

Validación perceptiva de dos procedimientos de representación de la melodía aplicados al español, al portugués brasileño y al español como lengua extranjera

Perceptual validation of two pitch representation procedures applied to Spanish, Brazilian Portuguese and Spanish as a foreign language

Cristiane Conceição Silva

Universidade Estadual de Campina
Brasil

Juan María Garrido Almiñana

Universitat Pompeu Fabra
España

ONOMÁZEIN 34 (diciembre de 2016): 242-260
DOI: 10.7764/onomazein.34.15



Cristiane Conceição Silva: Departamento de Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas (Unicamp), Brasil. | Correo electrónico: cris.silva.unicamp@gmail.com

Juan María Garrido Almiñana: Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra (UPF), España. | Correo electrónico: juanmaria.garrido@upf.edu

Fecha de recepción: agosto de 2015
Fecha de aceptación: enero de 2016

Resumen

Existen muchos procedimientos para describir, en términos fonético-acústicos, los contornos de Fo. Los sistemas de Garrido (1996, 2001, 2010) y los de normalización z-score, por ejemplo, tienen la ventaja de que pueden emplearse mediante procedimientos automáticos. Por esta razón, el objetivo de este trabajo es analizar ambos sistemas con el objeto de determinar la posibilidad de su posterior utilización para el análisis de la entonación del español hablado por brasileños (ELE). Para ello, se analizaron las curvas melódicas de declarativas, interrogativas totales y parciales en el español del centro peninsular, en portugués brasileño del estado de São Paulo (PB)

y, por último, en español como lengua extranjera (ELE). La validez de ambos sistemas se efectuó a través de una prueba de percepción con 10 sujetos brasileños y otros 10 españoles. Se solicitó a los oyentes que puntuaran el grado de semejanza entre la entonación de las señales originales y de las señales estilizadas automáticamente. Los resultados demostraron que el procedimiento de estilización y anotación automático es realmente capaz de crear representaciones perceptivamente equivalentes a las originales no solo en español, sino también en PB y en ELE, y mejores en todo caso que las obtenidas con el procedimiento basado en z-score.

Palabras clave: fonética acústica; fonética perceptiva; curvas melódicas; portugués brasileño; español como lengua extranjera.

Abstract

There are many procedures to describe Fo contours in acoustic-phonetic terms. Garrido's (1996, 2001, 2010) and z-score normalization systems have the advantage that they are automatic procedures. Therefore, this paper aims at analyzing both systems in order to determine the possible future application for the analysis of Spanish as a Foreign Language spoken by Brazilians (SFL). To do so, we analyze the melodic curves in declaratives, yes-no questions and wh-questions in Madrid Spanish, in Brazilian Portuguese (BP) spoken in the city of São Paulo and finally,

in SFL. The validity of both systems was tested through a perception test with 10 Brazilian and 10 Spanish listeners. Listeners rated the degree of similarity between the signals with original intonation and signals with stylized Fo contours. The results showed that Garrido's automatic stylization method is really able to perceptually create representations equivalent to the original ones in Spanish, BP and SFL and also better than those obtained with the procedure based on z-score normalization.

Keywords: acoustic phonetics; speech perception; melodic curves; Brazilian Portuguese; Spanish as a Foreign Language.

1. Introducción

Existen en la actualidad muchos estudios que se ocupan de la entonación tanto del español peninsular (Navarro Tomas, 1944; Quilis, 1988, 1993; Sosa, 1999; Estebas-Vilaplana y Prieto, 2010) como del portugués brasileño (Moraes, 1998; Lucente, 2006, 2012). Por el contrario, el número de trabajos dedicados a la entonación del español como lengua extranjera es exiguo (Pinto, 2009; Sá, 2008; Silva, 2009). Además, la mayoría de los estudios sobre la entonación del ELE se basan en presupuestos teóricos de la Fonología Métrica Autosegmental, es decir, se analizan los contornos desde un punto de vista fonológico, que intenta encontrar los patrones relevantes a nivel lingüístico a partir de un conjunto limitado de etiquetas.

Este trabajo se enmarca en otro más amplio que pretende comparar la entonación del portugués de Brasil (PB)¹, la del español hablado por brasileños (ELE)² y la de hablantes nativos de español³. La razón por la que se ha elegido el PB, el español y el ELE como objetivo del análisis se debe al hecho de que, a pesar de que el PB y el español son lenguas muy cercanas, presentan un sistema entonativo distinto, lo que para los brasileños representa una dificultad a la hora de hablar español, tal y como demuestran los estudios de (Pinto, 2009; Sá, 2008; Silva, 2009; Oliveira, 2011; Dias y Alves, 2012). La comparación de los rasgos melódicos del ELE con los del PB y los del español como lengua materna es, de esta forma, de mucha utilidad tanto para los estudios de lingüística teórica como para la enseñanza del ELE.

Como paso previo, se estableció el objetivo de encontrar un sistema de representación de los contornos entonativos de las tres variantes que permitiera la descripción de los tres siste-

mas entonativos y facilitara su comparación. La utilización de un modelo de representación de los contornos de entonación común permite una comparación directa de la entonación del ELE en relación con la del PB y la del español como lengua materna, con el fin de poder determinar las características propias del ELE que lo diferencian del español y del PB.

Existen muchos procedimientos para describir las curvas melódicas; algunos lo hacen en términos fonéticos (Alessandro y Mertens, 1995; Bagshaw, 1993; 't Hart, 1991; Hermes, 2006; Hirst y Espesser, 1993), otros en términos más fonológicos (Silverman y otros, 1992). En la comparación entre sistemas entonativos, la principal ventaja de los procedimientos que tienen como fundamento criterios fonéticos es que se basan en la descripción de los contornos de Fo, un elemento común a todos los sistemas prosódicos de las lenguas, con independencia de sus diferencias fonológicas. Este estudio pretende analizar la validez para describir la entonación del PB y del español hablado por brasileños de dos métodos de representación de los contornos entonativos basados en criterios fonético-acústicos: el modelo de análisis melódico propuesto por Garrido (1996, 2001, 2010) y el sistema de normalización *z-score* (Peres y Silva, 2012).

El objetivo es comprobar hasta qué punto estos dos procedimientos automáticos de estilización y anotación son capaces de proporcionar representaciones perceptivamente equivalentes a los contornos originales en ELE y PB, por medio de la realización de una prueba de percepción. En los apartados siguientes, se explica en detalle el funcionamiento de la herramienta de anotación y estilización automática (análisis melódico), así como el procedimiento de normalización *z-score*.

1 Se ha elegido la variedad del portugués hablado en el estado de São Paulo.

2 El ELE aprendido en la ciudad de Madrid.

3 Se ha elegido la variedad del español hablado en la ciudad de Madrid.

1.1. Descripción del modelo de análisis melódico

El modelo de Garrido (1996, 2001, 2010) se inspira en gran medida en la aproximación IPO ('t Hart y otros, 1990), pero a diferencia de esta, el análisis melódico parte de las medidas acústicas de los puntos de inflexión para llevar a cabo la reducción de los datos acústicos. Según este modelo, los contornos de Fo son el resultado de la superposición de patrones globales y locales. Los patrones globales son la evolución global de un contorno de Fo a lo largo de un Grupo Entonativo (GE). El autor define el GE como la porción del enunciado en la que empieza y termina la declinación, es decir, sus límites están en cada reajuste de la curva melódica. Los GE se modelan como líneas de referencia que predicen valores de Fo como una función del tiempo (figura 1 a la derecha). En función de la posición que ocupan dentro del enunciado, dichas líneas pueden ser iniciales, intermedias o finales.

Los patrones locales son formas típicas de Fo que ocurren a nivel del Grupo Acentual (GA). El GA es la porción del enunciado comprendida

entre el inicio de una sílaba tónica y el inicio de la siguiente sílaba tónica, si es que existe, o el final del GE. Estos patrones se superponen a un patrón global, de forma que se determina su altura relativa dentro del rango de Fo de cada hablante. Según su posición dentro del GE, los mismos pueden ser iniciales, intermedios o finales. Por último, se modelan como conjuntos de puntos de inflexión (puntos en que se produce un cambio de dirección del movimiento) de Fo fijados con el objeto de especificar las partes de las sílabas que integran el GA (figura 1 a la izquierda).

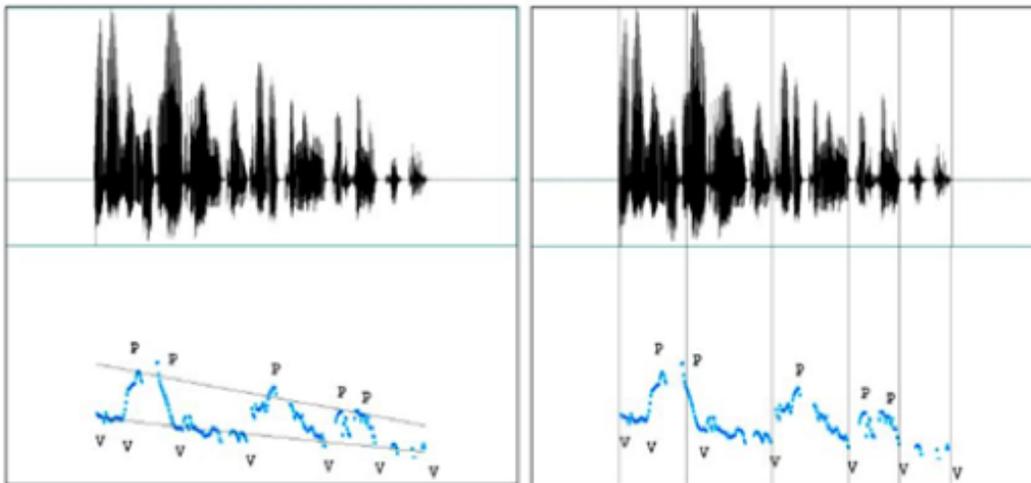
Cada patrón se identifica con una etiqueta que incluye información sobre:

- el nivel de los puntos de inflexión de Fo que componen el patrón: (P, P+, V y V-);
- la sílaba que contiene el punto: 0 (tónica), 1 (postónica), -1 (pretónica), etc.;
- la posición del punto de inflexión dentro de la sílaba: I (inicial), M (media) y F (final).

De esta manera, el objetivo del análisis melódico es definir los patrones de Fo a partir del análisis fonético de los contornos de Fo. Para ello, el pro-

FIGURA 1

Forma de onda y contorno de Fo del enunciado "Aragón se ha reencontrado como motor del equipo", pronunciado por una hablante española. Las líneas verticales (a la derecha) marcan las fronteras de los grupos acentuales y los puntos de inflexión están anotados con las etiquetas (P) y (V) (Garrido, 2010)



cedimiento incluye las siguientes etapas, que se explican en detalle en los apartados siguientes:

- estilización de los contornos;
- anotación de los contornos estilizados;
- definición de los patrones globales y locales a partir de los contornos anotados.

En Garrido (2010) se describe MelAn, una herramienta que realiza el procedimiento de estilización de manera automática en vez de manual, tal y como sucede en el modelo IPO. Una serie de experimentos de percepción (Garrido, 2010; Yao y Garrido, 2012) mostró que la herramienta automática es capaz de ofrecer una representación simplificada de la entonación con la que se pueden capturar los movimientos e inflexiones relevantes del español peninsular, del catalán y también del chino. En el ámbito de estas pruebas, se comparó el grado de semejanza de la entonación entre los contornos originales y sus respectivos contornos estilizados. Los promedios obtenidos fueron altos, lo que es una evidencia de que la herramienta realmente genera representaciones simplifica-

das de la entonación en español, en catalán y en chino, las cuales son perceptivamente equivalentes a los contornos originales. MelAn tuvo que pasar por algunas adaptaciones para que pudiera aplicarse al análisis de los datos de PB.

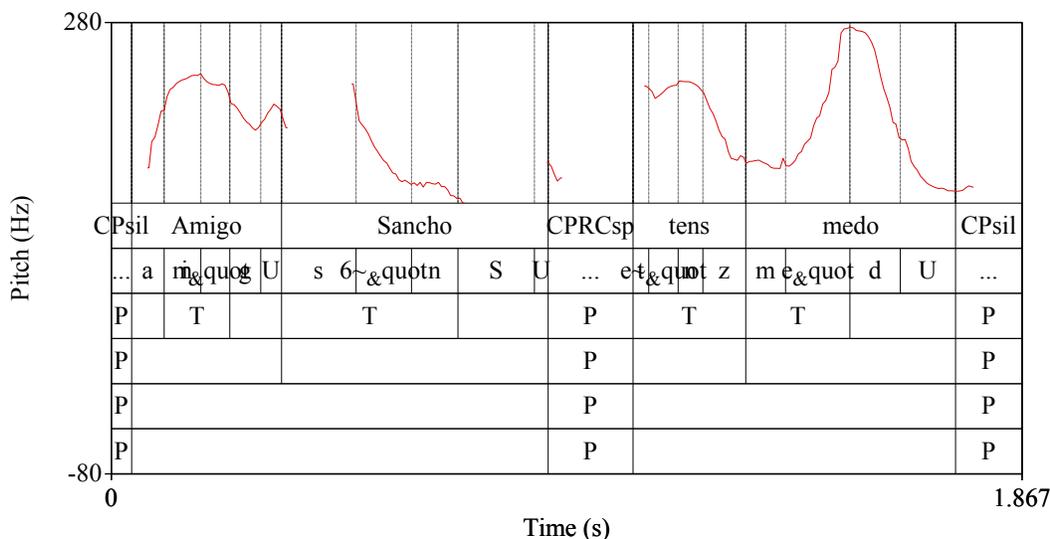
1.1.1. Estilización de contornos de Fo

Para la aplicación de todo el procedimiento, es necesario tener instalado el programa Praat, que puede descargarse de forma gratuita desde el sitio www.praat.org (Boersma y Weenink, 2013). En primer lugar, son necesarios un fichero *wav* y un fichero *TextGrid* con dos capas, una para la transcripción de los segmentos en formato SAMPA y la otra con la marcación de las sílabas tónicas. Se ejecuta el *script* Segproso de Garrido (2013b). Este script efectúa la segmentación de las unidades prosódicas (sílabas, grupo acentual, sintagma entonativo y grupo entonativo) y obtiene como resultado un nuevo fichero *TextGrid* con las capas correspondientes a cada unidad prosódica (figura 2).

Posteriormente, se realiza el procedimiento de estilización. El objetivo de la estilización es

FIGURA 2

Contorno de Fo y *TextGrid* del enunciado “Amigo Sancho, tens medo?”, pronunciado por la hablante brasileña G. Las capas: *Syllables*, *StressGroups*, *IntonationPhrase* e *IntonationGroup* han sido creadas por el *script* Segproso



Se utilizan también las etiquetas P+ y V- para marcar los puntos de inflexión con valores de Fo, claramente por encima o por debajo de los niveles medios de P y V en un determinado punto del contorno. La figura 5 ofrece un ejemplo del contorno de una interrogativa parcial en portugués brasileño, con el punto de inflexión etiquetado como P+ al comienzo de la vocal tónica de la palabra “gigantes”.

MelAn asigna automáticamente estas etiquetas a los diferentes puntos de inflexión obtenidos durante la fase de estilización, mediante un procedimiento que consta de cinco pasos:

- I. Cálculo de las rectas de regresión de los valores de Fo de cada GE para la definición de los límites entre los niveles P y V.
- II. Una primera interacción en el proceso de anotación para la asignación de etiquetas P y V a cada punto de inflexión del contorno estilizado.

III. Cálculo de las rectas de regresión delimitadas por los puntos P y V de cada GE.

IV. Una segunda anotación con las etiquetas P+ y V-, respectivamente, de aquellos puntos P y V que se desvían demasiado de las rectas de regresión calculadas.

V. Eliminación de las etiquetas redundantes.

El resultado final del proceso es una capa en el fichero *TextGrid* de Praat con las etiquetas asignadas y alineadas en el tiempo con los puntos de inflexión correspondientes (figura 6).

Para cada GE se obtienen, además, dos rectas de regresión, una para los puntos etiquetados con P y otra para los puntos etiquetados con V. Esos puntos representan la evolución global de los niveles P y V a lo largo del GE (figuras 7 y 8).

La herramienta calcula de forma automática las rectas de regresión para cada uno de los GE del enunciado de entrada. Para cada una de estas rectas, también calcula el valor de Fo inicial

FIGURA 5

Forma de onda, contorno de Fo y *TextGrid* del enunciado “¿Qué gigantes?”, pronunciada por la hablante brasileña G. El punto de inflexión P+ está anotado en rojo, los demás puntos, en negro

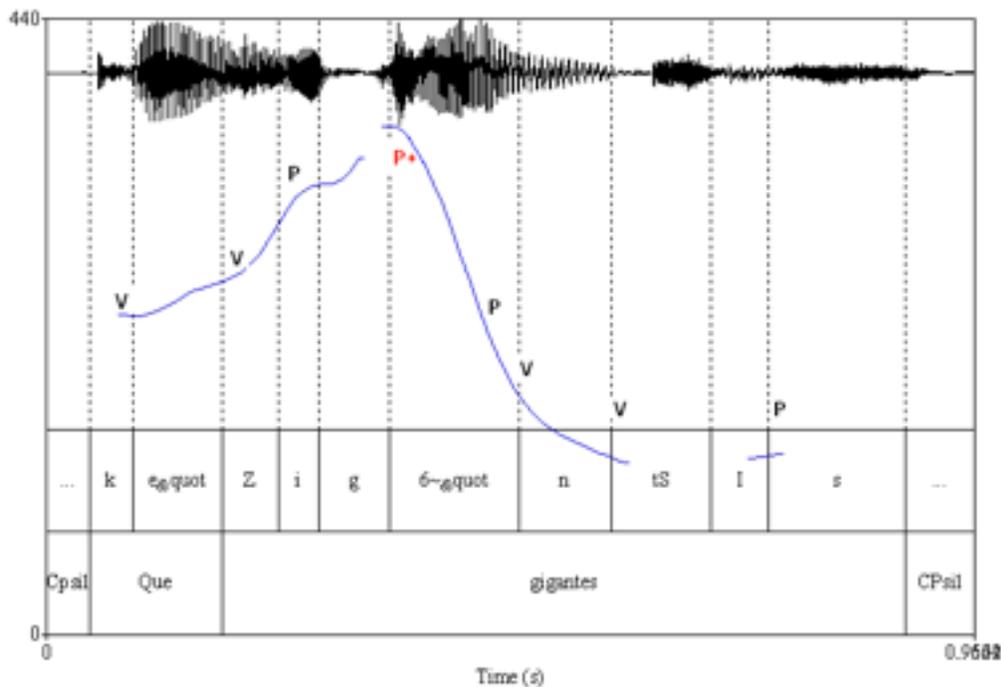


FIGURA 6

Contorno de Fo y *TextGrid* del enunciado “Amigo Sancho, ¿tienes miedo?”, pronunciado por la hablante de español L. Las tres últimas capas (8, 9 y 10) muestran la salida de los diferentes pasos del proceso de anotación

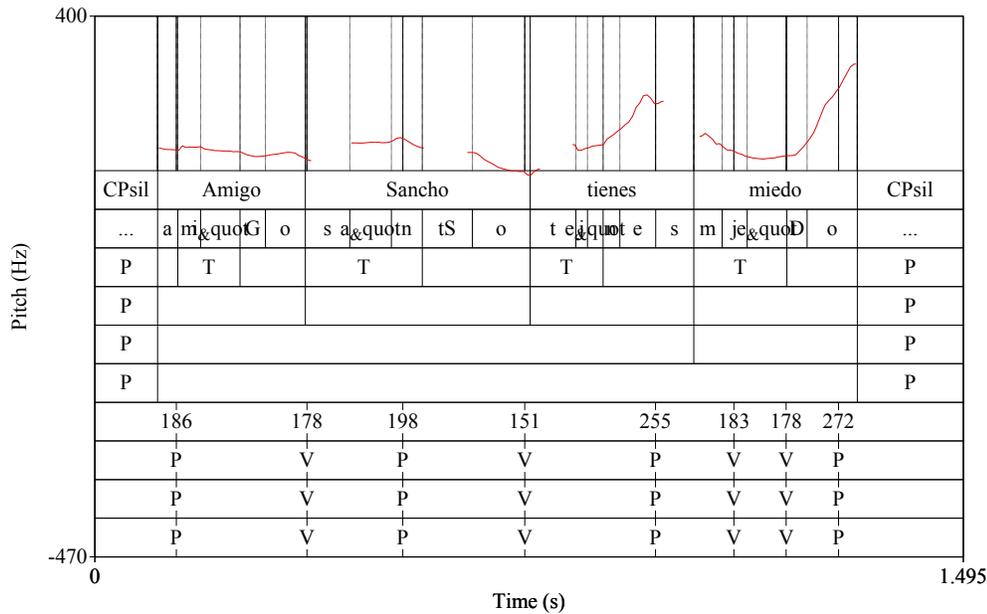
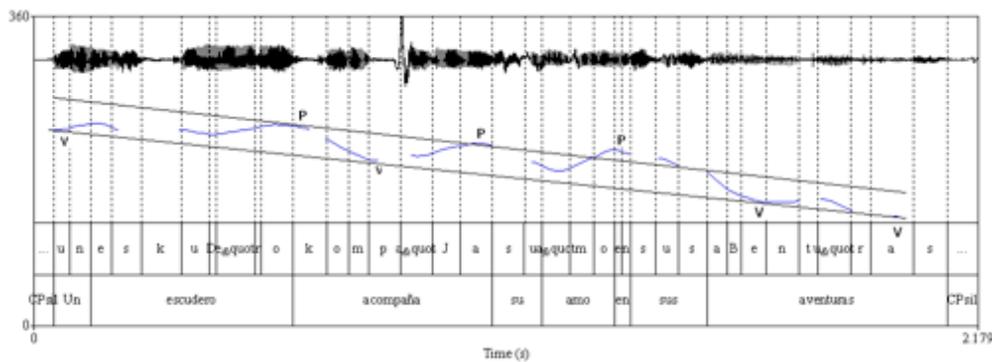


FIGURA 7

Forma de onda, contorno de Fo y *TextGrid* del enunciado “Un escudero acompaña a su amo en aventuras”, pronunciado por la hablante de español L. Las dos rectas superpuestas representan las líneas de referencia para los niveles P y V



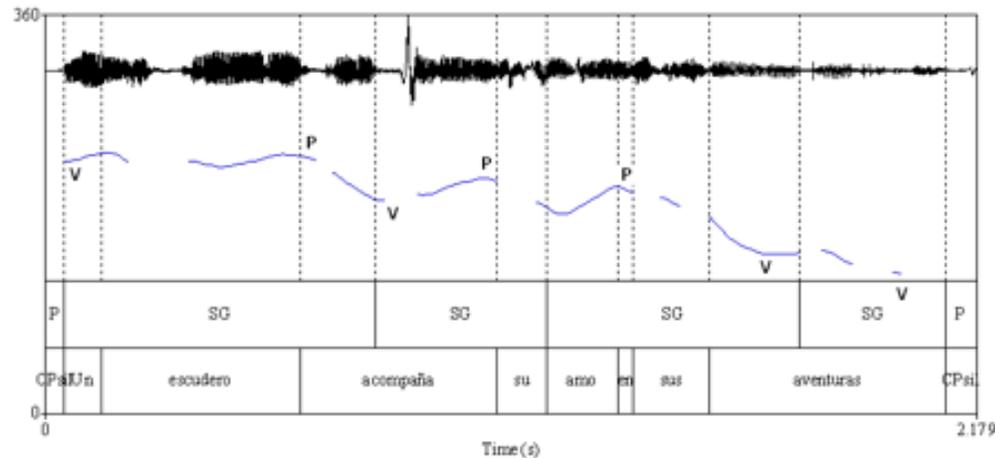
(en Hz), y el de pendiente (en Hz/s). Estas rectas permiten predecir los valores de Fo correspondientes a estos dos niveles tonales en cualquier punto del mismo.

Una vez obtenido con MelAn el listado de patrones del enunciado de entrada, este puede sintetizarse a partir de las etiquetas prosódicas utilizando una versión adaptada de *ModProso* (Garrido, 2013a). Para este fin, son necesarios: (1)

el fichero wav original, (2) el fichero *TextGrid* generado con los *scripts* Segproso y MelAn, (3) un fichero de texto *contour* que contiene las etiquetas prosódicas y (4) los ficheros con las rectas de regresión P y V, los cuales contienen los valores de las líneas de referencia globales de Fo (todos creados de forma automática por MelAn). El procedimiento de síntesis se realiza con el método Overlap-Add de Praat.

FIGURA 8

Forma de onda, contorno de Fo y *TextGrid* del enunciado “Un escudero acompaña a su amo en aventuras”, pronunciado por la hablante de español L. Dentro de cada grupo acentual (SG), hay una marca de todos sus puntos de inflexión

**1.2. Descripción de z-score**

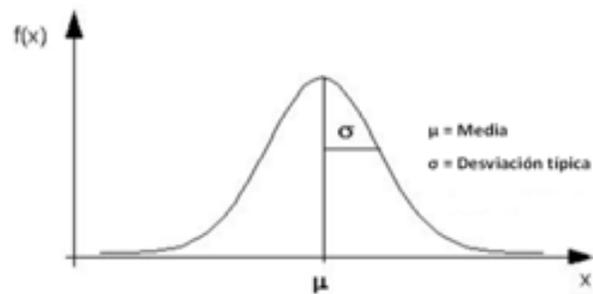
Peres y Silva (2012) observaron que la calidad de los estímulos resintetizados con la técnica de normalización *z-score* es muy buena⁴ y, por esta razón, se decidió añadir estímulos resintetizados con este sistema al test de percepción.

El *z-score* es un procedimiento de normalización bastante utilizado en estadística. Para entenderlo mejor hace falta volver a la distribución de probabilidades conocida como distribución normal o distribución de Gauss. La gráfica de su función densidad tiene una forma acampanada (figura 9).

La ecuación de la curva normal⁷ se expresa en términos de media y desviación estándar de

FIGURA 9

Gráfico de una función gaussiana. La línea vertical al centro de la distribución indica la media (μ)⁵ y la línea horizontal que se inicia en la media indica la desviación estándar (σ)⁶



la distribución, y por esa razón se obtendrán distintas curvas normales según cambien estas dos

4 En su estudio, los investigadores no realizaron una prueba de percepción para cuantificar dicha eficiencia del sistema *z-score*.

5 Se estima la media poblacional (μ) a partir del cálculo de la media de una muestra. Hay muchos tipos de cálculos de medias. La más común es la media aritmética: $\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$

6 Se estima la desviación estándar poblacional (σ) a partir del cálculo de la desviación estándar de una

muestra: $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$

7 La probabilidad de un evento continuo distribuirse según una curva normal sigue una ecuación: $Y = \frac{n}{s\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2s^2}}$ donde Y es la variable dependiente, n es el número de casos de la distribución, s es la desviación estándar de la distribución, π es 3,1416... (constante matemática), e es 2,7183... (constante matemática), X es la variable independiente y es la media de la distribución.

estadísticas, lo que dificultaría la comparación entre curvas normales distintas. De esa manera, se ha desarrollado un procedimiento que permite comparar directamente distintas curvas normales denominado *z-score*, puntuación *z* o puntuación estándar.

El *z-score* indica el grado de desviación estándar de un determinado valor (valor bruto) respecto a la media de una distribución. Los grados de desviación se determinan en una escala de unidades de desviación estándar, tal y como se presenta en la fórmula:

$$(1) \quad z = \frac{x - \bar{x}}{s}$$

donde x es una puntuación cruda para ser normalizada, \bar{x} es la media de la distribución y s es la desviación estándar de la distribución.

La distribución normal estándar, o distribución z , es una distribución normal, pero con media y desviación estándar fijas. Es decir, la distribución z tiene media de la población ($\mu=0$) y desviación estándar de la población ($\sigma=1$). Por ejemplo, si se quiere encontrar la probabilidad de que un valor y esté entre 3 y 6 en una distribución normal con media 2 y desviación estándar 2, se calcula en (2):

$$(2) \quad z = \frac{3-2}{2} = 0,5 \quad z = \frac{6-2}{2} = 2$$

El procedimiento de estandarización queda más claro en la figura 10.

Mediante la transformación de los valores de F_0 en valores de desviación estándar se obtiene una estandarización de las medidas de F_0 . A continuación, estos coeficientes se convierten a Hz, tal como se explica con más detalle a continuación.

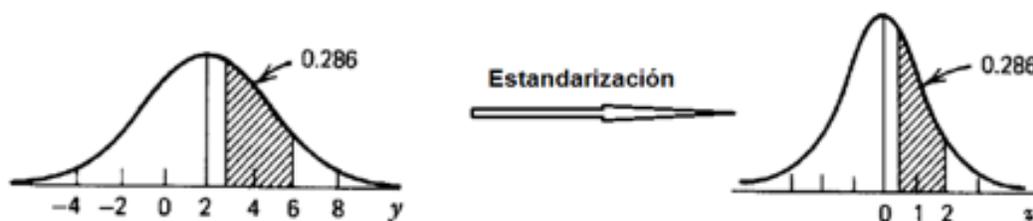
El procedimiento de normalización para este trabajo se efectuó en Praat. En primer lugar, se creó un objeto *Pitch* y a ese objeto se le aplicó la fórmula de la unidad tipificada a través de las opciones *Modify y Formula*.

De esta manera, se obtuvieron los valores tipificados brutos de media y desviación estándar para cada enunciado. Dado que no es posible utilizar esos valores obtenidos de la fórmula de puntuación estándar para resintetizar la curva, ya que se trata de valores normalizados, estos se convirtieron a Hz con los mismos valores de referencia para todos los enunciados. Como se trataba de dos mujeres, los valores de referencia han sido 200 Hz para la media y 50 Hz para la desviación estándar.

A partir de los valores obtenidos, se substituyó la curva original por el objeto *Pitch Tier* normalizado por el *z-score*, y para eso se utilizó la función *To Manipulation*. Finalmente, con la función *Publish resynthesis* de Praat, se procesó la síntesis a partir del nuevo contorno modificado.

FIGURA 10

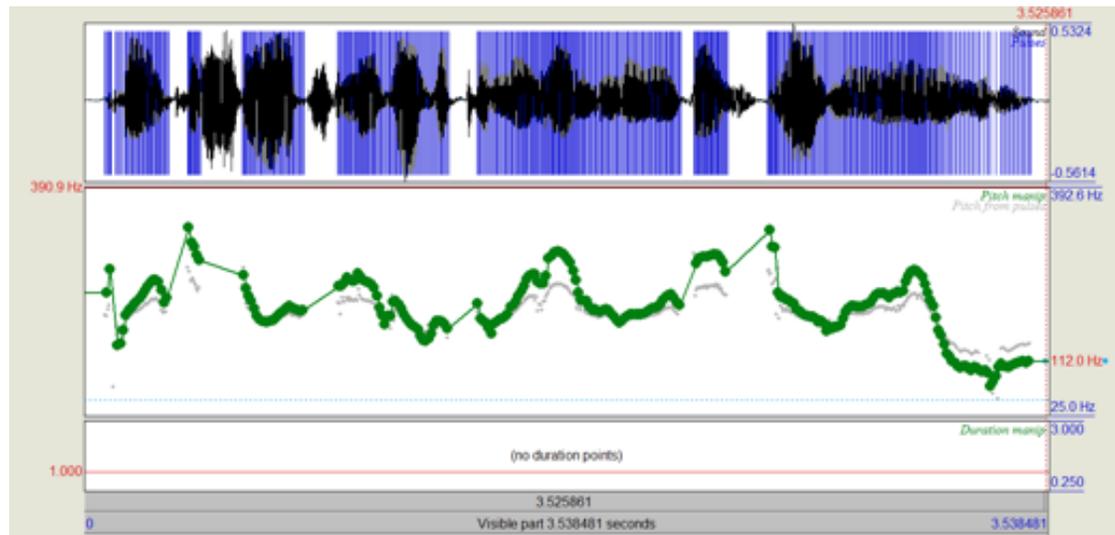
Gráfico de una función gaussiana y (a la izquierda) y función gaussiana estándar z (a la derecha). La línea vertical al centro de la distribución indica la media $\bar{x}=2$ (derecha) y $\bar{x}=0$ (izquierda). El eje de las abscisas en las dos distribuciones indica la distancia en unidades de desviación estándar $s=2$ (derecha) y $s=1$ (izquierda)



La estandarización preserva las áreas

FIGURA 11

Ilustración del procedimiento de normalización en Praat. Forma de Onda, Pulsos Glotales, Contorno de Fo original (puntos grises) y Contorno de Fo modificado (verde) del enunciado “*Dom Quixote levava camisas limpas e algum dinheiro*”, pronunciado por la hablante brasileña G

**2. Metodología****2.1. Corpus**

El material de base para la prueba fue un corpus de 12 enunciados (6 declarativos, 3 interrogativos totales y 3 parciales) leídos en PB, ELE y español lengua materna. Según Quilis (2003), el grupo fónico⁸ medio en español oscila entre ocho y once sílabas. Para observar los contornos melódicos de enunciados con más de un grupo fónico, se han elegido enunciados más largos. De esa manera, las declarativas escogidas contienen entre 18 y 40 sílabas, y las interrogativas totales y parciales, entre 7 y 13 sílabas. Todas las frases empleadas en el experimento pueden consultarse en el apéndice final del presente artículo. Este pequeño corpus está integrado en otro mayor, denominado corpus *Quijote*, que se compone de datos de la producción de 15 informantes brasileños y 5 españoles. En el presente estudio se ha analizado, en particular, una pequeña parte de los datos de dos informantes.

La informante L. es española, de Talavera de La Reina (región de Castilla la Mancha), tiene 22 años y es estudiante universitaria. La segunda informante G. es brasileña, de São Paulo, tiene 25 años, aprendió español con una profesora española de Salamanca y vivió en Madrid durante 11 meses. Considera que habla español con fluidez. La informante L. ha escuchado algunos segundos de la grabación de la informante brasileña y considera que habla con fluidez, aunque no supo decir cuál es la variante del español que utiliza la misma.

La grabación se realizó de forma individual, en casa de cada una de las informantes, en un ambiente de silencio y con las puertas y ventanas cerradas. Para efectuar la grabación se utilizó un netbook, un micrófono dinámico ShureLyric 8900 y una tarjeta de sonido SoundBlasterAudigyCreative. La grabación se realizó directamente con el software Praat, en formato wav a 22.050 Hz mono. Se mostró cada una de las frases en una diapositiva de PowerPoint de Microsoft Office. Se les pidió a las informantes que, en el caso de que

8 El grupo fónico corresponde a la porción de discurso comprendida entre dos pausas (Quilis, 2003).

considerasen que habían cometido cualquier equivocación durante la lectura, repitieran la lectura de la frase. La informante G. leyó las frases en español y en portugués, mientras que la informante L. solo las leyó en español.

Todo este material fue procesado con MelAn y con el sistema de normalización *z-score*, con el objetivo de obtener dos tipos distintos de representación fonético-acústica de sus contornos melódicos. Finalmente, los análisis estadísticos se realizaron con el Programa R, que puede descargarse de forma gratuita desde el sitio www.r-project.org (R Core Team, 2013).

2.2. Estímulos

Una vez concluidos los procedimientos de estilización y normalización descritos en la sección

anterior, se generó, a partir de cada una de las representaciones automáticas obtenidas, una versión resintetizada de los contornos. De esta manera, se dispuso de tres versiones de cada uno de los enunciados, una con el contorno original, otra con el contorno reconstruido a partir de la representación automática obtenida con MelAn y una tercera obtenida a partir del sistema *z-score*. Con esos estímulos se elaboraron pares original-sintetizado (siempre el mismo orden), que se emplearon para preparar la prueba de percepción.

La figura 12 muestra ejemplos de pares de contorno original-sintetizado (análisis melódico) de las dos locutoras grabadas en español, ELE y PB.

La figura 13 muestra dos ejemplos de pares de contorno original-sintetizado (sistema *z-score*).

FIGURA 12

Contornos de Fo del enunciado “¿Dónde están mis libros?”, pronunciado por la locutora española L. (Español) y la locutora brasileña G. (ELE) y del enunciado “Onde estão meus livros?”, pronunciado por la locutora brasileña G. (PB). El contorno negro corresponde al original y el rojo al reconstruido y sintetizado (análisis melódico)

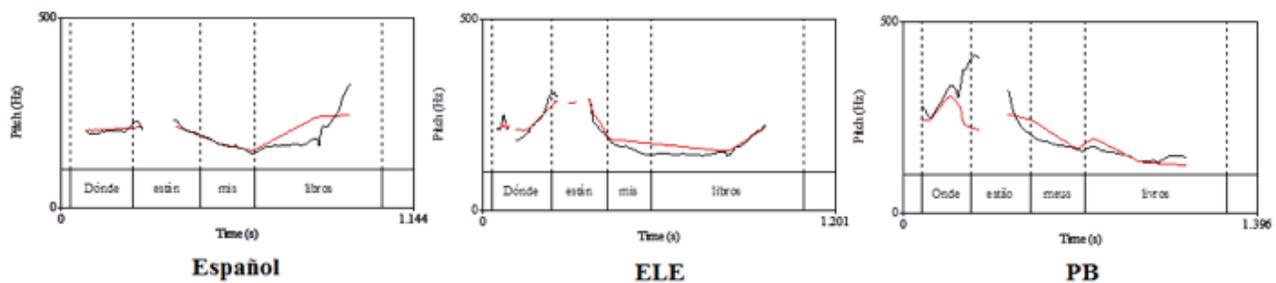
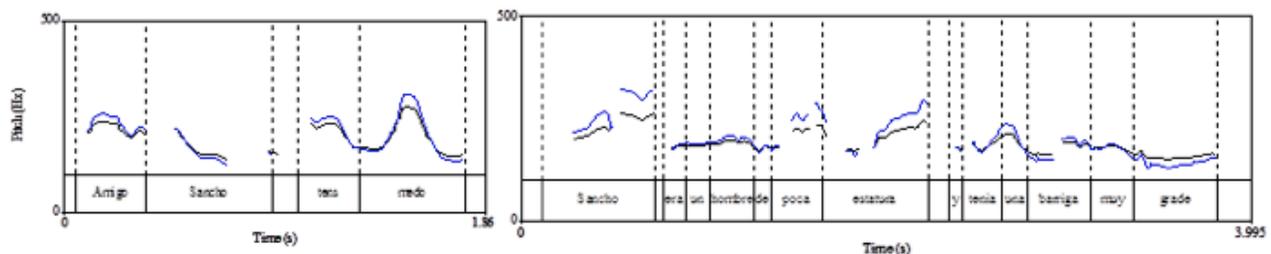


FIGURA 13

Contornos de Fo del enunciado “Amigo Sancho, tens medo?” y “Sancho era un hombre de poca estatura y tenía una barriga muy grande”, pronunciados por la locutora brasileña G. El contorno negro corresponde al original y el azul al reconstruido y sintetizado (*z-score*).



2.3. Prueba de percepción

Se solicitó a un conjunto de 10 oyentes españoles (6 mujeres y 4 hombres) y 10 oyentes brasileños (5 mujeres y 5 hombres) que evaluaran el grado de semejanza entre los contornos resintetizados y los originales. Los diez españoles provienen de diferentes regiones de España y tienen entre 25 y 49 años. Todos tienen familiaridad con temas relacionados con la entonación y la melodía del habla, son filólogos, lingüistas, profesores de español o logopedas. Los diez brasileños proceden de diferentes regiones de Brasil y su edad varía entre 24 y 49 años. Nueve de ellos son lingüistas y uno es estudiante de lingüística.

Se pidió a estos oyentes que evaluaran el grado de semejanza de los contornos melódicos de cada par de estímulos, dentro de una escala del 1 al 5, en la que 1 significaba que los contornos eran totalmente distintos y 5 que eran iguales. La prueba se realizó por internet, en el sitio SurveyGizmo (www.surveygizmo.com). Los interesados en participar en la prueba recibieron un enlace por correo electrónico y, de esa manera, pudieron escoger el mejor momento y lugar para la realización de la prueba. Se les dijo que podrían escuchar los estímulos tantas veces como considerasen necesario, además de que era imprescindible que utilizaran auriculares durante el desarrollo de la prueba y, por último, que deberían permanecer en un ambiente de silencio para que las condiciones de la prueba fueran lo mejor y más homogéneas posible.

En el caso específico de los españoles, la duración aproximada de la prueba fue de 15 minutos, pues tenían que escuchar los 24 pares de estímulos en español y, justo a continuación, los 24 pares de estímulos en ELE. En lo que respecta a

los brasileños, hicieron la prueba exclusivamente con los 24 pares de estímulos en portugués, de modo que la duración aproximada de la tarea fue de 8 minutos. Los pares de estímulos se presentaron en orden aleatorio.

3. Resultados y discusión

3.1. Medias – prueba de percepción

En la tabla 1 se presentan las medias globales para el conjunto de estímulos en PB, ELE y español (análisis melódico) y se comparan con las medias globales de los estudios anteriores para el español, catalán y chino (Garrido, 2010; Yao y Garrido, 2012).

La media global de los estímulos resintetizados con el análisis melódico fue en PB y ELE de 4,2 y, en el caso de español, de 4,1, lo que indica que los oyentes, en general, consideraron que los contornos originales y resintetizados eran bastante semejantes. La media es incluso ligeramente superior a la registrada en los estudios anteriores.

En la tabla 2 mostrada a continuación se presenta la puntuación media del análisis melódico de cada estímulo, así como la media total en PB, ELE y español, discriminadas por tipo de enunciado y cantidad de estímulos (6 declarativos, 3 interrogativos totales y 3 interrogativos parciales).

La tabla 2 muestra que, en el análisis melódico, la puntuación media de los enunciados declarativos e interrogativos no difiere mucho de la media total (4,2 en español y PB y 4,1 en ELE). Este resultado pone de manifiesto que, por término medio, la puntuación que se concedió a los contornos analizados tanto en las declarativas como en las interrogativas (totales y parciales) fue alta,

TABLA 1

Puntuación media total en la prueba de percepción (análisis melódico)

Lengua	Español	Chino	Catalán	PB	ELE	Español
Media	4,05	4,03	3,93	4,2	4,2	4,1

es decir, igual o mayor que 4,0, con excepción de las interrogativas parciales en PB y español, las cuales registraron unas medias de 3,6 y 3,8, respectivamente.

En la misma tabla, se han destacado en negrita los estímulos con medias más bajas (entre 2,9 y 3,5). Por ejemplo, en el caso de los estímulos

TABLA 2

Puntuación media en PB, ELE y español (análisis melódico)

Enunciados	Estímulos	PB	ELE	Español
Declarativas	1	4	3,5	4
	2	4,1	4,1	4,4
	3	4,6	3,2	3,5
	4	4,4	4,7	4,2
	5	4	4,5	4,4
	6	4,4	3,9	4,6
Media Declarativas		4,3	4,0	4,2
Interrogativas Totales	1	4,4	4,9	3,9
	2	5	4,5	3,9
	3	4,5	4,8	4,4
Media Int. Totales		4,6	4,7	4,1
Interrogativas Parciales	1	3,1	3,9	3,2
	2	2,9	3,8	3,8
	3	4,8	4,4	4,5
Media Int. Parciales		3,6	4,0	3,8

1 y 3 de las declarativas en ELE y español, la estilización no encontró los picos que estaban en un nivel más alto, tal y como se observa en el caso de las palabras *limpias* y *algún* (figura 14) del estímulo 3 en español. Las flechas azules indican tanto la posición de los picos en el contorno original como su ausencia en el caso de estas dos palabras, en el contorno resintetizado.

En las interrogativas parciales (estímulos 1 y 2 en PB y 1 en español), es posible que las medias más bajas se deban a que los valores mínimos y máximos de la curva no eran lo suficientemente bajos y altos en la parte final de estos enunciados, a pesar de que el procedimiento logró determinar de modo correcto la posición de los picos y valles en el contorno (figura 15).

En lo que respecta a los contornos en PB, las puntuaciones más altas se registraron en las declarativas e interrogativas totales, mientras que las más bajas correspondieron a dos interrogativas parciales (2,9 y 3,1, respectivamente). Este hecho parece que guarda relación, por un lado, con la altura y la posición del pico de F_0 en el pronombre interrogativo (figura 16). Las

FIGURA 14

Contorno de F_0 del enunciado en español “Don Quijote llevaba camisas limpias y algún dinero”, pronunciado por la locutora española L. El contorno en negro corresponde al enunciado original y el rojo al resintetizado

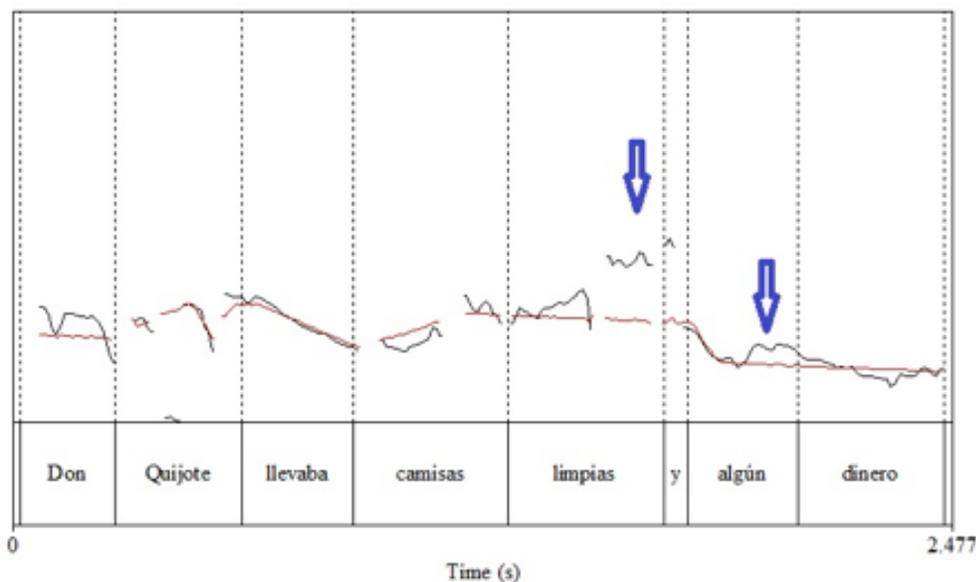
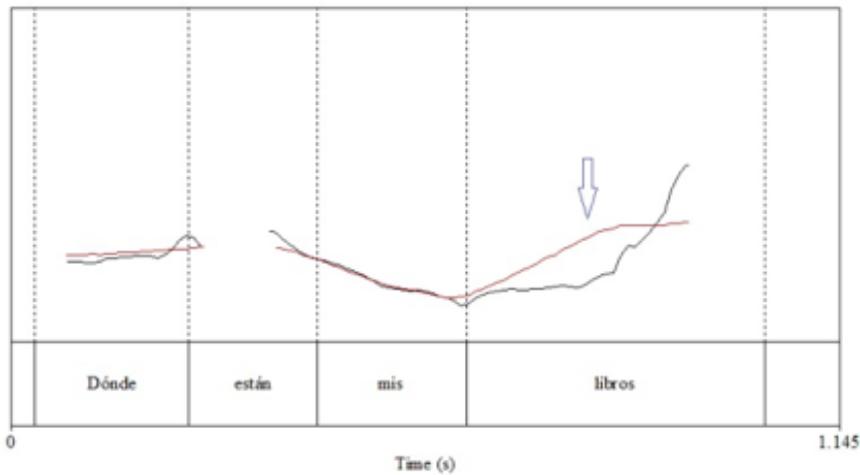


FIGURA 15

Contorno de Fo del enunciado en español “¿Dónde están mis libros?”, pronunciado por la locutora española L. El contorno en negro corresponde al enunciado original y el rojo al resintetizado

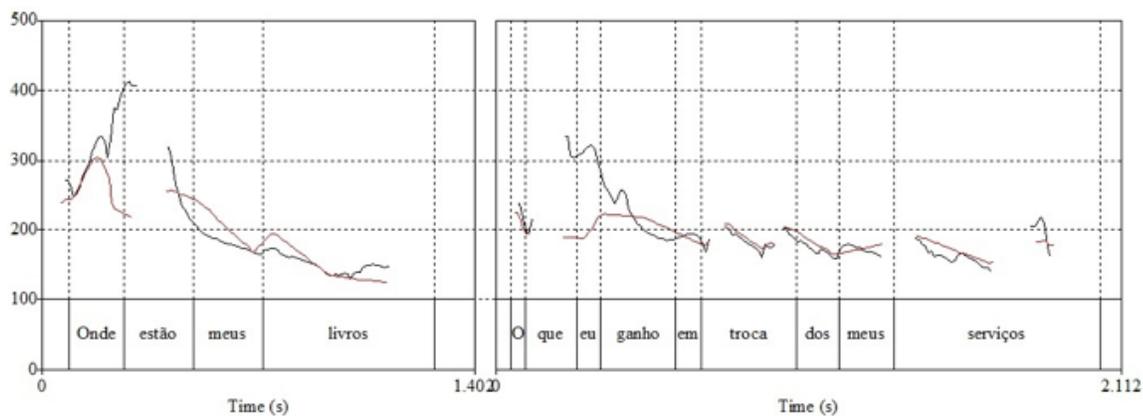


rectas paralelas al eje de abscisas señalan la altura y posición de los picos en los pronombres interrogativos de las dos preguntas. A pesar de que el modelo de análisis melódico ha conseguido identificar los picos en los dos contornos, su altura relativa es muy distinta a la de los contornos originales, en los que el pico de Fo del pronombre interrogativo se sitúa a más altura y, por esta razón, tarda más en completar todo el movimiento, mientras que el descenso en el contorno resintetizado empieza mucho antes. Por otro lado, en los contornos resintetizados

la curva de Fo baja al final, mientras que la curva de Fo original sube un poco. Estas pequeñas diferencias, que son evidentemente errores del sistema de estilización, son muy probablemente las responsables de que estos contornos tengan una puntuación media más baja, e indican que, en las interrogativas parciales en PB, el pronunciado ascenso de Fo en el pronombre interrogativo, seguido de un descenso más lento, parece ser importante para la percepción de este tipo de interrogativas, así como la pendiente de la curva al final del enunciado.

FIGURA 16

Contorno de Fo de los enunciados: “Onde estão meus livros?” e “O que eu ganho em troca dos meus serviços?”, pronunciados por la locutora brasileña G. El contorno en negro corresponde al enunciado original y el rojo al resintetizado



En la tabla 3, mostrada a continuación, se presenta la puntuación media de cada estímulo y la media total en PB, ELE y español, discriminadas por tipo de enunciado y número de estímulos (6 declarativos, 3 interrogativos totales y 3 parciales), para los pares de estímulos resintetizados con *z-score*.

La tabla 3 muestra que, con el procedimiento de normalización *z-score*, las interrogativas totales y parciales en PB, ELE y español registran las medias más altas (entre 4,0 y 4,7), lo que demuestra que las interrogativas resintetizadas con *z-score* fueron consideradas casi iguales a las originales. La explicación para ello es que, dado que se utilizan los valores del propio hablante para el procedimiento de normalización, la media y desviación estándar de *F₀* son más altas en las interrogativas, es decir, están más cerca de la media de 200 Hz y de la desviación estándar de 50 Hz, que se utilizaron como valores de referencia, y, consecuentemente, los cambios en el contorno resintetizado con respecto al original son mínimos. En lo que refiere a las declarativas, la media y la desviación estándar de *F₀* son más bajas y, por esta razón, los contornos resintetizados presentan más diferencias respecto a los originales. Esto explica por qué las declarativas registraron las medias más bajas (entre 2,6 y 3,3) a lo largo de toda la prueba de percepción. Así, a pesar de

dos resultados buenos en las interrogativas, las declarativas registraron los peores resultados de todo el test.

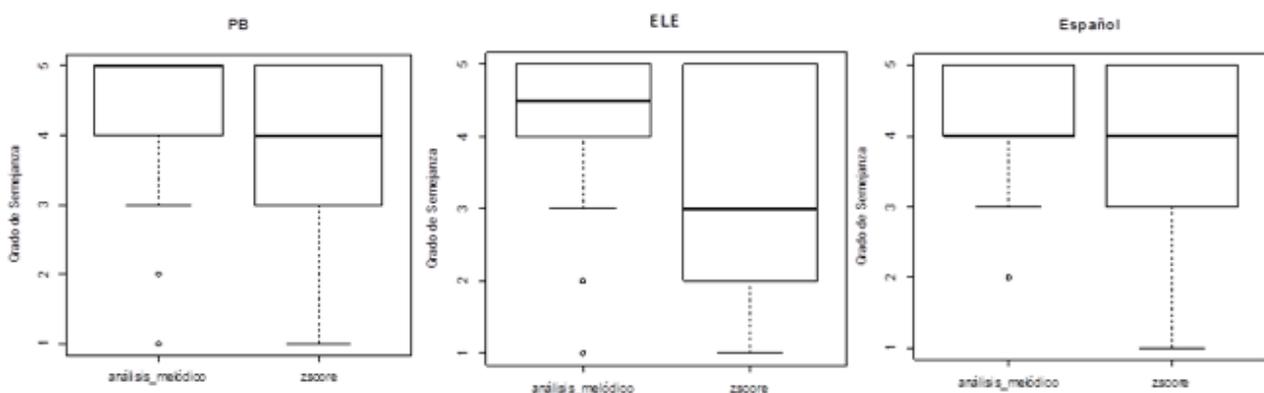
Los diagramas de caja a continuación evidencian la superioridad de los estímulos sintetizados con el análisis melódico con respecto a los estímulos sintetizados con el *z-score* (figura 17). Se observan distribuciones diferentes de los datos en los dos métodos. En los estímulos sintetizados con el análisis melódico, además de tener media y mediana más altas, la diferencia entre los valores mínimos y máximos es más pequeña y, por lo tanto, la desviación estándar es menor. Se observa que tanto en PB como en español y ELE los estímulos sintetizados con el análisis melódico son más semejantes a los estímulos originales. Una ANOVA de dos factores con variable dependiente, grado de semejanza y variables independientes, método e informante demostró diferencia significativa en el grado de semejanza según el método en PB ($F(1,9) = 0,03$ $p < 0,05$), en ELE ($F(1,9) = 1,68 \cdot 10^{-08}$ $p < 0,05$) y en español ($F(1,9) = 0,002$ $p < 0,05$).

4. Conclusión

Este estudio demuestra que el modelo de análisis melódico genera representaciones de los contor-

FIGURA 17

Diagramas de cajas del grado de semejanza de los estímulos sintetizados con el análisis melódico y el *z-score*. De la izquierda hacia la derecha se presentan los diagramas del PB, ELE y español



nos en PB, ELE y español que pueden considerarse perceptivamente equivalentes a los contornos originales (valoración media de 4,2 en ELE y PB y 4,1 en español, sobre un máximo de 5), lo que permite afirmar que las representaciones estilizadas generadas por el modelo de análisis melódico son lo suficientemente fiables como para comparar directamente las curvas melódicas en español con la de sus variantes en ELE y PB.

Los resultados de este experimento no solo refuerzan la idea de que el sistema de estilización y anotación es independiente de la lengua (la valoración obtenida para el PB y ELE es muy semejante a la obtenida para los contornos del español y del catalán), sino también la de que el sistema es independiente del locutor (no se producen diferencias significativas con respecto a los contornos resintetizados de la hablante española L. y los respectivos contornos de la hablante brasileña G.). Este resultado difiere del encontrado por Yao y Garrido (2012), que mostraba un claro efecto atribuible al locutor.

La comparación entre los estímulos resintetizados con MelAn y z-score demuestra la superioridad del primero respecto a este último, ya que con el análisis melódico se registraron buenas puntuaciones en todos los tipos de enunciado tanto en español como lengua materna como en PB y ELE.

Los buenos resultados obtenidos por MelAn con los contornos sintetizados de ELE, PB y español como lengua materna permiten pensar que, en estudios posteriores, se podría emplear para realizar un análisis más amplio de las características de la entonación del español hablado por los brasileños, como por ejemplo un estudio del comportamiento de los patrones globales y locales a partir del análisis de los GE y de los GA que forman parte de los enunciados en ELE, en comparación con los del PB y del español. El objetivo sería elucidar si la organización de tales grupos, así como sus características, están más próximos a los patrones del español como lengua materna (lo cual determinaría un grado mayor de dominio

de la lengua extranjera) o si, por el contrario, tienen más cercanía con los del PB (lo que caracterizaría un menor grado de dominio de la lengua extranjera).

5. Bibliografía citada

D' ALESSANDRO, Christophe y Piet MERTENS, 1995: "Automatic pitch contour stylization using a model of tonal perception", *Computer Speech and Language* 9(3), 257-288.

BAGSHAW, Paul C., 1993: "An investigation of acoustic events related to sentential stress and pitch accents in English", *Speech Communication* 13(3), 333-342.

BOERSMA, Paul y David WEENINK, 2013: "Praat: doing phonetics by computer (version 5.3.55)" [<http://www.fon.hum.uva.nl/praat/>], fecha de consulta: 18 de marzo de 2015].

DIAS, E. C. O. y M. A. ALVES, 2012: "Análise de produção de sentença interrogativa totais em aprendizes brasileiros de espanhol como língua estrangeira", *Journal of Speech Sciences* 2(2), 43-63 [www.journalofspeechsciences.org], fecha de consulta: 18 de marzo de 2015].

ESTEBAS-VILAPLANA, Eva y Pilar PRIETO, 2010: "Castilian Spanish Intonation" en Pilar PRIETO y Paolo ROSEANO (coords.): *Transcription of Intonation of the Spanish Language*, Lincom Europa: München, 17-48.

GARRIDO ALMIÑANA, Juan María, 1996: *Modelling Spanish Intonation for Text-to-Speech Applications*. Tesis de doctorado, Universitat Autònoma de Barcelona en Barcelona.

GARRIDO ALMIÑANA, Juan María, 2001: "La estructura de las curvas melódicas del español: propuesta de modelización", *LEA: Lingüística Española Actual* 23(2), 173-210.

GARRIDO ALMIÑANA, Juan María, 2010: "A tool for auto-

matic Fo stylisation, annotation and modelling of large corpora”, *Speech Prosody*, Chicago [http://speechprosody2010.illinois.edu/papers/100041.pdf, fecha de consulta: 18 de marzo de 2015].

GARRIDO ALMIÑANA, Juan María, 2011: “Análisis de las curvas melódicas del español en habla emotiva simulada”, *Estudios de Fonética Experimental* 20, 205-255.

GARRIDO ALMIÑANA, Juan María, 2013a: “ModProso: A Praat-Based tool for Fo Prediction and Modification”, *Proceedings of TRASP*, 38-41.

GARRIDO ALMIÑANA, Juan María, 2013b: SegProso: “A Praat-Based tool for the Automatic Detection and Annotation of Prosodic Boundaries”, *Proceedings of TRASP*, 74-77.

HART, Johan, 1991: “Fo stylization in speech: straight lines versus parabolas”. *Journal Acoustic Society of America* 90(6), 3368-3370.

HART, Johan, René COLLIER y Antonie COHEN, 1990: *A perceptual study of intonation: an experimental-phonetic approach to speech melody*, Cambridge University Press.

HERMES, Dik J., 2006: “Stylization of pitch contours” en Stefan SUDHOFF y otros (eds.): *Methods in empirical prosody research*, Berlin: Walter de Gruyter, 30-61.

HIRST, Daniel y Robert ESPESSER, 1993: “Automatic modelling of fundamental frequency using a quadratic spline function”, *Travaux de l'Institut de Phonétique d'Aix* 15, 75-85.

LUCENTE, Luciana, Leandro SILVEIRA, Plínio Almeida BARBOSA, 2006: “Declarativas em PB: downstepping ou nova combinação bitonal? H+!H* e H+L*”, comunicación presentada en el IV Congresso Nacional e III Congresso Internacional de Fonética e fonologia. Belo Horizonte, Brasil.

LUCENTE, Luciana, 2012: *Aspectos Dinâmicos da Fala e da Entoação no Português Brasileiro*. Tesis

de doctorado, Universidade Estadual de Campinas en Campinas.

MORAES, João Antonio, 1998: “Intonation in Brazilian Portuguese” en Daniel HIRST y A DI CRISTO (eds.): *Intonation Systems: a survey of twenty languages*, Cambridge University Press.

NAVARRO TOMÁS, Tomás, 1944: *Manual de entonación española*, cuarta edición, Madrid: Guadarrama.

OLIVEIRA, Aline Fonseca y Francisco José CANTERO, 2011: “Características da entoação do espanhol falado por brasileiros”, *VII Congresso Internacional da Abralín Curitiba*, 84-98.

PERES, Daniel y Cristiane Conceição SILVA, 2012: *Estudo piloto sobre os métodos de normalização de Fo: uma comparação entre IPO, Momel e Z-Score*, Trabajo final de asignatura, Universidade Estadual de Campinas en Campinas.

PINTO, Maristela da Silva, 2009: *Transferências prosódicas do português do Brasil/LM na aprendizagem do espanhol/LE: enunciados assertivos e interrogativos totais*. Tesis de doctorado, Universidade Federal do Rio de Janeiro en Rio de Janeiro.

QUILIS, Antonio, 1988: “Estudio comparativo entre la entonación portuguesa (de Brasil) y la española”, *Revista de Filología Española LXVIII*, 33-65.

QUILIS, Antonio, 1993: *Tratado de fonética y fonología españolas*, Madrid: Gredos.

QUILIS, Antonio y Joseph A. FERNÁNDEZ, 2003: *Curso de fonética y fonología españolas para estudiantes angloamericanos*, décima octava reimpresión, Madrid: Consejo Superior de Investigaciones Científicas.

R CORE TEAM, 2013: *A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria [http://www.R-project.org/, fecha de consulta: 18 de marzo de 2015].

SA, Priscila Cristina Ferreira, 2008: *Análise entoacional de enunciados assertivos, continuativos e interrogativos lidos em piadas em espanhol/LE e espanhol/LM*. Tesis de maestría, Universidade Feraldo do Rio de Janeiro en Rio de Janeiro.

SILVA, Cristiane Conceição, 2009: "Análisis melódico de declarativas e interrogativas absolutas em espanhol/LE". *Phonica* 05, 92-113.

SILVERMAN, Kim, John PITRELLI, Mary BECKMAN, Mari OSTENDORF, Colin WIGHTMAN, Patti PRICE, Janet PIERREHUMBERT, Julia HIRSCHBERG, 1992: "TOBI: A standard for labelling English prosody" en *Proc. International Conference on Spoken Language Processing*, 867-870.

SOSA, Juan Manuel, 1999: *La entonación del español*, Madrid: Cátedra.

YAO, Junming y Juan María GARRIDO ALMIÑANA, 2012: "Validación perceptiva de un sistema de anotación automático de contornos de Fo aplicado al chino mandarín" en E. RIDRUEJO y otros (eds.): *Tradición y progreso en la lingüística general*, Área de Lingüística General, Departamento de Filología Española, Universidad de Valladolid, 457-471.

6. Anexos

FRASES DEL EXPERIMENTO

Declarativas en español

- 1) Don Quijote llevaba camisas limpias y algún dinero.
- 2) Un escudero acompaña a su amo en sus aventuras.
- 3) Así que aceptó el oficio de escudero sin pensárselo dos veces.
- 4) Sancho era un hombre de poca estatura y tenía una barriga muy grande.
- 5) Este hidalgo amaba los libros de caballería y fue de esa manera que perdió la razón.

- 6) Don Quijote y Sancho salieron en plena noche sin despedirse de nadie y se pusieron en camino en busca de aventuras.

Declarativas en portugués

- 1) Dom Quixote levava camisas limpas e algum dinheiro.
- 2) Um escudeiro acompanha seu amo nas aventuras.
- 3) Por isso aceitou o ofício de escudeiro sem pensar duas vezes.
- 4) Sancho era um homem de baixa estatura e tinha uma barriga muito grande.
- 5) Este fidalgo amava os livros de cavalaria e foi dessa maneira que perdeu a razão.
- 6) Dom Quixote e Sancho saíram em plena noite sem se despedir de ninguém e seguiram caminho em busca de aventuras.

Interrogativas totales en español

- 1) Amigo Sancho, ¿tienes miedo?
- 2) ¿Quieres ser mi escudero fiel?
- 3) ¿Ves aquellos gigantes de allí abajo?

Interrogativas totales en portugués

- 1) Amigo Sancho, tens medo?
- 2) Queres ser meu escudeiro fiel?
- 3) Vês aqueles gigantes ali embaixo?

Interrogativas parciales en español

- 1) ¿Dónde están mis libros?
- 2) ¿Qué hace un escudero?
- 3) ¿Y qué gano yo a cambio de mis servicios?

Interrogativas parciales en portugués

- 1) Onde estão meus livros?
- 2) Que faz um escudeiro?
- 3) E o que eu ganho em troca dos meus serviços?