

Revista Española de Lingüística

Órgano de la Sociedad Española de Lingüística

RSEL

45|1

Enero-Junio
2015

Edita
SeL

REVISTA ESPAÑOLA DE LINGÜÍSTICA
(RSEL)
45/1

Edita

SeL

REVISTA ESPAÑOLA DE LINGÜÍSTICA (RSEL)

ISSN: 0210-1874 · eISSN: 2254-8769

Depósito Legal: M-24.769-1971

DIRECTOR DE HONOR: D. Francisco Rodríguez Adrados (RAE y RAH).

DIRECTOR: Juan Antonio Álvarez-Pedrosa Núñez (UCM).

SECRETARIO: Luis Unceta Gómez (UAM).

CONSEJO DE REDACCIÓN: Montserrat Benítez (CSIC), José Antonio Berenguer (CSIC), Joaquín Garrido (UCM), Juana Gil Fernández (CSIC), Salvador Gutiérrez Ordóñez (U. León y RAE), Antonio Hidalgo (U. Valencia), Patricia Infante (CSIC), Manuel Leonetti (U. Alcalá), Eugenio Luján (UCM), Victoria Marrero (UNED), Ventura Salazar (U. Jaén), Esperanza Torrego (UAM).

CONSEJO ASESOR: Alberto Bernabé (UCM), Margarita Cantarero (SEL), Ramón Cerdá (UB), Victoria Escandell (UNED), Marina Fernández Lagunilla (UAM), José Manuel González Calvo (U. Extremadura), Emma Martinell (UB), Juan Carlos Moreno Cabrera (UAM), Gregorio Salvador (RAE), José Carlos de Torres (SEL), Jesús de la Villa (UAM).

A partir del número 38 (2008) la Revista Española de Lingüística ha recuperado el formato de dos fascículos al año, con periodicidad semestral. Los trabajos enviados para su publicación han de dirigirse al Secretario de la revista. Deberán ser originales e inéditos y ajustarse a las normas que aparecen en el número 38/2, así como en la página web de la Sociedad Española de Lingüística. Todos los trabajos son sometidos al dictamen de al menos dos evaluadores designados por el Consejo de Redacción, mediante informes de carácter confidencial.

Los derechos de publicación y difusión, bajo cualquier forma, son propiedad de RSEL. Todo texto publicado en la revista obliga a sus autores a no cederlo a terceros, sin autorización previa de la revista, quien sí queda autorizada a comercializarlo, debiendo entregar, en este caso, el 50% de los beneficios obtenidos a sus autores.

REDACCIÓN: Sociedad Española de Lingüística, Centro de Ciencias Humanas y Sociales del CSIC, c/ Albasanz, 26 - 28, 28037 Madrid. Correo electrónico: secretarioRSEL@gmail.com. <<http://www.sel.edu.es/>>

DISEÑO Y COMPOSICIÓN: Carmen Chíncoa & Carlos Curiá (produccionRSEL@gmail.com)

SERVICIOS DE INFORMACIÓN: Los contenidos de la RSEL son recogidos sistemáticamente en *Bibliographie Linguistique / Linguistic Bibliography*, *CINDOC – Base de datos Sumarios ISOC, CSA – Linguistic and Language Behavior Abstracts, Dialnet, Francis, Modern Language Association (MLA) Bibliography*.

ÍNDICE 45/1 (2015)

INTRODUCCIÓN

Nuevas aportaciones al estudio de la percepción del habla7
JUANA GIL - EUGENIA SAN SEGUNDO

ARTÍCULOS

Implicaciones perceptivas de la variación: la fricativa labiodental25
BEATRIZ BLECUA FALGUERAS - ASSUMPCIÓ ROST BAGUDANCH

La percepción de la cualidad de voz y los estereotipos vocales45
MARIANELA FERNÁNDEZ TRINIDAD

La prosodia audiovisual de la ironía verbal: un estudio de caso73
SANTIAGO GONZÁLEZ FUENTE

¿Son distintos el creak y la voz creaky?: Estudio perceptivo preliminar105
PATRICIA INFANTE

La percepción del habla en ruido: un reto para la lingüística129
y para la evaluación audiológica (estudio experimental)
VICTORIA MARRERO-AGUIAR

Percepción de los enunciados interrogativos de duda y de enfado153
sin apoyo visual en alumnos chinos de ELE
SÍLVIA PLANAS-MORALES

Percepción y evaluación de la pronunciación del español como L2175
ENRIQUE SANTAMARÍA BUSTO

ARTÍCULOS

LA PERCEPCIÓN DE LA CUALIDAD DE VOZ Y LOS ESTEREOTIPOS VOCALES*

MARIANELA FERNÁNDEZ TRINIDAD

Consejo Superior de Investigaciones Científicas

RESUMEN

El presente trabajo realiza un estudio experimental sobre las asociaciones fonosimbólicas que suscita en español la cualidad de voz derivada de los ajustes laríngeos (F0 global y tipos de fonación *breathy* y *pressed*). Se ha llevado a cabo un test perceptivo de identificación de respuesta forzada. Para entender la manera en que las respuestas de los participantes estuvieron influidas por los atributos acústicos de las diferentes cualidades de voz, se realizó un análisis cuantitativo y cualitativo. Los resultados indican que las correspondencias entre significados y voces son uniformes para todos los jueces. También se ha podido comprobar que los oyentes siguen pautas de combinación bastante sistemáticas para establecer asociaciones perceptivas a partir de una determinada cualidad de voz. Estas asociaciones están doblemente pautadas; por un lado, por las asociaciones establecidas independientemente para el tono y los tipos de fonación, y por otro lado, por las asociaciones fijadas y estereotipadas lingüística y culturalmente. Se ha visto también que de todas las posibles asociaciones motivadas por la cualidad de voz se pueden derivar factores comunes principales que subyacen a todos los adjetivos calificativos asociados a una voz y que se harán manifiestos de forma más o menos evidente en los juicios de los sujetos, dependiendo de los contextos situacionales, lingüísticos y extralingüísticos.

Palabras clave: cualidad de voz; voz *breathy*; voz *pressed*; asociaciones fonosimbólicas; estereotipos vocales.

ABSTRACT

An experimental study was carried out to explore the sound symbolic associations that voice quality conveys. In this case, voice quality was modified by changing the global F0 and the phonation type (*breathy* or *pressed* voice). A forced choice answer perception test was chosen to verify what associations were

* Este trabajo ha sido posible gracias a la financiación del Consejo Superior de Investigaciones Científicas (Beca JAE-Predoctoral, Junta para la Ampliación de Estudios, BOE 03-06-2011) y del Ministerio de Ciencia e Innovación de España (FFI10-21690/Cualidad de voz). Quiero agradecer también a mis compañeros de trabajo sus sugerencias, todas ellas recogidas en este artículo. Muchas gracias a Helena Alves, Patricia Infante, José María Lahoz, Rubén Pérez y, muy especialmente, a Juana Gil y a Carolina Pérez.

produced by these variations. In order to understand whether the participants' answers were influenced by different voice qualities both quantitative and qualitative analyses were undertaken. Results show that the linkage between meanings and voices was very similar for all the judges. In the same way, it was proved that certain voice qualities led systematically to the same patterns of association. Associations depend on two different aspects; on the one hand, the tone and the phonation type, and on the other hand, previously fixed linguistic and cultural associations. In all the cases, depending on situational, linguistic or extralinguistic contexts, common factors are found to underpin all the qualifying adjectives used by the judges to define the voices.

Keywords: voice quality; breathy voice; pressed voice; sound symbolic associations; vocal stereotypes.

RECIBIDO: 01/04/2014

APROBADO: 23/01/2015

1. INTRODUCCIÓN

Los oyentes extraen una gran cantidad de datos cuando escuchan la voz de una persona. La voz revela información de diverso tipo sobre los propios hablantes: características físicas (edad, sexo, estado de salud, apariencia), psicológicas (actitud, estado emocional, personalidad) y sociales (educación, estatus social, origen). De hecho, la voz es uno de los principales medios a través de los cuales los hablantes proyectan su identidad (Kreiman y Sidtis 2011). No obstante, el poder asociativo de una determinada cualidad de voz no solo se limita a suscitar en el receptor múltiples asociaciones referidas al propio hablante, sino que el tipo de voz puede también influir en la percepción de otros significados, como, por ejemplo, atributos de objetos (colores, formas, texturas, sabores, etc.).

Los significados e impresiones que los oyentes extraen a partir de las voces pueden no corresponderse con la realidad, pero lo más interesante es quizás que esas asociaciones –correctas o no– puedan darse de manera repetida entre los individuos, al menos entre los que comparten una misma lengua y cultura. Es decir, que existan estereotipos vocales.

Diferentes disciplinas se interesan por estudiar la producción y la percepción de la voz. En el campo de la publicidad, por ejemplo, las impresiones transmitidas por las voces son esenciales para definir la imagen de una marca. Kreiman y Sidtis 2011, p. 3, recuerdan las típicas características proyectadas por las voces que publicitan coches lujosos.

Low pitch, *breathy* quality¹, and a fairly rapid speaking rate produce the image of an intimate message from a mature but energetic male who possesses authority, sex appeal, social status, and coolness.

1. Como se verá enseguida, *breathy voice* es un término inglés que hace referencia a un tipo concreto de fonación caracterizado desde el plano fisiológico por presentar mucho escape de

Los hablantes de español también utilizan de forma voluntaria distintos tipos de voces para despertar en los oyentes sensaciones y significados concretos². La voz *breathy*, por ejemplo, «suele dotar al mensaje al que da cuerpo de una carga muy marcada de sensualidad, pero también puede denotar intimidad, sorpresa o simple fatiga física» (Gil 2007, p. 220). La percepción de la cualidad de voz en español comienza a suscitar interés entre los estudiosos, pero lo cierto es que, en este ámbito, está prácticamente todo por decir.

El presente trabajo estudia algunas asociaciones posibles de las que es susceptible la cualidad de voz en español y analiza si es factible reconocer patrones vocales comunes vinculados con significados específicos y generales. Antes de proseguir, es conveniente realizar algunas aclaraciones terminológicas y conceptuales.

1.1. *Precisiones terminológicas*

1.1.1. *Voz y cualidad de voz*

Suele hacerse una distinción entre «voz» y «cualidad de voz»; además, estos términos pueden hacer referencia a conceptos más o menos amplios según los enfoques y propósitos (cf. Kreiman y Sidtis 2011).

Es posible, de una parte, considerar la voz desde el punto de vista articulatorio y acústico y, por tanto, referirse únicamente a la fisiología de su producción y al resultado acústico, la señal sonora. Si lo que interesa, en cambio, es resaltar la impresión perceptiva que provocan en el oyente determinados mecanismos fisiológicos y sus correspondientes resultados acústicos, suele preferirse el término «cualidad de voz» o simplemente «cualidad». Es decir, el término «voz» concierne estrictamente al ámbito fisiológico y físico del sonido, mientras que «cualidad» pertenece propiamente al ámbito de la percepción y alude a cómo suena y se percibe una determinada voz para los oyentes.

En este trabajo, la «cualidad de voz» se concibe, siguiendo a Kreiman y otros 2014, y Kreiman y Sidtis 2011, como el resultado de la interacción entre una señal acústica y un oyente, y se intenta analizar la relación entre los dominios fisiológico, acústico y perceptivo puesto que todos ellos interactúan en cualquier proceso comunicativo.

De otra parte, el término «cualidad de voz» puede hacer referencia únicamente al resultado perceptivo del proceso de fonación o, concebido en términos más globales, a la respuesta de los oyentes frente al sonido

aire por el espacio glótico, debido al cierre incompleto de los pliegues vocales. Se ofrecen más detalles a este respecto de la Sección 1.3.

2. Incluso en algunas lenguas, entre las que no se encuentra el español, determinados tipos de fonación, por ejemplo, se explotan lingüísticamente para significar (cf. Gordon y Ladefoget 2001).

final del habla (esto implica considerar la influencia perceptiva de aspectos laríngeos, supralaríngeos y temporales del habla, puesto que todo ello contribuye al efecto perceptivo global del habla de un locutor). Esta última perspectiva, mucho más completa que la primera, supone concebir la «cualidad de voz» como una compleja interacción entre oyente y señal acústica.

Este trabajo no diferencia los términos «voz» y «cualidad de voz» y la aparición de uno u otro alterna en el texto como expresiones equivalentes a «cualidad de voz». Esta se concibe en el sentido más amplio que se acaba de exponer aunque, como se verá más adelante, se analizan exclusiva y deliberadamente las asociaciones perceptivas provocadas por el tono y los tipos de fonación (cf. Sección 3), es decir, los correlatos perceptivos de algunos aspectos laríngeos. No se han considerado en el análisis otros aspectos del habla asociados con el comportamiento laríngeo, supralaríngeo o con la organización temporal, puesto que han sido previamente controlados en el experimento.

1.1.2. Frecuencia fundamental y tono

El término «frecuencia fundamental» (F0) hace referencia a una propiedad física, la frecuencia de vibración de los pliegues vocales; mientras que «tono» alude a la impresión perceptiva que produce en los oyentes la frecuencia fundamental de la onda sonora. De manera muy general y simplista, si la frecuencia de vibración de los pliegues vocales es baja, el tono se percibirá como grave; si es alta, se percibirá como agudo.

1.1.3. Ajustes laríngeos y tipos de fonación

El hablante puede realizar, de manera más o menos controlada según el caso, ciertos ajustes laríngeos para configurar la fuente sonora. De esta manera, el uso que el locutor hace de su laringe originará un resultado acústico determinado, es decir, un tipo de fonación. Existen varios tipos de fonación: *creaky voice*, *harsh voice*, *whispery voice*, *breathy voice* (Laver 1980) y *pressed voice* (Sundberg 1994), entre otros que pueden verse mencionados en la bibliografía. Cada uno de ellos presenta características acústicas y sonoras propias y conceden a la voz diferentes matices. Este estudio analiza los dos mencionados en último lugar. La voz *pressed*³ se caracteriza, en el plano fisiológico, por un cierre abrupto

3. Las traducciones al español que se han propuesto para los distintos tipos de fonación han sido muchas y variadas, por ejemplo, la voz *breathy* suele traducirse como 'soplada', 'de hálito', 'jadeante' o 'aérea' (cf. Poyatos 1994). Para evitar confusiones, en este artículo se ha considerado pertinente evitar la traducción de los adjetivos *breathy* y *pressed*. En la bibliografía, además de *pressed voice*, también se pueden encontrar los términos *pressed phonation* (por ejemplo, Ní Chasaide y Gobl 2005) y *tense voice* (Laver 1980), todos ellos traducibles al español como *voz tensa*.

y fuerte de los pliegues vocales. Puesto que los pliegues vocales contactan durante una gran parte del ciclo vibratorio, sale menos aire durante la fonación.

Es por tanto, un tipo de voz que presenta un elevado grado de tensión general y supone un importante esfuerzo laríngeo. La fuerte y rápida aducción glótica⁴ hace que el resultado acústico sea un ascenso de las altas frecuencias.

En la voz *breathy*, en cambio, la aducción de los pliegues vocales es incompleta. Este contacto y cierre ineficaces de los pliegues vocales provoca un continuo escape de aire a gran velocidad por el espacio glótico. Los pliegues vibran con poca masa de cuerpo y cubierta y presentan baja tensión longitudinal y poca fuerza de aducción, por ello casi nunca llegan a juntarse. La glotis permanece algo abierta a lo largo de la mayor parte de su longitud, a menudo, con una abertura en forma de ojal (Laver 1980). Este tipo de fonación, considerada hipofuncional, se produce con un bajo esfuerzo laríngeo, y el resultado perceptivo es una voz suave, débil, con muy poca sonoridad, que por lo general presenta una leve fricción glotal audible a causa del aire que escapa a gran velocidad (Bele 2005). Catford 1977, p. 99 explica sobre este tipo de fonación:

The effect is somewhat like that of sighing. [...] The vocal folds are vibrating, but without ever closing or, indeed, coming anywhere near closing. They simply flap in the breeze of the high velocity air-flow...

2. TEORÍA E HIPÓTESIS

2.1. Marco teórico

Distintos estudios han establecido ciertas asociaciones entre la altura de la F0 y algunos significados no lingüísticos, y asimismo, se han podido determinar asociaciones semejantes entre esta clase de significados y los tipos de fonación. Como se comprueba en la bibliografía, se han buscado, por una parte, coincidencias en la codificación de la frecuencia fundamental con ciertos significados y, por otra, concordancias entre las asociaciones previamente establecidas por los tipos de fonación. Respecto de la F0, han existido varios intentos por vincular determinados usos frecuenciales con la transmisión de significados. Por ejemplo, Bolinger 1964; Ohala 1983, 1984 y Morton 1994 han establecido que la altura tonal (F0 de una voz) transmite señales sociales para la dominación y la sumisión⁵. En diversas lenguas y culturas se ha observado que una F0 alta o una subida de tono se

4. Cierre de los pliegues vocales, por oposición a *abducción* (abertura).

5. El trabajo de Huron, Kinney y Precoda 2006 muestra una situación muy semejante en la música puesto que melodías con características acústicas equiparables provocan las mismas asociaciones.

asocian con la amabilidad, la deferencia, la sumisión, la falta de confianza en uno mismo, el miedo, la falta de defensa; mientras una F0 baja o un descenso en el tono se vinculan con la autoridad, la asertividad, la amenaza, la agresión, la confianza en uno mismo, etc.

Es posible que muchas de las asociaciones que se establecen a partir de patrones acústicos tengan una base fisiológica y biológica que, según Ohala, derivarían en parte de un código de frecuencia universal (*frequency code*) que compartimos incluso con los animales. Debido al tamaño de las cavidades de resonancia, los animales más pequeños, por ejemplo, emiten sonidos más agudos, al tiempo que los más grandes, más graves. Es factible que el cerebro humano recoja regularidades estadísticas como estas, presentes en la naturaleza, para realizar asociaciones o inferencias iniciales tales como tono agudo y tamaño pequeño, y tono grave y gran tamaño⁶. Luego, estas asociaciones primarias y muy icónicas probablemente susciten asociaciones secundarias, derivadas metafórica o metonímicamente. Por ejemplo, que el tono agudo se asocie con los significados de sumiso, indefenso, miedoso, infantil, superficial, y muchos otros; y que el tono grave se asocie más fácilmente con significados como amenazante, dominante, seguro, convincente, etc. Las asociaciones secundarias, si se dan de manera repetida en una comunidad hablante, pueden convertirse fácilmente en un hábito generalizado y pasar a constituir estereotipos.

Se ha intentado también vincular la altura global de la F0 con emociones y estados de ánimo. Por ejemplo, se ha visto que una altura tonal baja suele corresponderse con la tristeza y el aburrimiento; mientras que el miedo, el enfado y la alegría se identifican con alturas tonales elevadas (Fonagy 1978; Scherer 1974; Scherer y Oshinsky 1977; Frick 1985, etc.).

Varios estudios han observado que las emociones y los estados de ánimo tienen un correlato fisiológico en el hablante y, por consiguiente, en la fonación. Una determinada emoción o estado anímico conlleva inevitablemente una respuesta fisiológica o varias encadenadas, que provocan un resultado acústico que luego tendrá a su vez una interpretación perceptiva. Por ejemplo, un estado de excitación puede tener como respuesta una respiración agitada y esta desencadenar un aumento de la presión subglótica, la cual a su vez puede provocar un aumento en la frecuencia fundamental y en la intensidad, en contraste con los estados más relajados que suelen caracterizarse por valores más bajos de la F0, menor intensidad y menor energía en las altas frecuencias (Kreiman y Sidtis 2011).

Algunos investigadores sostienen que mucho de lo que los hablantes expresan con sus voces y los oyentes perciben –principalmente en lo que se refiere a emociones y estados de ánimo– es el nivel de excitación expresado

6. Esta correspondencia asociativa se ha comprobado en muchísimas lenguas del mundo, lo que, sin duda, constituye una pista acerca de su posible carácter universal (cf. Jakobson y Waugh 1979; Hinton, Nichols y Ohala 1994 y Tsur 2006, entre muchos trabajos).

en una dimensión «activo» frente a «pasivo». Las emociones y los estados anímicos de «activación» implican la excitación por parte del hablante y, en esta medida, exigen de este un mayor esfuerzo laríngeo, y, por tanto, ajustes de fonación más tensos (Lindsley 1951; Lieberman 1967; Scherer 1974; Laver 1980, entre otros). Al tiempo que los «pasivos» –o de baja «activación»– no requieren prácticamente de ningún esfuerzo fisiológico por parte del locutor, lo que supondría ajustes laríngeos más distendidos. Por ejemplo, la activación de emociones como el miedo, el enfado o la alegría incontrolada presenta unos correlatos acústicos similares: valores altos de F0, mayor rango tonal, intensidad más elevada, mayor energía en las altas frecuencias, e, incluso, a nivel supralaríngeo, mayor velocidad de elocución (Kreiman y Sidtis 2011).

De este modo, es posible postular cierto grado de iconicidad entre los ajustes laríngeos de mucha o poca tensión y los significados relacionados con los conceptos de «activación» frente a «baja activación (relajación)». Un tono de voz agudo y los tipos de fonación en los que el hablante necesite, por ejemplo, mayor esfuerzo laríngeo, mayor tensión de los pliegues vocales, etc., podrán asociarse más fácilmente a sensaciones, estados mentales y emocionales que impliquen precisamente algún tipo de tensión o activación. De otra parte, un tono de voz grave y los tipos de fonación que, al contrario, presenten ajustes articulatorios relajados y no supongan, por tanto, grandes esfuerzos para los hablantes, se asociarán perceptivamente con sensaciones, estados anímicos y emocionales distendidos o relajados, «pasivos» o «de baja activación» (Kreiman y Sidtis 2011).

Como se expone a continuación, en la bibliografía aparecen diversos trabajos de investigación que postulan una relación entre la forma en la que se producen distintos tipos de fonación y determinados significados. Por ejemplo, se ha observado, a partir de los juicios de los oyentes, que en inglés existe una asociación de la fonación *pressed* o *tense* con el enfado o la ira (Scherer 1986; Laukkanen y otros 1996; Buckhart y Sendmelier 2000; Gobl y Ní Chasaide 2003), el estrés (Gobl y Ní Chasaide 2003), el miedo (Scherer 1986), la actividad y excitación (Scherer 1974, Laver 1980, Frick 1985, entre otros), la alegría extrema (Scherer 1986; Gobl y Ní Chasaide 2003), la formalidad, la seguridad, y el interés (Gobl y Ní Chasaide 2003).

La voz *breathy* ha sido asociada a la tristeza (Scherer 1986; Laukkanen y otros 1996, Burkhardt y Sendlmeier 2000; Ní Chasaide y Gobl 2003, etc.), a la debilidad e inseguridad (Laver 1968; Gobl y Ní Chasaide 2003), a la intimidad y la confidencialidad (Laver 1980; Ní Chasaide y Gobl 2003), a la satisfacción y la amabilidad (Gobl y Ní Chasaide 2003), al relax (Gobl y Ní Chasaide 2003; Moos y otros 2011) y a connotaciones placenteras (Reich y Lerman 1978; Moos y otros 2013).

A su vez, en cada uno de estos grupos se pueden identificar asociaciones con emociones, actitudes, sensaciones o significados tanto «positivos» como «negativos», dependiendo fundamentalmente del contexto comunicativo

(Gobl y Ní Chasaide 2003; Kreiman and Sidtis 2011). Por ejemplo, una voz *breathy* se ha asociado con emociones y estados «pasivos» como la debilidad, inseguridad, tristeza (generalmente interpretados como «negativos»), pero también con estados «pasivos» de valoración «positiva», tales como relajado, distendido, satisfecho, íntimo, confiado y amable (Gobl y Ní Chasaide 2003).

Finalmente, si se hiciera abstracción de los significados particulares a los que las distintas características vocales se han asociado, sería posible encontrar algunos componentes básicos muy genéricos que aparecen como rasgo semántico subyacente a muchas asociaciones. Es decir, sería posible reducir todos los atributos particulares asociados con una voz a unos pocos factores comunes principales. Estos factores comunes estarán presentes de manera subyacente en las asociaciones y se harán manifiestos de forma más o menos evidente en los juicios de los sujetos, dependiendo de los contextos situacionales, lingüísticos y extralingüísticos.

Metodológicamente, existen al menos dos maneras de enfocar el análisis. Una es proponer *a priori* cuáles pueden ser los factores comunes a partir de unas escalas semánticas que el investigador marca de antemano en función de su propia intuición –lo que se conoce como «análisis factorial»–. Otra es tratar de derivar esos factores comunes principales de las asociaciones que los sujetos establezcan entre los estímulos propuestos «análisis multidimensional» (cf. Kreiman y Sidtis 2011 para más detalles sobre estos y otros modelos de análisis). Un enfoque similar al que se lleva a cabo mediante el análisis multidimensional es el que se realizará para este trabajo. Por tanto, y como se verá en la metodología, se someterá a los oyentes a distintos estímulos y posteriormente se intentará deducir los factores comunes que subyazcan a las asociaciones más o menos explícitas que los oyentes hayan establecido entre cualidad de voz y significado, en función de las respuestas obtenidas en el experimento.

2.2. Hipótesis

A partir de estos argumentos, la primera hipótesis de este trabajo queda formulada de la siguiente manera:

H1: Detrás de las múltiples y variadas asociaciones que los sujetos hacen a partir de las propiedades físicas de una voz (F0 y tipos de fonación) están presentes componentes básicos muy genéricos que aparecen como rasgos semánticos (factor común principal) de múltiples calificativos dados a una determinada cualidad vocal. Estos factores comunes principales se podrán deducir a partir de las asociaciones que establezcan los sujetos de forma más o menos explícita.

En este trabajo además de intentar descubrir qué asociaciones suscitan entre los oyentes, por un lado, la frecuencia fundamental y, por otro, los tipos de fonación, se pretende estudiar el resultado asociativo derivado de

la combinación de ambos. Se postula que dicha combinación tampoco será azarosa, sino que estará pautada doblemente, por un lado, por las asociaciones previas que los oyentes hayan hecho de manera independiente a partir de los parámetros físicos de la voz (valores de la F0 y tipos de fonación) y, por otro, por la lengua y la cultura de los oyentes. Esto implica postular que los oyentes conocen las asociaciones que de forma aislada evocan la F0 y los tipos de fonación, y que son capaces de establecer relaciones entre ambos combinando los significados de manera reglada por pautas culturales. Por tanto, la segunda hipótesis que se pretende comprobar es:

H2: Las asociaciones perceptivas que los oyentes establezcan a partir de una determinada cualidad de voz (resultado de la combinación de los valores de la F0 y de los tipos de fonación) estarán pautadas en función de los factores comunes asociados a los parámetros físicos por separado y seguirán pautas combinatorias claramente establecidas por una misma lengua y cultura para transmitir significados muy concretos.

3. DISEÑO EXPERIMENTAL

3.1. Estímulos

Se diseñó un experimento perceptivo con el fin de observar las asociaciones que distintas cualidades de voz provocan en los oyentes, en concreto, qué tipos de voces se vinculan preferentemente con determinadas características o atributos. Para ello se enmarcó el experimento en el ámbito de la publicidad y se pidió a los jueces que evaluaran el grado de adecuación de diferentes voces (grabadas por una locutora a este propósito) con distintos productos que también se crearon *ad hoc*.

Para el estudio se seleccionaron cuatro categorías de productos:

1. Un parque acuático con emocionantes atracciones
2. Un parque temático de ciencias naturales
3. Un helado de chocolate con trocitos de caramelo y trufa
4. Un zumo de frutas, con vitaminas A y C, 100% natural y 0% de materia grasa

Los estímulos sonoros consistieron en dieciséis frases grabadas. Para anunciar cada producto se registró la misma frase publicitaria locutada con 4 tipos de cualidad de voz diferentes. Las distintas voces combinaban dos tipos de fonación (*breathy voice* y *pressed voice*) y dos alturas tonales (F0 global elevada y F0 global baja).

Se presentan a continuación las frases y los diferentes tipos de cualidad de voz utilizados para el experimento:

1. Es el parque acuático con las más emocionantes atracciones
2. Es el nuevo parque temático de ciencias naturales

3. Es el nuevo helado de chocolate cremoso con trocitos de caramelo y trufa
4. Es el zumo de frutas, con vitaminas A y C, 100% natural y con 0% de materia grasa

- a. *breathy voice* + tono agudo
- b. *breathy voice* + tono grave
- c. *pressed voice* + tono agudo
- d. *pressed voice* + tono grave

El experimento fue realizado con habla femenina y se utilizó una única locutora. Las frases se grabaron, en el Laboratorio de Fonética del CSIC, por una locutora profesional, lingüista y fonetista especializada en voz⁷. Además, las muestras de voz se sometieron a una validación acústica y perceptiva por parte de cinco fonetistas especializados en calidad de voz, quienes coincidieron en que las voces respondían a los modelos buscados.

Puesto que el objetivo fundamental de este estudio es observar la asociación sonido-idea con respecto a diferentes cualidades de voz (tipo de fonación + tono global), la locutora intentó controlar al máximo otros aspectos fonéticos que hubieran podido influir en las asociaciones perceptivas de los sujetos. Por ello, en cada frase mantuvo constante la intensidad global (como se verá luego, alrededor de los 73 dB), mantuvo una velocidad de articulación semejante (con una media de 5,4 sílabas por segundo⁸), no realizó pausas y repitió los mismos movimientos entonativos. En la Tabla I se ofrecen las mediciones realizadas para cada frase en cuanto a la velocidad de articulación, la frecuencia fundamental y la intensidad.

3.1.1. Tono e intensidad

Las voces agudas que han servido de estímulo para este experimento están en torno a los 272 Hz de media y los 24,6 Hz de desviación estándar. Por su parte, las voces graves presentan 180,0 Hz de media y 8,3 Hz de desviación estándar. Las voces que en este trabajo se han considerado graves o agudas pertenecen a dos grupos de frecuencia diferente puesto que excitan una zona distinta de la membrana basilar. Las que fueron clasificadas como graves entran todas dentro de la banda crítica centrada

7. Para la grabación de todas las frases se utilizó un micrófono (Audio-technica AT3031) de condensador de respuesta plana y patrón polar cardioide. Se colocó a 15 cm de la comisura derecha de la boca, fuera de eje, para evitar las explosiones generadas por el aire impactando de forma directa sobre la cápsula. El micrófono estaba sujeto a un pie de micro con suspensión para evitar ruidos provocados por posibles golpes. Se utilizó también una tarjeta de sonido externa Tascam US 122 MK II que proporciona la alimentación Phantom al micrófono y un cable microfónico marca Vovox equilibrado y protegido contra interferencias, de 3 metros. Se obtuvo una señal monoaural recogida con la aplicación WavePad, con una frecuencia de muestreo de 44100 Hz.

8. Las diferencias de velocidad de articulación que pueden apreciarse se encuentran dentro de márgenes aceptables.

en los 150Hz; mientras que las voces clasificadas como agudas pertenecen a la siguiente banda crítica, la de los 250Hz, cuyo ancho de banda en ambos casos es de 100Hz. En cuanto a las bandas de tercio de octava, las frecuencias fundamentales catalogadas como bajas (tono grave) están repartidas en las bandas centradas en 160Hz y 200Hz; mientras que las altas (tonos agudos) se hallan en bandas superiores, la de 250Hz y la de 315Hz.

	VA (s/s)	F0 (Hz)	TF	Int. (dB)
zumو_breathy_high	5,9	249	<i>breathy voice</i>	73
zumو_breathy_low	5,6	170	<i>breathy voice</i>	74
zumو_pressed_high	5,6	289	<i>pressed voice</i>	73
zumو_pressed_low	5,9	182	<i>pressed voice</i>	71
	VA (s/s)	F0 (Hz)	TF	Int. (dB)
helado_breathy_high	5,0	232	<i>breathy voice</i>	75
helado_breathy_low	5,1	173	<i>breathy voice</i>	75
helado_pressed_high	5,0	277	<i>pressed voice</i>	73
helado_pressed_low	5,1	182	<i>pressed voice</i>	69
	VA (s/s)	F0 (Hz)	TF	Int. (dB)
ciencias_breathy_high	5,7	267	<i>breathy voice</i>	70
ciencias_breathy_low	6,1	182	<i>breathy voice</i>	76
ciencias_pressed_high	5,5	279	<i>pressed voice</i>	72
ciencias_pressed_low	5,9	178	<i>pressed voice</i>	72
	VA (s/s)	F0 (Hz)	TF	Int. (dB)
ciencias_breathy_high	5,7	264	<i>breathy voice</i>	69
ciencias_breathy_low	5,4	175	<i>breathy voice</i>	72
ciencias_pressed_high	5,5	313	<i>pressed voice</i>	75
ciencias_pressed_low	5,7	197	<i>pressed voice</i>	72

TABLA 1. Cálculo de la velocidad de articulación, la F0 y la intensidad de todas las frases que conforman el corpus. Abreviaturas utilizadas: VA: velocidad de articulación, s/s: sílabas por segundo; F0: frecuencia fundamental, Hz: hercios; TF: tipo de fonación; Int.: intensidad; dB: decibelios.

Las diferencias de intensidad entre las distintas secuencias presentadas como estímulos en el experimento no son perceptivamente relevantes. La media de intensidad de todas las voces es de 72,6dB. Por tanto, si se observan los datos recogidos en la Tabla 1, todos los valores están dentro del rango +/-4dB. Serían necesarias variaciones mayores para percibir una clara diferencia de intensidades, sobre todo en altas frecuencias (cf. Plack 2005).

3.1.2. Tipos de fonación

Para comprobar que los estímulos presentados a los oyentes responden a los modelos de fonación descritos previamente como voz *breathy* y voz *pressed*, se llevaron a cabo análisis acústicos, además de la comprobación perceptiva por parte de los expertos. Las medidas espectrales que pueden extraerse a partir del sonido suelen proporcionar información útil sobre la cualidad de voz (Gobl y Ní Chasaide 2003). El análisis basado en el Espectro Medio de Largo Alcance (LTAS)⁹ permite estudiar la variación en la cualidad de voz a partir del producto acústico vocal. En concreto, la inclinación espectral aporta información sobre los tipos de fonación, puesto que permite observar, por ejemplo, la relación que se establece entre la amplitud de la frecuencia fundamental (F0) y la de los demás armónicos del espectro. La comparación de los niveles de amplitud entre la F0 y el segundo armónico (H2) se utiliza muy frecuentemente para inferir el comportamiento de la fuente sonora. La técnica del LTAS resulta muy útil principalmente para diferenciar la cualidad de voz en la dimensión tenso-laxo, una F0 con menos amplitud que el H2 indicaría una voz más tensa –*pressed voice*–, al tiempo que una F0 con mayor amplitud que el H2 sería índice de ajustes fonatorios más laxos –*breathy voice*– (cf. por ejemplo, Sundberg. Titze v Scherer 1993: Gobl v Ní Chasaide 2003).

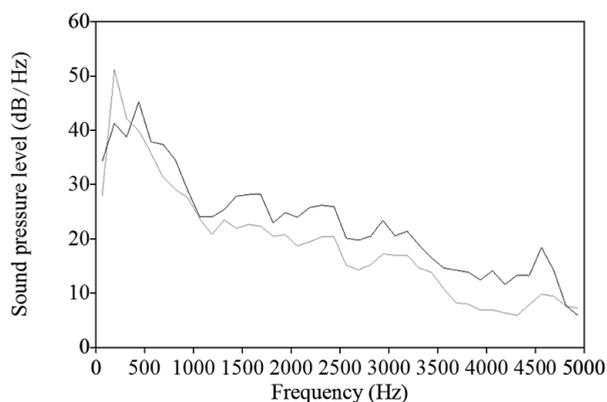


FIGURA 1. Análisis Espectral Medio de Largo Alcance (LTAS) de la frase *Es el parque de ciencias naturales*, locutada con *pressed voice* (color negro) y con *breathy voice* (color gris). Las frecuencias (Hz) están representadas en el eje de abscisas (x), y el nivel de presión sonora (dB) que alcanza cada frecuencia está representado en el eje de las ordenadas (y).

9. LTAS (*Long-Term Average Spectrum*) es un método de análisis acústico que permite visualizar cómo se distribuye la energía promedio del sonido (dB) en diferentes frecuencias (Hz) a partir de una muestra de voz continua (Leino 2009).

En las curvas superpuestas que se muestran en la Figura 1 se representa cómo se distribuye la energía sonora en función de todas las frecuencias entre 0 y 5000 Hz, en dos cualidades de voz diferentes, emitidas por la misma locutora. Con bastante claridad se observa que la voz *breathy* (en color gris) presenta una mayor declinación espectral que la voz *pressed* debido a la mayor pérdida de aire. La voz *pressed* exhibe, en cambio, una inclinación espectral bastante menos marcada y se observa, además, un refuerzo de la energía en los armónicos de altas frecuencias, esto último debido a que, en este tipo de fonación, las paredes del tracto vocal están tensas y permiten que rebote mejor la onda obteniendo así mayor energía.

En la voz *breathy* (color gris) se aprecia que el primer pico de energía¹⁰, correspondiente a la F0, alcanza una amplitud de unos 50 dB en torno a los 250 Hz, mientras que esa misma frecuencia, pero en la voz *pressed* (color negro), se ubica prácticamente 10 dB más abajo. La amplitud del segundo pico de energía, es decir, del segundo armónico (H2), se ubica en torno a los 45 dB para la voz *pressed* y en torno a los 40 dB para la voz *breathy*.

3.2. Jueces

Participaron en el experimento perceptivo 96 jueces españoles (61 mujeres y 35 hombres, con edades comprendidas entre 13-50 años, todos ellos hablantes nativos de castellano y de nivel socio-cultural medio, sin diferencias evidentes que pudieran afectar a los resultados.

3.3. Procedimiento general

El experimento se llevó a cabo en distintas sesiones; en cada una de ellas los participantes permanecieron sentados alrededor de una mesa dentro de un aula. Se les dijo que una empresa estaba interesada en recoger sus impresiones y opiniones sobre las voces de la publicidad de potenciales productos y que su participación sería totalmente voluntaria y anónima. Todos dieron su consentimiento y aseguraron que no tenían problemas de audición.

Se les entregó un folleto-encuesta en el que se podía leer: «Una marca está interesada en recoger vuestras reacciones sobre la publicidad de productos que saldrán pronto al mercado». La tarea consistiría en escuchar una serie de voces que anunciaban diferentes productos y contestar sobre el grado de ajuste y de agradabilidad entre las voces y los productos. Los participantes leyeron la consigna del experimento:

10. El primer pico de cada una de las curvas corresponde a la F0 media obtenida analizando todos los segmentos sonoros de cada emisión.

«A continuación te presentamos cuatro nuevos productos. Luego de escuchar parte de los distintos anuncios publicitarios debes indicar qué tipo de voz te resulta más apropiada (te gusta más) para anunciar cada producto.»

Se llevó a cabo un test perceptivo de identificación de respuesta forzada. Las voces fueron presentadas con un ordenador, utilizando el programa Praat. El orden de presentación fue aleatorio, y entre una voz y la siguiente transcurrieron 0,5 segundos. Los participantes escucharon los estímulos a través de los altavoces a un nivel de intensidad suficiente como para apreciar los matices de las diferentes cualidades de voz. Una vez terminada la escucha de toda la serie con la repetición, señalaron la voz que consideraban más adecuada para publicitar el producto. Los sujetos tenían también la posibilidad de escoger la misma voz para asociarla a productos diferentes.

Puesto que pueden ser muchos y diversos los motivos por los cuales se elige una voz para un determinado producto, se les pidió que describieran en un espacio dedicado a comentarios cómo percibían cada voz, por qué motivos la asociaban a un determinado producto y qué características o atributos tenía ese producto para ser asociado preferentemente a un tipo de voz.

Este mismo procedimiento se repitió para cada categoría de producto. El experimento duró aproximadamente 15-20 minutos.

4. RESULTADOS

Los participantes del experimento lograron identificar con rapidez y facilidad un tipo de voz con un producto. Como se verá enseguida, los jueces utilizaron la misma distribución de frecuencias entre voces y productos independientemente de la edad y del sexo.

4.1. Voces y productos

Los resultados indican uniformidad en la asignación de las distintas voces (cualidades de voz) a los productos para todos los jueces. Al parecer, los hablantes no seleccionan las propiedades perceptivas de las voces al azar, ya que identifican nítidamente cada producto con un tipo de voz.

Al calcular los porcentajes de asociación dentro de cada producto, se observa que el 76% de los sujetos eligió para el helado de chocolate la cualidad de voz que combina la fonación *breathy* con un tono grave. El 17,7% asoció ese mismo producto con la voz *breathy_aguda*, el 5,2% con la voz *pressed_aguda* y solamente el 1% lo vinculó con *pressed_grave*. El parque de ciencias naturales fue asociado a la voz *pressed* con tono grave en un 57,4% de las respuestas. El 22,3% lo asoció con la voz *breathy_grave*, el 10,6% con *breathy_aguda* y, finalmente, el 9,6% de los sujetos lo

vinculó con la cualidad de voz *pressed_aguda*. El 67,4% relacionó el parque acuático también con la voz *pressed*, pero combinada con tono agudo.

		Producto			
		helado chocolate	parque ciencias naturales	parque acuático	zumo frutas
Cualidad de voz	<i>pressed_grave</i>	1	54	13	16
	<i>pressed_aguda</i>	5	9	64	28
	<i>breathy_grave</i>	73	21	1	7
	<i>breathy_aguda</i>	17	10	17	40

TABLA 2: Asociación de productos con cualidad de voz (recuento numérico de casos).

Al zumo se le asociaron preferentemente dos tipos de fonación, *breathy* y *pressed*, pero siempre combinados con un tono agudo (44% y 30,8% respectivamente). El 17,6% identificó el zumo con la voz *pressed_grave* y apenas un 7,7% lo hizo con la voz *breathy_grave*. Nótese que ciertas cualidades de voz fueron desestimadas para ciertos productos. Los tipos de fonación *pressed* prácticamente nunca se asociaron al helado de chocolate, y la voz *breathy_grave* no se relacionó con el parque acuático (una única respuesta).

Al considerar los porcentajes de asociación dentro de la cualidad de voz se aprecia que, de entre todas las ocasiones en que los sujetos prefirieron la voz *pressed_grave*, el 64,3% de las veces estuvo asociada con el parque de ciencias naturales. El 19% se asoció con el zumo, el 15,5% con el parque acuático y el 1,2% con el helado de chocolate. De todas las ocasiones en que los jueces optaron por la cualidad de voz *pressed_aguda*, el 60,4% se correspondió con el parque acuático, el 26,4% con el zumo, el 8,5% con el parque de ciencias y solo el 4,7% de las veces se asoció con el helado. Del total de selecciones para la voz *breathy_grave*, el 71% correspondieron al helado de chocolate, el 20,6% al parque de ciencias, el 6,9% al zumo y el 1% al parque acuático.

Por último, de entre todas las ocasiones en las que los sujetos escogieron la cualidad de voz *breathy_aguda*, el 47,6% de las veces fue para asociarla con el zumo, el 20,2% con el parque acuático, otro 20,2% para vincularla con el helado y el 11,9% con el parque de ciencias. Para comprobar que el conjunto de asociaciones establecidas entre cualidad de voz y producto no se realizaron por azar, se aplicó el test del chi-cuadrado y sus resultados ($X^2(9)=305,440$ $p < 0,001^{***}$) permiten afirmar con seguridad que tal distribución no es fortuita, con una medida de probabilidad de exactitud del 80%.

Por último, se ha hecho un análisis de correspondencias entre cualidad de voz y producto mediante escalamiento multidimensional, para representar gráficamente la relación asociativa entre las distintas categorías de productos y tipos de voces.

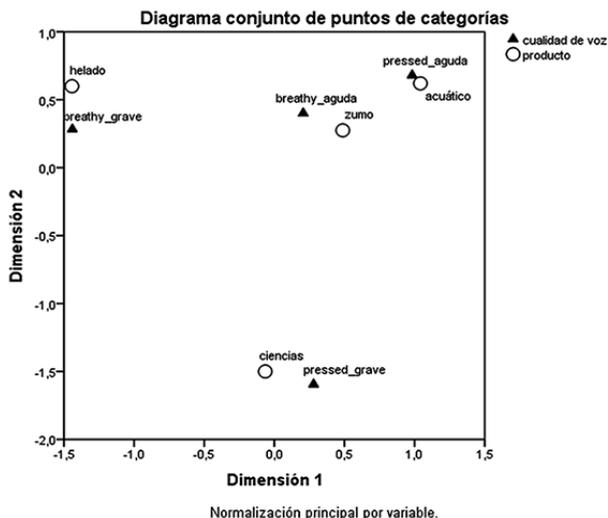


FIGURA 2. Representación gráfica de las categorías de las variables (cualidad de voz y tipo de producto) mediante escalamiento multidimensional.

Los puntos representan las categorías (4) de las variables (tipo de voz y producto) y la cercanía de los puntos indica el grado de asociación entre ellos: a mayor proximidad, mayor asociación. El parque acuático fue asociado a la voz *pressed_aguda*, el parque de ciencias a la voz *pressed_grave*, el helado de chocolate a la voz *breathy_grave* y, para el zumo, la voz *breathy_aguda* resultó ser la favorita. Resulta interesante señalar que las asociaciones más fuertemente establecidas, es decir, las que podrían estar más claras entre los oyentes, son las que involucran las cualidades de voz *pressed_aguda* y *breathy_grave*, es decir, cualidades vocales que combinan articulariamente tensión o distensión. En estas cualidades vocales, ambos ajustes laríngeos van en la misma dirección y se refuerzan en la dimensión tenso/laxo.

4.2. Tono de voz y productos

Al calcular los porcentajes de asociación dentro del producto, el 77,1% de las personas que realizaron el experimento eligió para el helado de chocolate el tono grave, y el 22,9% prefirió el agudo. Para el parque de ciencias, el 79,8% escogió el tono grave y el 20,2% el tono agudo. El 85,3%

de los jueces asoció el parque acuático con la voz aguda y solo el 14,7% con la voz grave. El zumo se vinculó con el tono agudo para un 77,4% de las personas y con el tono grave para el 25,3% restante.

		Producto			
		helado	parque de ciencias	parque acuático	zumo
Tono	grave	74	75	14	23
	agudo	22	19	81	68

TABLA 3. Asociación entre productos con tono grave y agudo (recuento numérico de casos).

Al atender a los porcentajes de asociación dentro del tono, se observa que, considerando todas las ocasiones en las que los jueces escogieron el tono grave, en un 39,8% de las veces lo hicieron para vincularlo con el helado; en un 40,3% de las veces, para vincularlo con el parque de ciencias; en el 12,4%, con el zumo; y en el 7,5% restante, con el parque acuático. Finalmente, de todas las veces que las personas prefirieron la voz aguda, el 42,6% de estas elecciones se realizaron para asociarla con el parque acuático, el 35,8% se correspondió con el zumo de frutas, mientras que únicamente el 11,6% de las veces asociaron el tono agudo con el helado de chocolate y el 10% con el parque de ciencias.

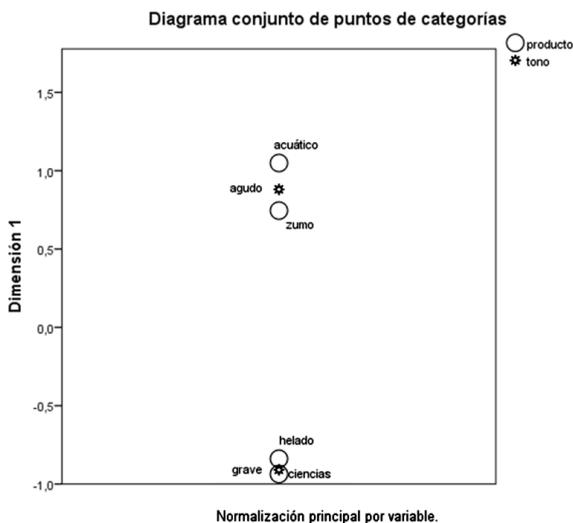


FIGURA 3. Representación gráfica de las categorías de las variables (frecuencia fundamental y tipo de producto) mediante escalamiento multidimensional.

Una prueba chi-cuadrado demostró que estos resultados son estadísticamente significativos, $X^2(3) = 131,006$, $p < 0,001^{***}$. El parque de ciencias y el helado de chocolate se vincularon fuertemente al tono grave. El zumo y el parque acuático se asociaron con el tono agudo.

4.3. Tipos de fonación y productos

El cómputo de los porcentajes de asociación dentro de cada producto indica que el helado de chocolate se identificó en un 93,8% con la fonación *breathy* y apenas en un 6,3 % con la *pressed*. Para el 67% de los sujetos, el parque de ciencias se vinculó a la voz *pressed* y para el 33% restante, a la voz *breathy*. El parque acuático fue asociado a *pressed* para el 81,1% de los jueces y a *breathy* para tan solo el 18,9%. Para el zumo, la asociación con un tipo de fonación no aparece numéricamente tan clara, puesto que el 51,6 % lo asocia con *breathy* y el 48,4%, con *pressed*.

		Producto			
		helado	parque ciencias	parque acuático	zumo
Tipo de fonación	<i>breathy</i>	90	31	18	47
	<i>pressed</i>	6	63	77	44

TABLA 4. Recuento de casos para la asociación entre productos y tipos de fonación (*breathy/pressed*).

Al evaluar los porcentajes relativos a las asociaciones establecidas por los oyentes dentro de cada tipo de fonación se observa una marcada preferencia por la voz *breathy* para el helado de chocolate, que alcanza casi la mitad de todas las respuestas (48,4%). El 25,3% de todas las veces que se eligió este tipo de fonación fue para vincularlo al zumo; el 16,7%, para relacionarlo con el parque de ciencias; y el 9,7%, para asociarlo con el parque acuático. La voz *pressed* se asoció más fácilmente con el parque acuático (40,5%) y con el parque de ciencias (33,2%), menos fuertemente con el zumo (23,2), y prácticamente nunca con el helado (3,2%).

Todos los resultados correspondientes a las asociaciones entre tipo de fonación y productos resultaron ser estadísticamente relevantes, $X^2(3) = 121,106$ $p < 0,001^{***}$.

En la asociación entre tipos de fonación y productos, se observa una clara preferencia por la voz *breathy* para el helado de chocolate. La voz *pressed* se asoció más fácilmente con el parque de ciencias naturales y con el parque acuático. Las asociaciones entre voces y productos se dividen entre ambos tipos de fonación para el zumo, apreciándose una ligera preferencia por la fonación *breathy*.

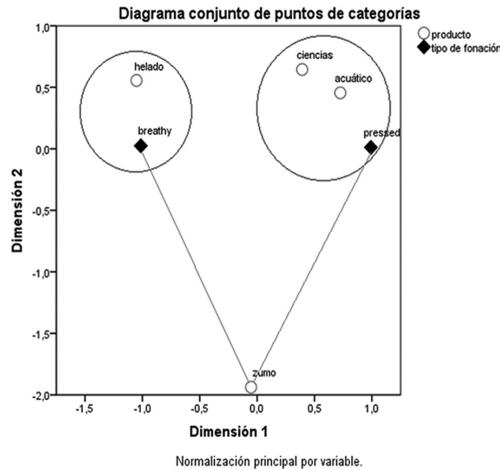


FIGURA 4. Representación gráfica de las categorías de las variables (tipo de fonación y tipo de producto) mediante escalamiento multidimensional.

4.4. Calidad de voz y sexo del oyente

Las pruebas de chi-cuadrado indican que no hay razones estadísticamente significativas para afirmar que, entre los sexos, existan diferencias en la asignación de voces a productos ($X^2= 3,780$ $p < 0,286$); dicho de otro modo, los hombres y las mujeres no asocian distintos tipos de voces para cada producto.

4.5. Calidad de voz y edad de los oyentes

Tampoco se han encontrado motivos estadísticamente significativos para asegurar que existan diferencias en la asociación de una calidad de voz a un producto entre los grupos de edades¹¹ de los sujetos ($X^2= 9,017$ $p < 0,436$).

5. DISCUSIÓN

En la tarea perceptiva realizada en este experimento los sujetos habían de elegir cuál era el tipo de voz que mejor se ajustaba a un producto u otro. Los oyentes debieron realizar, por tanto, una implicación encadenada. Primero analizaron, al menos de manera inconsciente, qué atributos o sensaciones les evocaba cada voz y, luego, si tales atributos y sensaciones

11. Para la variable demográfica edad se establecieron 3 grupos: de 12-18 años (44 personas), de 19-30 (36 personas), de 31 a 50 (16 personas).

eran los que en su conocimiento del mundo mejor cuadraban con cada producto.

El experimento obtiene datos muy sólidos y coherentes: o todos los sujetos asocian un mismo producto con un tipo de voz o, si las respuestas se dispersan, normalmente se dividen en dos cualidades de voz que siempre presentan algo en común, o el tono o el tipo de fonación.

Al parque temático de ciencias naturales se le asoció la voz *pressed* y grave, y los sujetos explicaron en los comentarios que esa cualidad de voz transmitía seriedad, rigurosidad, convencimiento, fiabilidad, objetividad, veracidad, didáctica y enseñanza. La voz *pressed* y aguda se eligió para el parque acuático por ser una voz que transmitía significados tales como diversión, entusiasmo, acción, emoción, excitación y tensión. La cualidad de voz que combina la fonación *breathy* con un tono grave resultó ser la favorita para el helado de chocolate por comunicar mejor los significados de placer, seducción, tranquilidad, serenidad, cremosidad y succulencia. Finalmente, las asociaciones vocales se dividen para el zumo (30,8% de respuestas para la voz *pressed* y aguda, y 44,0% para la voz *breathy* y aguda). Todas las respuestas han coincidido en que la voz aguda era más apropiada para transmitir la idea de que el zumo es algo ligero y fluido. Con respecto al tipo de fonación, los sujetos que respondieron que la cualidad de voz más apropiada para un zumo era la voz *pressed* y aguda comentaron que les transmitía significados como energía, vitaminas, enganche. Los que eligieron para el zumo la voz *breathy* y aguda asociaron ese tipo de voz con significados del tipo sano, saludable y natural. Por tanto, los oyentes han dado prioridad, en un caso, al concepto general de «energético» y, en el otro, al de «saludable». Lo que ha ocurrido es que los oyentes han cambiado de factor, el factor que actuaba para el zumo natural y para el helado no fue el mismo que actuó para el zumo con vitaminas. Esto podría significar que la elección de una determinada cualidad de voz para un mismo producto haya dependido de la relevancia cognitiva otorgada a cada factor en el proceso de percepción.

Si se analizan las respuestas obtenidas aislando las asociaciones entre, por un lado, producto y tono y, por otro, producto y tipo de fonación, se observa que un único parámetro físico, por ejemplo, la frecuencia fundamental, cuando obtuvo valores bajos, se asoció perceptivamente tanto con el helado de chocolate como con el parque de ciencias naturales. Del mismo modo, un tono agudo, correlato de una frecuencia fundamental más elevada, se asoció con el zumo y con el parque acuático. Algo análogo ocurre al analizar las asociaciones evocadas por los tipos de fonación. A un mismo mecanismo de fonación, la voz *breathy*, se le ha asociado un zumo y un helado; y a la voz *pressed*, un parque de ciencias naturales y un parque acuático. ¿Cómo es posible que un mismo parámetro haya sido asociado a la vez con productos en apariencia tan diferentes? ¿Es posible

encontrar algún punto de unión entre las asociaciones concretas que los oyentes han establecido?

Como se explicará en las líneas que siguen, las cualidades de voz no suelen estar asociadas con un único atributo, sino que más bien se asocian en la mente de los oyentes a una constelación de atributos (Gobl y Ní Chasaide 2003). Se argumentará que estos atributos están relacionados entre sí y que les subyace un factor semántico común.

5.1. *Significados asociados al tono de voz*

Un tono de voz grave se ha asociado a adjetivos tales como cremoso y succulento (para el helado)¹², y a serio, riguroso, fiable, objetivo y veraz (para el parque de ciencias). Aunque la primera impresión es que la voz grave puede tener asociados adjetivos que en principio están muy alejados entre sí, luego es posible observar que esta serie de adjetivos se mueve en una misma línea, la línea de la «densidad», tanto en una idea (parque de ciencias) como en una textura (helado). En el otro extremo, y en cierto modo contrario, el tono de voz agudo se ha vinculado con adjetivos como ligero y fluido, para el zumo, y con calificativos como divertido y entusiástico para el parque acuático.

Al hacer abstracción de las asociaciones concretas que se han hecho para el tono, y si se llevara a cabo un análisis multidimensional, entre todos esos adjetivos sería posible derivar un factor común muy general de «formalidad» para el tono grave, y un significado básico común de «informalidad» para el tono agudo. Luego, esos significados básicos de «formalidad»/«informalidad» podrían interpretarse en escalas semánticas distintas, relativas a calificativos asociados como profundo, denso, reflexivo, minucioso, serio, maduro, sabio, tranquilo, juicioso, aburrido, exacto, riguroso, veraz, severo... en el primer caso; y en el segundo, a adjetivos como superficial, ligero, fluido, infantil, tonto, divertido, alegre, entusiástico, engañoso..., y otros muchos dentro de la misma escala. Es decir, es posible postular que lo que hay detrás de las asociaciones concretas que se han hecho para

12. El tono grave asociado a calificativos como cremoso y succulento para un helado se había encontrado ya en un estudio experimental que analizaba el efecto fonosimbólico que provocan en los oyentes las características acústicas de los segmentos del habla (Yorkston y Menon 2004). Se ha visto que la estructura acústica segmental influye en la percepción de los atributos y de los significados provocando en los oyentes asociaciones fonosimbólicas de distinto tipo (cf. Jakobson y Waugh 1979). Un número creciente de estudios actuales aplican estos conocimientos al ámbito de la publicidad, en concreto, al diseño del nombre de marcas –*brand naming*–, y demuestran que el simbolismo sonoro constituye un factor importante para inferir significados específicos a partir de los sonidos que conforman los nombres de marcas desconocidos (cf. Spence 2012 para una síntesis). El estudio de Yorkston y Menon demostró que, frente a dos nombres de marca ficticios, *Frish* y *Frosh*, los sujetos prefirieron, de manera inconsciente y automática, el nombre con la vocal posterior o (sonido grave) para un helado en el que se resaltaba la *cremosidad* y la *suculencia* (*richness*) como atributos positivos y principales.

el tono y unos determinados productos –a su vez estos, con propiedades también específicas– es un componente básico, muy genérico, que aparece como rasgo semántico (factor común) de una serie de adjetivos distintos.

Se ha visto antes (cf. Sección 2.1) que distintos estudios han establecido correspondencias asociativas en la dimensión «activo-pasivo» para el tono de la voz (valores de la F0) y los tipos de fonación. En este trabajo, la interpretación perceptiva de la altura tonal también ha presentado patrones asociativos comunes en esa dimensión. El tono agudo de la voz se asoció con sensaciones y estados de «activación positivos» como la diversión y el entusiasmo. No parece ser tan evidente la asociación entre tono grave y los significados de seriedad y rigurosidad como pertenecientes al ámbito de la «pasividad», si bien se podría apuntar que todo aquello que es opuesto a la improvisación, a la precipitación o a la activación inmediata y, que, al contrario, implica reacciones sopesadas, es meditado y requiere tiempo, tiene que ver con sensaciones y estados «pasivos» y «positivos».

5.2. Significados asociados al tipo de fonación

La voz *breathy* se ha asociado tanto al helado de chocolate como al zumo de frutas. ¿Qué es lo que puede unir al zumo con el helado para que los sujetos hayan elegido una voz *breathy* en ambos casos? ¿Es posible encontrar un factor común asociado a ese tipo de fonación que subyazca en las propiedades de los dos productos? Al vincular la voz *breathy* con el helado de chocolate, los sujetos manifestaron que el motivo por el cual elegían esa voz y no otra era porque les transmitía las ideas de algo placentero, sensual, tranquilo y sereno. Cuando fue el zumo el que se asoció a este tipo de fonación, se destacaron los significados de sano, saludable y natural. Un posible punto de conexión entre estas asociaciones concretas podría ser la idea bienestar. Por tanto, el factor principal que subyace implícitamente en las asociaciones que en este experimento se han establecido para la voz *breathy* podría formularse en términos de «bondad».

La voz *pressed*, independientemente de que se combinara con un tono grave o agudo, se ha elegido para tres productos bien diferentes: un parque de ciencias naturales, un parque acuático emocionante y un zumo de frutas. Cuando este tipo de fonación fue preferentemente asociado al parque de ciencias, los sujetos indicaron que esa voz era la más apropiada por transmitir significados relacionados con la «enseñanza» y la «didáctica». Al asociarlo a la voz *pressed*, los sujetos interpretaron el parque de ciencias como una actividad intelectual, de formación, que requiere atención y concentración, en definitiva, esfuerzo o tensión intelectual. Cuando se asoció con el parque acuático, se resaltaron atributos como emocionante, excitante, y tenso. La voz *pressed* asociada al zumo transmitió las sensaciones de que el zumo era algo energético, lleno de vitaminas y enganchado.

Del mismo modo que las múltiples asociaciones hechas para la voz *breathy* han podido quedar reducidas al factor común «bondad», la voz *pressed* podría tener asociado el significado genérico de «vigor».

Por tanto, las respuestas dadas por los sujetos de este experimento parecen indicar una asociación repetida entre los tipos de fonación voz *breathy*/voz *pressed* y los significados generales y básicos de «bondad»/«vigor». Luego, estos significados básicos podrían ser interpretables, en el primer caso, como escalas semánticas en las que ocupan distintas posiciones los calificativos del tipo: beneficioso, generoso, amable, agradable, amoroso, compasivo, sacrificado, tolerante, obediente, inocente, noble... y otros muchos dentro de esta misma escala. Y, en el segundo, como enérgico, esforzado, poderoso, resistente, musculoso, animoso, enfático, activo, osado, violento, dinámico, joven, etc.

Nuevamente, se observa que existen significados primordiales muy genéricos asociados a diferentes voces. Dependiendo del producto u objeto que se someta a juicio, esos significados comunes y principales que subyacen implícitamente a todas las asociaciones se manifestarán en una serie de adjetivos distintos. Puesto que las asociaciones implícitas que realizan los oyentes son procesos complejos, en cierta medida, latentes, subyacentes e inconscientes, los factores comunes del análisis multidimensional normalmente se alejarán mucho de las distintas manifestaciones superficiales.

Por último, cabe señalar que las correspondencias asociativas entre los tipos de fonación y los significados obtenidas en este experimento se encuentran también justificadas en la interpretación «activo-pasivo» mencionada anteriormente en la bibliografía (cf. Sección 2.1). En este trabajo, los jueces del experimento han asociado la voz *breathy* con emociones y sensaciones de «baja activación» (o «pasivas»), siempre de valoración «positiva» (salud, bienestar, placer, sensualidad, tranquilidad, serenidad). En el extremo opuesto, la voz *pressed* se asoció preferentemente con sensaciones y estados mentales de «activación», interpretados contextualmente también en términos «positivos» (concentración, esfuerzo intelectual, atención, tensión, emoción, excitación, energía).

6. CONCLUSIONES DEL ESTUDIO Y FUTURAS LÍNEAS DE INVESTIGACIÓN

En este trabajo se ha estudiado a qué tipo de voces son preferentemente asociados distintos productos, y qué características y atributos exhiben esos productos para haber sido asociados preferentemente con un tipo de voz.

La cualidad de voz proveniente de la combinación de ajustes articulatorios tensos y una F0 baja (*pressed voice* y tono grave) se ha preferido para publicitar productos como el parque de ciencias naturales, cuando la seriedad, densidad, profundidad y concentración fueron considerados atributos positivos. Por su parte, la cualidad de voz que combina ajustes articulatorios tensos y una F0 alta (*pressed voice* y tono agudo) se ha

preferido para publicitar productos como un zumo de frutas y un parque acuático, para los cuales la tensión, acción, energía, emoción, excitación, diversión, entusiasmo fueron considerados atributos positivos. La cualidad de voz que presenta ajustes articulatorios relajados y una F0 baja (*breathy voice* y tono grave) se eligió para publicitar el helado de chocolate cuando se resaltaron como positivos los significados de placer, seducción, tranquilidad, serenidad, cremosidad y succulencia. Por último, la cualidad de voz resultante de ajustes articulatorios relajados y una F0 alta (*breathy voice* y tono agudo) fue escogida para publicitar el zumo cuando se pusieron de relieve significados como *saludable y natural*.

Analizando los resultados asociativos particulares, se ha visto que a los distintos adjetivos, que pueden incluso estar superficialmente muy alejados entre sí, subyacen primitivos semánticos o factores básicos comunes asociados a las características fonéticas de las voces. El tono, correlato perceptivo de la frecuencia fundamental, se ha vinculado con los significados básicos de «formalidad/informalidad», para el tono grave y agudo, respectivamente. La fonación *breathy* ha sido asociada con el significado básico de «bondad», al tiempo que la voz *pressed* se asoció con el significado general de «vigor». Asimismo, a partir de la interpretación de los resultados se han podido establecer patrones asociativos en cierta medida icónicos entre las configuraciones laríngeas tensas o relajadas y las emociones, los estados de ánimo y las sensaciones de «activación» frente a las de «baja activación» o «pasivas».

También se ha podido comprobar en los resultados que el oyente, y en consecuencia el hablante, sigue unas pautas de combinación bien establecidas, que no son aleatorias, para establecer asociaciones perceptivas a partir de una determinada cualidad de voz. Estas asociaciones están doblemente pautadas, por un lado, por las asociaciones establecidas independientemente para el tono y los tipos de fonación, y también, por las asociaciones fijadas por la lengua y la cultura. Dicho de otro modo, la combinación de las asociaciones establecidas de manera independiente para el tono y para la cualidad de voz sigue unas pautas que no son gratuitas, sino que están marcadas lingüísticamente y que son compartidas por el oyente y por el hablante.

Ahora bien, los resultados de este experimento también ponen de manifiesto que, situacionalmente, un mismo tono o un mismo tipo de fonación (e incluso, una misma cualidad de voz) puede suscitar asociaciones diversas. Se ha visto que los significados que se perciben a partir de una misma muestra de voz pueden variar fundamentalmente por dos razones. Por un lado, porque las asociaciones con las diferentes cualidades de voz son un reflejo directo de la distinta importancia o relevancia cognitiva que cada oyente dé a las características de los objetos que se someten a juicio. Y por otro, porque los oyentes atienden perceptivamente a diferentes aspectos de una misma cualidad de voz según los contextos. Es decir, los oyentes no

siempre dan la misma prioridad perceptiva al tono de la voz y a los tipos de fonación, aunque ambos aspectos aparezcan unidos simultáneamente cuando ellos perciben una cualidad de voz.

En cualquier caso, es interesante constatar que, aunque la imagen mental de las cualidades que se valoran o resaltan en cada producto a partir de la cualidad de voz no es exactamente igual en todos los jueces, los estereotipos que sugiere cada voz sí lo son.

En este experimento, cuando los sujetos asociaron significados con un tipo de voz, lo hicieron pensando en un producto en concreto. Por tanto, es posible que haya sido la voz asociada a un producto determinado lo que haya evocado esas asociaciones. Sería interesante estudiar en el futuro las asociaciones de la voz con atributos no vinculados a un producto en particular. Esto permitiría clarificar, por un lado, si las asociaciones fonosimbólicas de las voces con ciertos atributos dependen de la vinculación con un producto determinado y, además, permitiría valorar cuán generales y robustas pueden ser las asociaciones establecidas por este estudio.

Como primer intento por descubrir los significados y sensaciones que subyacen a las asociaciones que los oyentes realizan a partir de la cualidad de voz, este estudio no está exento de limitaciones. La primera de ellas es que los resultados no permiten deslindar la capacidad fonosimbólica de la cualidad de voz del contenido segmental. Es decir, este estudio no examina en qué medida las características propias de los sonidos (vocales y consonantes) y el léxico pueden estar influyendo en la provocación de las asociaciones. Otra limitación se deriva de no aclarar hasta qué punto las asociaciones que se han establecido entre voz y significados son exclusivas del español. Es evidente que se necesita mucha más investigación para dilucidar cuestiones como estas. Algunas posibles preguntas para avanzar más en esta temática podrían ser:

- ¿Qué es lo que hace que un oyente experimente asociaciones de tipo fonosimbólico? ¿Es suficiente la cualidad de la voz, en su vertiente laríngea –tipos de fonación y F0 global– para provocar y transmitir asociaciones?
- ¿Es posible identificar algún tipo de vinculación entre cualidad de voz y significados que pueda generalizarse a varias lenguas, del mismo modo que la música o la entonación suscitan muchas veces reacciones asociativas universales?

Por último, resulta también interesante señalar que el hecho de que determinadas características sonoras o acústicas de las voces despierten asociaciones repetidas y sistematizables entre los oyentes implica postular que existe un vínculo icónico en algún grado, no totalmente arbitrario, entre la forma lingüística y el sentido que transporta, en concreto, entre sonido y significado. Este hecho, no solo debate el alcance del principio

saussureano de la arbitrariedad de la lengua (Jespersen 1922; Sapir 1929; Benveniste 1939; Bolinger 1964; Jakobson 1968, entre otros) sino que además contribuye a la discusión teórica de la existencia de iconicidad al menos, a nivel fónico (Fischer 1999) y de la posible universalidad del uso de ciertos sonidos para vehicular significados específicos (cf. por ejemplo, Darwin 1872).

7. REFERENCIAS BIBLIOGRÁFICAS

- BELE, I.V. (2005): «Reliability in perceptual analysis of voice quality», *Journal of Voice* 19, pp. 555-573.
- BENVENISTE, E. (1939): «Naturaleza del signo lingüístico», *Problemas de lingüística general*, México, Siglo XXI, 1971.
- BOLINGER, D. (1964): «Intonation across languages», en Greenberg, J., Ferguson, C. y Moravcsik, E. (eds.), *Universals of human language*, Vol. 2, *Phonology*, Stanford, Stanford University Press, pp. 471-524.
- BURKHARDT, F. y SENDLMEIER, W.F. (2000): «Verification of acoustical correlates of emotional speech using formant-synthesis», en Cowie, R., Douglas-Cosie, E. y Schröder, M. (eds.), *Proceedings of the ISCA workshop on speech and emotion: A conceptual framework for research*, Belfast, Queen's University, pp. 151-156.
- CATFORD, J.C. (1977): *Fundamental problems in phonetics*, Edimburgo, Edinburgh University Press.
- DARWIN, C. (1984): *La expresión de las emociones en los animales y el hombre*, Madrid, Alianza [1872].
- FISCHER, A. (1999): «What if anything, is phonological iconicity?», en Nänny, M. y Fischer, O. (eds.), *Form, miming meaning: Iconicity in language and literature*, Filadelfia, John Benjamins.
- FONAGY, I. (1978): «A new method of investigating the perception of prosodic features», *Language and Speech* 21, pp. 34-49.
- FRICK, R.W. (1985): «Communicating emotion: The role of prosodic features», *Psychological Bulletin* 97, 3, pp. 412-429.
- GIL FERNÁNDEZ, J. 2007: *Fonética para profesores de español: de la teoría a la práctica*, Madrid, Arco/Libros.
- GOBL, C. y NÍ CHASAIDE, A. (2003): «The role of voice quality in communicating emotion, mood and attitude», *Speech Communication* 40, pp. 189-212.
- GORDON, M., y LADEFOGED, P. (2001): «Phonation types: a cross-linguistic overview», *Journal of Phonetics* 29, 4, pp. 383-406.
- HINTON, L., NICHOLS, J. y OHALA, J. (eds.) (1994): *Sound symbolism*, Cambridge, Cambridge University Press.
- HURON, D., KINNEY, D., PRECODA, K. (2006): «Influence of pitch height on the perception of submissiveness and threat in musical passages», *Empirical Musicology Review* 1, 3, pp. 170-177.
- JAKOBSON, R. (1968): *Child language, aphasia and phonological universals*, La Haya, Mouton.
- , y WAUGH, L.R. (1979): «El encanto de los sonidos del habla», *La forma sonora de la lengua*, México, Fondo de Cultura Económica, 1987, pp. 172-222.

- JESPERSEN, O. (1922): «Sound symbolism», *Language, its nature, development and origin*, Londres, George Allen, pp. 396-411.
- KREIMAN, J. y SIDTIS, D. (2011): *Foundations of voice studies. An interdisciplinary approach to voice production and perception*, Malden, Wiley-Blackwell.
- KREIMAN, J., GERRATT, B.R., GARELLEK, M., SAMLAN, R., y ZHANG, Z. (2014): «Toward a unified theory of voice production and perception», *Loquens* 1, 1, e009. doi: <http://dx.doi.org/10.3989/loquens.2014.009>
- LAUKKANEN, A.M., VILKMAN, E., ALKU, P., y OKSANEN, H. (1996): «Physical variation related to stress and emotionally state: a preliminary study», *Journal of Phonetics* 24, pp. 313-335.
- LAVER, J. (1968): «Voice quality and indexical information», *International Journal of Language & Communication Disorders* 3, 1, pp. 43-54.
- , (1980): *The phonetic description of voice quality*, Cambridge, Cambridge University Press.
- LEINO, T. (2009): «Long-term average spectrum in screening of voice quality in speech: Untrained Male University Students», *Journal of Voice* 23, pp. 671-676.
- LIEBERMAN, P. (1967): *Intonation, perception, and language*, Cambridge, MA, MIT Press.
- LINDSLEY, D.B. (1951): «Emotion», en Stevens, S.S. (ed.), *Handbook of experimental psychology*, Nueva York, Wiley, pp. 437-516.
- MOOS, A., SMITH, R. y SIMMONS, D. (2011): «Voice quality perceptions by synaesthetes, phoneticians and controls», *Proceedings of the 17th International congress of phonetic sciences (ICPhS)*, pp. 1426-1429.
- MOOS, A., SIMMONS, D., SIMNER, J. y SMITH, R. (2013): «Colour and texture associations in voice-induced synaesthesia», *Frontiers in Psychology* 4, 568, pp 1-12.
- MORTON, E. (1994): «Sound symbolism and its role in non-human vertebrate communication», en Hinton, L., Nichols, J. y Ohala, J. (eds.), *Sound symbolism*, Cambridge, Cambridge University Press, pp. 348-365.
- NÍ CHASAIDE, A. y GOBL, C. (2005): «On the relation between phonatory quality and affect», en Hardcastle, W. y Mackenzie Beck, J. (eds.), *A figure of speech: A festschrift for John Laver*, Mahwah, Lawrence Erlbaum pp. 323-346.
- OHALA, J.J. (1983): «Cross-language use of pitch: An ethological view», *Phonetica* 40, pp. 1-18.
- , (1984): «An ethological perspective on common cross-language utilization of F0 of voice», *Phonetica* 41, pp. 1-16.
- PLACK, C.J. (2005): *The sense of hearing*, Nueva York, Lawrence Erlbaum.
- POYATOS, F. (1994): *La comunicación no verbal*, Madrid, Istmo.
- REICH, A.R. y LERMAN, J.W. (1978): «Teflon laryngoplasty: an acoustical and perceptual study», *Journal of Speech and Hearing Disorders* 43, pp. 496-505.
- SAPIR, E. (1929): «A study in phonetic symbolism», *Journal of Experimental Psychology* 12, 3, pp. 225-239.
- SCHERER, K.R. (1974): «Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences», en Weitz, S. (ed.), *Non-verbal communication*, Nueva York, Oxford University Press, pp. 105-111.
- , (1986): «Vocal affect expression: A review and a model for future research», *Psychological Bulletin* 99, pp. 143-165.

- SCHERER, K.R. y OSHINSKY, J.S. (1977): «Cue utilization in emotion attribution from auditory stimuli», *Motivation and Emotion* 1, pp. 331-346.
- SPENCE, C. (2012): «Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism», *Journal of Consumer Psychology* 22, pp. 37-54.
- SUNDBERG, J., TITZE, I.R. y SCHERER, R. (1993): «Phonatory control in male singers: a study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source», *Journal of Voice* 7, pp. 15-29.
- SUNDBERG, J. (1994): «Vocal fold vibration patterns and phonatory modes», *STL-QPSR*, Estocolmo, Royal Institute of Technology.
- TSUR, R. (2006): «Size-sound symbolism revisited», *Journal of Pragmatics* 38, pp. 905-924.
- YORKSTON, E.A. y MENON, G. (2004): «A sound idea: Phonetic effects of brand names on consumer judgements», *Journal of Consumer Research* 31, pp. 43-51.

Edita
SeL

