

Revista Mexicana de Ingeniería Biomédica

Volumen
Volume 25

Número
Number 1

Marzo
March 2004

Artículo:

Estudio del efecto coarticulatorio en el habla

Derechos reservados, Copyright © 2004:
Sociedad Mexicana de Ingeniería Biomédica, AC

Otras secciones de
este sitio:

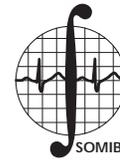
- 👉 Índice de este número
- 👉 Más revistas
- 👉 Búsqueda

*Others sections in
this web site:*

- 👉 *Contents of this number*
- 👉 *More journals*
- 👉 *Search*



Medigraphic.com



Estudio del efecto coarticulatorio en el habla

Martínez Liconá A.E.,*
Martínez Liconá F.M.,* Vidal
Cabrerá O.O.,* Goddard
Close J.C.*

* Departamento de Ingeniería Eléctrica,
Universidad Autónoma Metropolitana,
Iztapalapa, México.

Correspondencia:
Alma Edith Martínez Liconá.
Universidad Autónoma Metropolitana,
Av. San Rafael Atlixco 186 Col. Vicentina,
Iztapalapa D.F. CP 09340,
Tel: 58044635 ext. 212,
fax: 58044640
email:aaml@xanum.uam.mx

RESUMEN

La producción del habla es un proceso complejo de estudiar; dentro de las aproximaciones implementadas la espectrográfica muestra características que pueden reflejar diferencias importantes dentro de las señales de la voz. Dentro de estas características se encuentra la coarticulación, la cual presenta variaciones fáciles de identificar acústicamente, pero difíciles de caracterizar dentro del contexto del reconocimiento automático del habla. Este trabajo presenta un estudio del efecto coarticulatorio en el español. Partiendo de la definición de coarticulación y de una breve explicación de la representación espectral, se presenta la caracterización de la pronunciación de hispanohablantes de América Latina, así como de las diferencias en la pronunciación de frases obtenidas de la base de datos Latino40.

Palabras clave:

Coarticulación, representación espectral del habla.

ABSTRACT

Speech production is a complex process for study; among the approximations implemented, the spectrographic one shows characteristics that may reflect important differences in speech signals. One of these characteristics is coarticulation which shows acoustic variations easily identified, but at the same time, hardly described in an automatic speech recognition frame. This is a study of the coarticulatory effect in Spanish speech. From the coarticulation definition and a brief explanation of spectral representation, pronunciation characteristics of Spanish speakers from Latin America are shown as well as pronunciation differences in statements of Latino40 database.

Key Words:

Coarticulation, Speech spectral representation.

INTRODUCCIÓN

La producción del habla se puede analizar desde diversos puntos de vista dentro de los cuales el acústico es de los más importantes. Acústicamente la voz se ve como una fuente de sonido cuyas propiedades se determinan mediante un filtro selectivo, el cual representa al tracto vocal. Las propiedades del filtro varían en el tiempo durante el proceso de la produc-

ción del habla ya que están en función de la geometría del tracto vocal, el cual a su vez está en función de los movimientos y posiciones de los órganos de articulación como la lengua o los labios.

El efecto de coarticulación se observa en la gama de variaciones que existe en la pronunciación de palabras, siendo interesante la caracterización de esas diferencias, en particular, entre las personas de habla normal y las que

presentan problemas en la emisión de sonidos o en la articulación de palabras. El presente trabajo presenta un estudio del efecto coarticulatorio desde la perspectiva espectral de los sonidos del español. A partir del concepto de coarticulación se determinan algunas características regionales interesantes del español presentándolas en el dominio de la frecuencia y se plantean algunas reglas de diferenciación tomando como base hispanohablantes de diferentes países de Latinoamérica.

EL CONCEPTO DE COARTICULACIÓN

El fonema es la unidad mínima sonora de que se componen las palabras¹, la emisión de fonemas se forma por el movimiento de los articuladores de la voz que incluyen: cuerdas vocales, paladar, lengua, dientes, labios, úvula y quijada. Las limitantes mecánicas y neuronales en el movimiento de un articulador pueden causar variedad de influencias en los fonemas vecinos. Por otro lado se dice que los fonemas tienden a ser abstracciones implícitamente definidas por la pronunciación de palabras en un lenguaje. La forma acústica de un fonema depende fuertemente del contexto acústico en el que sucede, es decir, la manera en que se colocan los órganos de articulación hace que un mismo sonido se pronuncie de diferente manera; a este efecto se le denomina coarticulación.

Desde el punto de vista espectral sucede lo siguiente: una vez elegido el fonema como unidad de reconocimiento, su espectro dentro de la palabra o enunciado se ve influido por los fonemas adyacentes. La representación espectral presentará variaciones dependiendo de su entorno; lo mismo ocurre en el caso de fonemas aislados o sílabas. Dentro de la articulación de la voz en el lenguaje natural, la reducción u omisión de fonemas e incluso sílabas es uno de los efectos que se presentan con mayor frecuencia¹. Esto se debe a que el paso de un fragmento a otro no se da de manera espontánea ya que se articula el segundo antes de finalizar el primero; esta transición ligera entre fragmentos se debe al efecto coarticulatorio. Lo que para una persona en una conversación no representa ninguna dificultad, para un sistema automático supone conocer todas las posibles variantes de pronunciación; de no ser así habrá que asumir la imposibilidad de identificar una palabra y por tanto perder su información².

CARACTERÍSTICAS DEL IDIOMA ESPAÑOL

El idioma español está formado por diferentes clases de sonidos. Existen sonidos oclusivos que se articulan mediante una obstrucción total del aire en un primer momento para después liberarlo de manera repentina. Los sonidos fricativos son aquellos que se articulan mediante la obstrucción parcial de aire, el cual sale gradualmente mientras que los nasales se caracterizan porque la salida del aire se hace por la nariz siendo la única clase de sonidos que no se articulan en la boca. Hay sonidos similares a los fricativos en donde el aire sale con mayor libertad y se presentan algunas resonancias; este tipo de sonidos recibe el nombre de líquidos. El español se caracteriza por la sonoridad de sus fonemas; a diferencia de otros idiomas, los sonidos "sonoros" se asocian a cinco vocales principales cuyas características de sonoridad son muy particulares. Las vocales son sonidos emitidos sin ningún tipo de obstrucción por parte de los articuladores; es muy fácil distinguirlas atendiendo a la localización de las formantes, primera y segunda frecuencias que aparecen en la representación espectral, ya que su separación es apreciable.

Si bien existen diversos sistemas para la representación de los sonidos en los diferentes idiomas como IPA³ y WORDBET⁴, para la transcripción de los símbolos del alfabeto se hizo uso de los símbolos codificados en SAMPA (Speech Assessment Methods Phonetic Alphabet) para el español⁵. SAMPA es un alfabeto fonético; en una primera etapa se aplicó a varios idiomas europeos para posteriormente incorporar el español. Asociado con la codificación se presentan las pautas para la transcripción de los idiomas a los que SAMPA se ha aplicado. Los símbolos de transcripción de SAMPA se han desarrollado con la ayuda de expertos nativos de cada idioma y se regulariza internacionalmente⁶.

En las diversas versiones del idioma español hay muchos fonemas que en otros idiomas no existen, por lo tanto hay palabras difíciles de tratar por SAMPA, en este sentido se tuvieron que crear reglas para dichos fonemas para elaborar las transcripciones como lo muestran los ejemplos del Cuadro 1⁷:

Otro de los aspectos importantes a tratar es el de los acentos. Mientras que en idiomas como el inglés éstos no existen, en español su emisión en la pronunciación hacen diferencias entre las palabras; además si se consideran los "acentos re-

Cuadro 1. Fonemas particulares del español mexicano.

Fonema	Ejemplo	Transcripción
S	Xocoyote	Soco"jjote
C	Cojin	coC'in
Jj	Ayer	a'jjer
T	Zona	Tona
W	Agua	AGwa

gionales" que se presentan en las diversas áreas de habla hispana, el efecto coarticulatorio se observa con mayor claridad. El estudio de la coarticulación está muy relacionado con el contexto de la palabra o frase, ya que la pronunciación de algunos fonemas dependen de los fonemas que le preceden o que le siguen. De esta manera es necesario definir reglas específicas para el estudio de la coarticulación en el idioma español tomando en cuenta aspectos como el contexto y el acento.

ANÁLISIS DE FONEMAS AISLADOS DESDE EL PUNTO DE VISTA ESPECTRAL

Dentro de las técnicas de análisis del habla, el análisis espectral encuentra tanto las frecuencias presentes en la señal como los tiempos en que éstas aparecen. En su definición se supone un conocimiento de la señal para todo instante de tiempo y que cualquier propiedad o característica que se está buscando mediante su aplicación (espectro, resonancias, etc.) se mantiene invariante para todo el tiempo que ésta dure. La señal de voz es muy variante en el tiempo, cada periodo de tiempo de unos pocos milisegundos puede presentar cambios significativos en la señal de voz. Esto hace que en la mayoría de las aplicaciones haya más interés en conocer las propiedades de la señal de voz en intervalos pequeños de tiempo que en toda la señal. En este caso se habla del análisis localizado⁸.

El resultado de este análisis es un espectro. Una vez calculado el espectro para una sección corta de señal o ventana, típicamente 5 a 20 milisegundos de conversación, se calcula el espectro para la ventana inmediata, y así sucesivamente hasta llegar al final de la señal. Los espectros vecinos pueden variar en función de los movimientos ocasionados por el contexto del tracto vocal y estos movimientos dependen de la ventana de análisis. La integración en un solo gráfico de los

espectros por ventanas da como resultado el espectrograma.

El espectrograma es una representación de las variaciones de la frecuencia, las cuales se muestran en el eje vertical, y la intensidad en la señal de voz a lo largo del tiempo, en el eje horizontal. El espectrograma sirve para analizar diversas características de la señal del habla como la sonoridad, la duración, y la estructura de las frecuencias formantes, el timbre, la intensidad, las pausas, el acento y el ritmo⁹. Normalmente se calcula y guarda en la memoria de la computadora como una serie bidimensional de valores de energía acústicos. Para un espectrograma dado S , la energía de un componente de frecuencia dado w en un momento dado t es representado por la tonalidad en la escala de grises o color del punto correspondiente $S(t, f)$. El cálculo del espectrograma da lugar a una función compleja cuya expresión es:

$$S_w = \frac{1}{P} \sum_{n=0}^{P-1} s_n e^{jwn}$$

donde al digitalizar la señal el tiempo se vuelve una dimensión de número de muestras (n) y donde P representa el número de muestras por ventana.

A continuación se presenta el análisis de dos hablantes de diferentes países latinoamericanos en función del espectrograma obtenido en ejemplos de palabras donde se realzan las particularidades del acento regional de cada uno. En estos casos el efecto de coarticulación puede notarse en las diferencias con respecto al español de la zona de Castilla como se reporta en⁷:

I. ARGENTINA

1. Fenómeno del yeísmo: en la mayoría de las regiones argentinas se presenta el fenómeno de yeísmo que consiste en pronunciar el sonido /l/ denominado lateral sonoro por el sonido /jj/, fricativa palatal sonora (ver⁵), lo que resulta en un fonema con características "ruidosas". En el caso del acento argentino, en particular el de Buenos Aires, ambos sonidos se transforman en fricativa palatoalveolar sonora /Z/, fenómeno denominado zeísmo. En la palabra "millares" de la Figura 1, el fonema /l/ ('ll') presenta características espectrales de energía considerable a frecuencias altas mientras que a frecuencias

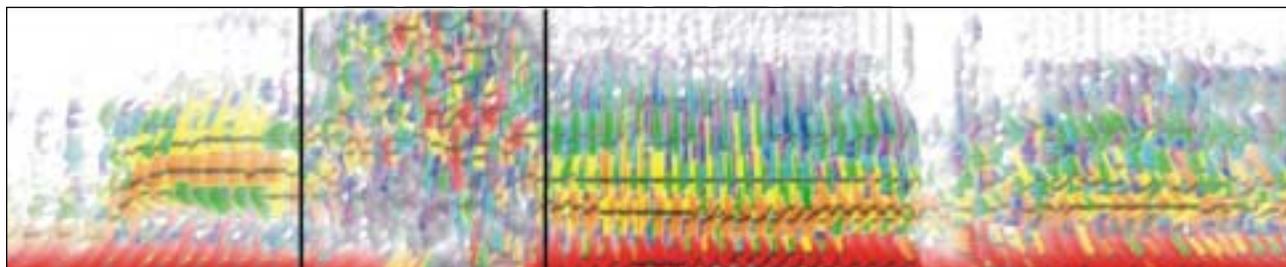


Figura 1. Espectrograma de la palabra "millares".

bajas se notan demasiado suaves. Sin llegar a ser completamente fricativa la presencia de estos componentes espectrales denotan la modulación dada por los articuladores.

2. El sonido fricativo alveolar sordo /s/ se vuelve fricativo glotal sordo /h/. No se presentan las componentes que dan la apariencia ruidosa de naturaleza alveolar en el espectro. Más bien se encuentran niveles bajos de energía obtenidos por la aportación de la glotis a la emisión del sonido. En la palabra "dos" de la Figura 2, la /s/ presenta estas características y dado que el fonema se presenta al final de palabra, se espera que la energía sea de menor intensidad.

II. COLOMBIA

1. La fricativa velar sorda /x/ se pronuncia como fricativa glotal sorda /h/. En el caso de la palabra "jornada" presentada en la Figura 3, el so-

nido /x/ (/j/ de jamón) tiene una apariencia similar a una fricativa sorda, es decir, con discretos niveles de energía a frecuencias bajas, y una ligera percepción. Dado que el siguiente sonido es una vocal se puede notar la diferencia, pero no siempre se da este caso.

2. Las oclusivas sordas /d/ y /b/ y /g/ se pronuncian como oclusivas excepto en la posición postnuclear o en el inicio de una sílaba que comienza con vocal. En estos casos se presentan los alófonos sonoros /B/, /D/, /G/. En la palabra "partidos" de la Figura 4 se encuentra que la /d/ de la sílaba "dos" presenta las características oclusiva y plosiva claramente.

LA BASE DE DATOS LATINO-40

El laboratorio Entropic Research diseñó la base de datos Latino-40 con el objetivo de obtener un conjunto de frases o enunciados que sirvieran de en-

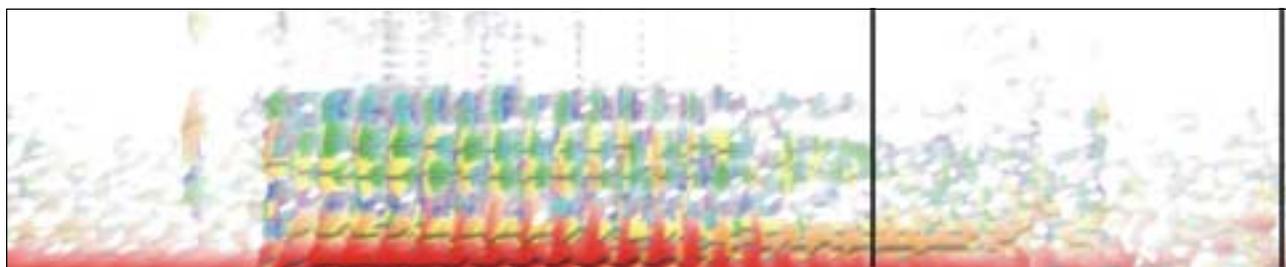


Figura 2. Espectrograma de la palabra "dos".

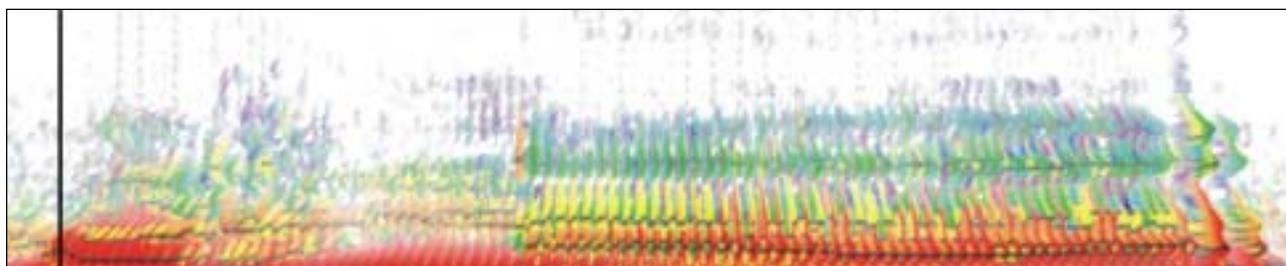


Figura 3. Espectrograma de la palabra "jornada", hablante de Colombia.

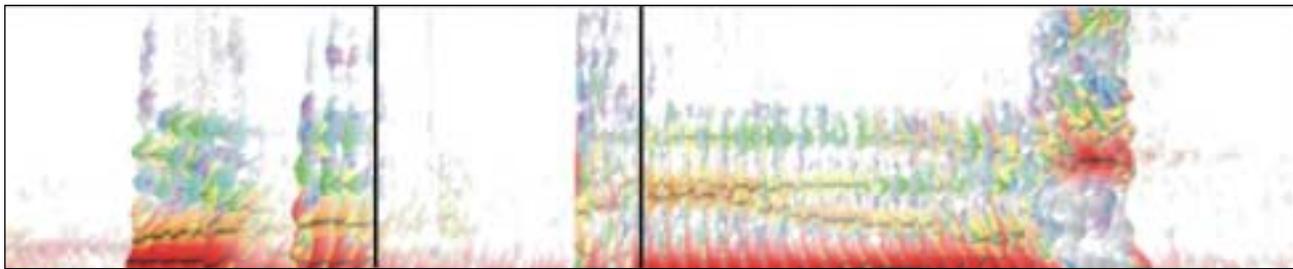


Figura 4. Espectrograma de la palabra “partidos”, hablante de Colombia.

trenamiento a sistemas de reconocimiento del habla independientes del hablante que reconozcan el español latinoamericano¹⁰. La base de datos resultante, denominada Entropic Latino40, consta de 40 conjuntos de 125 enunciados para un total de 5,000 archivos de voz. Cada conjunto de enunciados fue asignado a uno de los 40 hablantes, 20 femeninos y 20 masculinos, que formaron parte del proyecto. Las grabaciones se realizaron con un micrófono de alta calidad en un ambiente de oficina y la grabación de las emisiones se digitalizaron a 16-bits y una frecuencia de muestreo de 16 kHz.

Hablantes

Los hablantes participantes fueron voluntarios de diferentes países de Latinoamérica, cuyas edades iban entre 18 y 59 años y la lengua materna era el español. Siete de los hablantes eran de Perú, cinco de cada uno de los siguientes países: Argentina, Colombia, Guatemala, y Nicaragua, tres de Venezuela y dos de Chile, Costa Rica, Cuba, El Salvador y México.

Procedimiento de grabación

Los hablantes se sentaron en una silla frente a una estación de trabajo de Silicon Graphics, Indy (SGI).

Cada hablante fue instruido en el proceso de grabación; se le indicó como controlar el software utilizado, manejar apropiadamente el micrófono y evaluar la calidad de las frases grabadas. El control de las grabaciones se hizo mediante el uso de un botón que iniciaba la grabación al tiempo que desplegaba el texto en español que iba a ser leído. La grabación de la frase terminaba cuando se accionaba un botón que daba acceso inmediato al siguiente enunciado y reiniciaba el proceso. El hablante tuvo un control completo del proceso de grabación lo cual le permitía volver a grabar enunciados anteriores si así lo deseaba, así

como moverse a lo largo de la base de datos que estaban construyendo.

ANÁLISIS DE LA COARTICULACIÓN EN PALABRAS EMITIDAS POR DIFERENTES HABLANTES

A fin de apreciar el fenómeno de coarticulación con mayor claridad se describen las diferencias en la pronunciación de diversos hablantes en un conjunto de cinco palabras. Se presentan los espectrogramas de las palabras aisladas enunciadas en los contextos de las frases que se indican así como el país de cada hablante.

Transporte

1. Esta palabra tiene dos oclusivas sordas /t/ y /p/ cuyas características se ven con claridad en la hablante de Argentina (Figura 5).
2. La característica fricativa de la /s/ se observa en todas las pronunciaciones siendo la más fuerte la del hablante de Nicaragua (Figura 6).
3. En los hablantes de Nicaragua y Perú se nota levemente la presencia de la /n/ mientras que en la de Argentina no se ve nada (Figuras 6 y 7).
4. La /t/ de la segunda sílaba de la palabra transporte, /por/, se puede observar en todos los casos con cierta similitud, no así la /t/ de la sílaba /trans/ que dada su corta duración se ve opacada por los fonemas de mayor energía, en este caso la /a/.
5. Para segmentar fonéticamente algunos de los sonidos no permiten su identificación clara. En este caso es más fácil identificar la sílaba compleja /trans/ por la característica de la /s/ presente en todos los casos.

Gastos

1. En todos los casos no es posible ver la presencia de la oclusiva /g/; dada su naturaleza “sor-



Figura 5. Espectrograma de "transporte" en la frase "No se han completado los arreglos para el transporte", hablante femenino de Argentina.



Figura 6. Espectrograma de "transporte" en la frase "También se incluye el transporte espacial", hablante masculino de Nicaragua.

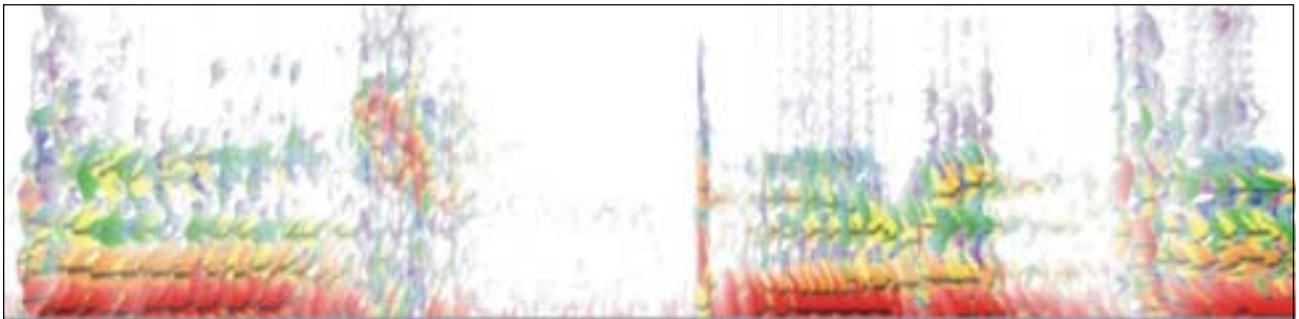


Figura 7. Espectrograma de "transporte" en la frase "También se incluye el transporte por vía lacustre", hablante masculino de Perú.

- da", dura muy poco tiempo y su espectro se opaca por el de la vocal /a/
2. La pronunciación de /a/ es más clara en la hablante de Colombia, mientras que las hablantes de Argentina y Guatemala la pronuncian de una manera muy suave y corta (Figura 9).
 3. La característica fricativa de /s/ es clara en las hablantes de Colombia y Guatemala, mientras que en la de Argentina el fonema se presenta muy suave y con energías a bajas frecuencias.
 4. El efecto de oclusión se nota en todas las hablantes variando la duración y la claridad ya que mientras que en las hablantes de Argentina y Guatemala se ve con claridad, éste en la hablante de Colombia se muestra más discreto (Figuras 8 y 10).
 5. Algo similar ocurre con la vocal /o/ donde para la hablante de Colombia se presenta con sus formantes claramente mientras que para las de

Argentina y Guatemala el sonido se muestra más difuso. En el caso de Argentina, la pronunciación de la palabra presenta un decremento en la energía de la última sílaba comparada con la primera (Figura 8).

6. Con respecto a la última /s/ hay casos en los que su emisión es muy débil y de corta duración, como en el caso de Colombia y Guatemala, mientras que en la emisión argentina se logra apreciar el efecto fricativo.

Viene

Esta palabra tiene como componentes una oclusiva sonora, un diptongo, una nasal y una vocal. Los espectrogramas muestran que es muy difícil separar los sonidos, ya que algunos de ellos duran poco tiempo mientras que otros tienen una transición tan ligera que no se llega a notar el cambio.

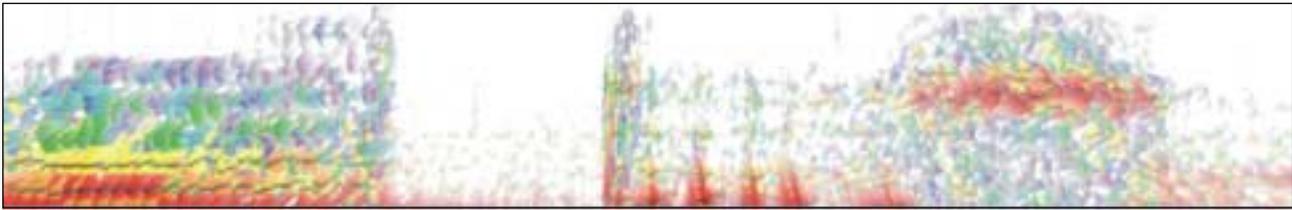


Figura 8. Espectrograma de “gastos” en la frase “ Como consecuencia hemos incurrido en grandes gastos”, hablante femenino de Argentina.

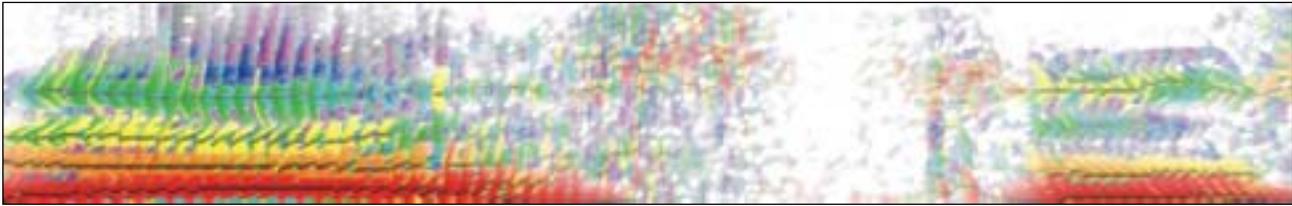


Figura 9. Espectrograma de “gastos” en la frase “Los gastos y los beneficios resultan importante”, hablante femenino de Colombia.

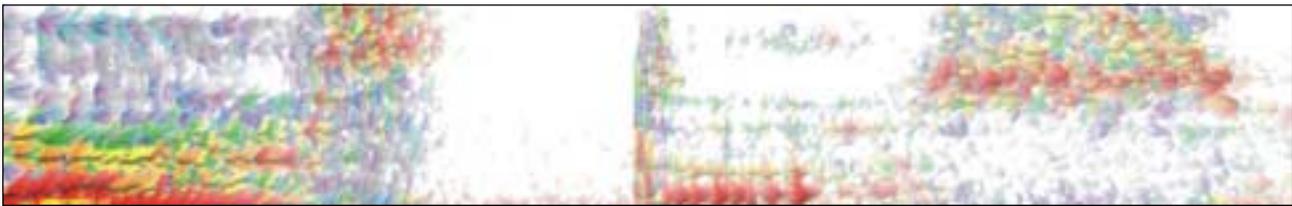


Figura 10. Espectrograma de “gastos” en la frase “Como consecuencia hemos incurrido en grandes gastos”, hablante femenino de Guatemala.

1. De los tres ejemplos el del hablante de Guatemala presenta mayor dificultad ya que, además de la baja energía que el espectrograma presenta, prácticamente no se logra observar ningún componente. Por otro lado el de Cuba también presenta dificultad para el análisis, aunque su nivel de energía es mayor (Figuras 11 y 12).
2. Para la hablante de Perú se presentan los mismos problemas aunque pareciera que lo único que se ve con claridad es el diptongo /ie/ (Figura 13).
3. En el caso del hablante de Chile se presentan dificultades para distinguir la /b/ y la /t/; las vocales sí logran diferenciarse (Figura 15).
4. La hablante de México muestra componentes claros en la segunda parte de la palabra. Se logran diferenciar los sonidos /a/, /s/, /t/ y /a/, aunque la /u/ se confunde con la /b/ (Figura 16).

Subasta

1. En los tres casos el espectrograma permite ver los límites entre los sonidos involucrados.
2. En el espectrograma de la hablante de Guatemala la /t/ no logra presentar la característica oclusiva ya que presenta leves niveles de energía en todo momento; esto es suficiente para diferenciarla de la /s/ y la /a/ (Figura 14).

Restaurados

Para esta palabra se presentan tres formas distintas derivadas de sendas pronunciaciones diferenciadas.

1. El espectrograma de la hablante de Nicaragua no presenta la componente de la /t/ inicial, el sonido queda cubierto por la /e/; el cambio entre vocales /au/ se ve con claridad y el sonido oclusivo sonoro /d/ se pierde de nuevo ante la presencia de la vocal (Figura 17).
2. En el caso del hablante del Salvador su gráfica es la más clara del conjunto ya que logran distinguirse todos los fonemas involucrados, tanto

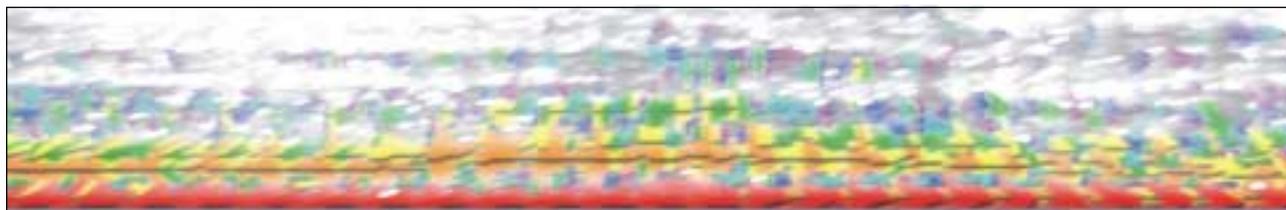


Figura 11. Espectrograma de "viene" en la frase "Cada vez que viene la registran, cuando llega y cuando se va", hablante masculino de Cuba.

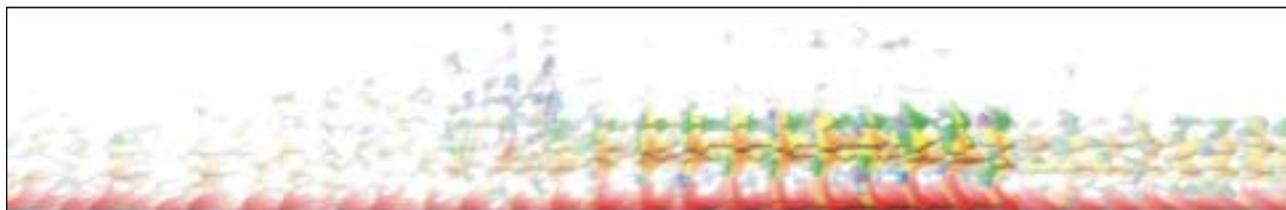


Figura 12. Espectrograma de "viene" en la frase "Esto no se va a quedar así, viene la revancha, agregó", hablante masculino de Guatemala.

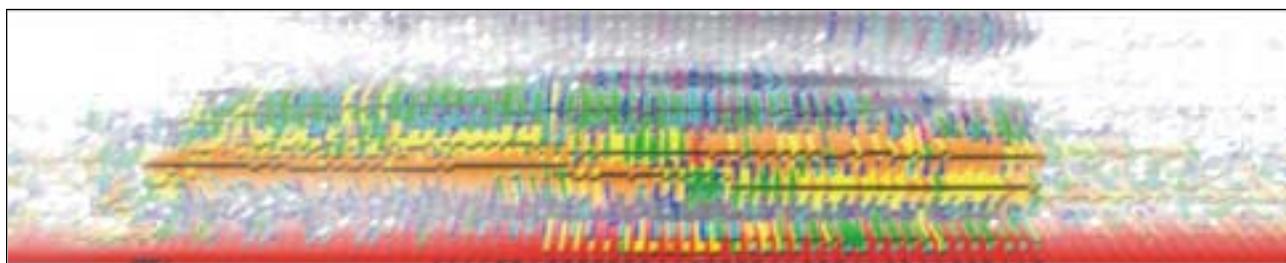


Figura 13. Espectrograma de "viene" en la frase "¿Por qué viene a un país a buscar fortuna?", hablante femenino de Perú.



Figura 14. Espectrograma de "subasta" en la frase "En la subasta de tierras no hubo cambio de dinero", hablante femenino de Guatemala.

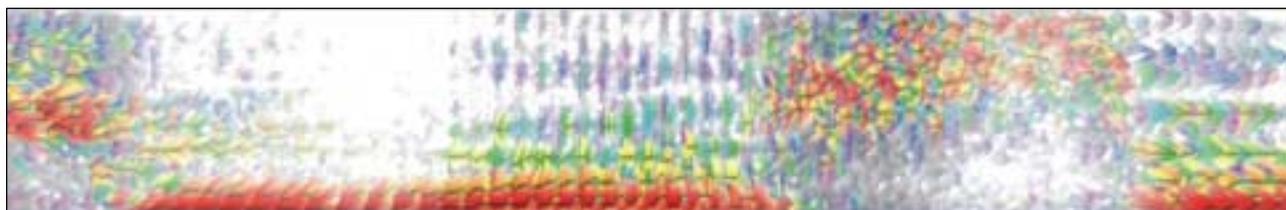


Figura 15. Espectrograma de "subasta" en la frase "Empresa peruana gana subasta de fábrica de cemento", hablante femenino de Chile.

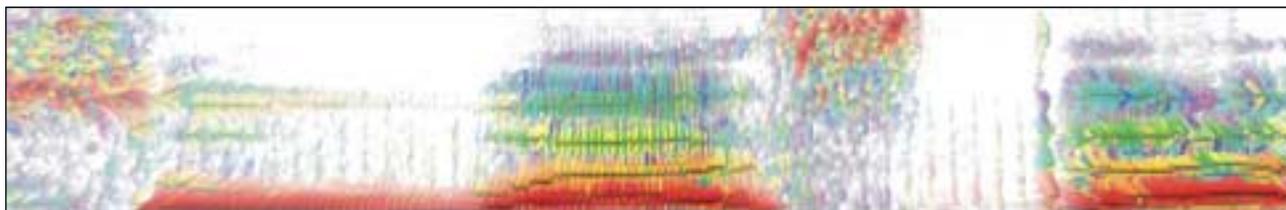


Figura 16. Espectrograma de “subasta” en la frase “Empresa peruana gana subasta de fábrica de cemento”, hablante femenino de México.

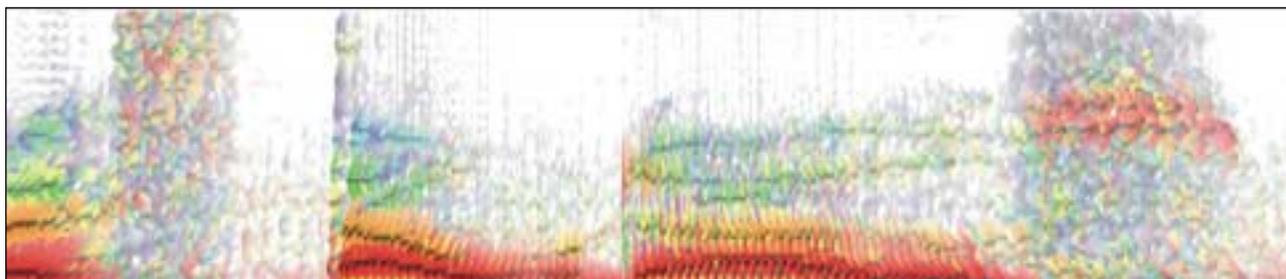


Figura 17. Espectrograma “restaurados” en la frase “Los templos dañados han sido restaurados”, hablante femenino de Nicaragua.

en características espectrales como en duración (Figura 18).

3. El hablante de Perú presenta el espectrograma más difuso de los tres puesto que no se ve ninguna estructura similar a las de los ejemplos anteriores, ni siquiera en los más evidentes como las vocales (Figura 19).

CONCLUSIONES

El estudio espectral de los hablantes pertenecientes a diferentes regiones latinoamericanas permite observar la existencia de ciertos patrones de pronunciación que pueden llegar a un buen grado de generalización.

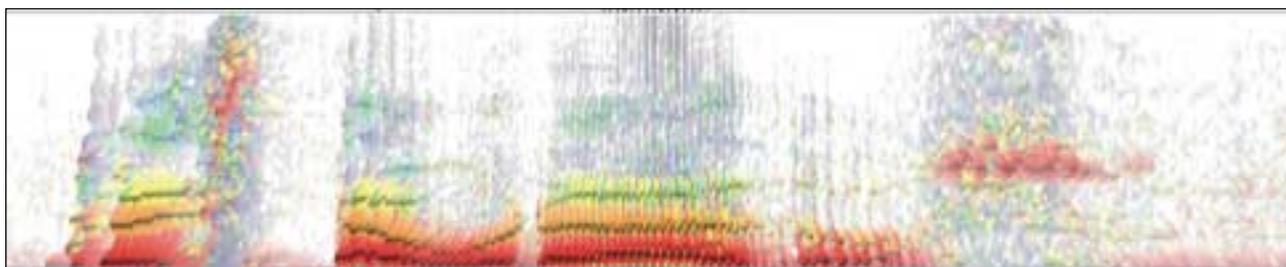


Figura 18. Espectrograma “restaurados” en la frase “Los templos dañados han sido restaurados”, hablante masculino de El Salvador.

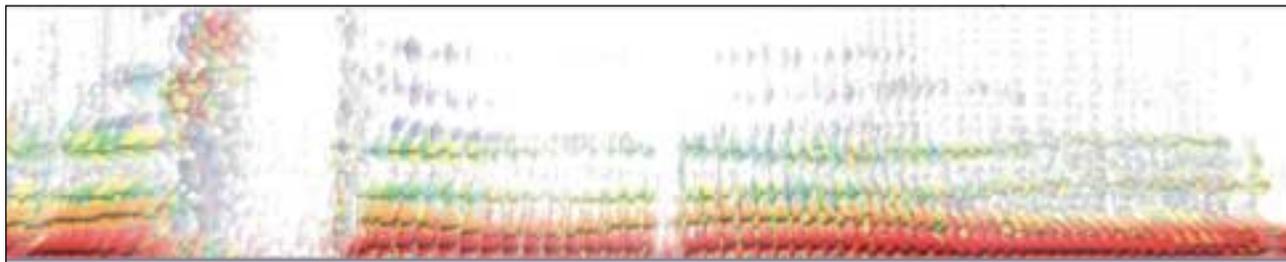


Figura 19. Espectrograma “restaurados” en la frase “Sin embargo, la paz no se ha restaurado aún”, hablante masculino de Perú.

Uno de los más claros se relaciona con la intensidad en la pronunciación de las palabras; en la mayoría de los casos el espectrograma presenta un decremento de la energía al final de la palabra cuando ésta termina el enunciado, lo cual indica una menor intensidad en la pronunciación de la misma. La intensidad en la emisión de la palabra es un factor importante para la determinación de los fonemas a partir de la representación mediante el espectrograma. Algunos fonemas, principalmente las consonantes, presentan características espectrales difíciles de identificar ya sea por su corta duración, como en el caso de las oclusivas, o porque se involucran varias frecuencias a la vez que dan una apariencia "ruidosa", como en las fricativas. Cuando la palabra se pronuncia a baja intensidad, aumenta la dificultad en la identificación y el espectrograma puede incluso no mostrar las características de algunos de los fonemas involucrados; la falta de información clara puede ocasionar que en las transcripciones estos fonemas no aparezcan.

La calidad de la emisión de la palabra también influye en la identificación de los fonemas. La base de datos Latino40 presenta una variedad de pronunciaciones que resultó valiosa para este estudio. Algunos hablantes se caracterizaban por una clara pronunciación y a una velocidad adecuada que facilitaban la correcta identificación fonética, mientras que otros pronunciaban los enunciados lentamente y con deficiencias, lo cual dificultaba el análisis espectral. Al observar los espectrogramas se encontraba que en algunos casos la representación no lograba evidenciar la presencia del fonema, mientras que en otros los fonemas se llegaban a confundir con otros cuyas características espectrales eran similares.

Los efectos de coarticulación más relevantes se encontraron en hablantes de los siguientes países: Argentina, Colombia, Cuba, Chile, Guatemala y El Salvador. Cabe mencionar que, aunque se incluye un ejemplo de hablante mexicano, no se tomaron en cuenta los sonidos particulares del español de México debido a que en la base de datos los enunciados de los hablantes mexicanos no presentan palabras que los ejemplifiquen. A continuación se presentan los efectos coarticulatorios encontrados.

Los hablantes de Argentina tienen el fenómeno denominado yeísmo muy marcado; la pronunciación del sonido /jj/ presenta gran cantidad de energía y una clara representación ruidosa en el espectrograma.

En el caso de los hablantes de Colombia, la pronunciación de la *j* (/x/) se escucha como fricativa glotal sorda donde la energía que se produce es leve a bajas frecuencias. Las vocales se escuchan claras mientras que la oclusiva /t/ es muy leve.

Para los hablantes de Cuba se encontró que cuando se pronuncia la fricativa velar sorda /x/, ésta sale como fricativa glotal sorda /h/ donde continúan los bajos niveles de energía al arrastrar el sonido. La nasal alveolar sonora /n/ se convierte en nasal velar sonora /N/ con una emisión uniforme alcanzando casi todo el espectro de frecuencias presentes. Los fonemas oclusivos sonoros /d/ son muy cortos y casi no se alcanzan a ver en los espectrogramas.

En el caso de Chile, la fricativa velar /x/ previo a las vocales /e/ e /i/, se convierte en fricativa palatal sorda. La característica sorda se alcanza a observar al no haber una concentración de energía importante en el rango de frecuencias; hay casos en que la fricativa velar sorda /x/ se convierte en fricativa glotal sorda /h/. Este efecto se puede observar en palabras bisílabas con diferentes contextos, en la primera sílaba el efecto es ruidoso mientras que en la segunda la velar se convierte en glotal. Es difícil distinguir la /b/ y la /t/ cuando en su contexto se encuentran vocales, las cuales logran ser bien diferenciadas.

Los hablantes de Guatemala se caracterizaron por tener un fonema fricativo /s/ muy bien definido el cual se observa en el espectrograma con claridad, aunque en ocasiones la emisión en ocasiones resulta corta. En el caso de la vocal /o/ el sonido se presenta difuso; algunas vocales que se encuentran al final de la palabra resultan difíciles de identificar debido a la disminuida intensidad con que se pronuncian. En los hablantes del Salvador las palabras se pronunciaron con claridad destacando las oclusivas sordas /t/ y /p/, y las vocales /o/ y /e/ cuyos espectrogramas presentaban las características bien definidas.

El uso del espectrograma permite una adecuada identificación de los sonidos debido a las correlaciones entre las características acústicas y espectrales. Se pueden identificar diversos efectos coarticulatorios comparando los espectrogramas de las palabras o enunciados y observando las diferencias en las formas y duraciones de las características espectrales. La información que presentan es de utilidad para la elaboración de las transcripciones fonéticas de las bases de datos siendo éstas necesarias para la evaluación de

los sistemas de reconocimiento automático del habla. Aunque representa una herramienta fundamental para el análisis, es importante señalar que se basa en una estadística de segundo orden, la cual no logra caracterizar completa y satisfactoriamente la señal del habla. Dado que se le considera una señal de naturaleza dinámica, la búsqueda de métodos de análisis y representación que caracterice los sonidos del habla con mayor exactitud y resuelva los conflictos coarticulatorios se convierte en un campo atractivo de investigación.

BIBLIOGRAFÍA

1. Quilis A. Fonética acústica de la lengua española, Editorial Gredos, España, 1988.
2. Bernal BJ, Bobadilla SJ, Gómez VP. Reconocimiento de Voz y Fonética Acústica, Editorial: Alfaomega, 2000.
3. The International Phonetic Alphabet, International Phonetic Association, 1993, <http://www.arts.gla.ac.uk/IPA/fullchart.html>.
4. Hieronymous JL. ASCII Phonetic Symbols for the World's Languages: Worldbet, Technical Report, Bell Labs., 1993.
5. Alfabeto fonético del Español de acuerdo a SAMPA, 1996, <http://www.phon.ucl.ac.uk/home/sampa/spanish.htm>.
6. Speech Assessment Methods Phonetic Alphabet (SAMPA) <http://www.phon.ucl.ac.uk/home/sampa/home.htm>.
7. Moreno A, Mariño J. Spanish Dialects: Phonetic Transcription, 5th International Conference on Spoken Language Processing, Sydney Australia, 1998.
8. Rabiner LR, Schafer RW. Digital Signal Processing of Speech Signals, Prentice-Hall, Englewood-Cliffs N. J., 1993.
9. Proakis JG, Rader CM, Ling F, Nikias CL, Moonen M, Ian K. Proudler, Algorithms for Statistical Signal Processing, Prentice Hall, 2000.
10. Bernstein J et al. The Latino40 Speech Database, Entropic Research Laboratory, Washington, DC. 1994.