

LA IMPORTANCIA DE LA INTENSIDAD COMO PISTA ACÚSTICA EN LA IDENTIFICACIÓN DE ORACIONES INTERROGATIVAS

María Cristina Armero¹, Ignacio Moreno-Torres¹, Paolo Roseano^{2,3}

¹Universidad de Málaga, ²Universitat de Barcelona, ³University of South Africa
mararmper@uma.es, imoreno@uma.es, paolo.roseano@ub.edu

ABSTRACT

La evidencia en varias lenguas sugiere que los usuarios de implante coclear no usan la F0 como una señal acústica para la identificación de las características prosódicas y, por lo tanto, confían en otras señales acústicas (por ejemplo, intensidad, duración de las vocales). Con el objetivo de determinar en qué medida estas señales secundarias proporcionan información confiable sobre el contraste entre oraciones declarativas e interrogativas, se diseñó un experimento de percepción en el que los oyentes tenían que clasificar un conjunto de oraciones naturales y manipuladas (por ejemplo, para simular el implante coclear) como aseveraciones o preguntas. Los resultados indican que la intensidad puede ayudar a los pacientes con implantes cocleares a diferenciar entre afirmaciones e interrogativas totales.

Palabras clave: fonética clínica, intensidad, prosodia, entonación, implantes cocleares

Evidence from various languages suggests that cochlear implant users fail to recognize the F0 as an acoustic cue for the recognition of prosodic characteristics, and therefore rely on other acoustic cues (e.g. intensity, vowel duration). With the aim of determining to what extent these secondary cues provide reliable information about the declarative / interrogative contrast, a perception experiment was designed in which listeners had to classify a set of natural and manipulated (e.g. to simulate cochlear implant) sentences as statements or questions. The results indicate that intensity can help patients with cochlear implants to differentiate between statements and yes-no questions.

Keywords: clinical phonetics, intensity, prosodic, intonation, cochlear implant

1. INTRODUCCIÓN

Los hablantes suelen asociar diversas pistas acústicas a un mismo fenómeno fonológico (Lisker, 1986; Giezen, Escuero y Baker, 2010). Las interrogativas absolutas en español se caracterizan por un patrón entonacional con una subida de la frecuencia fundamental (F0) al final del enunciado (Navarro Tomás, 1944: 135; Quilis, 1993: 414), mientras que en las declarativas se produce una bajada. De acuerdo con Dorta, Hernández y Cabrera (2009), la F0 es la pista acústica fundamental y, por esa razón, también es la más estudiada. Por otra parte, la duración e la intensidad serían pistas secundarias (Candía, Urrutia y Fernández, 2006). La hipótesis de este trabajo es, justamente, que existen otras pistas acústicas, independientes de la F0, que permiten diferenciar desde el punto de vista

prosódico las oraciones interrogativas totales de las declarativas.

Conocer la aportación de las diferentes pistas acústicas a la hora de vehicular información lingüística tiene interés no solo desde el punto de vista descriptivo, sino desde un punto de vista aplicado. En concreto, es sabido que las personas con implante coclear tienen dificultades para acceder a la F0 (Zeng, Tang y Lu, 2014), por lo que sería importante saber si hay otra información alternativa a la que pueden acceder. Este trabajo tiene el objetivo concreto de estudiar el papel de la intensidad como pistas acústica complementaria en la percepción de la modalidad oracional por parte de los normo-oyentes en condiciones de degradación espectral, para en un segundo momento extrapolar las conclusiones y aplicarlas a la población con implante coclear.

El estudio de las pistas secundarias de la modalidad oracional, sin embargo, es complejo por varios motivos. En primer lugar, en la medida en que la F0 es una pista acústica necesaria (Candía, Urrutia y Fernández, 2006) y presente en todas las emisiones con voz modal, que son las más habituales en el habla. Para poder garantizar que los oyentes no la están utilizando y que solo están usando las pistas secundarias, es necesario manipular la señal acústica mediante un proceso de vocodizado (v. apartado 2.3) que permite eliminar la F0 aun conservando otras propiedades espectrales y temporales de la señal acústica (DiNino, Wright, Winn y Bierer (2016). La vocodización permite, además, simular el proceso de percepción en implantados cocleares.

En segundo lugar, puede que la presencia de estas pistas secundarias sea variable entre locutores, y que igualmente varíe en función de factores como la variedad dialectal, contexto, etc. Es más, es posible que un mismo locutor las emplee de forma variable.

El presente estudio, de carácter preliminar, pretende comprobar si un grupo de hablantes efectivamente emplea la intensidad para diferenciar enunciados declarativos e interrogativos. Además, en el caso de que la empleen, nos preguntamos si la utilizan todos ellos.

Para ello se ha grabado a un grupo de sujetos de la localidad andaluza de Álora (apartado 2.1), y se han obtenido señales vocodizadas (apartado 2.2) que se han utilizado para llevar a cabo un experimento de percepción (apartado 2.3). Los resultados del experimento (detallados en el apartado 2.4) permiten comprobar el papel que desempeñan las pistas acústicas secundarias en la percepción de la modalidad oracional.

2. METODOLOGÍA

2.1. Locutores y grabaciones

Para la grabación de las frases que se han utilizado como estímulos en el experimento de percepción, se contó con la participación de 12 sujetos (4 niños, 4 hombres y 4 mujeres). Todos los sujetos han crecido y viven en Álora, una localidad de la comarca del Guadalhorce perteneciente a la provincia de Málaga. Para diferenciar los locutores les hemos asignado un número identificativo (101,102, 103, etc.).

Los estímulos se obtuvieron pidiendo a los locutores que leyeran el enunciado escogido en las diferentes modalidades de la forma más natural posible.

El dispositivo empleado para la grabación fue una grabadora ZOOM H4n Pro conectada mediante conector XLR a un micrófono dinámico Shure WH20XLR. Se siguieron las recomendaciones de grabación aportadas por el producto.

El enunciado de las frases que se han grabado es, tanto para las interrogativas como para las declarativas, "Marina limpiaba la nevera". El criterio seguido para la elección de esta oración fue el de que existiese el mayor número posible de segmentos sonoros. Utilizando este enunciado se grabó a los 12 sujetos produciendo en dos ocasiones el enunciado, tanto en modalidad declarativa como interrogativa, y se escogió la que resultaba más natural.

Los locutores fueron grabados a campo libre en una habitación sin interferencias acústicas del exterior, en mono, con una frecuencia de muestreo de 44.100 Hz, que posteriormente se redimensionó utilizando Praat (Boersma y Weenink, 2014) a 22.050 Hz y una resolución de 16 bits. Los archivos se generaron en formato WAV.

2.2. Estímulos vocodizados (VV)

Siguiendo DiNino et al. (2016), el proceso de codificación de voz se diseñó para simular el procesamiento del implante IC Fidelity 120 con las mismas asignaciones de banda de frecuencia que las utilizadas en muchos dispositivos Advanced Bionics (Advanced Bionics Corp., Valencia, CA). Se obtuvieron un total de 15 bandas de frecuencia contiguas pseudo-logarítmicamente espaciadas de 250 a 8700 Hz de cada estímulo original. La raíz cuadrada de la energía total en cada canal se obtuvo para calcular la envolvente y la señal resultante se filtró bajo los 68 Hz. La envolvente de cada canal se utilizó para modular una banda de ruido con una frecuencia central igual a la del canal correspondiente. Las pendientes de salida del filtro se ajustaron a 30 dB/octava.

2.3. Prueba de percepción

La prueba de percepción se administró mediante un script MFC de Praat. El script presentaba a los jueces las grabaciones (tanto interrogativas como declarativas) en orden aleatorio, por duplicado y con la condición de que no se reprodujesen dos grabaciones iguales seguidas. Para cada enunciado, los jueces debían hacer clic en la pantalla en la casilla de la modalidad oracional correspondiente (declarativa o interrogativa). Los jueces tenían la opción de hacer una pausa cada 16 ítems.

Con el fin de comprobar la estabilidad de los resultados, se dividió la tarea en dos pruebas (prueba 1 y prueba 2), donde cada prueba tenía 48 audios (6 locutores x 2 modalidades x 2 voces x 2 repeticiones).

Un total de 70 alumnos de primer curso de Grado de Logopedia de la Universidad de Málaga participaron en el estudio, donde 7 de ellos eran hombres y 63 mujeres. Todos los participantes tenían el español como lengua materna y residían en Andalucía. La realización del experimento formó parte de una práctica académica. Al Grupo 1 de alumnos (N = 26) se les pasó la prueba 1, al Grupo 2 (N = 24) la prueba 2 y al Grupo 3 (N = 20) se le pasaron las dos pruebas. En total, se recogieron 8640 respuestas.

2.4. Medidas acústicas y análisis de correlaciones

Usando Praat, de cada grabación se tomaron las siguientes medidas acústicas relacionadas con la intensidad: intensidad máxima de la antepenúltima, penúltima y última sílaba. A partir de estas medidas se obtendrían otras relativas como el incremento de intensidad entre la sílaba tónica y la sílaba postónica (ΔI).

Los datos de las respuestas de los jueces se pasaron a formato SPSS (versión 25). Con SPSS se calculó el índice de correlación de Spearman R entre el porcentaje de aciertos en estímulos vocodizados (VV) en interrogativas y las medidas acústicas señaladas anteriormente.

3. RESULTADOS

3.1. Resultados de la prueba de percepción con VM

Antes de calcular las medidas indicativas de posibles pistas acústicas se comprobó si el porcentaje de aciertos en VM era superior al 95% (con objeto de descartar posibles deficiencias auditivas, atencionales, etc.). Ello llevó a eliminar las respuestas de 8 jueces. La muestra final constó de 7600 respuestas válidas de 62 oyentes (7 hombres y 55 mujeres).

El porcentaje de aciertos fue muy similar en hombres (79,2%) y mujeres (80%), lo que indica que la variable sexo del oyente no era determinante en este estudio.

Centrándonos en la voz vocodizada, el porcentaje total de aciertos medio en declarativas fue del 88,9% y en interrogativas del 66,7%, tendencia que se observó por igual en los tres grupos. Ello muestra que los jueces no respondieron al azar, e indica que

la prueba puede servir para analizar las pistas acústicas empleadas por los jueces.

3.2. Resultados de la prueba de percepción con VV

En voz vocodizada, en 11 de los 12 locutores el porcentaje de aciertos en las declarativas era próximo al 100% y superior al porcentaje en las interrogativas. En uno de los locutores (con el código 203 en la figura 1), sin embargo, el porcentaje de aciertos fue superior en las interrogativas.

El porcentaje de aciertos en las interrogativas se localizaba, en la mayoría de los casos, entre el 30% y 70%, y existían 5 casos con valores superiores a 50%. El hecho de que haya un nivel de aciertos tan alto en las interrogativas en voz vocodizada (o sea, sin F0), indica claramente que los oyentes deben haber accedido a otras pistas acústicas (v. apartado 3.3).

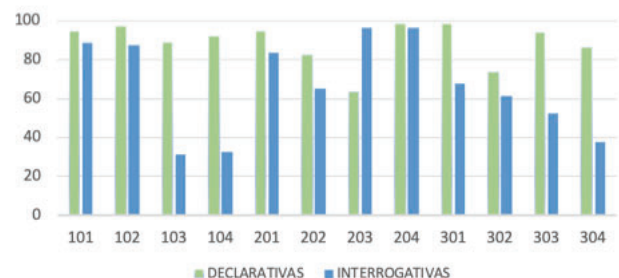


Figura 1: Porcentaje de aciertos por locutor en voz vocodizada en declarativas e interrogativas.

3.3. Correlación entre porcentaje de aciertos y medidas acústicas

Se calculó el índice de correlación de Spearman R, entre la variable “porcentaje de aciertos en VV en interrogativas” y las medidas acústicas indicadas en el apartado 2.4. Se observó una correlación negativa con la intensidad media de la penúltima sílaba del enunciado ($\rho = -0,60$; S.: < 0,05), y una correlación positiva con la intensidad media de la última sílaba del enunciado ($\rho = 0,64$; S.: < 0,05). Además, se observó una correlación positiva muy significativa con el incremento de intensidad entre las dos sílabas finales del enunciado ($\rho = 0,78$; S.: < 0,01). Esto quiere decir que aquellos locutores que resultaron más fáciles de reconocer redujeron la intensidad de la penúltima sílaba e incrementaron la intensidad de la última sílaba.

Con el fin de comprobar esta correlación observada entre ΔI y ratio de aciertos, se examinaron las curvas de intensidad de las interrogativas de los locutores con mayor y menor porcentaje de aciertos en VV. Los resultados mostraron que en los primeros hay un pico de intensidad en la postónica (101 y 204 en Fig. 3) que no se aprecia en las curvas de intensidad de los segundos (103 y 304 en Fig. 3).

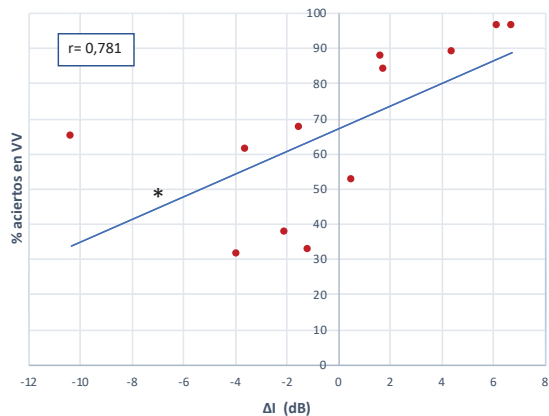


Figura 2: Correlación entre ΔI (entre la sílaba tónica y postónica) en las interrogativas y el porcentaje de aciertos en VV.

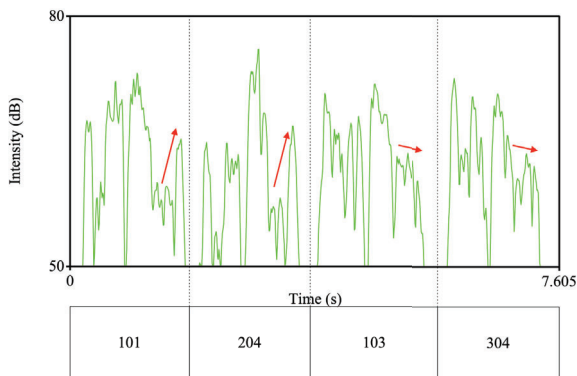


Figura 3: Los cambios de intensidad en las sílabas tónica y postónica de las interrogativas de los locutores con mayor y peor porcentaje de aciertos en VV.

4. DISCUSIÓN Y CONCLUSIONES

Son tres los resultados más destacables de este estudio. En primer lugar, y con la excepción del locutor 203, las interrogativas vocodizadas son más difíciles de reconocer que las declarativas vocodizadas.

En segundo lugar, el hecho de que los oyentes logren reconocer la modalidad en la mayoría de los casos (y con porcentajes cercanos al 100% en algunos) indica que debe haber pistas acústicas alternativas a la curva melódica. En análisis de resultados nos ha permitido observar que cuando se

produce un incremento final de intensidad entre la tónica y la postónica, los oyentes identifican la oración como interrogativa, lo que implicaría que el ΔI entre estas sílabas podría ser una pista acústica secundaria. Este resultado es compatible con otros trabajos (Dorta et al., 2009) que han observado cómo la intensidad puede tener una función prosódica relevante en español, por ejemplo, para marcar posiciones de frontera de la estructura sintagmática de interrogativas absolutas en Canarias.

En tercer lugar, este estudio muestra que existen diferencias individuales entre locutores. Todos los locutores usan la subida final de F0 para denotar una interrogativa, pero esto no ocurre con el ΔI : no todos los sujetos usan ΔI como pista acústica secundaria para denotar una interrogativa, lo que indica que es una pista variable. Este resultado parece compatible con el hecho de que las pistas secundarias no se observan de forma sistemática y pueden variar en función de factores de diversa índole (contextuales, individuales, etc.) (Candía, Urrutia y Fernández, 2006).

En general, los datos obtenidos en este estudio confirman la hipótesis sobre la existencia de una pista, diferente de la F0, que emplean algunos sujetos para marcar la modalidad interrogativa. Tal y como se ha visto, la pista en cuestión es el incremento de intensidad en la última sílaba del enunciado.

Este estudio, por lo tanto, puede ser el comienzo de nuevas líneas de investigación. Desde un punto de vista lingüístico se podría profundizar en la forma en la que varía ΔI entre dialectos, entre locutores, o incluso, en la variabilidad del uso que hace de este incremento un mismo locutor. Esta línea de investigación puede tener implicaciones también en otros ámbitos, como el de la logopedia, puesto que existe un grupo de hablantes, los usuarios de implante coclear, para los que –debido a la dificultad de acceder a la F0 (Driscoll, Gfeller, Kliethermes y Oleson, 2013) las pistas acústicas secundarias podrían ser esenciales. Futuros trabajos, vueltos a profundizar en las aplicaciones prácticas de los resultados obtenidos en población normo-oyente, se tendrán que plantear la pregunta de si y cómo es posible enseñar a utilizar el incremento de intensidad para vehicular la modalidad oracional interrogativa en el habla dirigida a personas con implante coclear.

5. REFERENCIAS

- Candía, L., Urrutia, H. y Fernández, T. (2006) Rasgos acústicos de la prosodia acentual del español. *Boletín de Filología*, 41, 11-14.
- DiNino, M., Wright, R. A., Winn, M. B. y Bierer, J. A. (2016). Vowel and consonant confusions from spectrally-manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces. *Acoustical Society of America*, 140(6), 4404–4418.
- Dorta, J., Hernández, B. y Díaz, C., (2009). Interrogativas absolutas: relación entre F0, duración e intensidad. *Estudios de Fonética Experimental*, 18, 123-144.
- Lisker, L. (1986). “Voicing” in English: A catalogue of acoustic features for Signaling /b/ Versus /p/ in Trochees. *Language and Speech*, 29(1), 1-13.
- Navarro Tomás, T. (1944): *Manual de entonación española*. Nueva York: Hispanic Institute in the United States.
- Quilis, A. (1981): *Manual de fonética acústica española*. Madrid: Editorial Gredos.
- Quilis, A y Fernández, J.A. (1989). *Curso de fonética y fonología españolas*. Madrid: Consejo Superior de Investigaciones Científicas.
- See, R.L., Driscoll, V.D., Gfeller, K., Kliethermes, S. y Oleson, J. (2013). Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing. *Otol Neurotol* 34(3), 490-498. doi: 10.1097/MAO.0b013e318287c985.
- Zeng, F., Tang, Q. y Lu, T. (2014). Abnormal Pitch Perception Produced by Cochlear Implant Stimulation. *PLoS ONE*, 9(2), p.e88662. doi:10.1371/journal.pone.0088662