

# *VOT como rasgo distintivo de las consonantes oclusivas españolas e inglesas*

*Jorge Ernesto Lemus Sandoval*

## **Resumen**

*En este artículo se demuestra a través de un experimento acústico que las vibraciones glotales por sí mismas no son suficientes para hacer la distinción entre consonantes oclusivas sordas y sonoras en inglés y en español. En español, las vibraciones glotales son importantes rasgos distintivos para las oclusivas pero en inglés no ya que se basan en otras claves acústicas para identificar a los sonidos. En inglés, las consonantes oclusivas se distinguen entre sí por la duración relativa del VOT positivo. Así, las consonantes sordas tienen un VOT positivo mayor que las sonoras.*

## **Introducción**

La presencia o ausencia de vibraciones glotales en la producción de consonantes oclusivas ha sido tomada como suficiente indicador para distinguir la producción y percepción de las mismas como sordas o sonoras. Esta diferencia se ha utilizado para describir los inventarios fonéticos de los diferentes idiomas del mundo. Las consonantes sonoras son aquellas que cuentan con vibraciones glotales y las sordas son las que carecen de éstas.

La sonoridad de las consonantes es un rasgo importante utilizado en el análisis fonológico de los idiomas. Por ejemplo, en inglés, la duración relativa de las vocales que preceden a las consonantes oclusivas se determina por la sonoridad de las mismas: las vocales son cortas antes de las consonantes oclusivas sordas [bæt] y largas antes de las consonantes oclusivas sonoras [bæ:d]. En el español ibérico también influye la sonoridad de las consonantes oclusivas en la pronunciación de sonidos adyacentes. El fonema alveolar fricativo sordo /s/ se convierte en el alófono [z] antes de consonante sonora [dézde] y se mantiene sordo ante consonante sorda [espína].

Sin embargo, aunque la sonoridad de las consonantes parece ser un principio básico universal, no siempre es una clave

suficiente o confiable para la percepción y producción de consonantes. Cualquier experimentado profesor de inglés habrá notado, por ejemplo, que sus estudiantes de habla hispana perciben las consonantes oclusivas sonoras del inglés como sordas (la palabra [bæt] probablemente sería percibida por el oído hispano como [pæt]); y, por otro lado, cualquier profesor de español habrá también notado que sus estudiantes de habla inglesa perciben las consonantes sordas españolas como sonoras (la palabra para será percibida en inglés como bara. Pero, ¿porqué sucede esto si en ambos idiomas las consonantes oclusivas tienen como rasgos distintivos sordas y sonoras?

Williams (1977) ya señala la insuficiencia del contraste sordo / sonoro para clasificar a las consonantes de los idiomas naturales. En inglés, por ejemplo, tanto las consonantes sordas como las sonoras se pueden producir sin vibraciones glotales. Pero, al contrario, en español la presencia de vibraciones glotales es suficiente para diferenciar las consonantes sordas (sin vibraciones) de las sonoras (con vibraciones). Por lo tanto, un sistema que se basa únicamente en la presencia o ausencia de vibraciones glotales será capaz de distinguir consonantes sordas y sonoras en español pero no podrá hacerlo en inglés o en otros idiomas en donde las pulsaciones

glotales no distinguen a las consonantes sordas de las sonoras.

En este artículo analizo las diferentes señales fonéticas necesarias para llevar a cabo el contraste de sonoridad entre las consonantes oclusivas españolas e inglesas. Demuestro que los hablantes de español e inglés se basan en diferentes señales acústicas para distinguir estas consonantes y que esta es la razón por la cual hablantes de ambos idiomas las perciben equivocadamente. Sobre la base de este análisis, propongo añadir el tiempo de inicio de las vibraciones glotales o VOT (voice onset time) como un nuevo rasgo necesario para una descripción fonética más exacta de los idiomas naturales y como un rasgo distintivo para el desarrollo de mecanismos automatizados de reconocimiento de la voz.

Aunque el experimento aquí descrito tiene más que ver con la producción de estos sonidos, sus resultados hacen interesantes predicciones con respecto a la percepción de los mismos por hablantes nativos de inglés y

español, y confirma los experimentos de percepción consonántica realizados por Williams (1977).

### Tiempo de inicio de la sonoridad

El tiempo relativo de las maniobras articulatorias para la producción del habla y sus consecuencias acústicas se han considerado como importantes claves para la discriminación de sonidos (Summerfield 1982). A este tiempo relativo se le conoce como VOT o tiempo de inicio de la sonoridad y fue propuesto por primera vez por Lisker y Abramson (1964, 1971). Este es el tiempo que transcurre desde la explosión y el inicio de la sonoridad, como se ilustra en la figura 1.

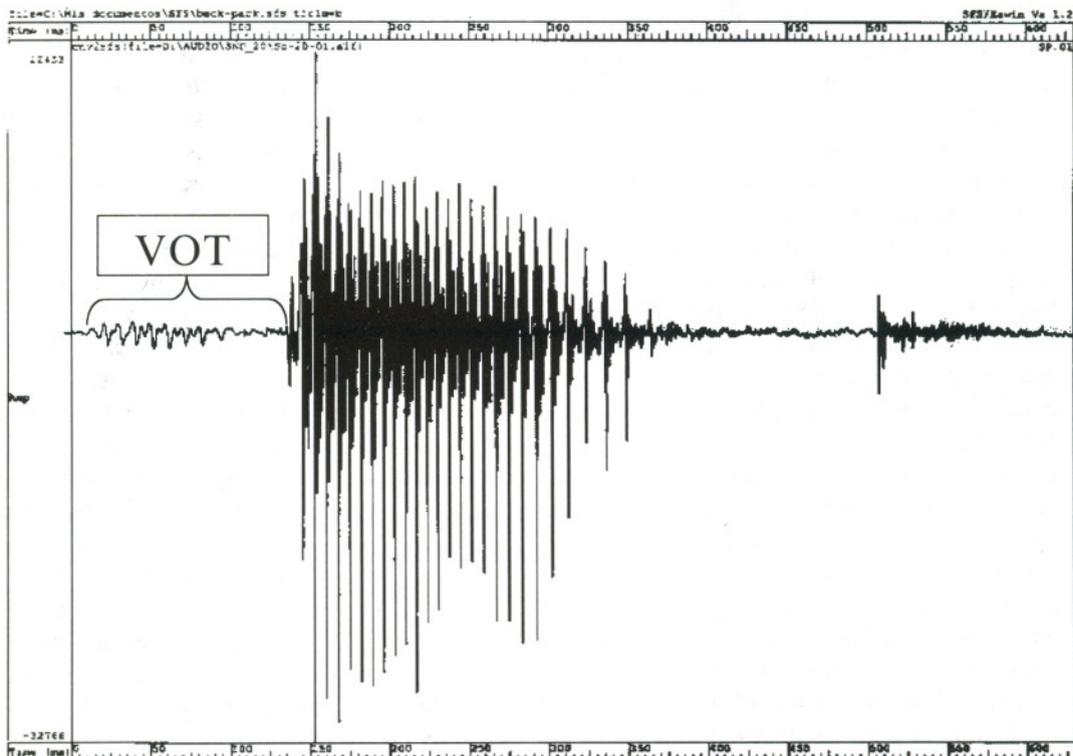


Figura 1

### Metodología y descripción del experimento<sup>1</sup>

El experimento consistió en grabar listas de palabras que contuvieran consonantes oclusivas sordas y sonoras de español e inglés. Los sujetos que colaboraron en el estudio eran todos hablantes nativos de inglés o español. A los sujetos se les entregaron palabras en su idioma en tarjetas independientes, a razón de una palabra por tarjeta. Luego se les pidió que las leyeran una por una a una velocidad normal de conversación frente a un micrófono y aparato de grabación de alta fidelidad. Después se transfirieron las palabras a una computadora para obtener una representación visual de las ondas sonoras de todas las palabras. Con dicha representación visual, se pudo medir con exactitud el VOT negativo (antes del inicio de vibraciones) y positivo (después del inicio de vibraciones). La figura 1 muestra longitud de onda de una palabra en inglés tal como se ha medido en el experimento.

A los hablantes de español se les pidió leer los siguientes pares de palabras: tos-dos, tomo-domo, pano-bano, peso-beso, casa-gasa, un domo-un tomo, un gato-un cato y un peso-un beso. Se les pidió que leyeran cuatro veces la misma palabra antes de pasar a la siguiente, y de leer toda la lista de palabras dos veces. Esto totalizó 144 palabras grabadas. A los hablantes de inglés se les presentaron los siguientes pares de palabras: barking-parking, dip-tip, bat-pat, cut-gut, one dip-one tip, one bat-one pot, one gut-one cut. Luego se les pidió que las leyeran de la misma manera que los hablantes de español. Esto hizo un total de 112 palabras para los hablantes de inglés.

Los sujetos de este experimento tenían todos como lengua materna el español o el inglés. Dos de los hablantes de español eran mejicanos y uno era salvadoreño. Esto tenía como objeto descubrir si existía alguna variación dialectal en la producción de las consonantes oclusivas. De los hablantes de inglés, solo uno era bilingüe inglés-español, los otros dos eran monolingües.

Se incluyó en el experimento el mismo grupo de palabras oclusivas precedidas de una nasal para descubrir el efecto de las consonantes nasales sobre el valor general del VOT.

### Resultados

El experimento demostró que las consonantes oclusivas del español y el inglés difieren grandemente en sus valores del VOT. Las consonantes oclusivas inglesas (sordas y sonoras) no tienen VOT negativo en posición inicial absoluta de palabra, demostrando que no es la vibración de las cuerdas vocales lo que distingue a las consonantes sordas de las sonoras en inglés. Más abajo se demuestra que los hablantes de inglés utilizan el valor positivo del VOT para diferenciar las consonantes sordas de las sonoras. Por otro lado, en español, los hablantes se basan exclusivamente en los valores negativos y positivos para distinguir entre consonantes sordas y sonoras. Por lo tanto, los hablantes de inglés y español dependen de diferentes señales acústicas para distinguir a las consonantes oclusivas sordas y sonoras.<sup>2</sup> A continuación describo los resultados obtenidos para cada uno de los idiomas. Estos resultados no difieren de otros experimentos similares llevados a cabo por otros investigadores.

<sup>1</sup> El experimento se llevó a cabo con el equipo para análisis acústico espectrográfico de la Universidad de Arizona bajo la supervisión del Dr. Carry Green y repetido con nuevos sujetos en la Universidad Don Bosco utilizando el programa SFSwin para análisis espectrográfico desarrollado por el departamento de fonética y lingüística del Mark Huckvale University College London

<sup>2</sup> Aunque en el experimento únicamente se estudiaron las consonantes oclusivas, se puede inferir que las mismas claves se utilizan para distinguir otras consonantes como sordas o sonoras.

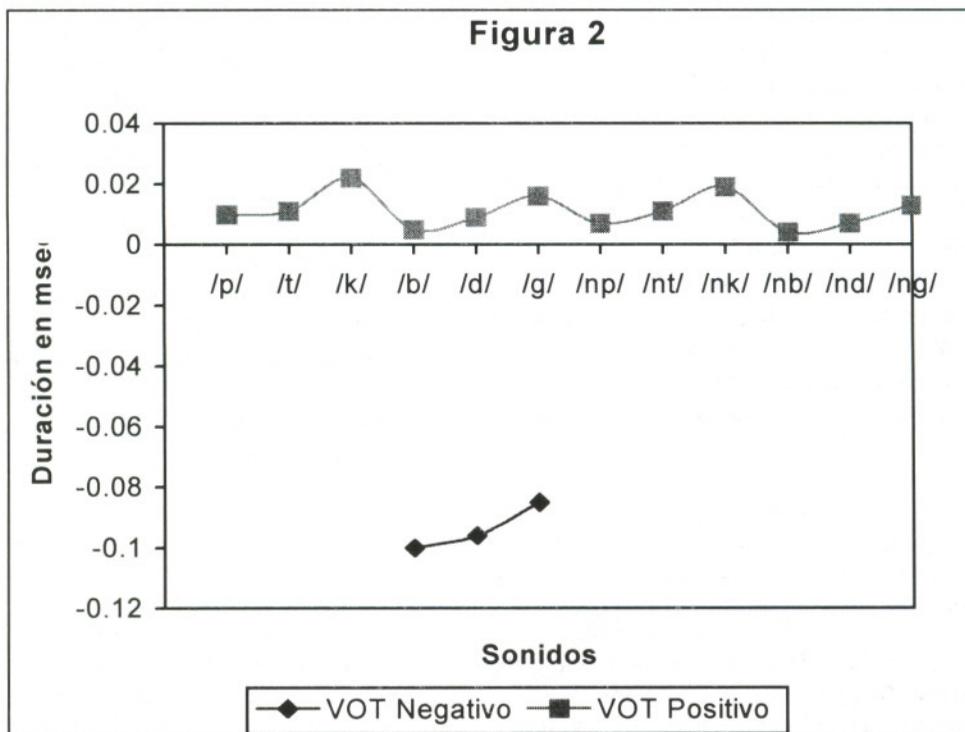
## Español

Los valores VOT de los hablantes españoles son cortos; estos oscilan entre los 0,004 y 0,027 msec, dependiendo del punto de articulación de la consonante evaluada (ver Anexo1). El valor del VOT aumentó con relación al punto de articulación. Así, las consonantes velares mostraron los valores VOT más largos mientras que las labiales mostraron los más cortos, aunque muy cerca de las dentales (0,001 msec de diferencia promedio entre la /p/ y la /t/). Sin embargo, en otros estudios con mayor número de sujetos (Williams 1977, Flege et al 1986) se han encontrado mayores diferencias entre las consonantes labiales y dentales del español.

Las consonantes oclusivas sonoras españolas muestran una presonoridad relativamente larga antes de la explosión de su inicio. El VOT negativo se midió desde el inicio de la primera baja frecuencia periódica visible que indica el principio de la sonoridad hasta la explosión de energía

espectral que marca la liberación articuladora. Esta liberación de energía consiste generalmente de una explosión aperiódica de energía. El VOT positivo se midió desde el principio de esta liberación articuladora hasta el principio de la frecuencia periódica. Las consonantes labiales mostraron VOT negativas más largas que las dentales y alveolares, demostrando que el VOT negativo disminuye a medida el punto de articulación de las consonantes se mueve hacia atrás del tracto articulatorio. Este resultado es lo contrario para el VOT positivo descrito arriba.

Los valores para el VOT negativo oscilaron entre  $-0,082$  a  $-0,115$ , mostrando las labiales el mayor valor y las velares el menor. El VOT positivo se comportó en forma similar al de las oclusivas sordas. Es decir, aumentó con relación al punto de articulación: menor para las oclusivas anteriores y mayor para las oclusivas posteriores. Los valores oscilaron entre 0,013 y 0,008 msec.



La presencia de una consonante nasal precediendo las consonantes oclusivas tiene un impacto mínimo en el valor promedio del VOT de estas consonantes. Aunque algunas veces la presencia de la consonante nasal disminuyó la duración del VOT, la diferencia no fue significativa y algunas veces inexistente (ver Apéndice). Sin embargo, fue imposible medir el VOT negativo. No se pudo determinar en los diagramas el punto exacto donde terminaba la periodicidad nasal y donde comenzaba la periodicidad del VOT negativo. Por esta razón, no se midió el VOT negativo de las consonantes precedidas de nasal. También cuando se leyeron listas de palabras, la vocal final de la palabra anterior dificultó la identificación del VOT negativo de las consonantes oclusivas sonoras. La Figura 2 claramente demuestra que lo que distingue la /p/, /t/ y /k/ de /b/, /d/ y /g/ es el VOT negativo y que el VOT positivo juega un papel mínimo para la discriminación de estas consonantes.

### Inglés

En general, los valores VOT de las consonantes oclusivas inglesas son significativamente más largos que los españoles. Los valores VOT de las consonantes oclusivas sordas oscilaron entre 0,053 y 0,095 msec, con las duraciones más largas para las velares y las más cortas para las labiales, la misma tendencia que en español.<sup>3</sup> El VOT de las oclusivas sordas se midió desde el principio del ruido aperiódico hasta el inicio de la aspiración, que prácticamente encierra todo el VOT. El grado de aspiración de las oclusivas sordas en inglés es una diferencia significativa entre las consonantes inglesas y españolas.

Las oclusivas sonoras inglesas poseen VOTs significativamente más cortos que los de las sordas. Su longitud oscila entre los 0,006 y 0,025 msec, siendo el menor valor

para las labiales y el mayor para las velares. Sin embargo, las oclusivas sonoras no mostraron ningún valor negativo de VOT.<sup>4</sup> Esta característica acústica del inglés es una diferencia importante con el español.

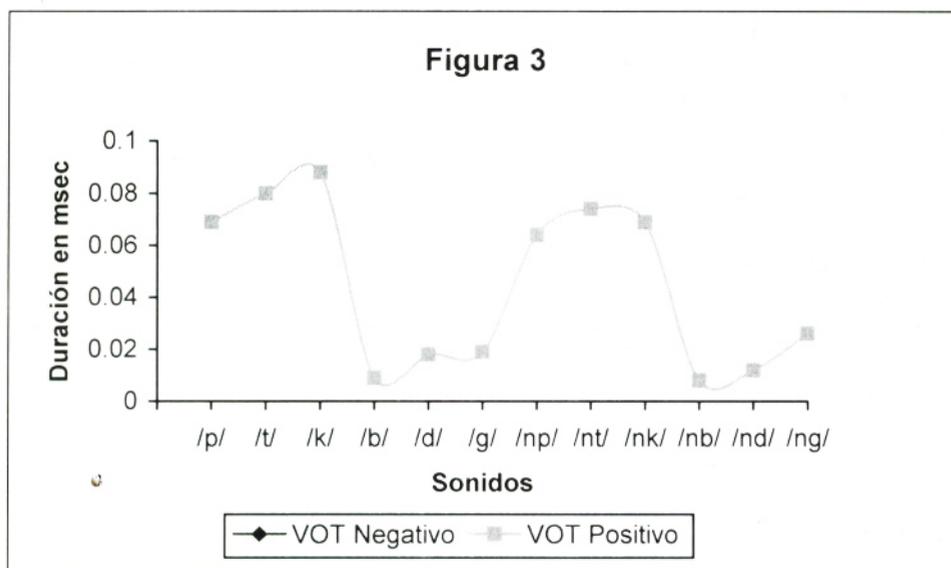
Al igual que en español, la presencia o ausencia de nasalidad antes de las consonantes oclusivas no tuvo ningún efecto significativo sobre el valor general del VOT. De hecho, en inglés se pudo notar que algunos sonidos mostraron VOTs más largos cuando les precedía una nasal. Las oclusivas sordas disminuyeron el valor promedio del VOT en 12.7% cuando les precedía una nasal y no seguían el patrón general de VOTs más largos para las consonantes velares y más cortos para las labiales. Las alveolares fueron las más largas seguidas de las velares y las labiales, respectivamente.

El experimento mostró que las consonantes oclusivas sonoras del inglés no tienen VOT negativo en posición absoluta inicial de palabra. Esto significa que en el momento en el que se pronuncian las oclusivas sonoras, las cuerdas vocales no están vibrando, contradiciendo lo que generalmente se ha aceptado como la señal que distingue las consonantes sonoras de las sordas.<sup>5</sup> La fig. 3 muestra como el VOT positivo de las consonantes oclusivas sordas es más largo que el de las sonoras. También se puede notar en la misma figura que la nasalidad tiene poco efecto sobre el valor total del VOT.

<sup>4</sup> En otras posiciones que no sean al inicio de la palabra, las oclusivas sonoras inglesas sí muestran valores negativos de VOT, aunque siempre más cortos que los de sus contrapartes españolas. Por lo tanto, no se midió el valor negativo VOT de las consonantes inglesas.

<sup>5</sup> Ladefoged (1982) clasifica las oclusivas inglesas como sonoras, parcialmente sonoras y sordas, y menciona el hecho de que en otros idiomas la sonoridad no es una señal suficiente para distinguir este tipo de consonantes. En Tailandés, por ejemplo, las consonantes /b, d, g/ no muestran ninguna sonoridad antes del inicio de la consonante por lo que suenan para el oído español como /p, t, k/.

<sup>3</sup> Estas mediciones contradicen la afirmación de Williams (1977) de que en inglés los VOTs son más largos para las consonantes velares que para las labiales.



### Español vs. Inglés

Se ha demostrado en las secciones anteriores que los hablantes de español e inglés se basan en diferentes señales acústicas para diferenciar sonidos tradicionalmente conocidos como sonoros y sordos. En inglés, lo que distingue a estos sonidos no es la vibración de las cuerdas vocales o VOT negativo el que señala la diferencia entre ambos tipos de consonantes sino el VOT positivo. En español es el VOT negativo el que diferencia a las consonantes sonoras de las sordas. El que hablantes de inglés y español utilicen diferentes claves acústicas para diferenciar entre las oclusivas sordas y sonoras explica porque los hablantes de ambos idiomas perciben erróneamente las oclusivas entre ellos. Así, los hablantes ingleses perciben a las consonantes españolas /p, t, k/ como /b, d, g/ debido a su similitud en valores VOT. De la misma manera, los hablantes españoles perciben a las oclusivas sonoras inglesas (i.e. /b, d, g/) como su contraparte sorda en español (i.e. /p, t, k/). Esto se debe a que los hablantes de español perciben el VOT negativo como clave para distinguir entre las consonantes sordas y sonoras, y en inglés, las oclusivas sonoras no tienen esta señal acústica, el oído español las percibe como

sordas. La siguiente tabla muestra los valores del VOT positivo de las oclusivas sonoras inglesas comparadas con las oclusivas sordas españolas.

**Tabla 1.** Comparación de VOTs de las consonantes oclusivas sonoras y sordas del español y el inglés

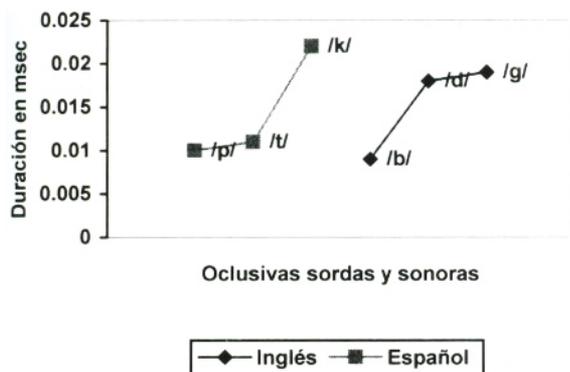
Inglés	VOT	Español	VOT
/b/	0.009	/p/	0.010
/d/	0.018	/t/	0.011
/g/	0.019	/k/	0.022

La tabla 1 muestra que la diferencia entre los valores del VOT positivo de las consonantes oclusivas sonoras del inglés comparadas con las consonantes sordas del español es mínima. La /b/ inglesa, por ejemplo, difiere de la /p/ española en solo una milésima de segundo.<sup>6</sup> La figura 4 muestra gráficamente las similitudes entre estas consonantes.

La figura 4 muestra gráficamente que las diferencias en duración del VOT positivo de las consonantes españolas /p, t, k/ son casi idénticas con los valores de las

<sup>6</sup> Recuerden que en los ejemplos considerados en este experimento se demostró que las oclusivas sonoras inglesas carecen de VOT negativo, al menos en posición absoluta de inicio de palabra.

Figura 4



consonantes oclusivas sonoras del inglés /b, d, g/. Esto explica los problemas de percepción de estas consonantes entre los hablantes de español e inglés. Así, los hablantes de inglés perciben erróneamente las oclusivas sordas españolas /p, t, k/ como la contraparte sonora inglesa /b, d, g/ inglesa, y viceversa, los hablantes de español perciben las consonantes sonoras españolas como las sordas españolas.<sup>7</sup>

En este artículo se ha comprobado que las vibraciones glotales por sí mismas no son suficientes para hacer la distinción entre consonantes oclusivas sordas y sonoras en inglés y en español. En español, las vibraciones glotales son importantes rasgos distintivos pero en inglés no para las oclusivas ya que se basan en otras claves acústicas para identificar a los sonidos. En inglés, las consonantes oclusivas se distinguen entre sí por la duración relativa

<sup>7</sup> Este problema de percepción es vivido diariamente por miles de hispanohablantes residentes en los EE.UU., por ejemplo, quienes tienen que repetir, en el mejor de los casos, las palabras que comienzan con oclusivas sordas en inglés ya que ellos las pronuncian como en español. La palabra "toro" en español suena como "doro" para el oído inglés. Un amigo mío cuyo apellido es Toro siempre tiene que decir "Toro, with 't' as in Thomas" a los hablantes ingleses para que puedan escribir correctamente su nombre. Los hablantes de inglés en países de habla hispana tienen problemas para que sus oclusivas sonoras sean percibidas correctamente en español.

del VOT positivo. Así, las consonantes sordas tienen un VOT positivo mayor que las sonoras.

También se puede ver en las figuras 2 y 3 que tanto en inglés como en español, los valores del VOT positivo aumentan con relación al punto de articulación de la consonante, de tal manera que las consonantes labiales tienen VOTs más cortos y las velares más largos. Williams (1977) encontró valores contrarios para las consonantes oclusivas inglesas. Según Williams los valores del VOT disminuyen con respecto al punto de articulación, de tal suerte que las velares tienen un VOT más largo y las labiales el más corto.

Las nasales precediendo a las oclusivas no afectaron significativamente el valor promedio del VOT en inglés y español, aunque su presencia volvió más difícil la medida del VOT. También se descubrió que en español existe un rasgo más que demostró ser significativo: el silencio. Después de las nasales, y antes del inicio de la articulación de la oclusiva, siempre había un periodo de silencio que duraba entre 0.097 y 0.156 msec. Este silencio se midió desde el final de la periodicidad nasal hasta el principio de la articulación de la consonante.

## Referencias

- Blumstein, S.E., Isaacs, E. & Mertus, J. (1982) "The Role of the Gross Spectral Shape as a Perceptual Cue to Place of Articulation of Stop Consonants" *JASA* 72: 43-50
- Flege, James & Wieke Eefting (1986) "Linguistic and Developmental Effects on the Production and Perception of Stop Consonants" *Phonetica* 43: 155-171
- Ladefoged, Peter (1982) *A Course in Phonetics*. Harcourt Brace Jovanovich Publishers, New York, N.Y.
- Mack, Molly (1989) "Consonant and Vowel Perception: Early English-French Bilinguals and English Monolinguals" *Perception and Psychophysics* 46 (2): 187-200
- Stevens, K.N., and Blumstein, S.E. (1978) "Invariant Cues for Place of Articulation in Stop Consonants" *JASA* 85: 2081-2087

Williams, Lee (1977) "The Voicing Contrast in Spanish", *Journal of Phonetics* 5, 169-184

\_\_\_\_\_ (1979) "The Modification of Speech Perception and Production in Second Language Learning" *Perception & Psychophysics* 26 (2): 95-104

## Apéndice 1

**Tabla 1.** Valores promedios del VOT negativo y positivo de las consonantes oclusivas españolas producidas por los tres sujetos participantes del estudio

Sonido	Hablante 1		Hablante 2		Hablante 3		Promedio VOT negativo	Promedio VOT positivo
	VOT Negativo	VOT Positivo	VOT Negativo	VOT Positivo	VOT Negativo	VOT Positivo		
/p/		0,013		0,008		0,010		0,010
/t/		0,013		0,009		0,011		0,011
/k/		0,027		0,018		0,021		0,022
/b/	-0,098	0,008	-0,087	0,005	-0,115	0,002	-0,100	0,005
/d/	-0,091	0,010	-0,099	0,006	-0,098	0,011	-0,096	0,009
/g/	-0,089	0,017	-0,082	0,011	-0,083	0,019	-0,085	0,016
/np/		0,006		0,006		0,008		0,007
/nt/		0,015		0,007		0,010		0,011
/nk/		0,018		0,020		0,018		0,019
/nb/		0,003		0,009		0,002		0,004
/nd/		0,006		0,006		0,008		0,007
/ng/		0,015		0,006		0,019		0,013

**Tabla 2.** Valores promedios del VOT negativo y positivo de las consonantes oclusivas inglesas producidas por los tres sujetos participantes del estudio

Sonido	Hablante 1		Hablante 1		Hablante 2		Promedio VOT negativo	Promedio VOT positivo
	VOT Negativo	VOT Positivo	VOT Negativo	VOT Positivo	VOT Negativo	VOT Positivo		
/p/		0,080		0,053		0,075		0,069
/t/		0,092		0,055		0,092		0,080
/k/		0,078		0,095		0,091		0,088
/b/		0,010		0,009		0,009		0,009
/d/		0,025		0,006		0,024		0,018
/g/		0,016		0,019		0,021		0,019
/np/		0,067		0,054		0,070		0,064
/nt/		0,078		0,065		0,078		0,074
/nk/		0,067		0,066		0,073		0,069
/nb/		0,011		0,008		0,006		0,008
/nd/		0,011		0,008		0,018		0,012
/ng/		0,030		0,029		0,019		0,026