

**PERCEPCIÓN DE LA VELOCIDAD DE HABLA  
EN EL ESPAÑOL DE MÉRIDA (VENEZUELA)**

**SPEAKING RATE PERCEPTION  
IN THE SPANISH OF MERIDA (VENEZUELA)**

DARCY ROJAS

*Universidad de Los Andes (Mérida, Venezuela)*  
darcyjra@gmail.com

HERNÁN MARTÍNEZ

*Universidad de Los Andes (Mérida, Venezuela)*  
hernan.martinez@ula.ve

*Artículo recibido el día: 24/01/2011*

*Artículo aceptado definitivamente el día: 15/11/2011*

*Estudios de Fonética Experimental, ISSN 1575-5533, XX, 2011, pp. 179-203*

### RESUMEN

La presente investigación constituye un estudio piloto que busca realizar un primer acercamiento al fenómeno de la percepción de la velocidad de habla en el español merideño. Este estudio tuvo como objetivo identificar los incrementos mínimos de estímulo que permiten diferenciar una velocidad de habla normal de una velocidad de habla rápida, y una velocidad de habla lenta en el español de Mérida. Para este fin, se realizaron dos tests de percepción (uno de identificación y otro de discriminación tipo ABX) que fueron presentados a un total de 119 informantes merideños con edades comprendidas entre los 20 y 40 años. Los resultados reflejan que los oyentes logran categorizar y discriminar entre distintos tipos de velocidades de habla y que es posible identificar las diferencias mínimas perceptibles que dividen la velocidad de habla *rápida*, *normal* y *lenta* en el español de Mérida a través de análisis perceptivos.

Palabras clave: *percepción del habla, velocidad del habla, español hablado en Mérida.*

### ABSTRACT

This is a pilot study that aims at approaching to the phenomenon of speaking rate perception in the Spanish of Mérida. The objective was to identify the minimum increments of stimulus that help distinguishing normal, low and high speaking rates. Two perception tests (identification test and ABX discrimination test) were applied to 119 informants from twenty to forty years of age. Results show that listeners are able to categorize and discriminate between different types of speaking rate and that it is possible to identify the minimal perceptible differences that divide the speaking rates in the Spanish of Mérida across perceptive analysis.

Key words: *speech perception, speech rate, Spanish speak in Mérida.*

## 1. INTRODUCCIÓN

La velocidad de habla, entendida como *el número de sílabas divididas entre el tiempo total de locución (aquel que incluye el tiempo de pausas)*<sup>1</sup> (Blondet, 2006: 19), ha sido un fenómeno lingüístico capaz de suscitar gran cantidad de dificultades al momento de ser sometido a análisis debido a que cualquier hablante puede variar la velocidad de habla según diferentes factores de orden lingüístico, paralingüístico o extralingüístico. Por ello, determinar el hecho que produce el cambio en la velocidad, y más aún, registrar parámetros acústicos que delimitan los diferentes tipos de velocidades de habla que perciben los hablantes puede resultar bastante arduo (Cf. Blondet, 2006).

Hechos, como por ejemplo, las distintas variaciones que sufre la velocidad de habla, no solo en función de la situación comunicativa en que se produzca la emisión, sino además en función de las distintas variaciones que ocurren debido al dialecto e incluso las diferentes variaciones que se registran de individuo a individuo, nos llevan a preguntarnos si realmente un indicio prosódico como la velocidad de habla es percibido según distintas categorías. Para Blondet (2006) la interacción que se produce entre emisores y receptores es siempre lo más importante. En este sentido, el emisor puede variar sus realizaciones en función de distintos factores (sean de orden lingüístico, paralingüístico o extralingüístico), lo que trae como consecuencia modificaciones, entre otras, temporales en la señal acústica. En consecuencia, todo hablante o perceptor fluido de la lengua posee la capacidad de manejar todas esas fuentes de variación de la señal, de manera que no interfieran en el proceso comunicativo. Desde esta perspectiva,

*cada velocidad posee un rango de variación enmarcado dentro de un margen de seguridad; cada hablante tiene conciencia de hasta dónde puede variar su velocidad de habla, bien para ser entendido o bien para codificar una determinada información que requiera de la utilización de la variación de la velocidad de habla para su realización [...] Sin duda habrá un indicio invariante en la señal sonora que lleve al oyente –y hablante– a reconocer –y codificar– una velocidad de habla determinada (Blondet, 2006:1).*

Esta última aseveración pareciera responder, en parte, a una de las principales interrogantes que suscita el estudio de la velocidad de habla, sin embargo, nos lleva

---

<sup>1</sup> La fórmula empleada para el cálculo de la velocidad de habla es la siguiente: sílabas + pausas/ tiempo total de habla.

también a preguntarnos ¿qué diferencia temporal debe existir entre distintos estímulos para que un oyente pueda diferenciar la velocidad de habla normal de la velocidad de habla rápida y lenta? Esta interrogante parece aún no respondida no sólo para el español venezolano sino para muchas otras variedades del español. Este estudio tiene por objetivo principal identificar los incrementos mínimos de estímulo necesarios para que un oyente pueda diferenciar la velocidad de habla normal de la velocidad de habla rápida, y de la velocidad de habla lenta en el español de Mérida.

Dar respuesta a este objetivo se considera relevante en nuestro ámbito de estudio ya que la percepción parece un tema casi inexplorado. En lo que respecta al español, el estudio perceptivo de la velocidad de habla, sólo parece haber sido abordado someramente en una investigación que determinó que la velocidad de habla normal para el español rioplatense era de 12,9 alófonos por segundo aproximadamente (Cf. Wainschenker *et al.*, (2002:103). La bibliografía restante, que logra determinar un patrón promedio de velocidad de habla para el español, lo hace desde el punto de vista de la producción. Así, Navarro Tomás (1982) calcula que la velocidad de habla promedio sería de 200 palabras por minuto, mientras que otros como Loperete (1984) calculan que la velocidad de habla normal estaría entre las 120 y las 150 palabras por minuto. De la misma manera, Blondet (2006) propone una serie de rangos que permiten distinguir entre tres temporemas básicos en el español venezolano: el temporema lento, el temporema normal y el temporema rápido. El primero se caracteriza por contar con un número de sílabas estimado entre las 3,49 y 4,1 sil/seg y por distribuciones regidas por reglas como: mayor duración de sílabas en fines de constituyentes textuales<sup>2</sup>, fines de constituyente textual mayor marcados doblemente (mayor duración silábica más pausa), y posibilidad de marcar fines de constituyente textual menor a través de pausa y de mayor duración silábica.

El temporema normal está caracterizado por contar con un número de sílabas estimado entre las 4,7 y 5,3 sil/seg y por distribuciones regidas por reglas como: mayor duración de sílabas en fines de constituyente textual, fines de constituyente textual menores marcados ocasionalmente por pausa o a través de mayor duración silábica. Por su parte, en el temporema rápido presenta un rango estimado entre 6 y 7,1 sil/seg y por distribuciones regidas por reglas como: mayor duración de sílabas en fines de constituyentes textuales, fines de constituyentes textuales mayores marcados con pausa, y fines de constituyente textual menores marcados sólo con mayor duración.

---

<sup>2</sup> Para Blondet (2006), los constituyentes textuales pueden ser mayores (párrafos) o menores (cláusulas y constituyentes de cláusulas).

## 2. PERCEPCIÓN: PRINCIPALES ENFOQUES

La percepción, en tanto que función que nos permite, por medio de los sentidos, interpretar, asignar significado y organizar las experiencias generadas por algún tipo de estímulo; se caracteriza por ser un fenómeno amplio e integral en el que influyen nuestras necesidades, valores y deseos. Cada uno de nosotros codifica y clasifica los datos que nos llegan del exterior, reduciéndolos a categorías de las que dispondremos para comprender luego el entorno. Estas clasificaciones y codificaciones son procesos intermediarios entre los estímulos y la conducta (Bruner, 1984). Desde esta perspectiva, se reconoce, en efecto, que la percepción es un proceso compuesto de sensaciones (objetivas) y de conocimientos previos (subjetivos) que se nos aparece como simple; esto es, todo lo percibido parece estar dado simultáneamente y en igualdad de condiciones, formando una unidad. Específicamente, la percepción del habla la definimos, siguiendo a Martínez (2009:35):

*como un proceso global, estructurado y orientado temporalmente, que involucra la sensación y la experiencia o el conocimiento previo del sujeto e implica la segmentación de las unidades lingüísticas y la interpretación o asignación de un significado.*

Desde el punto de vista lingüístico, varios han sido los estudios sobre percepción del habla. En ellos este proceso se ha vislumbrado bastante complejo, así lo indican, entre otros, autores como Matlin y Hugh (1996), Yeni-Komshian (1999), Marrero (2001) y Martínez (2009). Los trabajos de percepción del habla han explicado cómo las personas son capaces de percibir los límites entre las palabras y también los límites entre elementos menores como la sílaba o los segmentos, aún cuando estos carecen de pausas y se caracterizan por ser de naturaleza variable.

Esta realidad experimental ha sido fundamental para el desarrollo de esos trabajos, ya que con ellos se evidencia, en primer lugar, que el proceso de percepción se debe a procesos de segmentación realizados en un nivel superior, y en segundo lugar, que la variabilidad constituye un fenómeno fundamental en la percepción lingüística.

Tan relevante llega a ser el tema de la variabilidad en estos estudios que, como lo indica Marrero (2001), este constituye el eje sobre el que giran las diferencias entre las principales teorías que buscan dar explicación a la percepción del habla. Así, dos grandes vertientes teóricas se derivan de esta concepción: la primera,

---

conformada por los llamados modelos pasivos, centrados en la búsqueda de invariantes; la segunda, conformada por los llamados modelos activos, caracterizados por defender la existencia de mecanismos intermedios capaces de compensar las distintas distorsiones que sufre la señal (Cf. para mayores detalles Obediente (1998), Marrero (2001), Martínez (2009).

Entre las principales teorías desarrolladas siguiendo el modelo activo de percepción del habla se encuentra la teoría de la percepción categorial, formulada por investigadores de los laboratorios Haskins (Liberman, Harris, Hoffman y Griffith, 1957). Esta teoría, en términos generales, buscaba comprobar que mientras grandes cambios físicos en un rango determinado, de tiempos o de frecuencias, no producían diferencias perceptivas; diferencias muy pequeñas en otro rango de duración o de frecuencia provocaban cambios drásticos en la percepción, al punto de que se podía identificar un estímulo, antes clasificado dentro de una categoría, como un elemento lingüístico perteneciente a otra (Cf. Marrero, 2001:33).

A través de este enfoque los investigadores de los laboratorios Haskins buscaban comprobar que los seres humanos procesaban de manera distinta los sonidos lingüísticos y los sonidos no lingüísticos. Esta teoría en un inicio resultó sorprendente, ya que aparentemente probaba que (a pesar de que al percibir otros tipos de señales acústicas los oyentes podían establecer muchas distinciones entre los estímulos) al percibir los sonidos lingüísticos los oyentes sólo podían discriminar entre *los sonidos hallados en los lados opuestos de un límite del fonema y no entre los sonidos ubicados dentro de la misma categoría del fonema* (Cf. Liberman, Harris, Hoffman y Griffith, 1957: 358).

Para determinar la existencia de la percepción categorial los investigadores de los laboratorios Haskins partieron de dos pruebas prototípicas en el estudio de la percepción del habla. La primera, consistía en una tarea de identificación de consonantes oclusivas dentro de un continuo construido a partir de la manipulación paulatina de la frecuencia de inicio de la transición al segundo formante. La segunda prueba consistía en una tarea de discriminación llamada ABX, en la cual A y B eran dos estímulos diferentes del continuo y X podía ser A o B.

A través de la tarea de identificación se buscaba comprobar que cada sonido era clasificado dentro de una categoría hasta alcanzar un punto crítico a partir del cual se comenzaba a identificar el estímulo como perteneciente a otra categoría. Mientras que con la tarea de discriminación se buscaba comprobar la posibilidad de distinguir solo entre los sonidos lingüísticos pertenecientes a categorías diferentes.

Los resultados obtenidos por los investigadores de los laboratorios Haskins dieron lugar al establecimiento de tres de las características definitorias de la percepción categorial, a saber:

1. Los estímulos debían demostrar cuevas precipitadas, entre las categorías del fonema adyacente.
2. La discriminación de estímulos dentro una misma categoría del fonema debía ser pobre, mientras que la discriminación entre estímulos de diversas categorías de fonemas debía ser buena.
3. Los picos de discriminabilidad de pares de estímulo a lo largo de una serie continuada debían corresponder con los límites de la categoría obtenidos dentro del etiquetado o identificación de los estímulos (Cf. Jusczyk y Luce, 2002: 5).

A través de los experimentos realizados por estos investigadores se buscaba demostrar que la percepción categorial era un fenómeno en el que el procesamiento del lenguaje se realizaba de arriba-abajo, es decir, que la señal era interpretada en función de categorías preexistentes, las cuales permitían segmentar y ordenar el continuo fónico que ingresaba al oído.

Cabe destacar igualmente que, como lo señalan Marrero (2001), Matlin y Hugh (1996) y Jusczyk y Luce (2002), pese a que los investigadores defensores de la percepción categorial sostenían que esta constituía un módulo especial distinto al usado para procesar otro tipo de sonidos y que era una capacidad exclusiva del lenguaje; posteriormente se logró comprobar la invalidez de estas dos afirmaciones; pues, por una parte, la percepción categorial podía ser encontrada en otro tipo de sonidos, tales como los sonidos musicales complejos (Cf. Pisoni, 1978). Y por otra parte, se evidenció que ciertos sonidos del lenguaje como las vocales y algunos rasgos prosódicos no eran percibidos categorialmente (Cf. Fry, Abramson, Eimas y Liberman, 1962).

Pese a estos hallazgos la hipótesis categorial se siguió considerando atractiva debido a que lograba dar una explicación a la variabilidad característica de la señal acústica, lo cual parecía desplazar la hipótesis de la existencia de un sistema único de características invariantes para cada segmento fonético defendida por autores

centrados en estudiar la percepción del habla a través de la formulación de modelos pasivos<sup>3</sup>.

Actualmente, tanto la hipótesis de procesamiento de la información de arriba-abajo como la explicación de la ausencia de invarianzas defendida por la percepción categorial, logran encontrar una explicación alternativa en un enfoque que considera la percepción del habla como un fenómeno continuo. Quienes siguen este enfoque defienden (al igual que quienes defienden el enfoque de la percepción categorial) la existencia de categorías, en tanto que mecanismos usados para la organización de la información; y en tanto que método utilizado para simplificar lo real a partir de procedimientos como la generalización y la discriminación. Sin embargo, según este nuevo punto de vista, las categorías son entidades no discretas, y por tanto es una acción que *se realiza a partir de relaciones prototípicas que determinan límites difusos entre las categorías* (Cf. Martínez Celdrán, 2001:2). Una de las ventajas que parece tener este nuevo enfoque es que permite establecer gradaciones entre los estímulos pertenecientes a una misma categoría; lo cual permite diferenciar entre elementos centrales o prototípicos de las categorías y elementos periféricos que en ocasiones pueden dar a los límites o fronteras de las categorías un carácter difuso o borrosos (Cf. Martínez Celdrán, 2001).

Esta concepción, para Martínez Celdrán (2001), no necesariamente excluye la existencia de los fenómenos extremos registrados en el enfoque de percepción categorial, sino que intenta incluirlos dentro de una escala en la que también se consideran los términos intermedios.

### 3. METODOLOGÍA

#### 3.1. Las grabaciones y el tratamiento de la señal

Para llevar a cabo esta investigación se realizaron 9 grabaciones de audio con el programa PRAAT (versión 5.1.2.3) que fueron almacenadas en formato .wav. Para la conducción de la señal de audio hacia el computador se utilizó una consola Europower PMP1000 y un micrófono Tech. Las grabaciones se realizaron en un espacio cerrado y sin ruido.

---

<sup>3</sup> Para mayores detalles sobre los distintos modelos de percepción Cf. Marrero (2001).



Para las grabaciones se contó con la colaboración voluntaria de un hablante masculino nativo de la ciudad de Mérida que produjo la cláusula *veamos el ala del ave rapaz*, a velocidades consideradas por él como: normal, rápida y lenta. El resultado final fueron 9 grabaciones de audio conformadas por tres repeticiones de la cláusula a una velocidad normal, tres repeticiones a velocidad rápida y tres repeticiones a velocidad lenta<sup>4</sup>. A continuación, en las figuras 1, 2 y 3, se presentan tres de los espectrogramas de la cláusula indicada para cada una de las velocidades de habla consideradas.

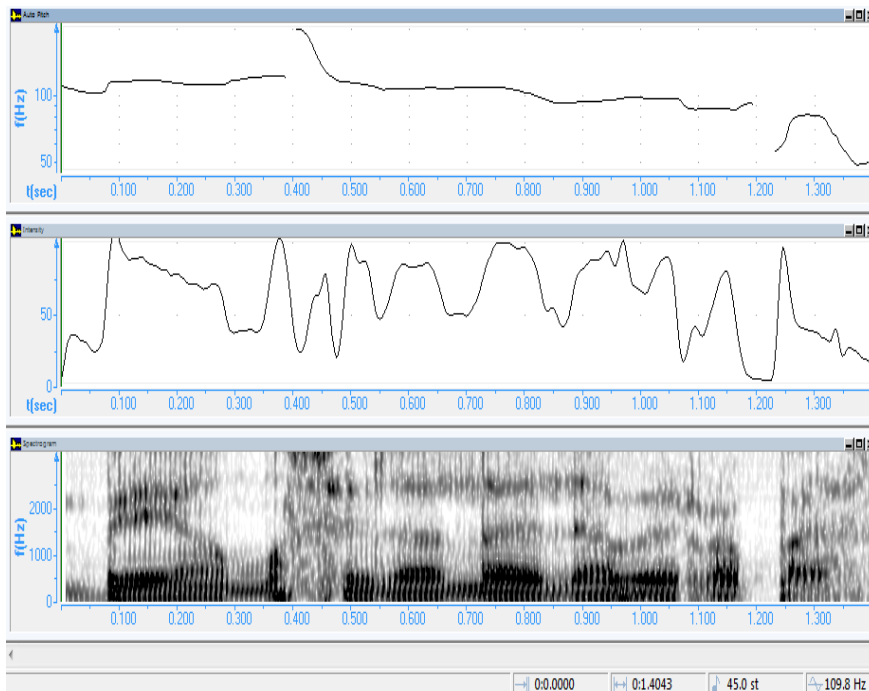


Figura 1. Curva de  $F_0$ , curva de intensidad y espectrograma de la muestra de habla rápida etiquetada como estímulo 3.

<sup>4</sup> Los valores de las cláusulas normales fueron: 5,66 sil/seg, 5,22 sil/seg y 5,00 sil/seg. Los valores de las cláusulas rápidas fueron: 4,17 sil/seg, 4,00 sil/seg y 3,87 sil/seg. Los valores de las cláusulas lentas fueron: 6,79 sil/seg, 7,14 sil/seg y 6,67 sil/seg.

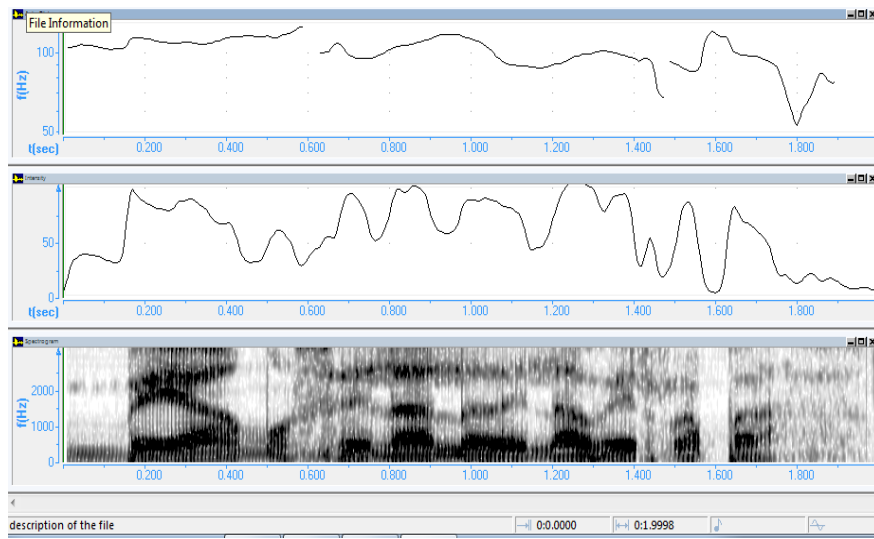


Figura 2. Curva de F0, curva de intensidad y espectrograma de la muestra de habla normal etiquetada como estímulo 9.

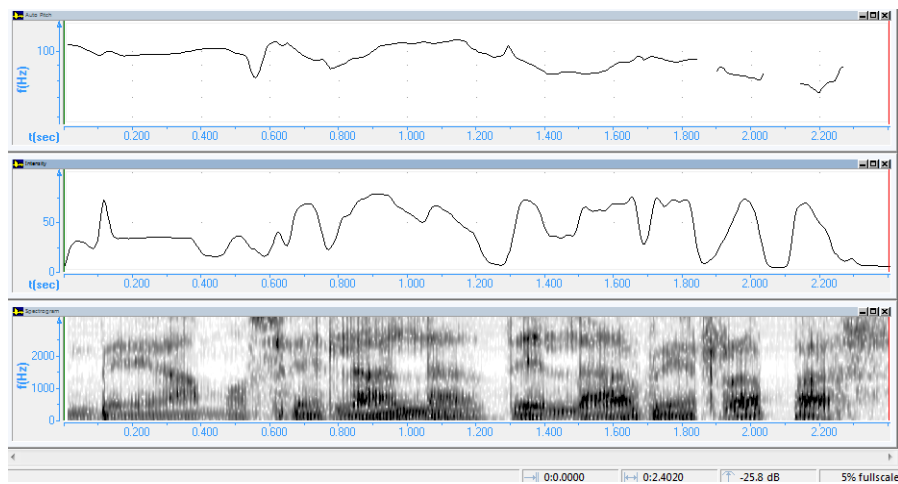


Figura 3. Curva de F0, curva de intensidad y espectrograma de la muestra de habla lenta etiquetada como estímulo 13.

Siguiendo los rangos que desde la producción registra Blondet (2006) para la velocidad del habla en español venezolano, de las 9 grabaciones se seleccionaron las tres más representativas de cada velocidad: normal 5,00 síl/seg (estímulo 9), rápida 7,14 síl/seg (estímulo 3), lenta 4,17 síl/seg (estímulo 13). Para la realización de esta tarea, se segmentó la señal sonora correspondiente a la voz del hablante y se utilizó el espectrograma de banda ancha a fin de determinar el número de cada sílaba fonética y su respectiva duración en milisegundos. Esto permitió calcular la velocidad de cada cláusula siguiendo los parámetros de medición usados por Blondet (2006:20), a saber: sílabas + pausas/ tiempo total de habla.

Una vez finalizada la medición se procedió a crear un continuo de 15 estímulos con duraciones que oscilaban entre 1,20 y 2,60 segundos. (Cf. Tabla1). Este continuo se logró creando cinco conjuntos de estímulos a partir de las muestras de habla rápida, normal y lenta producidas por el hablante. Para generar estos conjuntos se seleccionaron los periodos más semejantes de las vocales de cada sílaba; se seleccionó un intervalo de vocal a alargar o a acortar que cumpliera con las siguientes normas:

1. Que correspondiera a un segmento de características temporales y espectrales relativamente estables de la vocal,
2. Que tuviera un número entero de ciclos o pulsos glotales,
3. Que empezara y terminara en el mismo momento de un ciclo, cuando el nivel de la señal sea lo más cercano posible a cero. La parte seleccionada del sonido vocálico fue insertada al principio de la parte estable de la vocal correspondiente (Mora *et al*, 2007:18-19). Esta selección permitió evitar al máximo la presencia de distorsiones o ruidos comunes cuando se manipula la señal.

La creación del primer conjunto de estímulos, que formó parte del continuo, se logró restando un periodo a cada sílaba del estímulo 3 (1,40 segundos de duración). Este procedimiento se realizó paulatinamente hasta obtener como resultado la creación del estímulo 2 (1,30 segundos de duración) y la creación del estímulo 1 (1,20 segundos de duración) (Cf. Anexo 1).

La creación del segundo conjunto de estímulos se logró aumentando un periodo a cada sílaba del estímulo 3. Este procedimiento se realizó paulatinamente hasta ob-

tener como resultado la creación del estímulo 4 (1,50 segundos de duración) y la creación del estímulo 5 (1,60 segundos de duración).

La creación del tercer conjunto de estímulos se consiguió disminuyendo paulatinamente la duración de cada sílaba de la muestra de habla normal, de 2,00 segundos de duración. Para lograr esta tarea se restó un periodo a cada sílaba hasta conseguir crear un estímulo con una duración de 1,90 segundos, etiquetado con el número 8; un estímulo con una duración de 1,80 segundos, etiquetado con el número 7; y un estímulo con una duración de 1,7 segundos, etiquetado con el número 6 (Cf. Anexo 1).

La creación del cuarto grupo de estímulos se consiguió, en primer lugar, disminuyendo la duración de las sílabas de la muestra de habla lenta, de 2,40 segundos, y en segundo lugar, disminuyendo la duración de las pausas silenciosas presentes entre los fines de constituyentes menores<sup>5</sup>. Para lograr la tarea planteada se restó un periodo a cada sílaba de la muestra de habla lenta y se restó un milisegundo a la duración de las pausas silenciosas presentes entre los fines de constituyentes menores. A través de estos dos procedimientos se logró conseguir un estímulo de 2,20 segundos, etiquetado con el número 11, y un estímulo de 2,30 segundos, etiquetado con el número 12.

Para la creación del último conjunto estímulos del continuo se procedió a aumentar un periodo a cada sílaba de la muestra de velocidad de habla lenta (de 2,40 segundos de duración) y a aumentar un milisegundo a las pausas silenciosas presentes en los fines de constituyentes menores. A través de estos procedimientos se logró conseguir un estímulo de 2,50 segundos de duración, etiquetado con el número 14; y un estímulo de 2,60 segundos de duración, etiquetado con el número 15. (Cf. Anexo 1)

### **3.2. La cláusula**

Para esta investigación decidimos utilizar la cláusula *veamos el ala del ave rapaz*. Esta cláusula fue seleccionada debido a que tanto el sintagma nominal como el adjetivo *rapaz* que lo compone, no suelen ser representativos del español venezolano, lo cual impide que factores semánticos se vean involucrados en la percepción y comprensión de la cláusula presentada en los tests.

---

<sup>5</sup> Este procedimiento fue obviado en el caso de las muestras de habla rápida y normal debido a que ellas carecían de este tipo de pausas.

El poco empleo del sintagma *ave rapaz*, y del adjetivo que lo conforma, en el español venezolano lo evidencian los datos del CREA (Corpus de Referencia del Español Actual); pues de 22 ejemplos de la cláusula reportados por este corpus, ninguno pertenece al español de la región, y de los 198 ejemplos del adjetivo rapaz registrados en el corpus, solo cuatro se presentan en Venezuela (ninguno de ellos en lengua oral).

### 3.3. Los oyentes

En la tarea de identificación participaron voluntariamente 69 oyentes, sin problemas auditivos. La edad promedio de los informantes fue de 24,4 años. Todos son habitantes de la ciudad de Mérida (Venezuela).

En la tarea de discriminación participaron 50 oyentes, sin problemas auditivos. La edad promedio de los informantes fue de 21,5 años. Todos son habitantes de la ciudad de Mérida (Venezuela).

Vale la pena agregar que quienes participaron en la tarea de identificación no participaron en la de discriminación, y viceversa.

### 3.4. Los tests

La aplicación de ambos tests se hizo en el Laboratorio de Fonética de la Universidad de Los Andes. Para ello se utilizaron altoparlantes a 62dB.

El test de identificación tuvo una duración aproximada de aplicación de 20 minutos. Para su realización se distribuyeron aleatoriamente los 15 estímulos del continuo creado. Cada estímulo fue presentado en serie de dos repeticiones con pausas silenciosas de 200 ms. entre cada repetición. La tarea específica que se le solicitó a los oyentes, después de escuchar cada una de las repeticiones de cada estímulo, fue evaluar cada uno según lo consideraran rápido, normal o lento.

También se realizó un test de discriminación tipo ABX, en el cual A y B eran dos estímulos diferentes, tomados del continuo inicial de 15 estímulos, y X podía ser A o B. Para crear los grupos se seleccionaron tanto estímulos pertenecientes, según los resultados del test anterior, a una misma categoría, como estímulos pertenecientes, según los resultados del test anterior, a categorías diferentes. De esta manera, se logró construir un conjunto de 18 estímulos, que fue posteriormente dupli-

cado con el fin de corroborar que las respuestas obtenidas no obedecieran al azar. Esta tarea dio como resultado 34 grupos, luego distribuidos en orden aleatorio. Cada grupo fue presentado en serie de dos repeticiones con pausas silenciosas de 200 ms. entre cada repetición. La tarea específica que se le solicitó a los 50 oyentes que participaron en este test, después de escuchar cada uno de los estímulos de cada grupo, fue determinar si se percibía alguna diferencia entre los tres estímulos que conformaban el grupo.

Ambos tests fueron analizados estadísticamente. Para ello se realizó un análisis de ítems mediante la prueba t de Student para muestras independientes con el fin de determinar la discriminación de ítems entre los grupos extremos a la media. Los resultados de la prueba demuestran que todos los ítems de los dos tests presentan poder discriminativo ( $p < ,05$ ) entre los grupos extremos a la media.

Para determinar la consistencia interna de cada uno de los ítems y la correlación con el total de los ítems que conforman los dos test se utilizó la prueba Alfa de Cronbach. En todos los casos las correlaciones de cada ítem con el total fue superior a ,10 (promedio = ,18). La consistencia interna (Alfa de Cronbach) de los tests fue alta = ,956, con una desviación típica de 10,261, una media de 79,35 y una varianza de 105,280. Según estos datos, estadísticamente, ambos tests son adecuados, esto es, miden lo que se pretende.

## 4. RESULTADOS

### 4.1. Test de identificación

En la tabla 1 se pueden apreciar los resultados en porcentajes obtenidos luego de la aplicación del test de identificación. La tarea que tenían los informantes era indicar qué tipo de velocidad tenía el tipo de estímulo que escuchaban. Para una mayor comprensión de los datos apuntados en los apartados siguientes se anexó también información no dada a los participantes en la prueba, a saber: información relacionada con la duración y la velocidad de habla de cada estímulo.

A partir de los datos arrojados por la aplicación del test podemos observar, en primer lugar, la existencia de elementos menos representativos de una velocidad específica. Tal como es el caso de los estímulos 5 y 6, los cuales fueron seleccionados en porcentajes similares tanto en la velocidad *rápida* como para la veloci-

dad *normal*. En segundo lugar, podemos observar la presencia de estímulos mayormente asociados con alguna de las tres velocidades, como es el caso de los estímulos entre 1 y 4, asociados a la velocidad de habla *rápida*, de los estímulos 8 y 9, asociados a la velocidad de habla *normal*, y de los estímulos 10, 11, 12, 13, 14 y 15, asociados a la velocidad de habla *lenta*.

ESTÍMULO			CATEGORÍAS		
Núm	Duración (seg)	Síl/seg	Rápida	Normal	Lenta
1	1,2	8,33	97%	3%	
2	1,3	7,69	96%	4%	
3	1,4	7,14	97%	3%	
4	1,5	6,67	88%	12%	
5	1,6	6,25	55%	45%	
6	1,7	5,88	48%	52%	
7	1,8	5,56	20%	68%	12%
8	1,9	5,26	12%	81%	7%
9	2	5,00		91%	9%
10	2,1	4,76		4%	96%
11	2,2	4,55			100%
12	2,3	4,35			100%
13	2,4	4,17			100%
14	2,5	4,00			100%
15	2,6	3,85			100%

Tabla 1. Duración y número de síl/seg por cada estímulo y percepción de los distintos tipos de velocidad de habla.

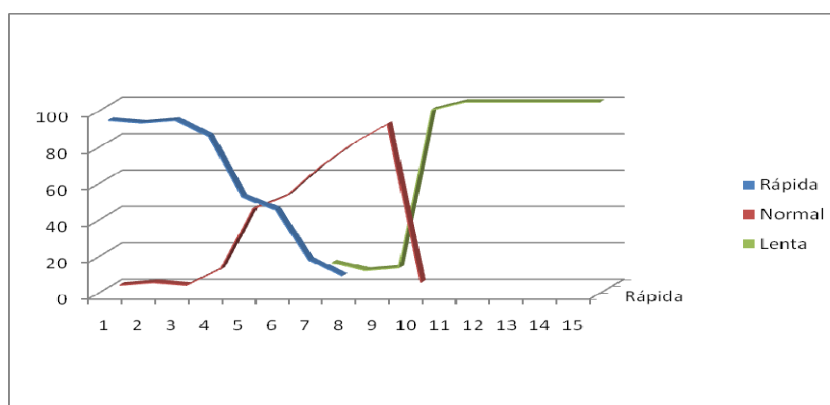


Figura 4. Clasificación de la velocidad de habla según las categorías: rápido, normal y lento.

La distribución de los estímulos antes descrita, y reflejada también en la figura 4, nos permiten sugerir que los oyentes clasifican los estímulos que oscilan entre las 8,33 síl/seg y las 6,67 síl/seg dentro de la categoría de velocidad de habla *rápida*. Los estímulos que oscilan entre las 5,56 y las 5,00 síl/seg dentro de la categoría *normal*. Y los estímulos que oscilan entre las 4,76 y las 3,85 síl/seg dentro de categoría de velocidad de habla *lenta*. En promedio, diremos que la velocidad de habla *normal* para el español de Mérida es aquella que tiene en su producción 5,27 sílabas por segundo<sup>6</sup>. La diferencia entre cada una de las categorías de velocidades es de 2,10 sílabas por segundo.

#### 4.2. Test de discriminación

En la tabla 2 se pueden apreciar los resultados en porcentajes obtenidos luego de la aplicación del test de discriminación. La tarea que tenían los informantes era indicar si percibían diferencias entre los tres estímulos que conformaban cada grupo. Para una mayor comprensión de los datos apuntados en los apartados siguientes se anexó también información no dada a los participantes en la prueba, a saber: información relacionada con la duración y la velocidad de habla de los estímulos A y B de cada grupo.

<sup>6</sup> Estos datos coinciden con los propuestos por Blondet (2006), desde la producción, para el español venezolano.



GRUPOS	DURACIÓN (SEG)		VELOCIDAD (SÍL/SEG)		NO DISCRIM. (%)
	ESTÍM. A	ESTÍM. B	ESTÍM. A	ESTÍM. B	
GRUPO 1: E1 E2 E2	1,2	1,3	8,33	7,69	96%
GRUPO 2: E3 E2 E3	1,4	1,3	7,14	7,69	98%
GRUPO 3: E3 E4 E3	1,4	1,5	7,14	6,67	94%
GRUPO 4: E4 E1 E4	1,5	1,2	6,67	8,33	17%
GRUPO 5: E1 E5 E1	1,2	1,6	8,33	6,25	4%
GRUPO 6: E5 E6 E5	1,6	1,7	6,25	5,88	13%
GRUPO 7: E3 E9 E3	1,4	2,0	7,14	5,00	4%
GRUPO 8: E6 E7 E6	1,7	1,8	5,88	5,56	91%
GRUPO 9: E7 E8 E8	1,8	1,9	5,56	5,26	91%
GRUPO 10: E8 E9 E8	1,9	2,0	5,26	5,00	94%
GRUPO 11: E9 E6 E9	2,0	1,7	5,00	5,88	2%
GRUPO 12: E9 E10 E10	2,0	2,1	5,00	4,76	4%
GRUPO 13: E10 E11 E10	2,1	2,2	4,76	4,55	94%
GRUPO 14: E11 E12 E11	2,2	2,3	4,55	4,35	94%
GRUPO 15: E12 E13 E12	2,3	2,4	4,35	4,17	91%
GRUPO 16: E13 E14 E13	2,4	2,5	4,17	4,00	91%
GRUPO 17: E15 E14 E15	2,6	2,5	3,85	4,00	100%
GRUPO 18: E10 E15 E10	2,1	2,6	4,76	3,85	85%

Tabla 2. Resultados del Test ABX que muestran la discriminación entre grupos de estímulos y las características temporales asociadas a cada grupo.

Los valores contenidos en el test de discriminación evidencian en primer lugar, que, como lo muestra la figura 5, los estímulos con menor posibilidad de ser discriminados corresponden a aquellos asignados en el test de identificación a una misma categoría y contiguos en una escala temporal, como por ejemplo los grupos 1, 2 y 3, asignados a la categoría *rápida*. Los grupos 8, 9 y 10 asignados a la categoría *normal* y los grupos 13, 14, 15, 16 y 17 asignados a la categoría *lenta*.

Como se muestra en la figura 5, los valores contenidos en el test de discriminación evidencian además la existencia de un conjunto de cinco grupos de estímulos discriminados por los oyentes. De dicho conjunto se puede además diferenciar entre dos subconjuntos, uno conformado por el grupo 7, el grupo 6 y el grupo 12; y otro subconjunto conformado por el grupo 4 y el grupo 11.

El primer subconjunto corresponde a tres grupos compuestos por estímulos identificados, en el primer test, como elementos pertenecientes a categorías diferentes. El grupo número 7, conformado por los estímulos 3 y 9, y el grupo 12, conformado por los estímulos 9 y 10, cuentan con la particularidad de estar integrados por elementos etiquetados como *rápidos* (en el caso del estímulo 3), *normales* (en el caso del estímulo 9) o *lentos* (en el caso del estímulo 10) por más del 90% de los encuestados del test de identificación. El último grupo de estímulos, etiquetado con el número 6, se caracteriza, por su parte, por estar integrado por elementos etiquetados como *rápidos* (en el caso del estímulo 5) y *normales* (en el caso del estímulo 6) por poco más del 50% de los encuestados del test de identificación.

Los dos grupos de estímulos pertenecientes al segundo subconjunto discriminado se componen, como lo evidencia la figura 5, de estímulos identificados en el test anterior dentro de una misma categoría. Cabe destacar que en uno de los casos (grupo 4) los elementos discriminados (estímulos 4 y 1) se caracterizan, por una parte, por ser estímulos identificados como pertenecientes a la categoría *rápida* por más del 88% de los encuestados, y por otra parte por constituir estímulos ubicados al extremo de las categorías. El grupo 11, por su parte, se compone tanto por un elemento (estímulo 9) etiquetado por más del 90% de los oyentes del test anterior como *normal*, y por un elemento (estímulo 6) asociado por solo el 52% de los oyentes a esta misma categoría.

La excepción a los fenómenos de discriminación descritos anteriormente corresponde al grupo 18, el cual, pese a contar con estímulos que se ubican en los extremos de la categoría, no son percibidos como diferentes por el 85% de los encuestados.

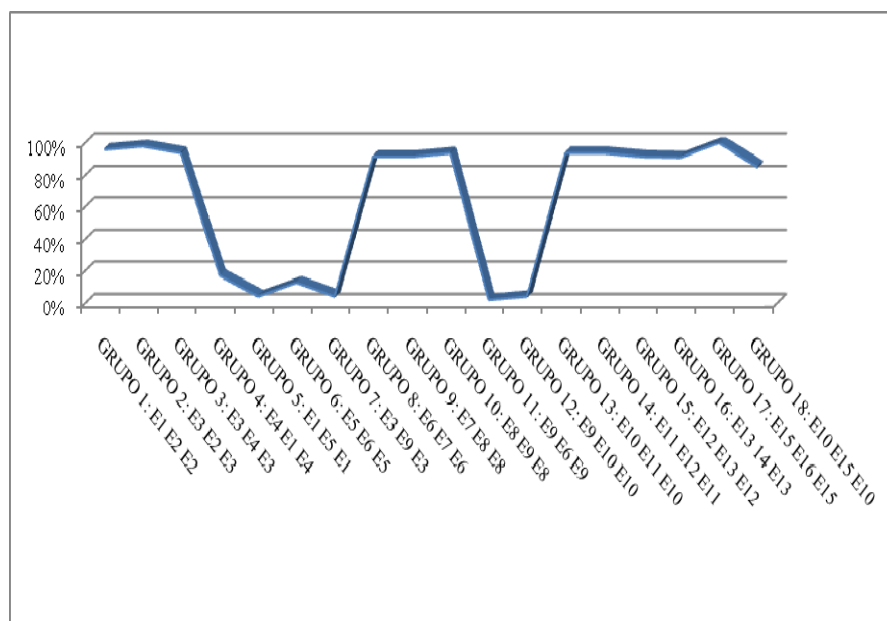


Figura 5. Resultados del test ABX

## 5. ANÁLISIS DE LOS RESULTADOS

Los datos arrojados por los tests presentados anteriormente, permiten sugerir tanto la presencia de elementos fuertemente asociados a una categoría, como la presencia de elementos asociados en menor medida a una categoría, lo cual deja ver que no todos los elementos del continuo obedecen a las características específicas de la percepción categorial. Pues, en primer lugar, como lo evidencian los datos de las categorías *rápida* y *normal* presentes en el test de identificación, no todos los estímulos ubicados al margen de las categorías alcanzan bruscamente un punto crítico a partir del cual se comienza a identificar el estímulo como perteneciente a otra categoría. Y en segundo lugar, como lo evidencia el test de discriminación además de existir la posibilidad de distinguir entre los sonidos lingüísticos pertenecientes a categorías diferentes también existe la posibilidad de distinguir entre estímulos ubicados al extremo de una misma categoría.

Por otra parte, cabe destacar que la existencia de elementos asociados en menor medida a una categoría específica, como el caso de los estímulos 5 y 6, parece sugerir más bien que, como afirma Martínez Celdrán (2001), existen dentro de las categorías elementos periféricos que en ocasiones pueden dar a los límites o fronteras de las mismas un carácter difuso o borrosos.

De igual manera, la presencia de estos fenómenos nos permite sugerir que existe una relación entre la cantidad de estímulo necesaria para distinguir entre categorías y el tipo de fenómeno presente en los límites de las mismas. Pues, cuando las categorías presentan límites difusos o borrosos se requiere de mayor cantidad de estímulo para distinguir entre ambas, mientras que cuando las fronteras presentan límites discretos la cantidad de estímulo requerido para distinguir una categoría de otra es menor.

Lo descrito anteriormente demuestra que las diferencias mínimas perceptibles varían en función de las categorías existentes. Así por ejemplo, mientras que las diferencias mínimas necesarias para distinguir entre los elementos hallados en el límite de las categorías *rápida* y *lenta* se sitúan en los 1,11<sup>7</sup> síl/seg aproximadamente, las diferencias mínimas perceptibles necesarias para distinguir entre los elementos hallados en el límite de las categorías *normal* y *lenta* se sitúan en los 0,24 síl/seg.

Otro dato arrojado en la investigación es el referente a la relación entre percepción y producción del habla. Dicha relación la podemos notar en los estímulos generalmente asociados a una categoría por la mayoría de los encuestados en el test de identificación y su semejanza con los distintos temporemas señalados desde la producción por la (Blondet, 2006). En este sentido, se puede observar que las velocidades identificados por nuestros oyentes para la categoría *rápida* (8,33 – 6,67 síl/seg), para la categoría *normal* (5,56 – 5,00 síl/seg) y para la categoría *lenta* (4,55 – 3,85 síl/seg); se enmarcan dentro de los márgenes de seguridad registrados por Blondet (2006) para el *temporema rápido* (7,1 - 6 síl/seg), para el *temporema normal* (5,3 - 4,7 síl/seg) y para el *temporema lento* (4,1 - 3,49 síl/seg). Un hecho que viene a confirmar lo dicho por investigadores como Liberman y Mattingly (1985) y Martínez (2009): existe una estrecha relación de dependencia entre la producción del habla y su percepción, y esta se refleja en la presencia del mismo conjunto de invariantes tanto en la percepción como en la producción del habla.

---

<sup>7</sup> Diferencia mínima del estímulo para diferenciar una categoría de otra: la resta entre el elemento final de una categoría y el elemento principal de la siguiente categoría.

El estímulo 10, con una velocidad de 4,76 síl/seg, constituye por su parte la excepción a la regla, pues a pesar de haber sido catalogado como *lento* por el 96% de los oyentes encargados de realizar la tarea de identificación se encuentra dentro de los rangos que rigen la categoría *normal* registrada por Blondet (2006).

## 6. CONCLUSIONES

De lo dicho anteriormente podemos concluir que, aunque este estudio constituye un primer acercamiento al fenómeno de la percepción de la velocidad del habla, es posible:

1. Ratificar la existencia de categorías, en tanto que mecanismos usados para la organización de la velocidad de habla, y en tanto que método utilizado para simplificar los distintos estímulos presentados a partir de procedimientos como la generalización y la discriminación.
2. Comprobar la existencia de categorías con límites no discretas que, como en el caso de los límites que dividen las categorías *normal* y *lenta*, pueden estar sujetos a distintas gradaciones hasta llegar a un punto óptimo en el que la mayoría de los oyentes comienza a detectar el estímulo como *normal*.
3. Comprobar la posibilidad de identificar las diferencias mínimas perceptibles que dividen la velocidad de habla *rápida*, *normal* y *lenta* en el español de Mérida a través de análisis perceptivos.
4. Sugerir, a partir de los datos registrados en el test de discriminación, que no parece ser condición indispensable para los oyentes la no discriminación de los estímulos para incluirlos dentro de una misma categoría.
5. Sugerir, como lo indican Liberman y Mattingly (1985) y Martínez (2009), la existencia de una estrecha relación entre los procesos de producción y percepción del habla.

*AGRADECIMIENTOS: Esta investigación ha sido financiada por el Consejo de Desarrollo Científico, Humanístico, Tecnológico y de las Artes (CDCHTA) de la Universidad de Los Andes a través del proyecto H-1330-10-06-EM.*

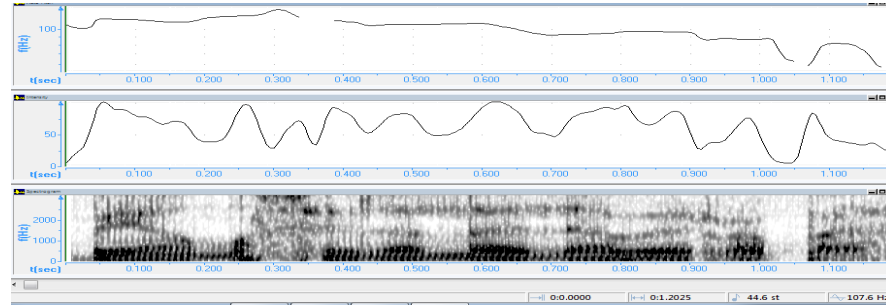
## 7. REFERENCIAS BIBLIOGRÁFICAS

- BLONDET, M. A. (2006): *Variaciones de la velocidad de habla en español: patrones fonéticos y estrategias fonológicas. Un estudio desde la producción*, tesis doctoral, Universidad de Los Andes, Mérida.
- BRUNER, J. (1984): *Acción, pensamiento y lenguaje*, Madrid, Alianza.
- JUSCZYK, P. y P. LUCE (2002): «Speech Perception and Spoken Word Recognition», *Ear & Hearing*.  
<http://www.cse.buffalo.edu/~rapaport/575/F07/JusczykLuce2002.pdf>  
[12/12/2010].
- LIBERMAN, A.; K. HARRIS; H. HOFFMAN y B. GRIFFITH (1957): «The discrimination of speech sounds within and across phoneme boundaries», *Journal of experimental psychology*, 65, pp. 497-516.
- LIBERMAN, A.; F. COOPER; D. SHANKWEILER y M. STUDDERT KENNEDY (1967): «Perception of the Speech Code», *Psychological Review* 74, 6, pp. 431-459.
- LIBERMAN, A. e I. MATTINGLY (1985): «The Motor Theory of Speech Perception Revised», *Cognition* 21, 1, pp.1-36.
- LOPERETE, C. (1984): *El lenguaje oral: fundamentos, formas y técnicas*, Buenos Aires, Plus Ultra.
- MARRERO, V. (2001): *Fonética perceptiva*, Madrid, Universidad Nacional de Educación a Distancia.
- MARTÍNEZ, H. (2009): *Evaluación de la percepción y de la producción del habla*, tesis doctoral, Universidad de Los Andes, Mérida.
- MARTÍNEZ CELDRÁN, E. (2002): «Aproximación a una fonética no discreta», *Actas del II Congreso de Fonética Experimental*, 2001, Sevilla, Universidad de Sevilla.
- MORA, E.; M. RODRÍGUEZ y C. CAVÉ (2007): «La duración de la vocal simple y de las homólogas en el español venezolano: el caso de pares mínimos», *Lenguaje*, 35, 1, pp. 17-25.

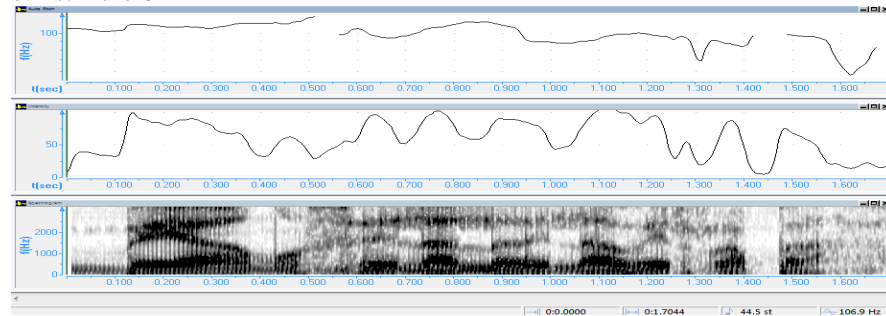
- MATLIN, M. y H. FOLEY (1996): *Sensación y percepción*, México, Prentice Hall.
- NAVARRO, T. (1982): *Manual de pronunciación española*, Madrid, Revista de Filología española.
- OBEDIENTE, E. (1998): *Fonética y fonología*, Mérida, Universidad de Los Andes.
- WAINSCHEKER, R.; J. DOORN y M. CASTRO. (2002): «Medición cuantitativa de la velocidad de habla», *Procesamiento del lenguaje Natural*, 28, pp. 99–104.
- YENI-KOMSHIAN, G. (1999): «Percepción del habla», en N. Brenstein, J. Berko Gleason, B. Narasimhan (eds.): *Psicolingüística*, Madrid, MacGraw Hill, pp. 113-167.

**ANEXO 1. CURVA MELÓDICA, CURVA DE INTENSIDAD Y ESPECTRO DE TRES ESTÍMULOS MODIFICADOS**

a. Estímulo 1



b. Estímulo 6



c. Estímulo 15

