

Forensic speaker comparison of Spanish twins and non-twin siblings: A phonetic-acoustic analysis of formant trajectories in vocalic sequences, glottal source parameters and cepstral characteristics

Tribunal:

Dr. Pedro Gómez Vilda (Universidad Politécnica de Madrid)

Dr. Daniel Ramos Castro (Universidad Autónoma de Madrid)

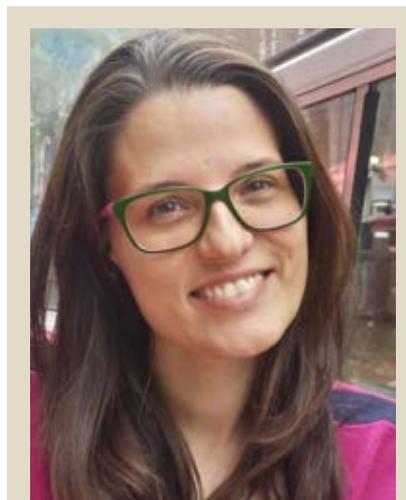
Dra. Lourdes Aguilar Cuevas (Universidad Autónoma de Barcelona)

Dra. Elena Battaner (Universidad Rey Juan Carlos)

Dr. Volker Dellwo (University of Zurich).

Resumen:

El objetivo de esta tesis ha sido estudiar las características fonético-acústicas de tres grupos principales de hablantes: gemelos monocigóticos, gemelos dicigóticos y hermanos no gemelares. Desde la perspectiva de la fonética judicial, el análisis del habla y de la voz de este tipo de hablantes es especialmente relevante por su extrema similitud física. Por tanto, distinguir las voces de estos hablantes supone un reto bien conocido en esta disciplina científica, como se desprende de la revisión bibliográfica llevada a cabo (cf. Capítulo 2). No obstante, existe un interés *per se* en este tipo de investigación, más allá de su aplicación forense. El estudio de hablantes genéticamente idénticos –como lo son los gemelos monocigóticos– y su comparación, por un lado, con hermanos que no son idénticos genéticamente –esto es, los gemelos dicigóticos y los hermanos no gemelares– y, por otro lado, con una población de re-



Autora:

Eugenia San Segundo Fernández

Director:

Dr. Joaquim Llisterrí Boix

Co-Directora:

Dra. Juana Gil Fernández

Exposición:

15 de diciembre de 2014

Lugar:

Centro de Ciencias Humanas y Sociales
del Consejo Superior de
Investigaciones Científicas (CSIC)

e-mail:

eugeniasansegundo@gmail.com

ferencia de hablantes sin parentesco, nos ha permitido arrojar cierta luz sobre una cuestión aún poco explorada; a saber: la contribución relativa de lo *innato* y de lo *adquirido* en las características vocales del ser humano. En otras palabras, ¿hasta qué punto nuestra voz está determinada por nuestro ADN y hasta qué punto influyen en ella

factores externos, como las influencias educativas o del entorno? Además, este estudio puede considerarse la primera investigación sobre las características fonéticas y acústicas de gemelos y hermanos hispanohablantes. Por ello, las voces de un total de cincuenta y cuatro hablantes han sido grabadas *ad hoc* para este estudio. De acuerdo con nuestros objetivos de investigación, hemos adoptado un enfoque centrado en tres análisis diferentes. Por un lado, hemos etiquetado y analizado las trayectorias formánticas de los tres primeros formantes (F1, F2 y F3) de diecinueve secuencias vocálicas del español (hiatos y diptongos). En segundo lugar, hemos extraído varias muestras del sonido [e:], sostenido de forma natural en las pausas llenas realizadas por los hablantes de forma espontánea y posteriormente hemos analizado ciertas características de la fuente glótica a partir de dicha vocal sostenida. Estos dos enfoques se han complementado utilizando el sistema de reconocimiento automático de hablantes *Batvox*.

Esta tesis está dividida en siete capítulos de la forma que explicamos a continuación. En el primer capítulo, a modo de introducción, se proporciona una definición de la «fonética judicial». Asimismo, se describen las principales tareas o aplicaciones de esta disciplina, incidiendo en la comparación forense de hablantes, y se señalan algunas controversias terminológicas que nos permiten justificar nuestra prefe-

rencia por el uso de «comparación forense de hablantes» frente a otros posibles términos. Por otro lado, en este primer capítulo describimos los principales enfoques metodológicos que existen actualmente en este campo, con un énfasis especial en las ventajas e inconvenientes de cada uno de ellos. Esta exposición del estado de la cuestión tiene como principal objetivo explicar la necesidad de adoptar un enfoque híbrido que combine los llamados parámetros fonéticos tradicionales con otros rasgos y métodos que hasta el momento solamente encontrábamos de forma mayoritaria dentro del ámbito del reconocimiento automático. Nos referimos, por ejemplo, a la caracterización biométrica mediante parámetros cepstrales o al uso de poblaciones de referencia.

El capítulo dos está dedicado a la revisión bibliográfica e incluye, por tanto, un análisis crítico de los estudios fonéticos existentes hasta la fecha que han abordado la cuestión de las voces gemelares, ya sea desde un punto de vista perceptivo, acústico o articulatorio, o bien desde un enfoque basado en sistemas automáticos de reconocimiento. Esta revisión está precedida de una introducción a las bases biológicas gemelares y al método gemelar, así como de una sucinta subsección dedicada a la relevancia forense de las voces de gemelos. A continuación, en el capítulo tres detallamos algunas cuestiones metodológicas relevantes para esta tesis. Esto incluye la descripción de las principales características –por ejemplo, edad y dialecto– de los distintos sujetos que han participado: 24 gemelos monocigóticos, 10 gemelos dicigóticos, 8 hermanos no gemelares y 12 hablantes no emparentados. Estos últimos constituyen la principal población de referencia, que aumentó en tamaño gracias al uso de técnicas de validación cruzada. Por

otro lado, se detalla el diseño y procedimiento de recogida del corpus *ad hoc*, que incluye cuatro tareas (conversación espontánea semidirigida entre las parejas de hablantes; intercambio de información para elicitación de las secuencias vocálicas; lectura de dos textos fonéticamente equilibrados y entrevista informal con el investigador), además de una técnica de control vocal. Este tercer capítulo incluye también detalles relacionados con el procedimiento de obtención de muestras sonoras, así como la descripción del material de grabación y sus características técnicas. Hemos dedicado una sección aparte a la explicación del filtrado telefónico aplicado sobre las grabaciones y, finalmente, hemos descrito el enfoque metodológico basado en relaciones de verosimilitud –en inglés *Likelihood Ratios, LRs*– pues dentro de este marco se inscriben los análisis llevados a cabo para la tesis. Es decir, los resultados de las distintas comparaciones entre hablantes se ofrecen en relaciones de verosimilitud, que nos informan no sólo sobre el grado de *similitud* entre las muestras de voz comparadas sino también sobre el grado de *tipicidad* con respecto a la población relevante.

Los capítulos cuatro, cinco y seis, respectivamente, incluyen toda la información relacionada con los tres tipos de análisis que se han realizado para esta tesis: análisis de trayectorias formánticas, análisis de la fuente glótica y análisis automático. Esto comprende, en primer lugar, la descripción de la metodología específica para cada análisis (por ejemplo, en el capítulo cuatro, el ajuste de las curvas de las trayectorias formánticas mediante funciones polinomiales y transformadas de coseno discretas) y, en segundo lugar, la exposición de los resultados y su posterior discusión. Finalmente, en el capítulo siete se ofrece un resumen de los resultados obtenidos

en los análisis previos y de las principales conclusiones extraídas a partir de ellos. El capítulo se cierra con la discusión de las implicaciones de esta tesis; asimismo se apuntan algunas direcciones para posibles trabajos futuros.

En lo que se refiere a los resultados, los parámetros analizados bajo los distintos enfoques metodológicos estarían en gran medida condicionados genéticamente puesto que se observa una escala decreciente en la siguiente *dirección* $MZ > DZ > \text{hermanos no gemelares} > \text{hablantes sin parentesco}$. En otras palabras, los valores de LRs que resultan de las comparaciones entre parejas de monocigóticos son mayores que los obtenidos al comparar parejas de dicigóticos; estos a su vez obtienen valores de LRs mayores que los hermanos no gemelares y, finalmente, los LRs de estos últimos están por encima de la media obtenida al comparar dos hablantes sin ningún tipo de parentesco. Los escasos resultados discordantes –para algunas comparaciones intrapareja concretas– se han discutido de forma pormenorizada y mayormente se han podido explicar por causas físicas como la existencia de ciertas patologías que podrían afectar a la voz, o bien por causas externas vinculadas a las idiosincrasias de las interrelaciones gemelares, que podrían haber ocasionado patrones de divergencia fonética frente a los casos más típicos de acomodación o mimetismo intragemelar. Pese a que estos aspectos de carácter sociolingüístico merecen ser investigados de manera más exhaustiva, la tendencia general observada en los resultados de esta tesis apunta a que las características de la voz que están más influenciadas por la genética que por aspectos ambientales deberían ser parámetros robustos para la comparación forense de voces.

Abstract:

The objective of this thesis has been to investigate the phonetic-acoustic similarities and differences in three main speaker groups: monozygotic (MZ) twins, dizygotic (DZ) twins and non-twin siblings. From a forensic-phonetic perspective, the study of this type of speakers is highly relevant, as they represent extreme examples of physical similarity. Distinguishing their voices poses a well-recognized challenge in the forensic realm. Yet, there is an interest in this investigation per se, as the study of genetically identical speakers (MZ twins) and their comparison with non-genetically-identical siblings (DZ twins and non-twin siblings), on the one hand, and with a reference population of unrelated speakers, on the other hand, allows gaining insight into the contribution of nurture and nature in the speech patterns of speakers in general. In other words: to what extent is our voice determined by our DNA and to what extent is it due to educational influences?

Besides, this study could be considered the first investigation into the phonetic and acoustic characteristics of Spanish-speaking twins and siblings. According to our research objective, a three-folded approach has been undertaken, as will be described below. For the 54 male speakers recorded ad hoc for this study, three different analyses have been carried out. On the one hand, we have labeled and analyzed the F1-F3 formant trajectories of 19 Spanish vocalic sequences. Secondly, several naturally sustained [e:] tokens have been extracted from the speakers' spontaneous vowel fillers and their glottal source characteristics have been analyzed. These two approaches have been complemented with an automatic speaker recognition analysis carried out with the software *Batvox*.