

## El VOT de las oclusivas del español en habla emocional simulada

### VOT in Spanish stops in simulated emotional speech

Juan C. Martínez-Belda<sup>a</sup>, Xose A. Padilla<sup>b</sup>

<sup>a</sup> Universitat d'Alacant (España), [jcmb@alu.ua.es](mailto:jcmb@alu.ua.es)

<sup>b</sup> Universitat d'Alacant (España), [xose.padilla@ua.es](mailto:xose.padilla@ua.es)

ARTICLE INFO	ABSTRACT
<i>Article history</i> Received: 30/06/2020 Accepted: 10/02/2021  <i>Keywords</i> Spanish stops Voice Onset Time (VOT) Burst Emotions Simulated speech	The goal in this article is to find out whether VOT in voiced and voiceless stops in Peninsular Spanish varies across different emotional dialogical situations. Ten pairs of female university students from Alicante aged 18-25 years old performed poli-functional simulated dialogue in three different situations: joy, anger, and neutrality. After this phase, we analyzed VOT in the three contexts and then compared our values in the neuter form with those obtained in previous studies. Our results show differences with those from previous research on VOT, which are based on the reading of word lists. The statistical analysis, however, only allows us to point out some inconclusive trends in the influence of the emotions examined.

### 1. Introducción

El Voice Onset Time (VOT) es descrito de forma general como una dimensión temporal con dos eventos articulatorios (Lisker y Abramson, 1964; Beckman et al., 2011): la apertura del cierre de la cavidad bucal y el comienzo de la sonoridad. Su realización forma parte de la configuración de los fonemas oclusivos y, como veremos más tarde, permite analizar la producción de estos fonemas dentro de una lengua en particular o comparar dos o más lenguas entre sí.

La realización fonética de los fonemas en el habla puede ser condicionada por factores diversos. Algunos de ellos son de naturaleza lingüística, como la coarticulación;<sup>1</sup> otros son de naturaleza

paralingüística, como la influencia en la articulación de los sonidos de las emociones y los estados de ánimo. El propósito de este trabajo es averiguar si el VOT de las oclusivas del español /p t k/ y /b d g/ sufre algún tipo de variación cuando comparamos diferentes situaciones emocionales en habla dialógica simulada. Para llevar a cabo este estudio, hemos confeccionado un diálogo polifuncional (Padilla, 2017) y hemos analizado las producciones de diez parejas de mujeres con edades comprendidas entre los 18 y 25 años, que lo han representado de tres formas distintas: la *alegría*, el *enfado* y la *neutralidad*. Hemos elegido estos tres contextos por dos razones: a) por el carácter básico de las emociones seleccionadas (Ekman, 1970, 1999; Plutchik, 1994; Padilla, 2020, en prensa); y b) por permitir establecer polos extremos

<sup>1</sup> Martínez Celdrán (1984) ejemplifica la influencia de las transiciones vocálicas de F2 de la vocal contigua. La direccionalidad de F2 varía dependiendo del índice

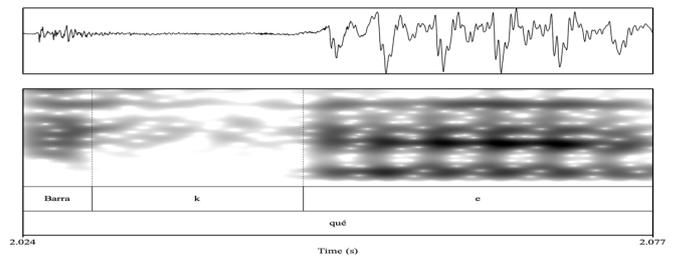
consonántico del locus, definido como “*el punto imaginario donde convergen o apuntan las transiciones vocálicas*”.

emocionales en un continuo (Scherer et al., 1984; Garrido, 2019). Finalizada la primera etapa, hemos analizado los valores del VOT en los tres contextos mencionados y hemos cotejado nuestros valores neutros con los resultados obtenidos en estudios previos. Hemos realizado, por último, un análisis estadístico de los valores extraídos del corpus con el propósito de examinar la significancia de los datos. A lo largo de las siguientes páginas presentaremos los resultados derivados de esta investigación.

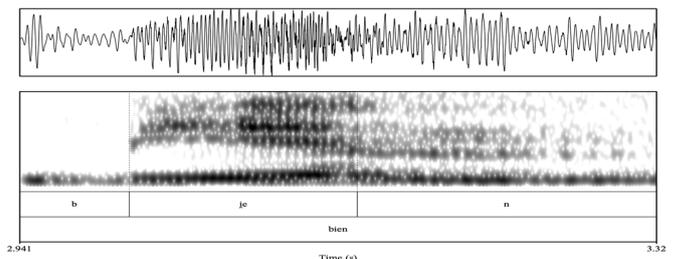
Nuestro trabajo forma parte de un proyecto de investigación más amplio, denominado *‘El habla con significado emocional y expresivo: análisis fonopragmático y aplicaciones’*, que tiene como objetivo general lograr un mejor conocimiento de la influencia de las emociones en el uso del lenguaje.

## 2. ¿Qué es el VOT?

Los primeros autores en mencionar el concepto VOT (o tiempo de inicio de la sonoridad) fueron Lisker y Abramson (1964) que propusieron una dimensión temporal para distinguir los fonemas oclusivos *homorgánicos*, es decir, aquellos que comparten el lugar de articulación, pero no la sonoridad ni el modo de articulación (Ladefoged y Johnson, 2006). A esta dimensión la llamaron VOT, y la midieron en milisegundos (ms). En las oclusivas sordas, el VOT cuantifica el intervalo de tiempo que transcurre desde la apertura del cierre de la consonante oclusiva hasta el comienzo de la sonoridad del sonido contiguo (Figura 1a). En las oclusivas sonoras, por el contrario, el VOT abarca desde el comienzo de la vibración de las cuerdas vocales, durante la oclusión de la consonante, hasta su liberación o apertura (Figura 1b). Estos dos momentos se marcan en relación con un punto cero (0 ms) que indica la frontera entre la oclusión y la apertura. Dicho de otra forma, la dimensión VOT da cuenta del “comienzo de la vibración laríngea antes o después de la explosión” (Martínez Celdrán, 2013, p. 254).



**Figura 1a.** Oclusiva velar sorda /k/: sílaba [ke] en ¿qué tal?

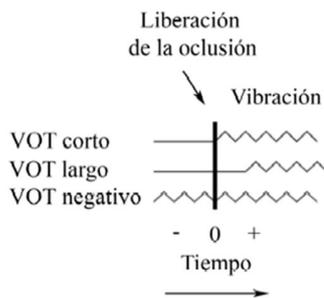


**Figura 1b.** Oclusiva bilabial sonora /b/: en [bien] en bien, gracias.

### 2.1 Rasgos clasificatorios

El comienzo de la sonoridad (o periodicidad) permite dividir el VOT en dos tipos: el VOT positivo y el VOT negativo. En las oclusivas sordas, el VOT se mide desde el punto cero (0 ms) hacia delante, y se le denomina VOT positivo. En las oclusivas sonoras, el VOT se mide desde el punto cero (0 ms) hacia atrás, y se le denomina VOT negativo.

Partiendo de lo anterior, se han descrito tres tipos de VOT (Lisker y Abramson, 1964): a) el VOT negativo (*lead* o *prevoiced*), cuando la vibración precede a la apertura; b) el VOT corto (*short-lag*), cuando la vibración se produce simultáneamente o inmediatamente después a la apertura; y c) el VOT largo (*long-lag*), si la vibración se produce aún más tarde, después de la apertura, en comparación con el VOT corto (ver Figura 2).



**Figura 2.** Explicación del comienzo de la sonoridad en relación con el punto 0:  
Lead = VOT negativo; Short-lag = VOT corto;  
Long-lag = VOT largo.

A la división anterior se suma, además, la posibilidad de que el VOT se vea acompañado o no, en su producción, de una aspiración o soplido laríngeo. Se añade, así, un nuevo rasgo para caracterizar el VOT y se establece la posibilidad de que el VOT se produzca con y sin aspiración.

Sumando todos los criterios expuestos, se diferencian, como mostramos en la Tabla 1, cuatro categorías, o fases, en el VOT.

El continuo VOT, por tanto, consta de dos polos, uno positivo y otro negativo, cada uno a su vez es subdividido en dos partes, con y sin aspiración. El VOT sin aspiración siempre está próximo al punto cero, mientras que el VOT con aspiración se sitúa en las partes extremas de los polos, lejos del punto cero.

## 2.2 ¿Por qué estudiar el VOT?

Uno de los objetivos de Lisker y Abramson (1964) al definir el VOT era crear una dimensión versátil que permitiese hacer cotejos entre unas lenguas y otras:

[...] to see how well a single dimension, voice onset time, serves to separate the stop categories of a number of languages in which both the number and phonetic characteristics of such categories are said to differ. (p. 388)

Así, tomando como punto de partida las cuatro categorías que mencionamos en el apartado anterior (vibración, momento de liberación de los articuladores y con sus consecuencias [+aspiración]), es posible, por ejemplo, diferenciar dos lenguas como el inglés y el español (Tabla 2).

<b>VOT negativo con aspiración<sup>2</sup></b>	Combinación de vibración de cuerdas vocales durante el cierre y el mismo tamaño de la glotis que en el VOT corto (Ladefoged y Johnson, 2006; Berkson, 2012).
<b>VOT negativo</b>	Las cuerdas vocales vibran durante el cierre ( <i>prevoicing</i> ).
<b>VOT corto</b>	La sonoridad comienza en el momento de la apertura de los articuladores que bloquean la salida del aire o brevemente después.
<b>VOT largo con aspiración</b>	La sonoridad se retrasa considerablemente después de la apertura de los articuladores que bloquean la salida del aire. La glotis está abierta considerablemente, lo que produce la aspiración (Piñeros, 2008, p. 80).

**Tabla 1.** Continuo VOT.

VOT negativo con aspiración	
VOT negativo sin aspiración	español [b d g]
<b>Barra de explosión</b>	<b>0 ms</b>
VOT corto sin aspiración	español [p t k] inglés [b d g]
VOT largo con aspiración	inglés [p <sup>h</sup> t <sup>h</sup> k <sup>h</sup> ]

**Tabla 2.** Comparación español-inglés.

<sup>2</sup> Se describe, en realidad, un cuarto tipo de VOT, denominado VOT negativo con aspiración, que se produce, por ejemplo, en lenguas como el marathi y el hindi (Lisker y Abramson, 1964), pues lo habitual es el VOT largo se produzca con aspiración, como de hecho, sucede en inglés.

La dimensión VOT permitió, por otra parte, que Lisker y Abramson (1964) pudiesen superar limitaciones descriptivas que presentaban otras categorías, como la sonoridad, la aspiración o la fuerza articulatoria, para llevar a cabo cotejos interlingüísticos.

La sonoridad, por ejemplo –siguiendo con la comparación entre el español y el inglés– no funciona adecuadamente, cuando se comparan los sonidos oclusivos de estas dos lenguas. Las oclusivas sonoras /b d g/ del inglés pueden ser descritas en realidad como falsas sonoras, porque las cuerdas vocales no siempre vibran durante el cierre. De hecho, sí vibran en posición intermedia,<sup>3</sup> pero no en posición inicial. En español, por el contrario, las oclusivas sonoras siempre son sonoras. La sonoridad no cambia con la posición. Es difícil, por tanto, comparar las dos lenguas de forma global atendiendo a este rasgo.

Algo parecido ocurre con la aspiración. La aspiración, en inglés, se produce en posición inicial de palabra y, en inicio de sílaba tónica, en posición intermedia. La aspiración separa, pues, las oclusivas sordas de las sonoras del inglés. En español, sin embargo, la aspiración no se produce en ningún caso, por consiguiente, esta dimensión tampoco permite hacer análisis contrastivos.

Todavía resulta más complicado analizar de forma objetiva la dimensión fuerza articulatoria (fortis/lenis, tenso/laxo, fuerte/débil) (Natarina, 2019, p. 40). Para comenzar, la fuerza articulatoria resulta un concepto ambiguo y difícil de medir (Slis, 1971). Generalmente, las oclusivas sordas tienen mayor tensión (presión oral) que las oclusivas sonoras, lo que explica su mayor energía

<sup>3</sup> Algunos estudios analizan el porcentaje de sonoridad de las oclusivas intervocálicas, como Beckman et al. (2013). Estos trabajos investigan qué rasgo define y distingue a las oclusivas en lenguas que sí tienen un contraste de sonoridad (*true voicing language*, como el español) con lenguas que contrastan por la aspiración (*aspirating language*, como el inglés). El español es una lengua con sonoridad activa y el inglés es una lengua con sonoridad pasiva. Es decir, la sonoridad pasiva proviene del contexto.

articulatoria. Sin embargo, la correlación entre sonoridad y fuerza articulatoria no parece universal (Ladefoged y Maddieson, 1996). Tampoco parece clara la independencia de ninguno de los rasgos que se han atribuido a la oposición fortis/lenis (Lisker y Abramson, 1964, p. 386). La fuerza articulatoria, por lo tanto, se presenta como un concepto ambiguo e interdependiente.

Frente a las dificultades de objetivación, falta de aplicabilidad y falta de independencia que presentan otras dimensiones en el análisis de las oclusivas, el VOT se revela como una medida objetiva, relativamente fácil de segmentar en el espectrograma y aplicable en casos de comparación fonológica intra e interlingüística.

### 3. ¿Cómo identificar y segmentar el VOT en el espectrograma?

No hay una postura unánime entre los autores sobre el momento preciso de segmentación del VOT en el espectrograma. En general, la mayor parte de los investigadores proponen segmentarlo a partir de la apertura del cierre (punto cero), señalando que el VOT positivo debe ser medido marcando el intervalo de tiempo que se produce entre la liberación de la oclusión (*release*) y el principio de la vibración (*periodicity*), esto es, de la voz (Lisker y Abramson, 1964; Hunnicutt y Morris, 2016; etc.). En cuanto al VOT negativo, la segmentación finalizaría en el momento de la apertura: “negative VOT was measured from the onset of voicing to the stop release” (Morris, 2018, p. 4).

Más allá de las coincidencias generales, los investigadores no siguen exactamente los mismos criterios cuando los análisis son más minuciosos. Por ejemplo, hay divergencias sobre cuándo comienza exactamente la sonoridad;<sup>4</sup> cuál es la

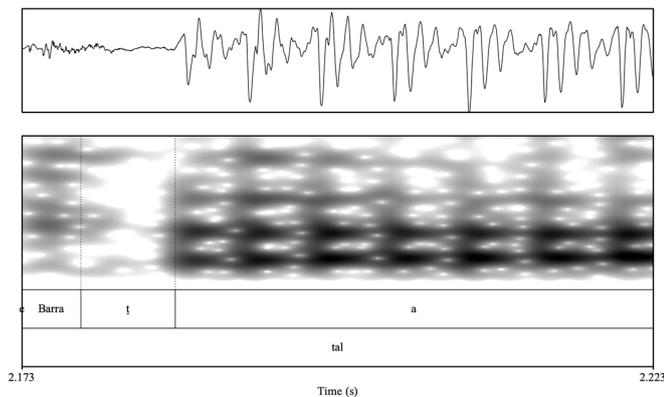
<sup>4</sup> Algunos autores, como Shea y Curtin (2006, 2011), usan para ello el segundo formante (F2); y otros, como Helgason y Ringen (2008, p. 610), miden el comienzo de la sonoridad (VOT) con el *modal voice onset*, es decir, incluyen la voz *susurrante* en las mediciones del VOT. Helgason y Ringen (2008, p. 610) siguen la metodología empleada por Ladefoged y Maddieson (1996, p. 70).

función del segundo formante en la segmentación; y, especialmente, si el VOT debe incluir o no, en sus mediciones, la barra de explosión (*burst of noise*) de la oclusiva (Shea y Curtin, 2006, 2011; Helgason y Ringen, 2008).

La justificación principal de estos desacuerdos, como veremos más adelante, es que, en los espectrogramas concretos, especialmente en contextos dialógicos, no siempre hay una frontera clara –esto es, visualmente observable– entre unos fenómenos y otros; por consiguiente, el análisis espectrográfico, en ocasiones, resulta complicado.

### 3.1. ¿Qué es la barra de explosión?

En la fase de explosión (o tercera fase de la producción de las oclusivas), los articuladores que bloquean la salida del aire se liberan, y, como consecuencia de ello, se produce una *barra* (de explosión) oscura (Figura 3) que es visible en los espectrogramas (Morales del Valle, 2016, p. 24).



**Figura 3.** Barra de explosión de la oclusiva dental sorda [t] en ¿qué tal?, que mide 4.696 ms. La sonoridad (VOT) tarda 7.529 ms en empezar. La barra de explosión y el VOT miden en total 12.225 ms.

Esta barra vertical observable (Fig. 3) ha sido descrita de forma general como el inicio abrupto de energía en el rango de la frecuencia de los formantes (Lisker y Abramson, 1964, p. 389). En el

oscilograma, se pueden observar, por ejemplo, sus picos (o *spikes*) característicos (vid. supra).

Martínez Celdrán y Fernández Planas (2007, p. 88) –usando los datos de Asensi et al. (1997)– apuntan, por otra parte, que la barra de explosión tiene unas duraciones medias específicas para cada sonido. En español, la duración media de la oclusiva bilabial sorda es 5,6 ms; 7 ms, la dental sorda; y 14,2 ms, la velar sorda. Ello permite ordenar los sonidos según la mayor o menor duración de la barra: [p] > [t] > [k]. La barra de explosión se produce, además, en una franja de frecuencias concreta, así, la bilabial se sitúa por debajo de los 1000 Hz, siendo la frecuencia media de la máxima intensidad de 535 Hz; la dental, por encima de los 2500 Hz, y la frecuencia máxima se sitúa en 5901 Hz; y la velar, entre los 1000 y 3000 Hz, con una frecuencia máxima de 1866 Hz (Martínez Celdrán y Fernández Planas, 2007, p. 88; Asensi et al., 1997).

Morales del Valle (2016, p. 24), por último, pone de relieve la relación entre fenómenos como el lugar de articulación y la duración de la barra de explosión. Según este autor, cuanto más posterior es la zona de contacto articulatorio, mayor es la duración de la barra.

### 3.2. ¿Debe incluirse o no la barra de explosión en la segmentación del VOT?

Una de las cuestiones más importantes al medir el VOT en los espectrogramas es decidir si se debe tener en cuenta o no la barra de explosión en el cómputo global y qué consecuencias tiene esta decisión en la obtención de datos.

En el caso del español, algunos autores, como Asensi et al. (1997), sí incluyen la medición de la barra en sus análisis, y aportan datos acústicos con y sin barra de explosión<sup>5</sup> (Tabla 3).

<sup>5</sup> Asensi et al. (1997, p. 227) mencionan, incluso, la posibilidad de que se produzca más de una barra de explosión en los sonidos velar y dental.

Autores	Variedad	Metodología	Oclusivas sonoras			Oclusivas sordas		
			[b]	[d]	[g]	[p]	[t]	[k]
Asensi et al. (1997)	Barcelona	Lectura de palabras				14,7 (9,1)	20,2 (13,2)	35,4 (21,2)

**Tabla 3.** Datos extraídos de Asensi et al. (1997).

	[p]/[b]	[t]/[d]	[k]/[g]
<b>Frecuencia máxima de la máxima intensidad</b>	600-800 Hz	4000 Hz	1800-2000 Hz

**Tabla 4.** Frecuencia de la intensidad donde se encuentra cada barra en inglés.

Otros, como Castañeda (1986), Rosner et al. (2000) o Déniz (2005), por el contrario, no la tienen en cuenta, y unen el VOT y la barra en sus mediciones.

Fry (1979, p. 122), en el caso del inglés, tampoco separa el VOT y la barra de explosión en la medición, sin embargo, sí proporciona información sobre sus valores<sup>6</sup> (Tabla 4).

Fry (1979) indica, además, que la frecuencia atribuible a la barra de explosión varía en función de la consonante. En las sonoras, “the intensity of the noise burst is much less than in the voiceless sounds” (p. 125), y apunta, asimismo, que la presencia de esta barra se ve influida por la precisión articulatoria y el estilo de dicción, ya sea cuidado o no. Esta puntualización de Fry (1979) podría justificar en parte, como veremos más tarde, la ausencia de la barra en algunas oclusivas sonoras de nuestros datos.

Martínez Celdrán (en prensa) menciona, para finalizar, que la segmentación del VOT en relación con la barra puede verse influida por su mayor o menor intensidad o por el silencio durante la oclusión.

<sup>6</sup> El orden que muestran los datos de Fry (1979), por otra parte, es idéntico al que señalaban Martínez Celdrán y Fernández Planas (2007) para el español: [p] > [k] > [t].

En nuestro estudio, proporcionaremos las dos mediciones: barra y VOT, cuando los datos del espectrograma lo permitan. Ahora bien, partiendo de la definición del VOT como el momento concreto de liberación de la oclusión a partir de un punto 0, la barra, en el caso de que sea visible, será una parte más del espacio temporal que va desde el mencionado punto 0 hasta la sonoridad contigua. Empezar a medir después de la barra supondría mezclar casos con barra visible con otros en los que su distinción es imposible.

#### 4. Algunos factores influyentes en la configuración del VOT en el habla

Independientemente de qué elementos decidamos tener en cuenta o no en la segmentación espectrográfica del VOT, es necesario apuntar, por último, que algunos autores indican que este puede verse influido por factores, en cierto modo, externos a su misma configuración.

##### 4.1. La velocidad de habla

Algunos autores (Kessinger y Blumstein, 1997; Magloire y Green, 1999), por ejemplo, ponen de relieve que la velocidad de habla puede tener efectos en la forma final del VOT.

Magloire y Green (1999), que comparan el inglés y el español, señalan que las consonantes oclusivas sordas del inglés modifican el comienzo de la sonoridad según la velocidad de habla adoptada.

Las sonoras, por el contrario, apenas sufren modificaciones. En español, por el contrario, según Magloire y Green (1999), las oclusivas sonoras modifican el VOT con la velocidad del habla, mientras que las sordas apenas sufren modificaciones. Estos autores señalan, igualmente, que el impacto de este factor se origina especialmente en el VOT largo, en el caso del inglés: [p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>]; y en el VOT negativo, en el español: [b̥ ɰ̥ g]. Es decir, el VOT aumenta cuando la dicción es rápida o lenta. Sin embargo, apuntan que el impacto de la velocidad de habla no es extrapolable al VOT corto de ambas lenguas: [b d g] en inglés y [p t̥ k] en español.

#### 4.2. La tonicidad y posición en la palabra

Castañeda (1986) examina, entre otras cosas, la relación entre el VOT y la tonicidad. Esta autora señala que tanto las oclusivas sordas como las sonoras, a excepción de la [g], toman valores más altos de VOT cuando la tonicidad recae sobre la vocal de la sílaba siguiente, es decir, que la sonoridad comienza más tarde. Castañeda indica, además, que hay diferencias en el VOT según la tonicidad de la sílaba y la sonoridad. Para las oclusivas sordas, sea la sílaba tónica o átona, el VOT es aproximadamente 1,5 ms; para las sonoras hay menos uniformidad, pues los valores medios de la [b] están en torno a los 9 ms.

Con respecto a la posición en la palabra, encontramos posturas diversas. Asensi et al. (1997, pp. 239-240), por ejemplo, indican que las oclusivas, en posición inicial e intervocálica, no presentan muchas diferencias; y afirman que, en la práctica, sus realizaciones pueden considerarse independientemente de su posición en la cadena sonora. Déniz (2005, p. 106), por el contrario, indica que el VOT se reduce en posición intervocálica. En esta misma última línea, Lewis (2000), que aporta datos de Bilbao, Medellín y Bogotá, informa de importantes variaciones ligadas a la posición en la frase, especialmente, en el caso de los sonidos [p] y [t] intervocálicos.

#### 4.3. El lugar de articulación y el VOT

El lugar de articulación de la oclusiva puede afectar también a la configuración del VOT. Castañeda (1986, p. 109) señala que cuanto más posterior es la consonante, más tarde comienza la sonoridad de la oclusiva e informa de que, en español, el orden de menor a mayor en el VOT de las oclusivas sordas es: [p] > [t̥] > [k].

Rosner et al. (2000), por su parte, indican que las labiales y las dentales sonoras tienen un VOT más largo que las velares; y que, en el caso de las oclusivas sordas, ocurre lo contrario. Roldán y Soto-Barba (1997), por el contrario, señalan que la dirección del incremento de la duración es el opuesto, esto es, que la duración aumenta, en las oclusivas sonoras, de la velar a la bilabial. Es decir, que el lugar de articulación funciona a la inversa y que cuanto más posterior es la consonante, más próximo es el valor VOT a cero.

Lisker y Abramson (1964, p. 399) señalan, por último, que hay algunos indicios de que la medida del VOT puede ser influida por el lugar del cierre de la oclusiva, ya que las velares parecen tener, según indican los autores, valores consistentemente más altos que el resto.

#### 4.4. Nuestro estudio

Como se pone de manifiesto en este breve repaso, no hay un consenso general sobre los factores externos concretos que influyen en la configuración final del VOT. No obstante, sí parece haber coincidencia en que la tonicidad, la velocidad del habla y el lugar de articulación podrían tener una cierta influencia.

En nuestro trabajo, como veremos más tarde, no tendremos en cuenta la tonicidad, pero sí la posición que ocupa el sonido analizado en la palabra, su lugar de articulación y la velocidad de habla, pues estos factores parecen estar relacionados con la emoción. No obstante, para nosotros, los factores externos serán siempre

analizados como dependientes del contexto emocional y no al contrario.

## 5. Estudios previos sobre el VOT en la lengua oral

Hasta la fecha el único trabajo –que conocemos– que estudia el VOT en la lengua oral es el de Déniz (2005), que analiza las oclusivas sordas en entrevistas en la norma culta del español de Las Palmas de Gran Canaria en hombres y mujeres. Las edades de los informantes oscilan desde los veinticinco años hasta mayores de cincuenta y seis. Esta autora evalúa las variables sociolingüísticas de la edad y el sexo de los informantes y las variables fonéticas del acento, el tipo de vocal dentro de la sílaba y el sonido que precede, es decir, si la oclusiva es intervocálica o si va precedida de otra consonante.

La relación del punto de articulación y el comienzo de la sonoridad ha sido analizada en varios trabajos. Autoras como Castañeda (1986) y Déniz (2005) indican que el VOT aumenta desde la zona labial hasta la velar: [p] > [t̥] > [k]. En este sentido, Rosner et al. (2000, p. 222) apuntan el mismo tipo de observación, a saber: las labiales y dentales tienen un VOT negativo más duradero que las velares; y menor VOT corto que las consonantes velares.

Asimismo, Déniz (2005) compara los valores medios de sus resultados con los de otros trabajos. En general, encuentra que la diferencia entre [p]/[t̥] es reducida, pero se incrementa con [t̥]/[k]. Déniz (2005) también apunta que la diferencia del VOT entre [p]/[t̥] y [t̥]/[k] aumenta, 6,9 ms y 11,6 ms, respectivamente. Castañeda (1986) solo obtuvo valores ampliamente mayores para la diferencia entre [t̥]/[k], 15,3 ms, mientras que observó una diferencia de 3,9 entre [p]/[t̥].

Déniz (2005), por otra parte, compara sus resultados con los obtenidos por Herrera (1997), que analizaba las oclusivas sordas del español de Las Palmas con informantes que leían una lista de palabras. Los valores medios de Herrera (1997)

arrojan la diferencia entre la lectura y un estilo más espontáneo con la oclusiva bilabial sorda, [p]. En general, el aumento del VOT de la bilabial hace que casi se solapen [p]/[t̥]. Este incremento en [p], en Herrera (1997), provoca que [p] y [t̥] lleguen “a tocarse” y se reduzca el VOT entre ellas hasta límites periféricos. Por tanto, se necesitan otras pistas para diferenciar estos sonidos.

## 6. Nuestra investigación

El objetivo de esta investigación es, como señalamos en la introducción, analizar el VOT de las oclusivas sordas y sonoras del español en un contexto dialógico emocional simulado. Con este fin, analizaremos cómo influyen dos emociones, alegría y enfado, sobre la producción del diálogo y qué diferencia estos dos contextos de la neutralidad. Examinaremos igualmente si hay coincidencia o no entre los datos extraídos de nuestro contexto neutro y los resultados proporcionados por otros autores (principalmente en contexto de palabras aisladas). Los cotejos emocionales, por último, serán sometidos a un análisis estadístico para averiguar la significación de los resultados obtenidos.

### 6.1. Cuestiones metodológicas

Para llevar a cabo la investigación creamos, en primer lugar, un *diálogo polifuncional* que incluyese los sonidos que queríamos analizar: [b ɔ̃] y [p t̥ k].

- (1)  
 A: Hola, buenas, ¿qué tal? [k] [t̥]  
 B: Bien, gracias. [b] [ɔ̃]  
 A: Dime, ¿te pongo algo? [ɔ̃] [p]  
 (...)

Este tipo de diálogos simula situaciones cotidianas (en nuestro caso, una interacción camarero-cliente en un bar), contiene un mismo número de elementos (sonidos, palabras, intervenciones, etc.), y está pensado como una herramienta de trabajo que permita obtener ejemplos(-situaciones) cotejables (Padilla, 2017).

Usando, pues, un mismo diálogo, pensado para ser representado de formas distintas, las actrices de nuestra investigación (en nuestro caso, no profesionales) interpretaron las diferentes situaciones (alegría, enfado y neutro) después de haber sido animadas a interactuar entre ellas con la mayor naturalidad posible. Antes de interpretar el diálogo, las actrices fueron informadas del objetivo de la representación (simular dialógicamente alegría, enfado y neutro en una situación cotidiana) y tuvieron tiempo para memorizar las distintas intervenciones. Partimos de la idea de que la forma dialogada y la interacción entre las participantes acerquen el resultado final –a pesar de las lógicas diferencias– a situaciones dialógicas reales.

El diálogo que aparece en (1) fue dramatizado en un entorno cerrado (silencioso y sin eco) por dieciocho parejas de mujeres universitarias de la ciudad de Alicante con edades comprendidas entre los 18 y 25 años. Cada pareja de mujeres grabó el diálogo de tres maneras distintas: neutra, alegre y enfadada, pudiendo ensayar la representación las veces necesarias para sentirse cómodas con el texto antes de la grabación final.

De las 18 parejas iniciales se seleccionaron finalmente diez. Para llevar a cabo esta selección, dos evaluadores externos (un hombre y una mujer) calificaron las grabaciones de las 18 parejas de mujeres por su mayor o menor proximidad a la naturalidad, utilizando para ello una escala de Likert 1-5, que evaluaba la capacidad de las actrices para diferenciar fónicamente las tres situaciones mencionadas. Los dos evaluadores debían indicar si las tres representaciones de cada pareja sonaban: 1) igual, 2) parecido, 3) medianamente parecido, 4) poco parecido y 5) nada parecido (diferente). Hecho esto, se seleccionaron para el análisis aquellas parejas que habían obtenido la mejor puntuación perceptiva según el criterio de los dos evaluadores.<sup>7</sup> El resultado final de todo el proceso fueron 30 ejemplos de diálogos emocionales (3 x

<sup>7</sup> Obviamente, los evaluadores perceptivos tienen un criterio subjetivo sobre qué es más o menos cercano a lo real. No obstante, la evaluación externa proporciona un criterio de selección medianamente objetivo.

10 parejas) que proporcionaron 180 ejemplos, o segmentos analizables.

Todos los sonidos que examinamos aparecen en principio de palabra (vid. supra) y las palabras seleccionadas para el análisis son las siguientes: [b] de *bien*, [d̥] de *dime*, [g] de *gracias*, [p] de *pongo*, [k] de *qué* y [t̥] de *tal*.

Como ya indicamos, proporcionamos las dos mediciones, barra y VOT, cuando los espectrogramas lo permiten. No obstante, es el punto 0 (liberación de la oclusión) el que nos sirve como punto de partida para obtener los datos de [b d̥ g] y [p t̥ k] que analizamos. Consideramos, por tanto, que, en el caso de las oclusivas sordas, el VOT es la suma de la barra de explosión y el resto de espacio anterior a la sonoridad del sonido contiguo. Las transiciones del segundo formante (F2) nos han servido para delimitar ese comienzo de la sonoridad. Todas las mediciones se han realizado usando el programa de análisis acústico *Praat* (Boersma y Weenink, 2019).

## 7. Análisis acústico

El primer objetivo de nuestra investigación era, como indicamos, analizar si la variación emocional, es decir, el cambio emoción (neutro, alegre, enfadado) puede influir en el VOT de los ejemplos de los diálogos.

Para responder a esta pregunta hemos extraído, en primer lugar, los sonidos [b] de *bien*, [d̥] de *dime*, [g] de *gracias*, [p] de *pongo*, [k] de *qué* y [t̥] de *tal* de los contextos mencionados (neutro, alegre, enfadado). Todos ellos en posición inicial de palabra. A continuación, hemos analizado la barra de explosión y el VOT de [p t̥ k] y el VOT de [b d̥ g], tomando la liberación de la oclusión como punto 0 de referencia y el segundo formante (F2) como finalización de esta segmentación.

Los datos de estos análisis aparecen en el Apéndice 1. Para facilitar su comprensión resumimos los resultados más importantes en la Tabla 5, que contiene: el número de parejas analizadas (N), la

media (VOT negativo, en el caso de las sonoras; y VOT positivo, en las sordas), la desviación estándar

(SD), los máximos y mínimos obtenidos y el error estándar (SE).

	N.	Media	Desv. est.	Mín.	Máx.	Err. est.
que-neutro	10	26,79	3,48	20,43	31,67	1,10
que-alegre	10	22,66	7,25	12,368	32,60	2,29
que-enfadado	10	23,58	5,18	16,305	34,77	1,64
tal-neutro	10	19,69	6,32	9,699	31,30	2,00
tal-alegre	10	14,22	3,45	9,536	20,52	1,09
tal-enfadado	10	14,86	4,01	9,679	21,27	1,27
pongo-neutro	6	15,72	3,38	9,665	19,05	1,38
pongo-alegre	6	19,94	11,75	10,448	40,25	4,80
pongo-enfadado	6	12,68	6,13	6,311	22,37	2,50
dime-neutro	10	-46,17	20,95	-83,313	-15,90	6,62
dime-alegre	10	-40,73	17,94	-64,007	-10,18	5,67
dime-enfadado	10	-41,87	18,01	-73,90	-12,76	5,70
gracias-neutro	10	-25,93	12,16	-47,65	-11,17	3,85
gracias-alegre	10	-23,91	11,21	-43,77	-7,06	3,54
gracias-enfadado	10	-39,05	26,05	-90,05	-13,08	8,24
bien-neutro	10	-47,03	19,41	-71,76	-20,67	6,14
bien-alegre	10	-41,14	24,72	-72,22	-15,37	7,82
bien-enfadado	10	-55,52	34,28	-136,31	-20,90	10,84

**Tabla 5.** Resultados de la estadística descriptiva de los sonidos [p, t̪, k] (barra+VOT) y [b, d̪, g] (VOT).

El número de muestras en ‘pongo’ [p] es, como se observa en la tabla, menor. Esto se justifica, como veremos más tarde, porque, en algunos casos, los informantes pronunciaron el sonido [p] bien como oclusiva [b] o como aproximante [β] invalidando su análisis.

Tomando como punto de partida los datos extraídos del análisis acústico, hemos comparado, en segundo lugar, los valores del VOT de nuestro corpus con los valores de otros trabajos (Tabla 6), que utilizaron corpus distintos, resultantes,

principalmente de la lectura de palabras aisladas. Aclaremos, además, que los valores de Poch (1984) que presentamos proceden de oclusivas en posición intervocálica; los resultados de Roldán y Soto-Barba (1997) son la combinación de pares mínimos de sintagmas nominales, *un peso-un beso*, y logotomos, *ancuna-anguna*. Por último, informamos de que Asensi et al. (1997) obtienen datos de oclusivas sordas en posición inicial de palabra y posición intermedia, no apreciando diferencias significativas entre ambas posiciones.

Autor	Variedad	Metodología	Oclusivas sonoras			Oclusivas sordas		
			[b]	[d]	[g]	[p]	[t]	[k]
Lisker y Abramson (1964)	Puerto Rico	Lectura de palabras	-138	-110	-108	4	9	29
Borzone (1980)	Argentina		-70	-50	-40	10	15	25
Poch (1984)	Barcelona	Lectura de palabras	---	---	---	18	17	32
Castañeda (1986)	Barcelona	Lectura de palabras	-69,8	-77,7	-58	6,5	10,4	25,7
Asensi et al. (1997)	Barcelona	Lectura de palabras	---	---	---	14,7 (9,1)	20,2 (13,2)	35,4 (21,2)
Roldán y Soto-Barba (1997)	Valdivia (Chile)	Lectura de palabras	-85,2	-75,6	-62,7	13,2	16,4	30
Herrera (1997)	Gran Canaria	Lectura de palabras	---	---	---	16,1	16,7	26,4
Machuca (1997)	Barcelona	Entrevistas semidirigidas	---	---	---	12,64	18,47	27,70
Rosner et al. (2000)	Madrid	Lectura de palabras	-91,5	-91,6	-73,7	13,1	14,0	26,5
Déniz (2005)	Las Palmas	Entrevista	---	---	---	9,7	16,6	28,2
<b>Promedio de los estudios previos</b>			-90,9	-80,98	-68,48	11,79	15,38	28,59
<b>Este estudio</b>	Alicante	Diálogo simulado (representación)	-47,02	-46,17	-25,93	15,72	19,69	26,79

**Tabla 6.** Tabla comparativa valores medios con otros estudios.

### 7.1. Resultados obtenidos sobre la realización fonológica

La comparación de nuestros datos en *contexto neutro* con los datos en promedio de los otros trabajos (Tabla 6) señala un comportamiento relativamente similar en el caso de las oclusivas sordas, pero muy distinto en el caso de las oclusivas

sonoras. En estas últimas, nuestras medias son mucho más bajas que los promedios de los resultados previos (vid. supra). No obstante, como mostramos en la Tabla 5, el grado de dispersión (SD) de los sonidos oclusivos sonoros es bastante mayor que en los sordos, por consiguiente, estas diferencias deben ser vistas con cautela.

A pesar de ello, nuestros resultados concuerdan con algunas de las informaciones expuestas en estudios anteriores, pues, Castañeda (1986), por ejemplo, mencionaba que la dispersión es mucho mayor en los casos de [b, d, g]; y Roldán y Soto-Barba (1997), por su parte, indicaban que “el contexto fonético incide en la variación de los valores del VOT solo en la serie de consonantes sonoras”. Los datos que presentamos, como decíamos, proceden de la representación de un diálogo (habla simulada), consecuentemente, puede deducirse que la forma de obtención de ejemplos, más próxima al habla natural, parece influir en la pronunciación de los informantes.

Con respecto a la aparición o no de la barra de explosión en el espectrograma, nuestros análisis acústicos manifiestan, como mostramos en el Apéndice 1, que, como presumíamos, en contextos dialógicos, no es tan frecuente su aparición. Los análisis acústicos con *Praat* han revelado que, en nuestro corpus, es una pista acústica débil. De 30 casos posibles por cada sonido, en las oclusivas sonoras no aparece 20 veces para /b/, 11 para /d/ y, en ningún caso, para /g/. En las oclusivas sordas aparece siempre en /t, k/ y para /p/ no se encuentra en 12 casos. De un total de 180 casos, falta, pues, en 73 de ellos, siendo más acusado en /b d g/, con 61 casos de 90 posibles.

Con respecto a la caracterización acústica de los fonemas examinados, nuestros análisis en contextos neutros revelan que el fonema /p/ es la oclusiva sorda más inestable en contexto dialógico. Los fonemas /t, k/, sin embargo, se han mantenido más fieles a sus respectivos fonemas. Podríamos intentar justificar el debilitamiento de /p/, en primer lugar, por su posición en el grupo fónico, y, en segundo lugar, por la alta frecuencia de la frase ‘Dime, ¿te pongo algo?’ (que tiende a la gramaticalización)<sup>8</sup>. También podríamos pensar que /p/, al estar tan próximo al punto cero, puede ser más inestable

<sup>8</sup> Algo similar ocurre con la frase ‘lo que quieras’ (aunque no la hemos analizado) en nuestro diálogo. El primer sonido /k/, ‘lo que’, se pronuncia en la mayor parte de ocasiones como aproximante [ɣ] (solo en 4 ocasiones de 30 posibles se realiza como [k]).

(sucede también con /k/, es decir, su resultado en ocasiones es [g], categóricamente hablando está en el otro lado del punto cero, en la parte negativa del continuo). Es decir, se trata de un sonido dentro del grupo fónico y con mucha frecuencia de uso. Es un ejemplo de falta de correspondencia entre el fonema y el alófono esperable. En definitiva, parece que, según nuestros datos, la posición del sonido dentro del grupo fónico influye en su producción.

Los valores de nuestras oclusivas sonoras, por otra parte, son muy reducidos en comparación con los otros trabajos. El promedio del VOT para /b/ es muy bajo, situándose casi a la par con /d/. Los valores de dispersión (Tabla 7) de [b] junto con [d], igual que sus medias, sitúan estos dos sonidos en una franja equivalente.

	Mín.	Máx.
<b>bien neu</b>	-20,67	-71,76
<b>dime neu</b>	-15,90	-83,31
<b>gracias neu</b>	-11,17	-47,65
<b>pongo neu</b>	9,67	19,05
<b>tal neu</b>	9,70	31,30
<b>que neu</b>	20,43	31,67

**Tabla 7.** Rango de dispersión mínimo y máximo para cada sonido.

Quizá este fenómeno se deba a la inestabilidad de /b/ en posición inicial de palabra. De hecho, no es extraño que en español hablado este sonido pueda eliminarse o intercambiarse con otro con rasgos labiales, por ejemplo: ‘Virgen Santa’ pasa a ‘Gensanta’, ‘mahonesa’ a ‘bayonesa’, ‘buenas’ a ‘güenas’; e, igualmente, en posición media de palabra, ‘vagabundo’ pasa a ‘vagamundo’, etc. Podríamos tratar de explicar este comportamiento debido a que, en nuestros datos, /b/ y /d/ van seguidos de un segmento vocálico alto y anterior, la semivocal [i̯] y la vocal [i], respectivamente. Rosner et al. (2000) también obtienen datos similares entre la oclusiva bilabial y la dental, aunque diferenciados por la vocal contigua, /ba/, /bo/ y /da/, /do/ y la configuración de la palabra, bisilábica (consonante-vocal + consonante-vocal).

También resulta interesante observar que /b/ tiene unos valores muy reducidos –comparados con otros trabajos– y el VOT de /p/ llega incluso a sonorizarse. El fonema /g/, por último, en ‘gracias’ se ha visto afectado probablemente por la vocal esvarabática. Es el único sonido oclusivo que hemos estudiado que no está seguido de vocal, /gr/, y es, por otra parte, la palabra que contiene más segmentos.

Nuestros datos, por otra parte, confirman los estudios que analizan la sonorización de las oclusivas sordas. Hemos encontrado las sonorizaciones documentadas en Martínez Celdrán (2009), quien siguiendo la metodología del *map task*, grabó a una mujer de manera semiespontánea y concluyó que hay grados de sonorización en las oclusivas sordas. Consecuentemente, estos datos también han puesto de relieve que el contraste laríngeo entre sorda-sonora se llega a neutralizar en el punto de articulación bilabial [p, b]. Es decir, /p/ se puede realizar como [b]. Los alófonos de /p/ son, por tanto, [p], [b] y [β]. Los alófonos de /b/ son [b] y [β]. Hay neutralización. La distinción puede venir dada por la duración de la oclusión.

Con respecto al patrón preferido en las lenguas del mundo [p] > [t] > [k], nuestros datos, como se muestra en la Tabla 6, lo confirman: [p] (15,72) > [t] (19,69) > [k] (26,79). No obstante, como hemos

mencionado anteriormente, los resultados de /p/ hay que tomarlos con precaución atendiendo al número de ejemplos.

En el caso de las oclusivas sonoras, podemos concluir que la duración media del cierre de las oclusivas es menor que el de otros estudios. Es decir, que, en contexto dialógico (simulado), se reduce o bien la vibración de las cuerdas vocales o bien el tiempo de oclusión. Nuestros resultados informan, asimismo, de que el rango de dispersión del VOT en español es más estrecho que el propuesto por otros autores como Castañeda (1986).

## 7.2. Resultados sobre los cotejos emocionales

La Tabla 5 recoge, como habíamos señalado, las diferencias entre los estados emocionales de las informantes: neutralidad, alegría, enfado. Una primera mirada a los resultados nos proporciona la siguiente información (vid. supra): los dos contextos emocionales analizados suponen siempre una modificación del VOT del contexto neutro, bien aumentándolo, bien disminuyéndolo.

La direccionalidad de los valores, esto es la dirección que presenta el cambio, es también reseñable.

	Neutro	Alegre	Enfado
<b>Pongo</b>	15,73	<b>19,94</b>	12,68
<b>Tal</b>	19,69	14,22	14,86
<b>Qué</b>	26,79	22,66	23,58
<b>Bien</b>	47,03	41,14	<b>55,52</b>
<b>Dime</b>	46,17	40,74	41,87
<b>Gracias</b>	25,93	23,91	<b>39,05</b>

**Tabla 8.** Comparación promedios neutro, alegre, enfado por sonido.

Como mostramos en la Tabla 8, todos los casos de alegría, salvo ‘pongo’ [p] (que recordamos que tiene un número menor de ejemplos) suponen una disminución de la media. Los casos de enfado muestran más variación, especialmente en las oclusivas sonoras. Todas las oclusivas sordas en el

contexto de enfado muestran un descenso de la media. En el caso de las sonoras, la oclusiva sonora [d] desciende y las oclusivas [g] y [b] aumentan.

Si tenemos en cuenta el cómputo global, y comparamos emoción (sea alegría o enfado) y

situación neutra, es posible señalar que las emociones analizadas, en términos generales (un 75% de los casos: 9 de 12), disminuyen el VOT. Es decir, cuando se expresan emociones, se habla más rápido y las cuerdas vocales tardan menos tiempo en empezar a vibrar. Estos resultados, aunque no concluyentes, sugieren que, en la alegría y el enfado, hay una tendencia a aumentar la velocidad del habla, y, consecuentemente a disminuir el VOT. Estos resultados confirman, por otra parte, lo expuesto en trabajos previos sobre el efecto de la velocidad en el VOT de las oclusivas (Kessinger y Blumstein, 1997; Magloire y Green, 1999).

### 7.2.1. Análisis estadístico y cotejo de situaciones

Los datos acústicos sobre las situaciones emocionales de la Tabla 5 fueron sometidos posteriormente a un análisis estadístico, con el propósito de estudiar si existían diferencias significativas en el tiempo del VOT del mismo sonido en función del contexto. Los contextos particulares comparados fueron: [segmento neutro] vs. [segmento alegre], [segmento neutro] vs. [segmento enfadado] y [segmento alegre] vs. [segmento enfadado], para la muestra de diez conversaciones, salvo en el caso de *pongo*, en el que la muestra de parejas es más pequeña (N=6) por el motivo ya explicado anteriormente.

Para llevar a cabo este análisis, se realizó un contraste de hipótesis a cada par de muestras utilizando el software *SPSS Statistics* (versión 26), con un nivel de significación del 95%, es decir, que existe un 5% de probabilidad de cometer un error al afirmar la conclusión de cada prueba de contraste de hipótesis realizada. Como las muestras de datos que se comparan no son independientes, pues corresponden al VOT del mismo sujeto en dos contextos distintos, se aplicó, en primer lugar, la prueba paramétrica Paired Samples *t* Test (*t*-test de aquí en adelante), que consiste en una variación de la prueba *t* de Student en el caso de muestras no independientes. Previamente a la realización de este análisis, llevamos a cabo una prueba de normalidad, también en SPSS, para verificar que los datos cumplieran la hipótesis del *t*-test de estar

aproximadamente normalmente distribuidos. Al ser una muestra pequeña (menos de 50 datos), aplicamos la prueba de normalidad Shapiro-Wilk, habiendo obtenido como conclusión que se puede suponer que todas las muestras emparejadas siguen una distribución normal, salvo en el caso de *gracias-enfadado*. Recordamos, igualmente, que los datos de *pongo* deben ser mirados con cautela, puesto que no todos los segmentos fueron susceptibles de análisis espectrográfico, es decir, el número de ejemplos de [p] es menor que en el resto de sonidos. Mostramos los resultados del análisis en la Tabla 9.

El examen de los datos de la Tabla 9 proporciona la siguiente información. En principio, no podemos rechazar la hipótesis nula, es decir, no existen diferencias significativas en el tiempo medio de realización del VOT en función del contexto, salvo en el cotejo particular *tal neutro-alegre* (pair 4) ( $p < ,05$ ). En este caso, el *t*-test señala que los participantes de nuestro estudio, de media, y de forma estadísticamente significativa, realizan un VOT más largo cuando el contexto es neutro (M (media) = 19,69 ms, DE (Desv. Estándar) = 6,32, SEM (Error Medio Estándar) = 1,99) que cuando el contexto es alegre (M = 14,22 ms, SD = 3,45, SEM = 1,09),  $t(9) = 3,4$ ,  $p < 0.05$ ,  $r = 0,96$ )<sup>9</sup>. Teniendo en cuenta que un valor de  $r = 0,5$  es el umbral de un gran efecto,<sup>10</sup> en este caso concreto, el análisis presenta un gran efecto (*effect size*), consecuentemente, representa una conclusión sustantiva.

En una segunda fase, aplicamos la prueba no paramétrica Wilcoxon Signed-Ranks a los datos (Apéndice 2), pues esta prueba es adecuada cuando la muestra es muy pequeña (menos de 30 datos), como es nuestro caso, y también en los casos en los que no se puede suponer que los datos están normalmente distribuidos (como sucede en *gracias-enfadado*, por lo que los resultados de las parejas

<sup>9</sup> Para obtener  $r$ , es decir, el tamaño del efecto (*effect size*), se ha aplicado la siguiente fórmula (Rosenthal, 1991; Rosenthal et al., 2000; Rosnow y Rosenthal, 2005):  $r = \sqrt{t^2/(t^2 + df)}$ , donde  $t = 3,4$  y  $df = 9$ .

<sup>10</sup> Véase Field (2009, p. 334).

14 y 15 de la Tabla 9 no podrían ser interpretadas). El resultado de esta prueba nos indica que solo se puede concluir que existe una diferencia estadísticamente significativa en el tiempo medio de realización del VOT en los casos de *tal neutro-*

*alegre* ( $p < 0,05$ ). Nótese que este cotejo es el que mostraba el p-valor más pequeño en el t-test, por consiguiente, el test de Wilcoxon Signed-Ranks corrobora los resultados del t-test.

		Diferencias pareadas					t	df	Sig. (2-tailed)
		M	DE	SEM	95% interv. conf. df.				
					Inf.	Sup.			
Pair 1	que-neutro - que-alegre	4,13	6,96	2,2	-0,85	9,11	1,88	9	0,09
Pair 2	que-neutro - que-enfadado	3,21	5,09	1,61	-0,43	6,85	2	9	<b>0,08</b>
Pair 3	que-alegre - que-enfadado	-0,92	5,91	1,87	-5,15	3,31	-0,49	9	0,64
Pair 4	tal-neutro - tal-alegre	5,47	5	1,58	1,89	9,05	3,46	9	<b>0,01</b>
Pair 5	tal-neutro - tal-enfadado	4,84	7,01	2,22	-0,18	9,85	2,18	9	<b>0,06</b>
Pair 6	tal-alegre - tal-enfadado	-0,63	3,54	1,12	-3,17	1,9	-0,57	9	0,59
Pair 7	dime-neutro - dime-alegre	-5,43	20,21	6,39	-19,89	9,03	-0,85	9	0,42
Pair 8	dime-neutro - dime-enfadado	-4,3	31,42	9,94	-26,77	18,18	-0,43	9	0,68
Pair 9	dime-alegre - dime-enfadado	1,14	18,72	5,92	-12,26	14,53	0,19	9	0,85
Pair 10	pongo-neutro - pongo-alegre	-6,73	11,89	5,95	-25,66	12,2	-1,13	3	0,34
Pair 11	pongo-neutro - pongo-enfadado	3,44	4,06	1,82	-1,6	8,49	1,9	4	0,13
Pair 12	pongo-alegre - pongo-enfadado	9,77	9,37	4,69	-5,14	24,68	2,09	3	0,13
Pair 13	gracias-neutro - gracias-alegre	-2,02	11,57	3,66	-10,3	6,26	-0,55	9	0,59
Pair 14	gracias-neutro - gracias-enfadado	13,11	28,99	9,17	-7,62	33,85	1,43	9	0,19
Pair 15	gracias-alegre - gracias-enfadado	15,13	30,59	9,67	-6,75	37,01	1,57	9	0,15
Pair 16	bien-neutro - bien-alegre	-5,89	16,77	5,3	-17,88	6,1	-1,11	9	0,3
Pair 17	bien-neutro - bien-enfadado	8,49	40,84	12,91	-20,72	37,7	0,65	9	0,53
Pair 18	bien-alegre - bien-enfadado	14,38	38,94	12,31	-13,47	42,23	1,16	9	0,27

Tabla 9. Tabla con los resultados del t-test.

Como indicábamos en el apartado 7.2., todos los casos de alegría, salvo 'pongo' [p] suponían una disminución de la media con respecto a los contextos neutros. Con el propósito de evaluar el grado de significación de esta tendencia, aplicamos a las parejas de cotejos (pairs) dos pruebas estadísticas: el t-test y la prueba no paramétrica Wilcoxon Signed-Ranks. De todas las parejas analizadas, solo el pair 4 *tal neutro-tal alegre* obtiene una diferencia en la media del VOT estadísticamente significativa: p-valor < 0,05, en ambos casos. Los resultados obtenidos en el análisis estadístico solo confirman, por tanto, el caso del pair 4 *tal neutro-tal alegre*, pues los tests de contraste de hipótesis realizados nos informan de que, en el resto de los casos, no existe evidencia estadística suficiente para poder rechazar que las medias sean iguales en las emociones examinadas.

## 8. Conclusiones

Nuestro estudio, como adelantábamos en la Introducción, tenía como objetivo el análisis del VOT en contextos dialógicos simulados, introduciendo, además, la variable emoción para comprobar si el VOT se modifica cuando se expresa alegría y enfado. Es decir, hemos cotejado los resultados medios de neutro-alegría, neutro-enfado y alegría-enfado.

Hemos comparado igualmente nuestros valores neutros con los valores obtenidos en otros trabajos en el ámbito hispánico. Con esta metodología hemos querido comprobar la fortaleza de la categoría, es decir, hemos examinado si las propiedades acústicas que definen el VOT son modificables.

Hemos defendido el papel de la barra de explosión, pero no la hemos incluido en el análisis, porque, como ya indicamos, en los diálogos no es tan frecuente su aparición. Los análisis acústicos con *Praat* han revelado que, en contexto dialógico, es una pista acústica débil. De 30 casos posibles por cada sonido, en las oclusivas sonoras no aparece 20 veces para /b/, 11 para /d/ y en ningún caso para /g/. En las oclusivas sordas aparece siempre en /t, k/ y para /p/ no está en 12 casos. De un total de 180 casos, falta en 73 de ellos, siendo más acusado en /b d g/, 61 casos de 90 posibles.

Con respecto a la caracterización acústica de los fonemas examinados, nuestros análisis en contextos neutros revelan que el fonema /p/ es la oclusiva sorda más inestable en contexto dialógico. Los fonemas /t, k/, sin embargo, se han mantenido más fieles a sus respectivos fonemas. Aunque el primer sonido /k/ en ‘lo que’, se pronuncia principalmente como aproximante [ç], siendo analizado mejor como un fenómeno de gramaticalización.

Los valores de nuestras oclusivas sonoras, por otra parte, son muy reducidos, vistas desde la perspectiva de los otros trabajos. El promedio del VOT para /b/ es muy bajo comparado con los de otros autores, situándose casi a la par con /d/. Los valores de dispersión (Tabla 7) de [b] junto con [d], igual que sus medias, sitúan estos dos sonidos en una franja equivalente.

También resulta interesante observar que /b/ tiene unos valores muy reducidos –comparados con otros trabajos– y el VOT de /p/ llega incluso a sonorizarse. El fonema /g/, por último, en ‘gracias’ se ha visto afectado probablemente por la vocal esvarabática, es el único sonido oclusivo que hemos estudiado que no está seguido de vocal, /gr/, y es la palabra con más sonidos.

Finalmente, nuestros datos confirman los estudios que analizan la sonorización de las oclusivas sordas (Hualde, 2014).

Con respecto al patrón preferido en las lenguas del mundo es [p] > [t] > [k], por su relación con el

punto de articulación, nuestros resultados, en contextos neutros, confirman el orden: [p] (15,72) > [t] (19,69) > [k] (26,79).

En relación con la fortaleza de la relación entre el VOT y la variable *emoción*, hemos comprobado que los dos contextos emocionales analizados modifican los resultados del VOT. En los dos contextos, o bien aumenta o bien disminuye; no obstante, el grado de la variación es distinto en función del caso analizado. La direccionalidad de los valores, esto es la dirección que presenta el cambio, es también reseñable. Todos los casos de alegría, salvo ‘pongo’ [p] (que recordamos que tiene un número menor de ejemplos) suponen una disminución de la media. Los casos de enfado muestran más variación, especialmente en las oclusivas sonoras. Todas las oclusivas sonoras en el contexto de enfado muestran un descenso de la media; la oclusiva sonora [d] desciende y las oclusivas [g] y [b] aumentan. Si tenemos en cuenta el cómputo global, y comparamos emoción (sea alegría o enfado) y situación neutra, es posible señalar que las emociones analizadas, en términos generales (un 75% de los casos: 9 de 12), disminuyen el VOT. Es decir, cuando se expresan emociones, se habla más rápido y las cuerdas vocales tardan menos tiempo en empezar a vibrar. Estos resultados, aunque no concluyentes, sugieren que, en la alegría y el enfado, hay una tendencia a aumentar la velocidad de habla, y, consecuentemente a disminuir el VOT. Nuestros resultados confirman, por otra parte, lo expuesto en trabajos previos sobre el efecto de la velocidad de habla en el VOT de las oclusivas. El efecto en el VOT ha sido la reducción. Interpretamos que la velocidad de habla ha sido rápida. Kessinger y Blumstein (1997) y Magloire y Green (1999) encontraron este cambio en el VOT para el VOT negativo y VOT largo, pero nunca para el VOT corto. La conclusión en esos estudios en torno al efecto de la velocidad de habla incide en que el VOT en los polos opuestos (VOT negativo y largo) son modificables, excepto el VOT corto.

Con el propósito de evaluar el grado de significación de estas tendencias reseñadas,

aplicamos a las parejas de cotejos (pairs) dos pruebas estadísticas: la prueba *t* de Student para muestras dependientes y la prueba no paramétrica Wilcoxon Signed-Ranks. De todas las parejas analizadas solo en el pair 4 *tal neutro-tal alegre* se obtiene como resultado que las medias son distintas (pues el nivel de significación es menor de 0,05), en ambos casos.

Los resultados obtenidos en el análisis estadístico, sin embargo, no son concluyentes (salvo en el caso del pair 4 *tal neutro-tal alegre*), pues los test de contraste de hipótesis realizados informan de que no existe evidencia estadística suficiente para poder rechazar que las medias sean iguales en los contextos analizados. Esto puede ser debido, quizás, al número de muestras analizadas. Un posible desarrollo futuro de nuestra investigación sería, por tanto, aplicar nuestro análisis a una muestra mayor.

### Agradecimientos

Este trabajo se ha llevado a cabo gracias al proyecto del Ministerio de Economía, Industria y Competitividad (MINECO) *El habla con significado emocional y expresivo: Análisis fonopragmático y aplicaciones* (FFI2017-88310-P/MINECO). Agradecemos a los revisores de este trabajo sus acertados comentarios y consejos.

### Referencias bibliográficas

- Asensi, L., Portolés, S., & Del Río, A. (1997). Barra de explosión, VOT y frecuencia de las oclusivas sordas del castellano, *Estudios de Fonética Experimental*, IX, 221-242.
- Beckman, J., Helgason, P., McMurray, B., & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification, *Journal of Phonetics*, 39(1), 39-49.
- Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages, *Journal of Linguistics*, 49(2), 259-284.
- Berkson, K. H. (2012). Capturing breathy voice: Durational measures of oral stops in Marathi, *Kansas Working Papers in Linguistics*, 33, 27-46.
- Boersma, P., & Weenink, D. (2019). *Praat v.6.0.37* [10/05/2020].
- Borzzone, A. M. (1980). *Manual de fonética acústica*. Hachette.
- Castañeda Vicente, M. L. (1986). El V.O.T. de las oclusivas sordas y sonoras españolas, *Estudios de Fonética Experimental*, II, 91-110.
- Déniz, M. T. (2005). El VOT de las oclusivas sordas en la norma culta de Las Palmas de Gran Canaria, *Boletín de Lingüística*, 24, 92-107.
- Ekman, P. (1970). Universal facial expressions of emotions, *California Mental Health Research Digest*, 8(4), 151-158.
- Ekman, P. (1999). Basic emotions. En T. Dalgleish, & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45-60). John Wiley & Sons.
- Field, A. (2009). *Discovering statistics using SPSS*. Sage.
- Fry, D. B. (1979). *The physics of speech*. Cambridge University Press.
- Garrido, J. M. (2019). Análisis de las curvas melódicas del español en el habla emotiva simulada, *Estudios de Fonética Experimental*, XX, 205-55.
- Helgason, P., & Ringen, C. (2008). Voicing and aspiration in Swedish stops, *Journal of Phonetics*, 36(4), 607-628.
- Herrera, J. (1997). Estudio acústico de /p, t, c, k/ y /b, d, y, g/ en Gran Canaria. En M. Almeida, & J. Dorta (Eds.), *Contribuciones al estudio de la lingüística hispánica: Homenaje al profesor Ramón Trujillo* (pp. 73-86). Montesinos.
- Hualde, J. I. (2014). Lenición de obstruyentes sordas intervocálicas en español: Estado de la cuestión. En Y. Congosto Martín, M.<sup>a</sup> L. Montero Curiel, & A. Salvador Plans (Eds.), *Fonética Experimental, Educación Superior e Investigación* (vol. I, pp. 113-136). Arco/Libros.

- Hunnicut, L., & Morris, P. A. (2016). Prevoicing and aspiration in Southern American English, *University of Pennsylvania Working Papers in Linguistics*, 22(1), 24.
- Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English, *Journal of Phonetics*, 25(2), 143-168.
- Ladefoged, P., & Johnson, K. (2006). *A course in phonetics* (5<sup>th</sup> edition). Thomson Wadsworth.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.
- Lewis, A. M. (2000). Acoustic variability of intervocalic voiceless stop consonants in three Spanish dialects. En *Hispanic linguistics at the turn of the millennium: papers from the 3rd Hispanic Linguistics Symposium 2000* (pp. 101-114), Cascadilla Press.
- Lisker, L., & Abramson A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements, *WORD*, 20(3), 384-422.
- Machuca, M. J. (1997). *Las obstruyentes no continuas del español: relación entre las categorías fonéticas y fonológicas en el habla espontánea* (Tesis doctoral inédita). Universitat Autònoma de Barcelona, España.
- Magloire, J., & Green K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish, *Phonetica*, 56(3-4), 158-185.
- Martínez Celdrán, E. (1984). *Fonética*. Teide.
- Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: Articulaciones y sonidos de español*. Ariel.
- Martínez Celdrán, E. (2009). Sonorización de las oclusivas sordas en una hablante murciana: Problemas que plantea, *Estudios de Fonética Experimental*, XVIII, 254-271.
- Martínez Celdrán, E. (en prensa). Fonética y fonología descriptivas de la lengua española. En J. Gil, & J. Llisterri (Eds.), *Fonética y fonología descriptivas de la lengua española*. Washington.
- Martínez Celdrán, E. (2013). Los sonidos obstruyentes en la cadena hablada. En M. A. Peñas Ibáñez (Ed.), *Panorama de la fonética española actual* (pp. 253-289), Arco/Libros.
- Morales del Valle, H. H. (2006). *La estratificación social de las variantes del grupo consonántico /ks/ en el español de estado de Guanajuato* (Trabajo final de grado inédito). Universidad de Guanajuato, México.
- Morris, P. A. (2018). Rate effects on Southern American English VOT, *Proceedings of the Linguistic Society of America 2018*, 3(1), 60:1-10.
- Natarina, A. (2019). A VOT study of the acquisition of English stop contrasts by Marathi speakers: the mutual influence of L1 and L2. En A. A. P. Belda, H. Galbraith, K. Josephs, A. P. Pinto, E. Pulkowski, K. Walker-Cecil, & C. Wuxiha (Eds.): *Research approaches to second language acquisition: Proceedings of the 2018 Second Language Acquisition Graduate Student Symposium (CARLA Working Paper Series)* (pp. 39-64). Minneapolis.
- Padilla, Xose A. (2017). Prosodia y (des)cortesía en contexto de diálogo: La creación y la negociación del ámbito tonal, *LEA*, 39(2), 243-268.
- Padilla, Xose A. (2020). Prosodia emocional y conversación espontánea: Bases para el establecimiento de un protocolo de identificación perceptiva, *Phonica*, 16, 4-35.
- Padilla, Xose A. (en prensa). La voz como reacción emocional: de qué nos informa la prosodia, *Spanish in Context*.
- Piñeros, C. E. (2008). *Estructura de los sonidos del español*. Pearson.
- Plutchik, R. (1994). *The psychology and biology of emotion*. Harper Collins College Publishers
- Poch, M. D. (1984). Datos acústicos para la caracterización de las oclusivas sordas del español, *Folia Phonetica*, 1, 89-106.
- Roldán, Y., & Soto-Barba, J. (1997). El VOT de /ptk/ y /bdg/ en el español de Valdivia: Un análisis acústico, *Estudios Filológicos*, 32, 27-33.
- Rosenthal, R. (1991). *Metaanalytic procedures for social research* (2nd edition). Sage.

- Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioural research: A correlational approach*. Cambridge University Press.
- Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops, *Journal of Phonetics*, 28(2), 217-224.
- Rosnow, R. L., & Rosenthal, R. (2005). *Beginning behavioural research: A conceptual primer* (5th edition). Pearson/Prentice Hall.
- Scherer, K. R., Ladd, D. R., & Silverman, K. E. (1984). Vocal cues to speaker affect: Testing two models, *The Journal of the Acoustical Society of America*, 76(5), 1346-1356.
- Shea, C., & Curtin, S. (2006). Learning allophonic alternations in a second language: Phonetics, phonology and grammatical change. En M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 124-131), Cascadilla Proceedings Project.
- Shea, C. E., & Curtin, S. (2011). Experience, representations and the production of second language allophones, *Second Language Research*, 27(2), 229-250.
- Slis, I. H. (1971). Articulatory effort and its durational and electromyographic correlates, *Phonetica*, 23, 171-188.

## Apéndice 1

Conversación	que (neutro)			que (alegre)			que (enfadado)		
	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT
1	10,69	12,46	23,15	6,83	5,88	12,71	5,87	10,44	16,31
2	13,41	12,27	25,67	16,71	15,89	32,60	20,32	8,41	28,73
3	17,03	14,33	31,36	7,25	13,31	20,56	14,25	6,71	20,96
4	12,42	8,01	20,43	10,62	9,72	20,34	6,17	12,44	18,61
5	13,08	11,95	25,03	15,15	10,87	26,02	6,33	15,13	21,46
6	15,67	15,999	31,67	9,08	23,30	32,38	5,74	17,58	23,32
7	16,93	10,05	26,98	8,69	3,68	12,37	17,30	6,47	23,77
8	19,42	9,47	28,89	6,43	13,68	20,10	16,66	7,56	24,21
9	17,63	10,44	28,07	6,66	22,46	29,12	13,78	20,99	34,77
10	18,66	8,00	26,66	17,06	3,35	20,41	15,37	8,27	23,64
Media	15,49	11,30	26,79	10,45	12,21	22,66	12,18	11,40	23,58

**Tabla i.** Análisis de la barra de explosión y del VOT de la oclusiva velar sorda [k] en contexto neutro, alegre y enfado.

Conversación	tal (neutro)			tal (alegre)			tal (enfadado)		
	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT
1	4,85	7,08	11,93	4,48	5,05	9,536	2,91	10,40	13,31
2	5,55	14,83	20,38	5,31	15,21	20,519	3,47	16,27	19,74
3	7,28	14,34	21,62	5,19	6,96	12,149	2,48	7,20	9,68
4	4,07	5,63	9,70	3,04	7,67	10,714	5,31	7,21	12,52
5	3,27	28,02	31,30	5,57	10,84	16,411	6,86	10,73	17,58
6	8,46	16,45	24,91	6,79	9,76	16,545	4,70	7,53	12,23
7	11,47	7,16	18,62	7,15	6,48	13,620	8,88	12,39	21,27
8	7,70	16,37	24,07	5,74	10,39	16,129	8,95	4,28	13,23
9	5,65	12,05	17,70	5,70	5,07	10,765	4,72	6,07	10,79
10	10,02	6,70	16,72	8,97	6,85	15,820	5,58	12,63	18,21
Media	6,83	12,86	19,69	5,79	8,43	14,22	5,39	9,47	14,86

**Tabla ii.** Análisis de la barra de explosión y del VOT de la oclusiva dental sorda [t] en contexto neutro, alegre y enfado.

Conversación	pongo (neutro)			pongo (alegre)			pongo (enfadado)		
	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT	Barra	VOT	Barra+ VOT
1				2,48	7,97	10,45			
2									
3	2,65	11,50	14,15	5,94	7,09	13,03	4,61	2,11	6,72
4	6,59	12,46	19,05	5,77	34,48	40,25	13,44	8,93	22,37
5									
6	5,30	4,36	9,67				2,04	4,28	6,31
7	12,95	4,22	17,17				5,51	7,18	12,69
8	8,10	8,46	16,56	3,78	24,30	28,08	2,03	9,26	11,29
9	5,46	12,29	17,75	2,43	10,63	13,06			
10				9,30	5,48	14,78	4,39	12,30	16,69
Media	6,84	8,88	15,73	4,95	14,99	19,94	5,33	7,34	12,68

**Tabla iii.** Análisis de la barra de explosión y del VOT de la oclusiva bilabial sorda [p] en contexto neutro, alegre y enfado.

Conversación	gracias (neutro)			gracias (alegre)			gracias (enfadado)		
	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ barra
1	-14,81		14,81	-16,95		16,95	-19,43		19,43
2	-33,38		33,38	-22,47		22,47	-13,08		13,08
3	-47,65		47,65	-35,70		35,70	-38,37		38,37
4	-16,72		16,72	-16,59		16,59	-23,15		23,15
5	-21,30		21,30	-12,85		12,85	-24,11		24,11
6	-16,62		16,62	-43,77		43,77	-29,79		29,79
7	-11,17		11,17	-7,06		7,06	-90,05		90,05
8	-24,89		24,89	-27,94		27,94	-25,00		25,00
9	-31,43		31,43	-23,48		23,48	-79,06		79,06
10	-41,37		41,37	-32,32		32,32	-48,43		48,43
Media	-25,934			-23,913			-39,046		

**Tabla iv.** Análisis de la barra de explosión y del VOT de la oclusiva velar sonora [g] en contexto neutro, alegre y enfado.

Conversación	dime (neutro)			dime (alegre)			dime (enfado)		
	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ Barra
1	-63,78		63,78	-58,65		58,65	-53,26		53,26
2	-52,10	4,65	56,75	-27,71		27,71	-32,60		32,60
3	-22,89	4,08	26,96	-10,18		10,18	-37,99	4,63	42,62
4	-64,21	2,38	66,58	-58,42	2,26	60,68	-53,68	5,12	58,79
5	-41,66	4,08	45,74	-64,01	12,25	76,26	-73,90		73,90
6	-30,75	4,72	35,47	-35,13	4,21	39,34	-54,97	5,19	60,16
7	-83,31		83,31	-33,35	4,68	38,03	-12,76	4,29	17,05
8	-15,90		15,90	-21,86	5,98	27,84	-38,05		38,05
9	-34,37	4,54	38,91	-44,64	4,08	48,72	-41,78		41,78
10	-52,72	4,39	57,11	-53,40	4,13	57,53	-19,72	3,93	23,65
Media	-46,17	4,12	50,29	-40,74	5,37	46,105	-41,87	4,63	46,50

Tabla v. Análisis de la barra de explosión y del VOT de la oclusiva dental sonora [d] en contexto neutro, alegre y enfado.

Conversación	bien (neutro)			bien (alegre)			bien (enfado)		
	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ Barra	VOT	Barra	VOT+ Barra
1	-20,67		20,67	-41,48		41,48	-57,97		57,97
2	-71,12	4,27	75,39	-61,51		61,51	-45,30	6,64	51,94
3	-40,52		40,52	-22,69		22,69	-31,44		31,44
4	-58,56		58,56	-49,52	2,47	51,993	-24,89	4,808	29,70
5	-26,44	4,78	31,22	-27,00	4,68	31,68	-136,31		136,31
6	-71,76		71,76	-72,22	2,59	74,81	-65,18		65,18
7	-55,88		55,88	-55,73		55,73	-67,16		67,16
8	-24,85	3,16	28,02	15,37		15,37	-20,90		20,90
9	-37,73		37,73	-48,60	10,17	58,77	-30,86		30,86
10	-62,74	7,42	70,17	-48,01		48,01	-75,18		75,18
Media	-47,03	4,91	51,94	-41,14	4,98	46,12	-55,52	5,73	61,24

Tabla vi. Análisis de la barra de explosión y del VOT de la oclusiva bilabial sonora [b] en contexto neutro, alegre y enfado.

## Apéndice 2

	<b>Z</b>	<b>Asymp. Sig. (2-tailed)</b>
que-alegre - que-neutro	-.357 <sup>b</sup>	0,72
que-enfadado - que-neutro	-.153 <sup>c</sup>	0,88
que-enfadado - que-alegre	-.764 <sup>c</sup>	0,45
gracias-alegre - gracias-neutro	-1.274 <sup>b</sup>	0,2
gracias-enfadado - gracias-neutro	-1.376 <sup>c</sup>	0,17
gracias-enfadado - gracias-alegre	-1.274 <sup>c</sup>	0,2
tal-alegre - tal-neutro	-1.988 <sup>c</sup>	0,05
tal-enfadado - tal-neutro	-1.274 <sup>c</sup>	0,2
tal-enfadado - tal-alegre	-.764 <sup>b</sup>	0,45
dime-alegre - dime-neutro	-.561 <sup>b</sup>	0,58
dime-enfadado - dime-neutro	-.153 <sup>b</sup>	0,88
dime-enfadado - dime-alegre	-.255 <sup>c</sup>	0,8
pongo-alegre - pongo-neutro	-.730 <sup>b</sup>	0,47
pongo-enfadado - pongo-neutro	-.674 <sup>c</sup>	0,5
pongo-enfadado - pongo-alegre	-1.095 <sup>c</sup>	0,27
bien-alegre - bien-neutro	-.764 <sup>b</sup>	0,45
bien-enfadado - bien-neutro	-.255 <sup>c</sup>	0,8
bien-enfadado - bien-alegre	-.968 <sup>c</sup>	0,33

Tabla vii. VOT SOLO (test no paramétrico de Wilcoxon Signed-Ranks).

	<b>Z</b>	<b>Asymp. Sig. (2-tailed)</b>
que-alegre - que-neutro	-1.27 <sup>b</sup>	0,2
que-enfadado - que-neutro	-1.78 <sup>b</sup>	0,07
que-enfadado - que-alegre	-.46 <sup>c</sup>	0,65
tal-alegre - tal-neutro	-2.40 <sup>b</sup>	0,02
tal-enfadado - tal-neutro	-1.38 <sup>b</sup>	0,17
tal-enfadado - tal-alegre	-.36 <sup>c</sup>	0,72
dime-alegre - dime-neutro	-1.21 <sup>b</sup>	0,23
dime-enfadado - dime-neutro	.00 <sup>d</sup>	1
dime-enfadado - dime-alegre	-1.10 <sup>c</sup>	0,27
pongo-alegre - pongo-neutro	-.73 <sup>c</sup>	0,47
pongo-enfadado - pongo-neutro	-1.75 <sup>b</sup>	0,08
pongo-enfadado - pongo-alegre	-1.46 <sup>b</sup>	0,14

Tabla viii. BARRA+VOT (test no paramétrico de Wilcoxon Signed-Ranks).

- <sup>a</sup>. Wilcoxon Signed Ranks Test  
<sup>b</sup>. Based on positive ranks.  
<sup>c</sup>. Based on negative ranks.  
<sup>d</sup>. The sum of negative ranks equals the sum of positive ranks.

