, ejercen diferentes influencias en el receptor. Es decir, nos referimos por un lado, al sonido particular de la voz producido por cada persona y por otro, a los diferentes sonidos de la voz que se producen al momento de articular mensajes a través del lenguaje oral.



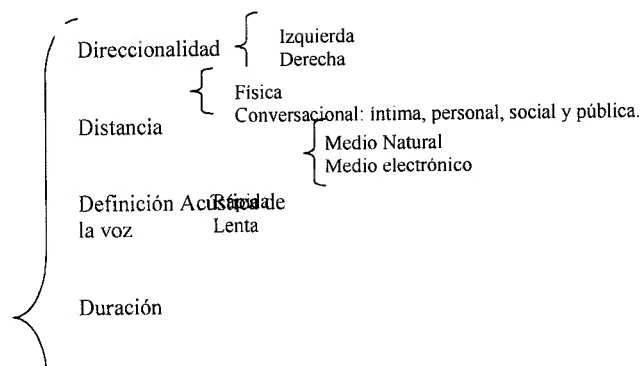
6.2 Señales vocales independientes del texto

Cuando decimos "independencia del texto" nos estamos refiriendo a todas las señales sonoras que transmiten información al oyente independientemente del mensaje verbal o lingüístico y que están vinculadas a la manera en como percibimos a través de nuestro sentido auditivo.

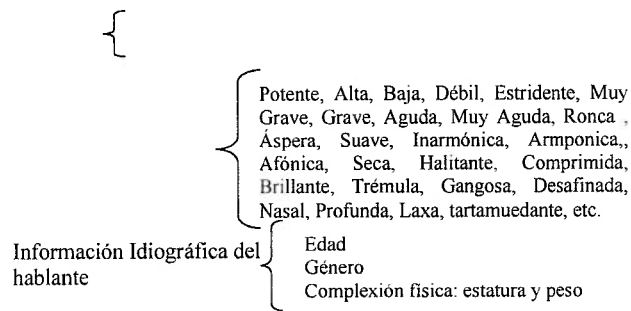
Todas las señales vocales independientes del texto se forman perceptivamente a partir de las características fisiológicas de cada hablante (tamaño de las cuerdas vocales, caja toráxica, salud física, edad, sexo, etc).

Dicho de otra manera, la producción de formas sonoras simples dependen de las características del hablante y la interpretación de esas señales dependen del sistema perceptivo del oyente, que no requieren de un aprendizaje especializado.

Podemos clasificarlas de la siguiente manera:



Señales vocales
Independientes del
texto



- Direccionalidad

Nos estamos refiriendo a una percepción básica auditiva del sonido que nos orienta sobre la dirección de donde proviene la voz a nuestros oídos, orientación izquierda-derecha, aunque cabe mencionar que el sistema auditivo es "capaz de procesar información circular en un radio de 360°" (RODRÍGUEZ BRAVO, 1998: 211).

La intensidad del sonido de una voz y el tiempo que tarda en ser percibida por un oído o por otro, nos da la sensación de direccionalidad.

- Distancia

Al igual que la direccionalidad, la distancia de la emisión de una voz con respecto al receptor depende de la intensidad y el tiempo en que el sonido es percibido por el sistema auditivo.

En la percepción sonora podemos distinguir entre la distancia física de la voz emitida con respecto al receptor de la distancia conversacional.

Por ejemplo, si una voz es emitida desde un aparato de radio, la percepción de distancia de esa voz dependerá del volumen al que se esté escuchando o de la distancia física del receptor con respecto del aparato.

En cambio, en una conversación, aún cuando sea por teléfono o se trate de dos actores desde la televisión, la voz del emisor puede marcar una distancia física o emotiva con la manipulación de su intensidad. La distancia conversacional puede marcarse según el contexto: íntimo, personal, social o público.

- Definición acústica de la voz

Como ya dijimos, el timbre puede darnos información sobre la definición auditiva, la impresión espectral y la armonicidad de una voz. Gracias a las dos primeras es que podemos distinguir entre una voz que es emitida desde un medio natural de la que

lo es por medio de un aparato electrónico o artificial: un altavoz, un teléfono, un CD, etc.

- Duración

En este esquema que proponemos, la rapidez o lentitud en la que es emitida una voz no guarda una dependencia del texto aun cuando sólo se distinga a través del habla. Es decir, depende de las características personales del emisor como la personalidad, el estado físico o la edad. Si bien es cierto que una persona al experimentar una emoción de euforia puede acelerar la velocidad de elocución también lo es que cada persona maneja una velocidad promedio según sus características propias.

Llamaremos ritmo cuando la duración guarda una dependencia directa con el texto, por ejemplo una voz melódica o plana al momento de hablar.

- Cualidad acústica

La cualidad acústica de una voz resalta aspectos más particulares del sonido vocal: por ejemplo, dos voces femeninas de la misma edad y complexión física pueden presentar características específicas muy diferentes: una más aguda que otra; una más débil o armoniosa y otra más fuerte o estridente.

Con la intención de nombrar con más especificidad las *cualidades acústicas* de la voz, proponemos una lista basándonos en conceptos propuestos por Fernando Poyatos (1994, 51), así como en adjetivos utilizados en la producción audiovisual y publicitaria y en términos que empleamos en la vida cotidiana cuando nos referimos a una voz, ofreciendo algunos sinónimos o descripciones para una mejor comprensión de la característica a la que nos estamos refiriendo.

- Fuerte: gritona
- Alta: medianamente fuerte
- Baja: medianamente baja en volumen o intensidad .
- Débil: muy baja en volumen o intensidad.
- Ronca: áspera.
- Seca: pastosa, sin saliva.
- Afónica: enferma
- Grave: tono bajo
- Muy grave: tono extremadamente bajo
- Aguda: tono alto

- Muy aguda: tono extremadamente alto, aflautado.
- Armónica: bonita, agraciada, agraciada de cierta proporción y belleza
- Estridente: fea, disonante, sucia, deslustrada, repelente, horrible.
- Suave: fina, delicada, ligera, tersa, uniforme, tranquila.
- Comprimida: oprimida, prensada, contenida, controlada, frenada, ahogada.
- Metálica: vibrante, brillante
- Trémula: temblorosa, temerosa.
- Gangosa: nasal.
- Profunda: baja, honda, reflexiva.
- Laxa: distendida, floja, relajada.

Aunque esta lista no es exhaustiva, creemos que resalta las características acústicas de más fácil identificación. Estas cualidades de la voz también son interpretadas según valores aprendidos socialmente.

- Expresión idiográfica

El término expresión idiográfica es propuesto por Rodríguez Bravo y está configurada "por aquellos rasgos acústicos de la voz que aportan al oyente información sobre el aspecto fisiológico del emisor" (2002: 163).

Al momento en que se escucha una voz, no cabe la menor duda de que es posible reconocer si se trata de la voz de un hombre o de una mujer, y si es de una persona joven, de un anciano o de un niño.

Nos atrevemos a poner estas señales vocales como formas simples independientes del texto ya que su emisión sólo depende de las características físicas del hablante y su interpretación no requiere una educación especializada más allá de la experiencia auditiva, aún cuando cada cultura pueda asignarles diferentes valores.

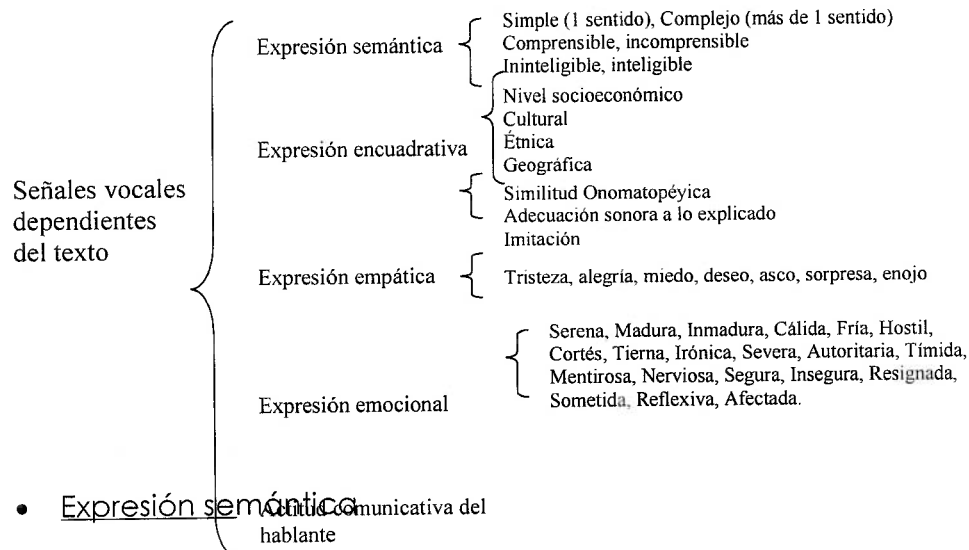
6.3 Señales vocales dependientes del texto

Cuando hablamos, producimos una señal acústica compleja que no es exactamente la misma aún cuando se trate de la misma palabra. Así como hemos venido mencionando que las formas sonoras simples y las formas sonoras complejas de una voz son percibidas e interpretadas como una sola unidad,

igualmente las señales vocales dependientes del texto están directamente relacionadas con las que no lo son, de hecho, muchas de las veces dependen de las características físicas del hablante y de su entorno físico y acústico. Sin embargo, el lenguaje oral, como un código complejo y especializado, pone en marcha una gran cantidad de señales complejas. Los diferentes significados que puede adquirir nuestro discurso los producimos jugando con la calidad de la voz: el volumen, el tono, la velocidad, la fluidez o perturbaciones en el habla (vacilaciones, falsos comienzos, repeticiones), diferenciadores, disfraces y alternantes.

Estas señales guardan una relación directa con la educación especializada y compleja que implica el manejo de una lengua así como de valores aprendidos culturalmente.

De esta manera, las señales vocales dependientes del texto pueden clasificarse de la siguiente manera:



El habla está basada en la variación, de ahí que podamos distinguir uno o más de un sentido en las palabras utilizadas y en las interpretaciones que de ellas se haga el receptor.

Por otro lado, un mensaje hablado puede resultarnos comprensible o incomprensible según nuestros propios referentes acerca del tema, y también, inteligibles o ininteligibles según las formas entonativas, rítmicas o tímbricas.

- Expresión encuadrativa

El nivel económico, cultural, el origen étnico y geográfico tanto del hablante como del receptor son importantes en la manera de interpretación y tipificación de una voz. Estas formas sonoras complejas sólo pueden ser aprendidas con el lenguaje y detectadas al momento de hablar.

- Expresión empática

La expresión empática puede lograrse por: similitud onomatopéyica de los sonidos descritos por el hablante y la fuente sonora que la produce, por la adecuación sonora de lo explicado y por imitación.

- Expresión emocional

Como ya vimos en páginas anteriores, la combinación de emociones hace que sea muy difícil de clasificarlas, así es que nos limitaremos a listar 8 emociones básicas con algunos sinónimos o grados de intensidad que nos ayuden a ejemplificar nuestra taxonomía:

- Triste: afligida, apenada, apesadumbrada, mustia, desconsolada.
- Alegre: contenta, jubilosa, festiva, bromista, entusiasta, divertida.
- Miedosa: temerosa, aterrada, pavorosa, espantada, horrorizada, asustada.
- Asqueada: repulsiva, antipática, harta.
- Molesta: enojada, disgustada, irritada, encolerizada, enfurecida, enfadada, con rabia.
- Sorprendida: asombrada, inesperada, extrañada, impresionada, sobrecogida.
- Deseosa: anhelante, apasionada, interesada en, ansiosa.
- Serena: tranquila, sosegada, calmada, moderada.

- Actitud comunicativa del hablante

Utilizamos nuestra voz con una gama infinita de matices comunicativos: voces que mandan, ordenan, imploran, hieren, sugieren, aman... y que refuerzan el valor de las palabras.

Los términos que nos clarifican estas intenciones pueden ser:

- Madura: Prudente, reflexiva.
- Inmadura: inestable, precoz, verde.
- Cálida: amistosa, próxima.
- Fría: Indiferente, sin gracia.
- Hostil: fastidiosa, antipática.
- Amable: cordial, afectuosa, cortés.
- Tierna: sensible, afable, cariñoso.
- Autoritaria: déspota, tirana.
- Irónica: Burlona, mordaz, sarcástica.
- Severa: rigurosa, inflexible, implacable, intransigente.
- Tímida: miedosa, encogida.
- Convinciente: persuasiva.
- Mentirosa: engañosa, embustera, falaz
- Nerviosa: inquieta, ansiosa, excitada.
- Calmada: sosegada, tranquila.

6.4 Propuesta conceptual de estereotipos vocales

Una vez elaborada nuestra *taxonomía de la voz*, estamos en condiciones de conformar conceptualmente estereotipos de voz. Sin embargo, nos enfrentamos a que existe la necesidad de "construir una biblioteca de voces, y de conceptos que faciliten la descripción de cualquier voz a través de sus rasgos más relevantes. En muchas ocasiones los oyentes no especializados quieren describir algún rasgo sonoro que recuerdan con claridad, pero su mayor dificultad radica en no poseer un repertorio de términos, un vocabulario claro, explícito y compartido" (ROGRÍGUEZ; LÁZARO; MONTOYA; BLANCO; BERNADAS; TENA; LONGHI; OLIVER;, 2003: 27).

Por estas razones es necesario construir conceptualmente grandes estereotipos vocales que reúnan más de dos señales vocales clasificadas en la taxonomía y asignarles un nuevo nombre.

Es decir, la taxonomía sólo hace referencia a las características particulares de una voz, ahora es necesario proponer algunos términos que faciliten la localización de las generalidades, es decir, de estereotipos vocales, los cuales se construyen de la percepción global de todas las características

particulares de una voz, otorgándoles sólo a unas algún valor positivo o negativo.

Por ejemplo, aunque la *tristeza* en una voz pueda ser una característica lo suficientemente preponderante como para considerarla un estereotipo, al momento de asignarle algún valor, una voz triste puede ser a su vez considerada una voz agradable o desagradable, tierna o insegura según la combinación con otras características particulares. Es por esto que las *características particulares* de una voz no son *estereotipos* y los términos que les hemos asignado no son lo suficientemente generales para nombrarlos.

Recordemos que según el estudio de los estereotipos se trata de un mecanismo de simplificación que nos ayuda a entender cómo nos formamos una idea del otro, es decir, siempre hay un sujeto emisor real o creado por un medio de comunicación al cual se le atribuyen los estereotipos. Es por esto que los conceptos que proponemos a continuación describen más las actitudes y personalidad del hablante, que a la voz en sí, y son a lo que Norminanda Montoya define como “estereotipos vocales de carácter”.

Proponemos 18 conceptos de *Estereotipos vocales de carácter* obtenidos de una primera observación y clasificación realizada por quien ahora escribe, poniendo mi propia experiencia como escucha.

Cuando escuchamos una voz, la persona que habla nos puede parecer:

1. Dulce: grata, afectuosa, apacible.
2. Agresiva: dura, severa, gruñona.
3. Simpática: atractiva, agradable, graciosa, que produce empatía.
4. Apática: aburrida, impasible.
5. Sumisa: obediente, subyugada, rendida, subordinada.
6. Dominante: represiva, fuerte y superior.
7. Sensual: incitante, excitante.
8. Repulsiva: repugnante, despreciable.
9. Anñada: de actitud y comportamiento semejante a los niños.

10. Afeminada: de actitud y comportamiento que resalta las cualidades positivas y negativas del sexo femenino.
11. Varonil: de actitud y comportamiento que resalta las cualidades positivas y negativas relativas al sexo masculino.
12. Jovial: de actitud y comportamiento juvenil.
13. Maternal: de actitud y comportamiento similar a una madre tradicional.
14. Paternalista: de actitud y comportamiento similar a un padre tradicional.
15. Elegante: distinguida, refinada y de buen gusto .
16. Tosca: poco refinada, sin pulimento, inculta.
17. Inteligente: instruida, enterada, preparada, sabia.
18. Tonta: pesada, molesta, que le falta intelecto.

Estos conceptos reúnen más de dos señales vocales de nuestra taxonomía pero intentando que cada uno fuera independiente del otro.

Por ejemplo:

Una voz *Dulce* puede estar conformada por los rasgos: armónico, cálido y sereno.

Y aunque para algunos, una voz *maternal* pueda tener rasgos similares a los de una voz *dulce*, quizá la característica *madurez* o *seguridad* puede determinar su conformación y diferencia.

Los 18 conceptos o adjetivos de estereotipos vocales los llamaremos *conceptos primarios* y a las señales particulares de la voz que los conforman, *conceptos secundarios, terciarios...etc.*

Esta propuesta está basada en una de nuestras hipótesis que presentamos al inicio, que dice que *los estereotipos vocales no se construyen de la suma de particularidades sonoras de cada voz, sino que, algunos cuantos rasgos sonoros fundamentales, son suficientes para que se conformen.* La construcción de esta idea se fundamenta en las investigaciones realizadas por Asch (FISCHER, 1990: 100) quien encontró que sólo algunos rasgos de una persona son suficientes para producir una *primera impresión* del otro: "sólo algunos rasgos desempeñan el papel de organizadores centrales" (Op. cit.; 101).

En el libro *Psicología Social*, G. N. Fischer resume los experimentos y conclusiones de Asch de la siguiente manera:

En un experimento ya clásico, Asch presentó a dos grupos de estudiantes una lista de rasgos característicos que venían a describir a una determinada persona: inteligente, hábil, trabajador, fogoso, decidido, práctico, prudente. En uno de los grupos <<fogoso>> fue remplazado por <<frío>>. Asch consideró los términos <<fogoso>> y <<frío>> como rasgos sumamente importantes, a partir de los cuales, los individuos podían hacer descripciones complejas de la personalidad de otra persona. Tras haber sido informados de esta serie de adjetivos, los estudiantes debían comunicar sus impresiones relativas a ese hipotético individuo y a continuación, juzgarle a partir de otras características elegidas por ellos.

Las conclusiones fueron:

1. En primer lugar, los estudiantes organizaron los distintos rasgos en un todo coherente.
2. A continuación se comprobó en la descripción de la persona una marcada diferencia, según que las listas incluyesen el término <<fogoso>> o el término <<frío>>.
3. Por último, se observó que los términos <<fogoso>> y <<frío>> ejercían una particular influencia sobre la percepción de conjunto.

La propuesta de nombrar *conceptos primarios, secundarios, terciarios... etc.*, es para facilitarnos la comprensión de cómo construimos los estereotipos, ya que implica una valoración: implica asignarles a unas características particulares de la voz, diferentes niveles de prioridad frente a otras que no son suficientemente importantes.

7. DISEÑO METODOLÓGICO PARA UNA CONTRSTACIÓN EMPÍRICA

*Tú y yo allí, rezando rezos interminables,
sin que ella oyera nada,
sin que tú y yo oyéramos nada,
todo perdido en la sonoridad del viento".
Juan Rulfo, Pedro Páramo*

Los 18 conceptos de estereotipos vocales propuestos, deben ser sometidos a la evaluación de personas ajenas a la investigación que corroboren si los términos empleados son los adecuados y suficientes para nombrar todos los posibles estereotipos que construyen en su mente al momento de escuchar diferentes voces.

El diseño del estudio empírico busca los siguientes objetivos:

1. Corroborar las características particulares de la voz propuestas en la Taxonomía .
2. Corroborar los conceptos de estereotipos vocales.
3. Encontrar a qué estereotipo pertenece una sola voz, es decir, cómo se percibe de manera colectiva y qué rasgos sonoros son los relevantes para su conformación.
4. Encontrar en distintas voces que hayan sido clasificadas como pertenecientes a un mismo estereotipo vocal, qué rasgos sonoros son similares y cuáles tienen más relevancia que otros.

Para cumplir con estos objetivos debemos antes resaltar un aspecto importante: el tipo de estudio que realizaremos.

Debido a que los estereotipos se construyen según las características culturales y sociales propias de cada región, no podemos hacer una generalización al momento de localizar estereotipos vocales, es decir, que los valores y patrones que una colectividad le asigna a una voz, pueden ser muy diferentes según la región o país en el que se realice el estudio. Sin embargo, nos puede ayudar a comprender los mecanismos mentales de simplificación y clasificación que realiza cualquier ser humano al momento de escuchar una voz.

Lo que nos interesa es diseñar una metodología que pueda ser aplicada en cualquier parte y que sea capaz de encontrar los estereotipos propios del lugar.

Lo siguiente es definir las características de nuestra muestra y del corpus, así como el instrumento que emplearemos para recopilar la información.

7.1 Diseño de la muestra

Ya que nuestro estudio no pretende generalizar los estereotipos vocales, nuestro muestreo será no probabilístico (ROJAS, 1981: 165) pero que nos ayude a ilustrar el proceso de construcción de los mismos.

El primer criterio para seleccionar al grupo de gente que evaluará nuestros conceptos es que éste debe estar conformado por un número lo suficientemente amplio que nos permita detectar coincidencias perceptivas más allá del azar. De 50 personas en adelante.

El segundo es que deben pertenecer a un mismo grupo cultural y que, por lo menos, hablen la misma lengua.

El tercer criterio es que debe ser un grupo de fácil acceso para el investigador y que acepte someterse a las pruebas que requiere el estudio.

Bajo estas condiciones, 30 alumnos de la licenciatura de Producción Audiovisual de la Universidad de Guadalajara y 30 alumnos de la licenciatura de Ciencias de la Comunicación del Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO) en México, conformarán nuestra muestra.

7.2 Construcción del corpus

En el Marco Conceptual mencionábamos que una de las dificultades para conseguir resultados contundentes en las investigaciones sobre la voz está en la construcción de un corpus adecuado.

Para nuestro estudio, es necesario reunir un amplio número de voces grabadas y su selección requiere de un análisis cuidadoso de las variables que pueden afectar y determinar el alcance de nuestros objetivos.

Intentaremos rescatar experiencias similares a la nuestra, de tal manera que las voces que grabemos reúnan características específicas.

En primer lugar, retomaremos las observaciones publicadas en el artículo *El control objetivo de la implicación de los informantes en el estudio del habla emocionada*, en el cual los autores plantean algunas de las dificultades detectadas al

recurrir a voces de actores profesionales para la construcción de su corpus:

Las condiciones suelen ser: 1) *predominio del criterio profesional del actor*; 2) *saturación semántica de los investigadores*; 3) *posibilidad de que el actor informante finja sus interpretaciones emocionadas*. En esta situación lo que conseguimos, en realidad, es un corpus muy incompleto basado simplemente en que un actor consiga que sus actitudes emocionadas sean reconocibles por un grupo de investigadores, normalmente no entrenados en las técnicas de interpretación, que además, están saturados de escuchar infinidad de veces las mismas frases con sutiles cambios sonoros. Obviamente, si el corpus de referencia a partir del cual esperamos modelizar acústicamente las emociones no contiene aquello que buscamos, o solo contiene en parte, es difícil que podamos obtener de él modelos realmente representativos a partir de su análisis acústico. (RODRÍGUEZ; LÁZARO; MONTOYA; BLANCO, BERNADAS; OLIVER;, 1998: 25).

Este problema en concreto nos hizo descartar la utilización de actores profesionales para la grabaciones de sus voces, ya que, aunque no sólo trabajaremos las emociones expresadas a través de la voz, creemos que la problemática puede extenderse a cualquier señal vocal que deseemos analizar.

Entonces, las características específicas de nuestro corpus serán:

1. Voces de personas de diferente género y de diferentes edades.
2. Voces que hablen el idioma castellano y sin un texto predefinido. Es decir, que expresen distintas ideas, emociones y actitudes.
3. Voces de personas que hablen lo más parecido a una conversación normal.
4. Voces en las que no se distinga un estilo de locución característico a los profesionales que trabajan en los medios de comunicación: locutores, presentadores, analistas y/o comentaristas de radio y televisión.
5. Voces monologadas que expresen una idea coherente por lo menos durante 1 minuto y 30 segundos aproximadamente. Esta característica es porque según un estudio (RODRÍGUEZ; LÁZARO, MONTOYA, BLANCO,

BERNADAS, TENA, LONGHI, OLIVER;, 2003: 31), el tiempo ideal para que una voz pueda ser evaluada es de un mínimo de 45 segundos de duración.

Siguiendo estos parámetros, hemos conformado un banco sonoro de doscientas voces de hombres y mujeres de distintas edades, grabadas de dos programas televisivos en los que la gente asiste para platicar o denunciar sus experiencias. Es decir, de gente no profesional al medio.

Aunque es indiscutible que hablar frente a un público en un set o plató de televisión repercute en las señales vocales por el nerviosismo y falta de naturalidad, decidimos grabarlas de estos dos programas ya que las voces reunían otras características importantes que se acercan a los parámetros planteados: por un lado, un registro sonoro bastante limpio, es decir, con poco ruido externo; una amplia variedad en los temas hablados que nos dan la oportunidad de registrar casi todas las emociones básicas en diferentes niveles de intensidad, así como actitudes e intenciones comunicativas; la posibilidad de registrar personas de distintos géneros y edades: lo mismo encontramos jóvenes que ancianos, profesores que carpinteros, etc.; y por último, la posibilidad de eliminar la voz del entrevistador por medio de una edición posterior, para evitar confusiones o distracciones al momento en que las voces deban ser calificadas.

7.3 Diseño del instrumento para recopilar la información

El instrumento con el que recopilaremos nuestra información será a través de un questionario que los jueces deberán contestar cada vez que escuchen una voz.

Su diseño debe ser claro, sencillo y de fácil manejo para los jueces o escuchas. Asimismo, debe estar pensado para someter a prueba las hipótesis generales de nuestra investigación, así como para lograr los objetivos planteados para el estudio empírico.

La propuesta que diseñamos es la siguiente:

7.4 Diseño del cuestionario. Primera propuesta

DATOS DEL ENCUESTADO

1. Edad: _____.
2. Sexo: F M

Con la respuesta de estos datos por parte de los encuestados podemos plantearnos realizar algunos cruces de información. Si los resultados arrojan datos significativos según la edad y/o el género, el análisis puede enriquecerse.

CUESTIONARIO

INDICACIÓN 1

1. En el primer recuadro, escriba el número de voz que indique el encuestador. Posteriormente, de los conceptos que aparecen en el ANEXO elija 5 opciones en orden de importancia que considere pertenecen o caracterizan a la voz escuchada:

voz

A través de esta indicación, intentamos encontrar las principales señales sonoras o *conceptos secundarios y terciarios* que el escucha percibe como importantes de cada una de las voces.

El ANEXO será una lista con todos los conceptos propuestos en la Taxonomía de la voz pero sin resaltar a qué tipo de señal pertenecen según nuestra clasificación. Cada concepto tendrá los sinónimos o descripciones que ya hemos señalado en el

capítulo anterior y que ayudarán en la comprensión de los términos.

INDICACIÓN 2

2. Describa tres aspectos en orden de importancia que le llamaron más la atención de la voz grabada:

Por ejemplo:

1. *La voz le temblaba mucho.*
2. *Lo que decía era muy triste*
3. *Hablaba muy lentamente*

VOZ 1. _____.

2. _____.

3. _____.

Estas descripciones nos darán a conocer si las características predominantes de la voz escuchada pertenecen a las señales sonoras independientes del texto o a las que son dependientes; si "lo que se dice" influye más que "el cómo se dice".

Esta pregunta cualitativa será sistematizada posteriormente, reasignándoles a las respuestas un nuevo valor del 1 al 3 según las señales sonoras de la voz a las que hagan referencia.

- 1 para las señales acústicas,
- 2 para las señales emotivas y
- 3 para las señales discursivas.

INDICACIÓN 3

3. Si tuviera que definir en una sola palabra a la persona escuchada, cuál de los siguientes conceptos utilizaría:

1. **Dulce:** grata, afectuosa, apacible.
2. **Agresiva:** dura, severa, gruñona.
3. **Simpática:** gtractiva, agradable, graciosa, que produce empatía.

4. **Apática:** aburrida, impasible.
5. **Sumisa:** obediente, subyugada, rendida, subordinada.
6. **Dominante:** represiva, fuerte y superior.
7. **Sensual:** incitante, excitante.
8. **Repulsiva:** repugnante, despreciable.
9. **Aniñada:** de actitud y comportamiento semejante a los niños.
10. **Afeminada:** de actitud y comportamiento que resalta las cualidades positivas y negativas relativas al sexo femenino.
11. **Varonil:** de actitud y comportamiento que resalta las cualidades positivas y negativas relativas al sexo masculino.
12. **Jovial:** de actitud y comportamiento juvenil.
13. **Maternal:** de actitud y comportamiento similar a una madre tradicional.
14. **Paternalista:** de actitud y comportamiento similar a un padre tradicional.
15. **Elegante:** distinguida, refinada y de buen gusto.
16. **Tosca:** poco refinada, sin pulimento, inculta.
17. **Inteligente:** instruida, enterada, preparada, sabia.
18. **Tonta:** pesada, molesta, que le falta intelecto.

OTRO

VOZ

Otro: _____.

Esta última indicación está pensada específicamente para medir los conceptos propuestos como estereotipos o *conceptos primarios*. Aquí se hace especial énfasis a que juzguen a la **persona escuchada** y no a la **voz escuchada**, debido a que el tipo de adjetivos hacen referencia más a estereotipos de

carácter o personalidad que a características físico-acústicas de la voz.

ANEXO 1

1. FUERTE: (gritona).
2. ALTA: (medianamente fuerte en intensidad).
3. BAJA: (medianamente baja en intensidad)
4. DÉBIL: (muy baja en intensidad).
5. MUY GRAVE: (tono extremadamente bajo).
6. GRAVE: (tono bajo).
7. AGUDA: (tono alto).
8. MUY AGUDA: (tono extremadamente alto, aflautada).
9. RONCA: (muy grave pero áspera).
10. AFÓNICA: (enferma).
11. SECA: (pastosa, sin saliva).
12. ARMÓNICA: (bonita, agraciada, agraciada de cierta proporción y belleza).
13. ESTRIDENTE: (fea, confusa, sucia, deslustrada, repelente, horrible).
14. SUAVE: (fina, delicada, ligera, tersa, uniforme, tranquila).
15. COMPRIMIDA: oprimida, prensada, contenida, controlada, frenada, ahogada).
16. METÁLICA: (vibrante, brillante)
17. TRÉMULA: (temblorosa, temerosa).
18. GANGOSA: (nasal, ininteligible).
19. PROFUNDA: (baja, honda, reflexiva).
20. LAXA: distendida, floja, relajada.
21. TRISTE: afligida.
22. NOSTÁLGICA: (apenada).
23. DEPRIMIDA: (Desconsolada, apesadumbrada).
24. ALEGRE: contenta.
25. DIVERTIDA: (entusiasta, festiva, bromista).
26. EUFÓRICA: (jubilosa).
27. ASUSTADA: (temerosa).
28. ESPANTADA: (miedosa).
29. ATERRADA: (horrorizada).
30. ANTIPÁTICA: (de rechazo).
31. FASTIDIADA: (harta).
32. ASQUEADA: (repulsiva).
33. MOLESTA: (enojada).
34. IRRITADA: (disgustada).
35. ENFURECIDA: (encolerizada).
36. ASOMBRADA: (extrañada).
37. SORPRENDIDA: (pasmada).

38. DESCONCERTADA: (sobrecogida, inesperada).
39. DESEOSA: (interesada en,).
40. AFECTUOSA: (cariñosa).
41. APASIONADA: (ansiosa).
42. SERENA: (tranquila, sosegada, calmada, moderada).
43. MADURA: (Prudente, reflexivo).
44. INMADURA: (inestable, precoz, verde).
45. CÁLIDA: (amistosa, próxima)
46. FRÍA: (Indiferente, sin gracia).
47. HOSTIL: (fastidiosa, antipática).
48. AMABLE: (cordial, afectuosa, cortés).
49. TIERNA: (sensible, afable).
50. IRÓNICA: (Burlona, mordaz, sarcástica).
51. SEVERA: (rigurosa, inflexible, implacable, intransigente).
52. AUTORITARIA: (déspota, tirana).
53. TÍMIDA: (Miedosa, encogida).
54. CONVINCENTE: (persuasiva).
55. MENTIROSA: (engañosa, embustera, falaz)
56. NERVIOSA: (inquieta, ansiosa, excitada).
57. SEGURA: (firme, sólida).
58. INSEGURA: (poca seguridad).
59. RESIGNADA: (conformada, sometida, abandonada).

7.5 Primera prueba piloto para la evaluación del cuestionario

Antes de llevar a cabo la contrastación de las voces grabadas a los 60 jueces o escuchas, decidimos someter a prueba el cuestionario así como la manera en que serían presentadas las 200 voces.

Es importante resaltar que en la prueba de recepción final no buscamos encontrar los estereotipos vocales de ese grupo específico de escuchas con la intención de *universalizar* los datos que se obtengan. El modelo metodológico que presentamos tan sólo pretende, por un lado, que los jueces validen las bases conceptuales planteadas en la taxonomía de la voz y los términos empleados para definir y acotar algunos estereotipos; por otro, nos servirá para clasificar el material sonoro según las coincidencias estereotípicas conformadas colectivamente para realizar más adelante una clasificación y modelización de las clases. Será al momento de correlacionar los datos numéricos con los conceptos propuestos, que podremos sacar conclusiones acerca de los mecanismos empleados por los escuchas al momento de configurar los estereotipos.

La prueba piloto tenía como objetivo saber:

- i. si el cuestionario era lo suficientemente claro y comprendido sin dificultad;
- ii. el tiempo aproximado que llevaba juzgar cada una de las voces; y
- iii. cuántas voces seguidas se pueden escuchar y evaluar sin mostrar saturación.

Se conformó un grupo integrado por tres personas, quienes, después de contestar el cuestionario y escuchar un grupo de voces, proporcionaron datos interesantes. Se realizaron dos sesiones.

En la primera sesión se observó que:

1. El grupo era capaz de escuchar voces de 1 minuto y de 1'30 seg. aproximadamente sin mostrar saturación.
2. El tiempo aproximado para escuchar el bloque de 12 voces y contestar el cuestionario fue de 1 hora. El tiempo para calificar cada voz fue variando y notoriamente disminuyendo conforme se fueron familiarizando con el cuestionario y las definiciones.
3. La presentación de voces del mismo género y con una edad aproximada parecida, produjo que las interpretaciones fueran también muy similares, aún cuando se trataran de voces acústicamente muy diferentes.
4. El contenido del discurso influía en la interpretación de la voz.
5. Una misma voz que cambiaba de estado emocional y/o de actitud comunicativa (por ejemplo, de alguien que a la mitad del discurso rompía a llorar) causaba confusión en la interpretación al momento de juzgarla .
6. Una voz de 20 segundos de duración y con un contenido confuso (es decir, que no se le entendiera la idea general del discurso) pudo ser juzgada sin ningún problema.
7. Cada una de las voces fue escuchada dos veces a petición de los integrantes del grupo.

Con esta información se pasó a conformar otro bloque de 12 voces distintas las cuales fueron modificadas reduciendo su duración (de 20 a 40 segundos) y dos voces fueron editadas para que el discurso no fuera lineal y con ninguna lógica pero respetando frases completas y pausas naturales. Asimismo se trató de dejar un solo estado emocional. El cuestionario escrito

tuvo algunas modificaciones de presentación que lo hacían más manejable pero respetando intacto el contenido.

Los resultados que se obtuvieron fueron:

1. El grupo fue capaz de juzgar las voces sin ningún problema aún cuando se redujo sustancialmente la duración.
2. Las dos voces editadas para que el discurso fuera incoherente pudieron ser evaluadas sin problema. Lo que no se midió fue la presentación posterior de esas mismas voces sin editar para comparar si las evaluaciones sufrían cambios notorios.
3. La duración total del experimento siguió siendo de 1 hora aproximadamente.

En esta segunda sesión la presentación de las voces se hizo intercalando el sexo, es decir: una voz femenina, una masculina, una femenina y otra masculina... etc.

Cuando terminó el bloque de 12 voces se les preguntó si podían seguir escuchando pero manifestaron cansancio más que saturación para seguir escuchando.

7.6 Segunda Prueba Piloto

A partir de las observaciones realizadas en estas dos sesiones se abrió claramente la necesidad de realizar algunas sesiones más que verificaran los datos presentados hasta el momento.

Pues bien, vimos necesario medir nuevamente algunas variables para determinar el protocolo final de la evaluación ya que otras experiencias similares a la nuestra señalan la importancia que tiene la presentación de un corpus adecuado para obtener resultados óptimos. Además, con un grupo tan reducido era imposible hacer algún tipo de lectura de los resultados.

Era necesario determinar si las voces debían ser mostradas clasificándolas por edad, género, temática, actitud emotiva o comunicativa, y si podían presentarse sin un discurso lineal y coherente, así como la duración promedio que debían tener cada una de las voces para ser evaluadas .

Si en nuestra Taxonomía de la voz hemos dividido las señales sonoras dependientes del texto de las independientes, no podíamos tomar la decisión de presentar voces sin un

contenido lingüístico coherente, aún cuando en las primeras sesiones los sujetos mostraran no tener problema para calificarlas. Si en las posteriores pruebas piloto esta variable continuaba sin afectar la percepción y evaluación, entonces podríamos determinar que la variable discurso no era lo suficientemente relevante.

Por otro lado, el cuestionario ya con las modificaciones tenía que ser nuevamente evaluado.

Para esta segunda prueba piloto elaboramos una serie de hipótesis para ser verificadas.

HIPÓTESIS 1

El contenido del discurso o texto lingüístico, influye en la percepción de una voz al momento de ser clasificada como perteneciente a un estereotipo vocal, de ahí que no deba ser eliminada su coherencia discursiva durante el estudio empírico final.

HIPÓTESIS 2

Dos voces seguidas que contengan señales sonoras de edad y sexo similares influyen en la percepción para clasificarlas como pertenecientes a un mismo estereotipo vocal, aún cuando en realidad puedan pertenecer a otro distinto. El orden de presentación puede afectar generando confusión.

HIPÓTESIS 3

La presencia de más de una emoción en una voz produce confusión para estereotiparla.

HIPÓTESIS 4

La duración de una voz para que sea juzgada sin dificultad puede ser menor de 40 segundos.

Para verificar estas hipótesis se conformó un grupo de discusión con 6 personas españolas, las cuales, en 8 sesiones de una hora, escucharon bloques de voces distintas conformadas y presentadas de diferente manera.

En la primera sesión se mostraron 10 voces de **45 segundos a 1 minuto** de edades distintas e intercalando género, cuyo **contenido o temática era coherente y similar.**

En la segunda se mostraron 9 voces de **30 a 45 segundos** de personas de distinta edad y género cuyo **contenido o temática era coherente pero diferente.**

En la tercera se mostraron 12 voces de personas del **mismo género cuyo contenido era coherente**. (se usaron voces femeninas de 30 a 40 segundos).

En la cuarta se mostraron primero, 7 voces de **30 a 40 segundos cuya expresión emocional no sufría variaciones y el contenido era coherente**. Después, se mostraron esas mismas 7 voces pero con **otra expresión emocional** respetando la coherencia del contenido.

En la quinta se mostraron 6 voces de **30 a 40 segundos cuyo contenido era confuso o incoherente**. Después se mostraron esas mismas voces pero respetando **la coherencia de su discurso**.

En la sexta, séptima y octava se mostraron voces de **30 a 40 segundos de personas de edades y sexos diferentes respetando la coherencia del discurso y la expresión emocional**.

7.7 Resultados y conclusiones de la segunda prueba piloto

Como dijimos, los objetivos de esta segunda prueba piloto eran, por un lado, someter el cuestionario y los conceptos utilizados a un grupo mayor de personas para afinar el instrumento; y por otro, verificar las hipótesis para una mejor y más adecuada presentación de las voces.

Para el primero, afinación del instrumento, en las tres primeras sesiones se pidió a los integrantes del grupo de discusión que contestaran una pregunta diferente cada día para que se fueran familiarizando con el cuestionario. En las sesiones siguientes, se les entregó el cuestionario completo.

Al principio se les indicó que contestaran el cuestionario de manera individual y al final de la ronda de voces programadas se entablaba una discusión a partir de preguntas formuladas por la encuestadora. Pero a partir de la cuarta sesión se vio más productivo que, después de escuchar una voz, el cuestionario fuera contestado por parejas y posteriormente se discutieran en grupo los resultados, y así hasta terminar la ronda de voces programadas. El papel de la encuestadora pasó a ser de observadora, pero en algunos casos fungió como moderadora

en las discusiones. También formuló preguntas para esclarecer dudas pero sin dar su punto de vista.

Los resultados obtenidos fueron los siguientes:

1. Conforme se fueron familiarizando con el cuestionario y el anexo de conceptos, las voces fueron evaluadas con mayor rapidez.
2. En las tres primeras sesiones se observó que no habían coincidencias en las clasificaciones, o muy pocas. Cuando se abrió la discusión por cada voz escuchada en vez de la evaluación individual, se detectó que muchos términos del anexo contenían demasiados sinónimos con diferentes niveles de intensidad o con matices que significaban una característica distinta. Esto permitió depurar y, en algunos casos ofrecer una descripción en lugar de un sinónimo para evitar ambigüedades. Por otro lado, en muchos casos, la aparente falta de coincidencia no se debió a otra cosa que a la enorme cantidad de características y a la limitación de opciones solicitadas para su clasificación. Asimismo, esos diferentes niveles de intensidad hacían que los números no coincidieran aún cuando la característica optada perteneciera al mismo tipo de señal: emotiva, de actitud o físico-acústica. Por ejemplo, para algunos una voz era POTENTE y para otros simplemente FUERTE, sin embargo, para todos, la intensidad de la voz era una característica prioritaria y diferencial con respecto de otras voces escuchadas. Esta observación debe tomarse en cuenta al momento de contabilizar los resultados.

En cada nueva sesión se incorporaban al cuestionario las modificaciones y correcciones propuestas durante la sesión anterior.

7.8 Propuesta final del cuestionario

El cuestionario quedó de la siguiente manera:

ANEXO

1. POTENTE: fuerte.
2. ALTA: medianamente fuerte en intensidad.
3. BAJA: medianamente baja en intensidad
4. DÉBIL: muy baja en intensidad.
5. ESTRIDENTE: con sonido agudo, desapacible y chirriante.
6. MUY GRAVE: tono extremadamente bajo.
7. GRAVE: tono bajo.
8. AGUDA: tono alto.
9. MUY AGUDA: tono extremadamente alto.
10. RONCA: afección laríngea que cambia la voz haciéndola bronca y un poco sonora.
11. ASPERA: insuave, inarmónica.
12. AFÓNICA: fallo de voz, enferma.
13. SECA: pastosa, sin saliva.
14. HALITANTE: que habla con soplos suaves.
15. ARMÓNICA: bonita, agraciada.
16. SUAVE: fina, delicada.
17. COMPRIMIDA: oprimida, prensada, contenida.
18. BRILLANTE: como un sonido metálico.
19. TRÉMULA: temblorosa.
20. GANGOSA: con resonancia nasal por algún defecto en los conductos de la nariz. Que ganguea.
21. NASAL: con resonancia nasal pero sin llegar a ser gangosa.
22. PROFUNDA: baja, honda.
23. LAXA: distendida, floja, relajada.
24. TARTAMUDEANTE: pronunciación entrecortada y repitiendo sílabas.
25. MELODICA: con acento muy cantado.
26. PLANA: (monótona)
27. APRESURADA: rápida
28. LENTA: de ritmo lento
29. MELANCOLICA: tristeza vaga, profunda, sosegada y permanente.
30. TRISTE: afligida.
31. DEPRIMIDA: Desconsolada, apesadumbrada.
32. NOSTÁLGICA: tristeza ocasionada por una dicha perdida.
33. ALEGRE: contenta.
34. DIVERTIDA: entusiasta, festiva, bromista.
35. EUFÓRICA: Sensación de bienestar casi jubilosa.
36. ASUSTADA: que está sobresaltada.
37. ESPANTADA: con miedo.
38. ATERRADA: horrorizada.
39. ANTIPÁTIA: sentimiento de rechazo en mayor o menor grado por algo o alguien.
40. FASTIDIADA: harta.
41. ASQUEADA: con impresión desagradable causada por algo que le repugna.
42. MOLESTA: disgustada.
43. IRRITADA: enojada.
44. ENFURECIDA: encolerizada.
45. ASOMBRADA: extrañada.
46. SORPRENDIDA: pasmada.
47. DESCONCERTADA: reacción ante algo inesperado.
48. DESEOSA: interesada en, pensativa.

2. Describa tres aspectos en orden de importancia que le llamaron más la atención de la voz grabada:

Por ejemplo:

4. *La voz le temblaba mucho.*
5. *Lo que decía era muy triste*
6. *Hablaba muy lentamente*

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

VOZ 1. _____
2. _____
3. _____

3. Si tuviera que definir en una sola palabra a la persona escuchada, cuál de los conceptos del ANEXO 2 utilizaría:

VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.
VOZ	<input type="checkbox"/>	<input type="checkbox"/>	Otro: _____.

ANEXO 2

1. Amable: grata, afectuosa.
2. Agresiva: dura, severa, gruñona.
3. Simpática: graciosa, que produce empatía.
4. Apática: aburrida, impasible.
5. Sumisa: obediente, subyugada, rendida, subordinada.
6. Dominante: represiva, fuerte y superior.
7. Sensual: incitante, excitante.
8. Repulsiva: repugnante, despreciable.
9. Convicente: Que convence, que persuade.
10. Sin carácter: falta de personalidad firme e inseguridad
11. Varonil: de actitud y comportamiento que resalta las cualidades positivas y negativas relativas al sexo masculino.
12. Jovial: de actitud y comportamiento juvenil.
13. Maternal: de actitud y comportamiento similar a una madre tradicional.
14. Paternalista: de actitud y comportamiento similar a un padre tradicional.
15. Elegante: distinguida, refinada y de buen gusto .
16. Tosca: poco refinada, sin pulimento, inculta.
17. Inteligente: instruida, enterada, preparada, sabia.
18. Tonta: que le falta intelecto.

Por otro lado, los resultados obtenidos sobre las distintas maneras de presentar las voces fueron los siguientes:

1. Las voces mostradas de 30 a 40 segundos son perfectamente analizables.
2. Bloques de voces del mismo género presentadas de manera seguida impide detectar diferencias al momento de evaluarlas.
3. Aunque una voz sin contenido discursivo coherente puede ser evaluada según las características y conceptos presentados en nuestro cuestionario, las interpretaciones sufren cambios sustanciales al momento en que esa misma voz es presentada sin cortes y con su discurso lógico y lineal. Esto nos hace concluir que la **variable discurso** influye en la tipificación de las voces y no debe ser eliminada de nuestra muestra.
4. La variación emocional en una voz también influye en la clasificación estereotípica por lo que debe mantenerse una sola expresión por voz.
5. La presentación de casi 70 voces durante estas 8 sesiones nos ayudó a depurar las voces que son más claramente estereotípicas que otras, las cuales utilizaremos en el protocolo final.

8. PROPUESTA DE UNA HERRAMIENTA DE MODELIZACIÓN Y MEDICIÓN DE LAS VOCES

*(...) Como que se van las voces. Como que se pierde su ruido.
Como que se ahogan.
Ya nadie dice nada. Es el sueño".
Juan Rulfo, Pedro Páramo*

Una vez que se haya realizado el estudio empírico necesitaremos de una herramienta de medición para determinar acústicamente las similitudes y diferencias de las voces juzgadas como pertenecientes a un mismo estereotipo vocal.

Nos basaremos en la *taxonomía de las formas sonoras simples* que realizó Rodríguez Bravo (1998, p. 171-191) ya que, ante

todo, la voz es un sonido y como tal está compuesta por: *intensidad, tono, duración y timbre*.

A través de estas cuatro características acústicas de la voz se estudiará el tono, entonación (curva melódica), intensidad y ritmo (tiempo) a través de instrumentos de medición (sonogramas, oscilogramas y espectogramas).

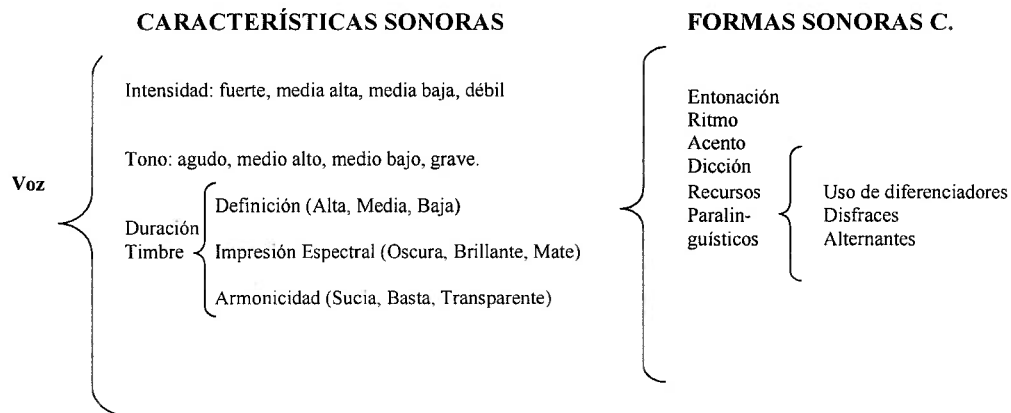
Mediremos:

- La Frecuencia Fundamental (F0)
- Presión Sonora (intensidad en dB).
- Duración de:
 - tiempo total del discurso
 - tiempo de fonación
 - tiempo de pausas
 - duración de cada grupo fónico
 - duración de cada pausa
 - duración media del grupo fónico
 - duración media de la pausa
 - relación pausa-fonación en %
 - nº de sílabas por segundo.

Pero también nos interesan medir el uso de los *recursos paralingüísticos: diferenciadores, disfraces y alternantes*.

A continuación presentamos la parte de la taxonomía que esquematiza las características del sonido y su relación en la conformación de formas sonoras complejas:

HERRAMIENTA DE MODELIZACIÓN Y MEDICIÓN DE LAS CLASES



9. CONCLUSIONES

*"Oía su voz, su propia voz, saliendo despacio de su boca.
La sentía sonar como una cosa falsa y sin sentido".*

Juan Rulfo, El llano en llamas.

Esta primera revisión teórica y conceptual nos ha permitido elaborar una propuesta de taxonomía de la voz y diseñar una prueba de recepción para evaluar las características particulares de la voz y los estereotipos propuestos. Pero también nos ha dado elementos para aportar algunas reflexiones sobre el tema.

En primer lugar, conceptualmente podemos decir que la definición de Expresión Fonoestésica debe ampliarse. Recordemos que Angel Rodríguez Bravo la define como:

La expresividad sonora transmitida mediante los rasgos de la voz que comunican acústicamente información sobre el gesto, la actitud, el estado emocional, el carácter, el aspecto físico y el contexto de un emisor; o bien sobre la forma, el tamaño, el color, la textura, el tipo de movimiento, etc., de aquello que describe oralmente el emisor. (2002: 160).

Esta definición surge después de que en la investigación sobre voz se recurre a términos como *rasgos lingüísticos* para la información sonora que aporta la voz al momento de emitir un mensaje oral y que no difiere del lenguaje escrito; *rasgos paralingüísticos* para la información sonora no lingüística que añaden información al texto, y *rasgos extralingüísticos* para la información sonora sobre el estado del locutor. El término Expresión Fonoestésica abarca estos tres niveles de información y expresión sonora.

Pues bien, proponemos agregar a la definición Expresión Fonoestésica aquellos rasgos sonoros --también dependientes del emisor más que del texto-- que no hacen referencia a las características particulares del emisor sino sobre las características de *otro sujeto* externo, gracias a los disfraces y capacidad de imitación vocal que tiene el ser humano.

La definición de Expresión Fonoestésica quedaría de la siguiente manera:

La expresividad sonora transmitida mediante los rasgos de la voz que comunica acústicamente información sobre el gesto, el carácter, el estado emocional, el aspecto físico y el contexto de un emisor **o bien sobre el**

gesto, el carácter, el estado emocional, el aspecto físico y el contexto de otro sujeto a quien hace referencia oralmente el emisor; así como la forma, el tamaño, el color, la textura, el tipo de movimiento, etc., de aquello que está describiendo.

Esta nueva definición contempla la capacidad que tenemos para disfrazar nuestra voz e imitar rasgos particulares y específicos de otras voces.

Por otro lado, metodológicamente, el estudio de los estereotipos vocales puede ser abordado desde diferentes perspectivas, como por ejemplo: *el uso de estereotipos vocales en los medios de comunicación audiovisual y sus repercusiones en un grupo social determinado, o, quizá, cómo el emisor modifica su voz según el contexto en el que se encuentra y la estereotipa, etc.*

Sin embargo, nosotros nos centramos en los *mecanismos mentales que el ser humano realiza al momento de percibir una voz para poder comprender, entre otras cosas, las simplificaciones que los medios de comunicación audiovisual realizan de la voz sin que se dificulte la comprensión de sus mensajes. El estudio del "cómo" nos llevará a entender los "por qué".* Creemos que esta perspectiva es necesaria y puede servir como sustento para otras investigaciones sociales.

Otro aspecto a resaltar sobre los aportes de esta investigación, es que para la contrastación empírica planteamos la presentación de un corpus vocal lo más análogo posible a la forma en como escuchamos una voz en la vida cotidiana, sea a través de un medio audiovisual o en una conversación. Es decir, que las voces a juzgar no son elaboradas desde un laboratorio con actores u otro tipo de profesionales, ni condicionando el contenido de sus mensajes pre-elaborando o eliminando la coherencia del texto.

Estas variables están presentes en nuestra investigación aún cuando la hacen más compleja.

La última reflexión que hacemos es que, al momento en que nos planteamos estudiar los estereotipos a partir de su percepción sonora de la voz, nos obliga a abordar nuestro objeto de estudio resaltando las particularidades sígnicas del sonido vocal semejante a lo que en procesos judiciales se hace de un retrato visual de una persona: tipo de nariz, boca, color de piel, estatura, etc. Si logramos encontrar y medir las características sonoras particulares y su relación para la conformación estereotípica de la voz, podremos crear un sistema para el retrato sonoro de una voz.

BIBLIOGRAFÍA

ÁLVAREZ-GAYOU JURGENSON, Juan Luis (2003) *Cómo hacer investigación cualitativa*, Editorial Paidós, México.

ALCOBA, Santiago (2000) *La expresión Oral*, Editorial Ariel, Barcelona.

BADDELEY, Alan.(1999) *Memoria Humana*, McGraww-Hill, España.

BALSEBRE, A. (1981) *"El lenguaje radiofónico*. Tesis de licenciatura, Universidad Autónoma de Barcelona, Dto. de Tecnología de la Información.

BARON, Robert y BYRNE, Donn. (1998). *Psicología Social*, Prentice Hall Iberia, Madrid.

BUNGE, Mario.(1989) *La ciencia. Su método y su filosofía*, Editorial Patria, México.

CÁCERES, María Dolores (2003) *Introducción a la Comunicación Interpersonal*, Editorial Síntesis, Madrid.

CASTARLENAS, Rafael. (1995) *La Locución: técnica y práctica*, Editorial Panapo, Venezuela.

CUESTA, Ubaldo.(2000) *Psicología Social de la Comunicación*, Ediciones Cátedra (Grupo Anaya), Madrid,).

CORNUT, Guy. (1985) *La voz*, Fondo de Cultura Económica, México.

CHION, Michel. (1998) *La audiovisión. Introducción a un análisis conjunto de la imagen y el sonido*, Paidós, Barcelona.

CHION, Michel. (1999) *El sonido. Música, cine, literatura*, Paidós, Barcelona.

ELLIS, R., MCCLINTOCK, A. (1993) *Teoría y práctica de la comunicación humana*, Ediciones Paidós Ibérica, Barcelona.

ESTRADA, Julio (1990) *El sonido en Juan Rulfo*, Universidad Nacional Autónoma de México, México.

FISCHER, G.N. (1990) *Psicología Social: conceptos fundamentales*, Narcea Ediciones, Madrid.

GARRIDO, Isaac. (2000) *Psicología de la Emoción*, Editorial Síntesis, Madrid.

IBÁÑEZ, Tomás. 1998 *Ideologías de la vida cotidiana*, Sendai Ediciones, Barcelona.

KIRK, G.S. (1999) *El mito: su significado y funciones en la antigüedad y otras culturas*, Ediciones Paidós Ibérica, Barcelona

KNAPP, Mark L. (1982) *La comunicación no verbal*, Paidós, Barcelona.

KONDRÁTOV, A. M. (1973) *Del sonido al signo*. Editorial Paidós. Buenos Aires.

LÓPEZ-HIGES, Ramón (2003) *Psicología del Lenguaje*, Ediciones Pirámide, Madrid.

MARTÍNEZ CELDRÁN, Eugenio (1984) *Fonética*, Teide

MARX, M. y HILLIX, W (1983) *Sistema y teorías psicológicas y contemporáneos*, Editorial Paidós, México.

MANZANERO, Antonio (1996) *La exactitud en la recuperación de sucesos complejos: la memoria de testigos*. En Sáiz, D, Sáiz, M, Baqués, J (Eds.), "*Psicología de la Memoria. Manual de Prácticas*", Eduard Fabregat Editor, Barcelona.

MÉNDEZ, I., NAMIHIRA, D., MORENO, L, y SOSA, C.(1986) *El Protocolo de investigación*, Editorial Trillas, México.

MIDDLETON, David, EDWARDS, Derek (1992) *Memoria compartida*, Ediciones Paidós Ibérica, Barcelona.

MIYARA, Federico: *La voz humana*
(<http://www.eie.fceia.unr.edu.ar/~acústica/biblio/fonatorio.pdf>).
(Fecha de consulta 22-06-2004).

MONTOYA, Norminanda. (1999) *El uso de la voz en la publicidad audiovisual dirigida a los niños y su eficiencia persuasiva*. Tesis doctoral. Universidad Autónoma de Barcelona. Dto. De Comunicación Audiovisual y Publicidad.

MONTOYA; LÁZARO; RODRÍGUEZ. (2000) *Estereotipos vocales de carácter en la publicidad infantil*. En *Comunicación y Pedagogía*, revista nº 171, Barcelona, diciembre de 2000, (pp. 69-76)

NAVARRO, Tomás (1946) *Estudios de fonología española*, Syracuse, Madrid.

PIÑUEL, J., y GAITÁN, J. (1999) *Metodología General: conocimiento científico e investigación en la comunicación social*, Editorial Síntesis, Madrid.

POYATOS, Fernando (1994) *La comunicación no verbal*, Istmo, Madrid.

QUILIS, Antonio (1982) *Curso de fonética y fonología españolas*, C.S.I.C. Madrid.

RODRÍGUEZ, LÁZARO, MONTOYA; BERNADAS; OLIVER; LONGHI; GONZÁLEZ; "El control objetivo de la implicación de los informantes en el estudio del habla emocionada". *Procesamiento del lenguaje Natural*, revista nº 23, Alicante, septiembre de 1998, (pp. 24-31).

RODRÍGUEZ; LAZARO; MONTOYA; BLANCO; BERNADAS; OLIVER; LONGHI; "Modelización acústica de la expresión emocional en el español". *Procesamiento del Lenguaje Natural*, nº 25, Lérida, septiembre de 1999, issn 1135-5948, (pp. 159-166).

RODRÍGUEZ; LAZARO; MONTOYA; (2003) *Identificación perceptiva de locutores para la acústica forense: las RRV*. En II Congreso Sociedad Española de Acústica Forense. Barcelona, 10-11 de abril.

RODRÍGUEZ BRAVO, A. (1984) *La voz en la radio*. Tesis de licenciatura. Universidad Autónoma de Barcelona. Dto. De Comunicación Audiovisual y Publicidad.

RODRÍGUEZ BRAVO, A. (1989) *La construcción de una voz radiofónica*. Tesis Doctoral. Universidad Autónoma de Barcelona, Dto. De Comunicación Audiovisual y Publicidad.

RODRÍGUEZ BRAVO, Ángel (1988-1989) *Máquinas que hablan y escuchan*. En *Telos*, revista nº 16 , pp. 117-125. Madrid.

RODRÍGUEZ BRAVO, Ángel. (1998) *La dimensión sonora del lenguaje audiovisual*. Ediciones Paidós Ibérica. Barcelona.

RULFO, Juan, (1993) *Pedro Páramo*, Fondo de Cultura Económica, novena edición, México.

RULFO, Juan (1993) *El llano en llamas*, Editorial Anagrama, séptima edición, Barcelona.

RUMSEY, Francis (2002) *Introducción al sonido y la grabación*, Instituto Oficial de Radio y Televisión, Madrid.

ROJAS, Raúl (1981) *Guía para realizar investigaciones sociales*. Universidad Nacional Autónoma de México. México.

SÁIZ, D. y BAQUÉS, J. (1996): La medición de la Memoria. En Sáiz, D, Sáiz, M, Baqués, J (Eds.), "*Psicología de la Memoria. Manual de Prácticas*" , Eduard Fabregat Editor, Barcelona.