



UNIVERSITAT DE
BARCELONA

Rueda de reconocimiento y Análisis acústico de voces femeninas en español

Treball de Fi de Grau

Autor: Carolina Fabra Navarro

Director del treball: Ana María Fernández Planas

Tutor del curs: Xavier Laborda

Grau de Lingüística

Facultat de Filologia

Universitat de Barcelona

Curs 2015-2016

Barcelona, 3 de juny de 2016.

Índice

1.	RESUMEN	3
2.	INTRODUCCIÓN	5
2.1.	Lingüística forense y Fonética forense	5
2.2.	Fonética perceptiva	6
2.3.	Rueda de reconocimiento	9
2.4.	Elementos segmentales y suprasegmentales	10
3.	MOTIVACIÓN	13
4.	OBJETIVOS	14
5.	HIPOTESIS:	14
6.	METODOLOGÍA	15
6.1	Corpus e informantes	15
6.2	Jueces	16
6.1.4.	Rueda de reconocimiento	21
6.1.5.	Análisis acústico	22
6.1.6.	Análisis estadístico	23
6.1.7.	Inconvenientes	23
7.	RESULTADOS.....	24
7.1.	Rueda de reconocimiento	24
7.2.	Análisis acústico de las voces	52
8.	DISCUSIÓN Y CONCLUSIÓN.....	63
9.	REFERENCIAS BIBLIOGRÁFICAS	64
10.	ANEXO	66
10.1.	Tablas de Edad de los jueces.	66
10.2.	Orden voces test perceptivos.	66
10.3.	Test 1.....	67
10.4.	Test 2.....	73
10.5.	Test 3.....	79

1. RESUMEN

Este estudio presenta el resultado de una Rueda de reconocimiento perceptiva. Para la realización de este trabajo se han creado tres test perceptivos online y a continuación se han justificado las respuestas de los jueces mediante un análisis acústico. Los resultados de los test confirman las hipótesis planteadas. Por una parte, resulta más fácil identificar una voz, que se ha escuchado previamente en registro espontáneo, entre un conjunto de voces en registro formal, aproximadamente un 80% de los jueces reconoció la voz indubitada frente a un 20% que discreparon. Por otra parte, ha resultado más difícil la identificación entre un conjunto de voces en discurso espontáneo de una voz en registro formal, sólo un 8% de los jueces identificó correctamente la voz indubitada. Por último, no ha resultado imposible la identificación de una voz en registro espontáneo enmascarada con ruido *party* entre el conjunto de voces sin enmascarar en registro formal, un 40% de los jueces fue capaz de identificar correctamente la voz.

Palabras clave: *lingüística forense, fonética acústica, rueda de reconocimiento, análisis acústico.*

RESUM

Aquest estudi mostra el resultat d'una Roda de reconeixement perceptiva. Per dur a terme aquest treball s'han creat tres test perceptius online i, a continuació, s'han justificat les respostes dels jutges mitjançant una anàlisi acústica. Els resultats del test confirmen les hipòtesis plantejades. D'una banda, resulta més fàcil identificar una veu, que s'ha escoltat prèviament en registre espontani, entre un conjunt de veus en registre formal, aproximadament un 80% dels jutges van reconèixer la veu indubitada mentre que només un 20% va discrepar. D'una altra banda, ha resultat més difícil la identificació entre un conjunt de veus en discurs espontani d'una veu en registre formal, només un 8% dels jutges va identificar correctament la veu corresponent. Per últim, no ha resultat impossible la identificació d'una veu en registre espontani emmascarada amb soroll *party* entre el conjunt de veus sense emmascarar en registre formal, un 40% dels jutges va ser capaç d'identificar correctament la veu.

Paraules clau: *lingüística forense, fonètica acústica, roda de reconeixement, anàlisi acústica.*

ABSTRACT

This study presents the results of a Perceptual Line-up. To carry out this study, we created three online perceptual tests and afterwards the judges' responses were justified by an acoustic analysis. The results of the tests confirm the hypotheses. On the one hand, it is easier to identify a voice, previously heard in spontaneous register, of a set of voices in formal register; approximately 80% of the judges recognized the correct voice compared to 20% who didn't. On the other hand, it is more difficult the identification of a voice, previously heard in formal register, of a set of voices in spontaneous speech; only 8% of judges identified the correct voice. Finally, it has not been impossible to identify a voice in spontaneous register, masked with party noise, of a set of voices in formal register without masking; 40% of the judges were able to correctly identify the voice.

Key words: *forensic linguistic, acoustic phonetics, perceptual line-up, acoustic analysis.*

2. INTRODUCCIÓN

2.1.Lingüística forense y Fonética forense

La lingüística forense es la aplicación de la lingüística que la relaciona con el ámbito jurídico. Su objetivo es el análisis de cualquier expresión del lenguaje –escrita u oral– para corroborar o descartar su autoría. Dentro de la lingüística forense encontramos diferentes áreas; fonética forense y reconocimiento de hablantes, discurso en “lenguaje” legal, autoría, plagio, traducción e interpretación. El presente trabajo se centra en el ámbito de la fonética forense y la identificación perceptiva del hablante.

Una de las tareas más comunes en la verificación forense del hablante es la comparación de una o más muestras de voces desconocidas con una o más muestras de voces conocidas. Para verificar si la misma voz está presente en dos o más muestras de habla, tiene que ser posible discriminar diferentes muestras de hablantes diferentes – pese a que la voz de un mismo hablante varía siempre puesto que nunca se produce el mismo sonido de manera idéntica. Es por esto que se debe de hablar sobre la variación intra-hablante, es decir, la variación que se produce en un mismo hablante, y la variación inter hablante, la variación que se produce entre diferentes hablantes. Tanto la variación intra- como inter-hablante se debe al hecho que la información que dan los sonidos del habla no sólo incluyen información lingüística, sino que también incluye información sobre el hablante. Por tanto, lo que hace que los sonidos que produce un hablante sean característicos es, por una parte, los factores internos como las propiedades anatómicas, y por otra parte, los factores externos tales como los factores sociales, geográficos e incluso los individuales.

Actualmente en España existen diferentes dificultades a la hora de expandir la disciplina de la lingüística forense. La naturaleza de nuestro sistema judicial se basa en la figura del juez y no tanto en el tribunal. En otros países como Reino Unido, la investigación policial suele apoyarse en los dictámenes periciales realizados por laboratorios independientes, por el contrario, en España, es la misma policía quien realiza estos peritajes y pocas veces –algunos casos ha habido– este peritaje ha llegado a laboratorios forenses universitarios. Existe una escasa presencia de peritos lingüistas en las salas de vistas de nuestro país, seguramente debido a la historia de nuestra cultura jurídica y, también, al desconocimiento por parte de los actores judiciales de la labora

que puede llevar a cabo estos peritos. Pese al desarrollo lento que existe en esta disciplina, el ámbito de la fonética y acústica forense están mucho más relacionados puesto que cuenta con información sobre la Referencia de Distribución Poblacional respecto a determinados parámetros de la voz.

2.2.Fonética perceptiva

La fonética perceptiva es el ámbito de la fonética que estudia la señal que entra en el oído del receptor para convertir dicha señal en mensaje lingüístico. La percepción es un ejercicio subjetivo; el material de estudio, por tanto, está formado por intuiciones y sensaciones que dependen de una mente individual. Así pues, entendida como ciencia multidisciplinar es una ciencia joven puesto que avanza paralelamente a una tecnología científica de la que depende. El objetivo principal de la fonética perceptiva es la descodificación de los sonidos lingüísticos, que permiten, como sistema de elementos, la comunicación oral. En la descodificación lingüística encontramos diferentes tareas y etapas del mensaje hablado. Por una parte, hay aspectos que dificultan la descodificación, como la variabilidad, la falta de correspondencia o los rasgos físicos de la onda sonora. La variabilidad corresponde a la falta de relación entre la señal física y la lingüística, como el contexto, las características del locutor y el ritmo del habla. Sin embargo, por otra parte también hay aspectos que facilitan la descodificación; claves acústicas y redundancia. Como clave acústica podríamos poner de ejemplo el acento, puesto que la frecuencia fundamental (F0), la duración y la intensidad juegan un papel importante en fonética perceptiva. La solución para los ruidos que afectan la descodificación -como el la articulación del emisor, las características del receptor y del código o las alteraciones en el canal- es la redundancia. La redundancia, en este caso, entendida como aquella información de más no necesaria que entrega el mensaje. Esto ocurre en todos los niveles del lenguaje tanto en el fonológico, morfosintáctico como semántico. Por ejemplo, en el nivel fonológico todos descodificaríamos **per*o** o **una*niñ** compr****. Hay dos factores más que influyen en la descodificación: a nivel de producción la inteligibilidad, y a nivel perceptivo la discriminabilidad.

De la misma manera que es necesario conocer los órganos articulatorios para la descripción articulatoria, o conocer conceptos básicos de física de ondas para la

descripción acústica, en fonética perceptiva también se requiere una descripción del sistema auditivo y, sobre todo, del funcionamiento de la corteza auditiva. A continuación, comentaré de forma esquemática el sistema auditivo y la participación de la corteza auditiva puesto que se trata de un ámbito mucho más extenso que no es necesario para el presente trabajo.

El sistema auditivo está compuesto por el oído, las vías auditivas, el nervio auditivo y la corteza cerebral (área auditiva). En el oído diferenciamos tres partes: el oído externo, el oído medio y el oído interno. El oído externo se encarga de codificar los cambios de frecuencia, intensidad y tiempo de la onda sonora aumentando ligeramente su intensidad. El oído medio recupera la intensidad que se pierde en el paso de ondas aéreas a líquidas o amortigua esta intensidad si es demasiado alta. El oído interno, por su parte, es el encargado de transmitir la señal acústica hasta la corteza cerebral, es decir, transforma las ondas en impulsos nerviosos que son enviados al cerebro. Las vías auditivas llevan la información de la cóclea hasta la corteza cerebral –pasa la información de sonidos a fonemas-, y la área auditiva es aquella zona especializada en la descodificación del sonido, tanto a nivel lingüístico como musical, etc. El nervio auditivo corresponde al VIII par craneal y existen fenómenos que afectan el funcionamiento de sus fibras: adaptación (cuando un estímulo es muy persistente y la fibra se “cansa” y deja de responder a partir de los primeros 15-20 ms.) y el enmascaramiento de un componente de la onda sobre otro (si hay un tono no ocurre nada pero cuando existen dos tonos a la vez estos interfieren entre sí y hacen la descodificación más complicada). Las diferentes funciones de la corteza cerebral son el análisis de los sonidos complejos, la localización del sonido, la discriminación de los patrones temporales y la memoria a corto plazo.

Para la descodificación del mensaje participan tres tipos de memoria: la memoria a corto plazo que dura de 2 a 3 segundos y es la que retiene las sílabas y esta información posteriormente va a la memoria a largo plazo, esta a su vez admite hasta 7 sílabas, se quedan en este punto y cuando se decide qué hacer con la información pueden ir al almacén léxico. Por último, la memoria semántica es la que permanece, donde se quedan las unidades léxicas. En psicoacústica –ciencia que estudia la psicología y las respuestas psicológicas asociadas con el sonido- existen cuatro etapas. La primera etapa es la detección, es saber si está presente o ausente un sonido. La

segunda etapa consiste en discriminar, buscar y comparar las diferencias entre los estímulos. La tercera etapa es la identificación, relaciona el estímulo que se presenta con la memoria. Por último, la cuarta etapa es la comprensión.

Tal y como explica Fernández Planas ([2005] 2011) el ser humano es capaz de percibir una banda de frecuencias muy amplia que va desde los 16Hz hasta los 20.000Hz. Esta sería la gama frecuencial que puede percibir el oído humano, ya que no podemos percibir ni infrasonidos (por debajo de los 16Hz) ni ultrasonidos (por encima de los 20.000Hz).

Existen dos tipos de percepción del sonido: percepción categorial y percepción continua. La primera consiste en la percepción de las diferencias de los estímulos y en su etiquetación. Además se categorizan los elementos según las propiedades que comparten, como el rasgo físico, el tiempo, la frecuencia, etc. La percepción continua por su parte es la percepción de las diferencias de los estímulos en la coarticulación del habla. Por ejemplo, en las vocales, cuando de /a/ se pasa a /o/ varía nuestra percepción. En las consonantes –cuyas diferencias son graduales- se perciben de una forma abrupta y no gradual como las vocales. Estas variaciones corresponden a la frontera fonológica. El cambio de una consonante o una vocal se percibirá cuando se alcance el punto crítico, es decir, cuando se deja de escuchar /a/ por /i/ o /b/ por /d/. La mayoría de estos estudios se realizaron mediante el análisis por síntesis (Halle Stevens). Este establece que el oyente produce inconscientemente una versión sintética de la señal. Si este modelo simplificado coincide con alguno previamente almacenado, se da por finalizado el proceso de identificación

Existen dos tipos de teorías sobre la percepción del habla: teorías activas y teorías pasivas. Por una parte, las teorías activas defienden que el oyente realiza una comparación con los modelos articulatorios que posee como hablante, cuando percibe un mensaje. Por el contrario, las teorías pasivas consideran que las ondas sonoras poseen índices invariantes que permiten la interpretación del mensaje y que, por tanto, al contrario que las teorías activas, no implican en el proceso de percepción el mecanismo articulatorio. Dentro de las teorías activas se encuentran diferentes teorías: la teoría motora y la teoría de análisis-por-síntesis, por ejemplo. La teoría motriz (Liberman Mattingly, 1985) establece que todo hablante y oyente dispone de patrones articulatorios que utiliza para la producción de sonidos o para la percepción. Para identificar perceptivamente un estímulo se tiene que identificar articulatoriamente de

manera previa, además, tanto la producción del habla como la percepción de ésta son capacidades motoras. La teoría del análisis-por-síntesis creada por Stevens y sus compañeros en el MIT (EE.UU.) “*asume la invariación acústica y expone que el oyente realiza diversos pasos en el proceso de percepción: analiza la señal acústica para extraer los índices invariantes y los compara con las instrucciones articulatorias que se necesitan para sintetizarlos. Si la comparación es satisfactoria se produce el reconocimiento*” (Fernández Planas, [2005] 2011: 126).

En los sonidos del habla, desde un punto de vista perceptivo, destacan tres cualidades: el tono, la sonoridad perceptiva y el timbre. Desde este punto de vista el tono se refiere a nuestra percepción del primer armónico de un sonido complejo. Las voces masculinas tienen un tono fundamental medio alrededor de 100 o 120 Hz mientras que el patrón femenino suele rondar entre los 180 y 220 Hz. La sonoridad perceptiva hace referencia a la percepción de la intensidad de un sonido –amplitud media global-. El nivel de sonoridad perceptiva presenta su propia unidad de medida: el fon (aunque la sonoridad perceptiva en sí misma la mide el son). Por último, la cualidad perceptiva del timbre hace referencia a “*la percepción de la particular distribución frecuencial de los formantes o de las resonancias de los diferentes sonidos*” (Fernández Planas [2005] 2011: 123).

2.3. Rueda de reconocimiento

En fonética perceptiva y dentro del ámbito de la lingüística forense, una de las pruebas más comunes es la rueda de reconocimiento. Una rueda de reconocimiento auditiva funciona del mismo modo que las ruedas de reconocimiento visual, ambas ocurren cuando existe un agresor que debe de ser identificado y que ha dejado una huella auditiva en la víctima. Aunque parezca muy obvio, es un procedimiento que se lleva a cabo cuando un testigo o una víctima ha escuchado la voz del acusado sin verle. Sin embargo, una rueda de reconocimiento de voz es una técnica compleja en la cual su éxito o fracaso depende del diseño realizado.

Para empezar, una rueda de reconocimiento puede realizarse mediante grabaciones, entre las cuales se presenta la voz del sospechoso junto con otras voces similares para que el testigo o la víctima puedan identificarla. Según H. Hollien (2002:

91) la selección de las voces de contraste o distractores debe ser entre 5 y 8. Los locutores deben de ser similares en edad, con variedad geográfica y variedad de acento y un estatus socio-económico y cultural parecido. Al menos una de las voces que se presenta debe ser parecida pero también debe de haber una voz bastante diferente a la del sospechoso.

Hay varios factores que influyen y condicionan una rueda de voz. Por una parte, si existe familiaridad entre los testigos y el sospechoso la rueda estará condicionada. Por otra, existe a veces la necesidad o la expectativa de los testigos de que la voz del sospechoso se encuentre entre las voces de la rueda. Además, podría ser que el sospechoso modificara la voz en el momento en el que cometió el delito que se le imputa. Del mismo modo que una rueda de reconocimiento visual también afecta mucho el estado emocional de los testigos –normalmente las ruedas de reconocimiento se hacen al poco tiempo de cometerse el acto “criminal”-. No podemos olvidar tampoco, que todo este procedimiento estará influido por la memoria auditiva de los testigos (o testigo).

Existen tres tipos de ruedas de voces; el primero consiste en la escucha de todas las voces consecutivamente y una única vez; el segundo consiste en la escucha de voces consecutivamente varias veces pero con el orden de aparición distinto; el tercer tipo, el que coincide con el objetivo de este trabajo, es la escucha una por una de cada voz, en el orden que solicite el testigo y con la posibilidad de escuchar más de una vez la misma grabación.

2.4.Elementos segmentales y suprasegmentales

Para un análisis acústico conciso de una o varias voces se comparan diferentes elementos acústicos, tanto segmentales como suprasegmentales. Los elementos segmentales son aquellos que producimos y adaptamos en la cadena fónica: vocales, oclusivas, fricativas, africadas, nasales, laterales, róticas, etc. La información suprasegmental es aquella que no se encuentra en un segmento determinado sino que afecta a varios y nos da información de matiz semántica. Los rasgos segmentales o prosódicos son el acento, el tono o entonación y el timbre.

El acento, desde una perspectiva fonética, es aquel que *recae sobre una sílaba y la destaca de las demás de una palabra* (Fernández Planas [2005] 2011:53). En las palabras encontramos un acento principal que es aquel que posee la sílaba tónica, por otra parte, en palabras compuestas y derivadas –por la adjunción de un prefijo o sufijo- se encuentra también un acento secundario, menos relevante, que recae frente a las sílabas átonas intermedias (léntamènte). Es importante también considerar el acento de frase –último acento principal de la frase- “*ya que supone entender la frase como un todo organizado y no como una lista de palabras aisladas concatenadas*” (Fernández Planas [2005] 2011:53). Respecto al acento principal de las palabras podemos encontrar lenguas de acento libre, como el inglés, en las que la aparición en una sílaba u otra es imprescindible y significativo. También hay lenguas de acento fijo o previsible, como el francés, donde el acento suele recaer siempre en una posición silábica concreta (en el caso del francés la mayoría de palabras son agudas). Cabe destacar que en ninguna lengua presenta que el acento sea en todos los casos libre o siempre fijo, por así decirlo, hay un tipo de acento –fijo o libre- que es mayoritario en una lengua.

El tono es aquel elemento suprasegmental que depende de la fisiología de las cuerdas vocales. En hombres la longitud de las cuerdas vocales es entre 1,7 y 2,4 cm, mientras que las de la mujer son más cortas, entre 1,4 y 1,7cm, además las cuerdas vocales masculinas tienen más grosor que las femeninas. Es por ello que las voces masculinas son más graves que las femeninas. Los niños tienen las cuerdas vocales más cortas y estrechas que las de mujeres adultas, por ello, no existe diferencia entre una voz de niño y una de niña hasta que no se experimenta el “cambio de voz” que consiste en el repentino crecimiento de la laringe y, por tanto, de las cuerdas vocales. El tono es el único rasgo que el hablante puede controlar a su voluntad (es lo que ocurre cuando alguien canta, que tensa, contrae o relaja los músculos para conseguir llegar al tono de la canción). En fonética el término tesitura de voz se ha adaptado del campo de la música, este se refiere a aquella gama frecuencial en la que el hablante está cómodo sin forzar la voz para conseguir un determinado tono, ya sea más grave o más agudo.

La utilización del tono –o tonos- también es un rasgo que clasifica los tipos de lenguas. Hay lenguas de tonales y lenguas entonativas. Las lenguas tonales son aquellas en las que el tono representa una diferencia léxica, distingue morfemas o palabras. Un ejemplo sería el chino. En cambio, las lenguas entonativas, como el español, el tono no sirve como diferencia léxica sino que los cambios tonales indican un cambio de estado

de ánimo o de intención del hablante y, también, diferencian las diversas modalidades de la oración. En español, como ejemplo general, cuando el tono final de la frase es ascendente suele informar que es una oración interrogativa, por el contrario, si el tono es descendente se trata de una oración afirmativa. Cabe observar que hay más rasgos que indican y diferencian las modalidades oracionales del español.

La duración se puede considerar tanto un rasgo segmental como suprasegmental que consiste en la prolongación en el tiempo de una determinada configuración articulatoria. Como elemento suprasegmental, además de ser mayor en las sílabas tónicas, es muy evidente en aquellas sílabas que constituyen el tonema (núcleo o final de las frases desde la última sílaba tónica) ya que puede aumentar respecto a aquellas sílabas anteriores al tonema, también llamadas pretonema.

3. MOTIVACIÓN:

El presente trabajo nace de mi curiosidad por un campo que no tiene un amplio estudio en el grado de Lingüística general de la UB: la lingüística forense. Pese a que se conoce que lingüística forense es aquella aplicación de la lingüística que trabaja en relación con el ámbito jurídico, no se hace más indagación en la carrera. Así pues, como desde el primer momento me interesó la fonética, conocí la labor de la fonética forense. Dicha área trabaja con la voz (o voces) para conseguir un acercamiento a pruebas que puedan servir en los casos judiciales para corroborar o descartar que aquella voz pertenece a una persona en concreto. Por los mencionados motivos y, sobre todo, por la falta de exploración en dicho ámbito, me pareció interesante enfocar mi Trabajo Final de Grado (TFG) sobre una cuestión relacionada con la lingüística forense, más concretamente con la fonética forense.

Una vez que tenía suficiente información sobre la fonética forense me planteé lo siguiente: ¿cómo puede ser que reconozcamos incluso aquellas voces que tan sólo hemos oído una vez en nuestra vida?. Es sencillo recordar las voces de nuestros familiares y amigos pero ¿qué pasa si es una voz que justo acabamos de escuchar?, ¿seríamos capaces de diferenciarla de otras voces? Está claro que discriminamos entre voces masculinas y femeninas pero, ¿qué pasa con las voces que no identifican claramente el género del emisor? Sabemos diferenciar en un aula si quien hace una pregunta al profesor es nuestra amiga, la compañera de al lado o si es una voz nueva para nosotros. Pese a que el primer pensamiento sea “qué hábil es la memoria” ésta necesita algún rasgo concreto que recordar. Por eso hablé con Ana María Fernández para plantearle mi enfoque de trabajo y que me asesorara sobre el trabajo. Cómo las cuestiones planteadas fueron bastantes como para abarcarlas todas en un TFG, decidimos centrarnos en la percepción y los rasgos segmentales y suprasegmentales de un corpus que había en el laboratorio. Por tanto, el presente trabajo nace de la curiosidad y se desarrolla con motivación y la ambición de conocer algo nuevo que pueda ser mi profesión en un futuro no muy lejano.

4. OBJETIVOS

El objetivo principal de este trabajo es la realización de una rueda de reconocimiento perceptiva. Para la justificación de los posibles resultados de las diferentes ruedas de reconocimientos el segundo objetivo es la realización de un análisis acústico de las 10 voces de las informantes.

5. HIPOTESIS:

Para la realización de la investigación se parte de dos hipótesis. La primera es la suposición que será más fácil identificar una voz, que sólo se ha escuchado una vez en discurso formal, entre un conjunto de voces similares en discurso espontáneo, que no el caso contrario. Es decir, que será más difícil si primero se escucha la voz en registro espontáneo y después se debe identificar entre un conjunto de voces en elocución formal. La segunda hipótesis es que no será imposible identificar una voz enmascarada con ruido, a pesar de que será siempre más fácil la identificación de la voz está limpia y no enmascarada.

6. METODOLOGÍA

6.1 Corpus e informantes

Para el desarrollo del trabajo, primeramente se utilizó un corpus del Laboratori de Fonètica de la UB recogido por Xènia Fernandez Pons para su trabajo de fin de grado “*Anàlisis de la variació intra i inter parlant de la fricativa alveolar /s/ del castellà amb finalitat forense*”, en el curso 2014-2015 dirigido por Ana María Fernández Planas. Se trata de un corpus oral donde todas las grabaciones se realizaron en una única sesión para cada informante -10 informantes en total- en el Laboratori de Fonètica de la UB. Las informantes eran 10 hablantes femeninas, de entre 18 y 25 años, alumnas de la facultad de letras de la UB. Una de las informantes era bilingüe de alemán e italiano pero hablaba habitualmente español. Las demás informantes eran bilingües de catalán y castellano y de estas nueve, siete tenían el catalán como L1 y las otras dos, el español. Las informantes fueron grabadas con un micrófono Audio Technica AT 2020 USB conectado al ordenador. El corpus estaba compuesto por 20 grabaciones: 10 en elocución formal y 10 en discurso espontáneo. Las informantes eran grabadas leyendo el texto “El viento norte y el sol” en castellano, y posteriormente eran grabadas en discurso espontáneo de tema libre. Las grabaciones del texto “El viento norte y el sol” tienen una duración media de 34.47 s de habla útil –sin silencios ni interrupciones- con una desviación estándar de 3.3 s; las grabaciones del discurso espontáneo tienen una duración media de 136.03 s de habla útil y una desviación estándar de 14.8s (Fernández Pons, 2015).

6.2 Jueces

Como la realización del test perceptivo ha sido totalmente por vía informática, para conocer las características de los jueces, la primera parte del cuestionario consistía en un conjunto de preguntas para conseguir dicha información. Así pues, la tipología de jueces de este trabajo se divide en 3 –una por cada rueda de reconocimiento–.

En primer lugar, véase la *Tabla 1* para observar los jueces que participaron en cada test:

Jueces	
Rueda de reconocimiento 1	71
Rueda de reconocimiento 2	41
Rueda de reconocimiento 3	30

Tabla 1 Jueces participantes en cada test perceptivo.

Así pues en las tablas siguientes se muestran las características de los jueces según sus respuestas:

Sexo:

Test perceptivo 1		Porcentaje
Mujer	48	67.60%
Hombre	23	32.40%
Total	71	100%
Test perceptivo 2		Porcentaje
Mujer	27	66%
Hombre	14	34%
Total	41	100%
Test perceptivo 3		Porcentaje
Mujer	19	63.30%
Hombre	11	36.70%
Total	30	100%

Tabla 2 Sexo de los jueces

Edad: la edad comprendida de los jueces era entre 16 y 59 años (véase tabla 42 pág.42) .En la primera rueda de reconocimiento el grupo de jueces 1 (22 a 35 años) era un total de 45 frente a los jueces del G2 (36 a 59 años) que eran 26. En el segundo test, el G1 estaba compuesto de 37 jueces (90%) y el G2 de 4 jueces (10%). Por último, en la tercera rueda de reconocimiento los jueces del G1 eran un total de 24 (80%) y en G2 estaba compuesto por 6 jueces (20%).

Para observar la lengua materna de los jueces véase *Tabla 3*.

Test perceptivo 1	Jueces	Porcentaje
Español	42	59.20%
Catalán	10	14.10%
Castellano y catalán	7	9.90%
Catalán y castellano	4	5.60%
Otros	8	11.30%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Español	12	29%
Catalán	11	27%
Castellano y catalán	10	24%
Catalán y castellano	4	10%
Otros	4	10%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Español	16	53.30%
Catalán	3	10%
Castellano y catalán	4	13.30%
Catalán y castellano	5	16.70%

Otros	2	6.70%
Total	30	100%

Tabla 3 Lengua materna de los jueces

¿Tienes estudios musicales?:

Test perceptivo 1	Jueces	Porcentaje
Sí	24	33.80%
No	47	66.20%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Sí	5	12%
No	36	88%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Sí	5	16.70%
No	25	83.30%
Total	30	100%

Tabla 4 Estudios musicales de los jueces

¿Sufres o has sufrido de problemas auditivos?

Test perceptivo 1	Jueces	Porcentaje
Sí	5	7%
No	65	91.50%
NS/NC	1	1.40%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Sí	1	2%
No	40	98%
NS/NC	0	0%

Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Sí	1	3.30%
No	28	93.30%
NS/NC	1	3.30%
Total	30	100%

Tabla 5 Problemas auditivos de los jueces

Para conocer más aspectos sobre el perfil de los jueces perceptivos véanse las tablas 6 a 9.

Test perceptivo 1	Jueces	Porcentaje
Sí	62	87.30%
No	9	12.70%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Sí	32	78%
No	9	4%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Sí	22	73.30%
No	8	26.70%
Total	30	100%

Tabla 6 Estudios universitarios de los jueces

¿Estás utilizando auriculares o altavoces?

Test perceptivo 1	Jueces	Porcentaje
Auriculares	24	33.80%
Altavoces	47	66.20%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Auriculares	21	51%
Altavoces	20	49%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Auriculares	12	40%
Altavoces	18	60%
Total	30	100%

Tabla 7 Uso de auriculares o altavoces por parte de los jueces.

¿Cursas o has cursado alguna asignatura de fonética?

Test perceptivo 1	Jueces	Porcentaje
Sí	38	54.30%
No	32	45.70%
Total	71	100%
Test perceptivo 2	Jueces	Porcentaje
Sí	11	27%
No	30	73%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Sí	11	39.70%
No	19	63.30%
Total	30	100%

Tabla 8 Conocimientos fonéticos de los jueces

Test perceptivo 2	Jueces	Porcentaje
Sí	12	29%
No	29	71%
Total	41	100%
Test perceptivo 3	Jueces	Porcentaje
Sí	17	56.60%
No	13	43.40%
Total	30	100%

Tabla 9 Realización del test anterior por parte de los jueces

6.1.4. Rueda de reconocimiento

Para la Rueda de reconocimiento de cada test se estudiaron varias posibilidades: o reunir a informantes y hacerles el test perceptivo en el Laboratori de Fonètica –lo que requeriría tiempo y esfuerzo extra tanto para los informantes como para el equipo del laboratorio- o bien, utilizar una plataforma online. Por cuestión de tiempo y facilidad, se optó por la segunda opción. Así pues, se procedió a la creación de tres test perceptivos a través del recurso online Surveygizmo¹.

Los test –que se pueden observar entre las páginas 67 y 84- están divididos en dos partes. La primera parte consta de unas preguntas básicas para conocer el perfil del juez como se ha visto en el apartado anterior. Estos datos sirvieron también como variables para el posterior análisis estadístico. La segunda parte es el test perceptivo en sí mismo; los jueces perceptivos primero oían la voz que hacía de voz criminal (o voz indubitada) en un registro concreto y posteriormente escuchaban las 10 voces (voces dubitadas), entre las cuales estaba la voz indubitada, en el registro contrario. Por último debían de escoger entre las 10 qué voz era la primera que habían escuchado y responder también en qué se habían basado (tono medio de la voz, acento, ritmo, NS/NC u otros).

¹ Agradezco la ayuda a Santiago González, investigador predoctoral del GrEP (Grup d'Estudis de Prosòdia), por prestarme los códigos para poder utilizar el programa.

En el primer test la voz indubitada (Voz de EDL, voz número 3 en el test) estaba en registro formal mientras que se debía reconocer entre las 10 voces en registro espontáneo. En el segundo test, la voz indubitada pertenecía a otra informante diferente del primero (voz NTL², número 10 en el test) y era en registro espontáneo, por tanto, se debía escoger entre 10 voces en registro formal. Por último, se procedió a la creación del tercer test con la misma plataforma online pero con la voz indubitada, oída en primer lugar, enmascarada en ruido *party*. Para ello, se subió la intensidad del archivo de la tercera informante (MM, número 8 en el test) en elocución formal, se resampearon los dos audios en cuanto a frecuencia de muestreo con Goldwave y se combinaron en un archivo estéreo con Praat³. Una vez que se consiguió dicho archivo se subió a la plataforma para poder crear el tercer test, con la misma estructura que los anteriores, pero con la voz indubitada enmascarada para reconocer posteriormente entre 10 voces sin enmascarar en registro formal. Tan sólo existe una pequeña modificación en el segundo y tercer test de la rueda de reconocimiento y es la incorporación de una pregunta más “¿Has realizado el test anterior?”, para tener una variable más y saber cuántas personas han realizado los test y ver su progresión.

La distribución de cada test era diferente para que no se aprendiera de memoria y evitar trampas en las respuestas. El orden era aleatorio aunque sí es cierto que se ha observado la tendencia que cuando la voz “criminal” estaba en las primeras posiciones parecía ser más fácil de acertar que si se hubiera encontrado en la posición media o final.

6.1.5. Análisis acústico

Para el análisis acústico se ha utilizado el programa libre Praat. Para empezar, se analizó los rasgos suprasegmentales: tono medio de la voz, entonación y acento. Para conocer el tono medio de la voz se observó la curva melódica de cada voz en registro formal e informal y se anotó aquel punto de la curva que representaba el valor medio en Hz. Para analizar la entonación y el acento se recurrió al mismo método de la observación de la

² Ver orden de las voces en la tabla del anexo pág. 66-67.

³ Paul Boersma & David Weenink (2013): *Praat: doing phonetics by computer [Computer program]*. Versión del programa: 6018 de 32 bit para windows. Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>

curva melódica o *pitch*. Para los rasgos segmentales se analizaron las diferentes producciones de nasales, oclusivas y fricativas en cada voz correspondida a análisis.

6.1.6. Análisis estadístico

Para el análisis estadístico se utilizó el paquete estadístico G-STAT versión 2.0.1. Primero se obtuvo una tabla de frecuencias para comparar los resultados con los resultados automáticos que proporcionaba el recurso online *Surveygizmo* y posteriormente se analizó con ANOVA cada variable explicativa (sexo, edad, L1, problemas auditivos, estudios musicales, estudios universitarios, utilización de auriculares o altavoces, conocimientos de fonética y test anterior para la rueda de reconocimiento 2 y 3) por la variable respuesta (Respuesta voz) para saber si los resultados eran estadísticos o no. Para ello se procedió a la recodificación de los datos.

6.1.7. Inconvenientes

No han sido muchos los inconvenientes que se han encontrado al realizar el presente trabajo, sin embargo, sí es cierto que en el caso de las ruedas de reconocimiento, no todas las respuestas han sido válidas. El motivo es que muchos jueces tan sólo contestaban la primera parte del cuestionario que correspondía a los datos personales y no realizaban la parte importante, el reconocimiento de la voz, es por ello que muchas respuestas tuvieron que ser descartadas. Pese a que las ruedas de reconocimiento se hacían por internet –desde un ordenador, móvil o tablet-, para la segunda y la tercera rueda de reconocimiento costó mucho más conseguir jueces dispuestos a dedicar 3 minutos para responderlas. Además, un porcentaje de jueces muy bajo fue el único en realizar los tres cuestionarios. Por lo tanto, no ha sido posible un seguimiento de las mismas personas en las ruedas de reconocimiento, así pues, se ha analizado al pequeño porcentaje de “repetidores” junto a los demás que sólo realizaron uno de los cuestionarios o dos de los tres.

7. RESULTADOS

7.1. Rueda de reconocimiento

7.1.1. Test perceptivo número 1 "Rueda de reconocimiento"

De 104 jueces que realizaron el test tan sólo 71 han sido válidos para analizar puesto que muchos no completaban la rueda de reconocimiento y sólo contestaban la primera parte sin realizar la de percepción.

De estos 71 la gran mayoría supo reconocer la voz correcta -que en el primer test perceptivo era la número 3 (voz de EDL)- como se puede observar en la siguiente tabla:

Respuesta voz	Frecuencias	Porcentajes
Voz 3	56	78.90 %
Voz 9	9	12.70%
Voz 6	3	4.20 %
Voz 2	2	2.80 %
Voz 1	1	1.40 %
Total	71	100 %

Tabla 10 Frecuencia voces "Rueda de Reconocimiento 1"

Tan sólo un 19.70% de los jueces no supo identificar la voz. Tal y como se observa la gran mayoría confundía la voz número 3 con la voz número 9. El porqué de esta confusión se analizará posteriormente en el apartado "Resultados del Análisis acústico". Veamos ahora la elección de la voz de los jueces según sus cualidades. En primer lugar, según el sexo se observa en la siguiente tabla que no hay mucha variación de acierto entre hombres y mujeres:

Respuesta voz	Hombre	Mujer	Total
Voz 3	20 (28.20%)	37 (52.10%)	56 78.90 %
Voz 9	1	8	9

	(1.40%)	(11.30%)	12.70%
Voz 6	2 (2.80%)	1 (1.40%)	3 4.20%
Voz 2	0 (0%)	1 (1.40%)	2 2.80%
Voz 1	0 (0%)	1 (1.40%)	1 1.40%
Total	22 31%	49 69%	71 100%

Tabla 11 Respuestas según sexo jueces "Rueda de reconocimiento 1"

Si bien es cierto que por lo que se observa en la anterior tabla las mujeres discrepan más respecto a la identificación de la voz, el análisis acústico mostró un p-valor de 0.24 y F-valor de 1.39, por lo que la variable de sexo no es representativa –ya que consideraremos que lo es cuando el p-valor es inferior a 0.05-. En segundo lugar, la siguiente tabla muestra la elección de la voz según la edad:

Respuesta voz	G1 (22 a 35 años)	G2 (36 a 57 años)	Total
Voz 3	36 (50.70%)	21 (29.60%)	56 78.90 %
Voz 9	6 (8.50%)	3 (4.20%)	9 12.70%
Voz 6	2 (2.80%)	1 (1.40%)	3 4.20%
Voz 2	1 (1.40%)	0 (0%)	2 2.80%
Voz 1	0 (0%)	1 (1.40%)	1 1.40%
Total	45 63.40%	26 36.60%	71 100%

Tabla 12 Respuesta de los jueces según edad "Rueda de reconocimiento 1"

Tal y como ocurría con el sexo, los resultados del análisis estadístico mostraron en esta variable un p-valor de 0.72 y un F-valor de 0.12 por lo que, de la misma manera que el anterior no resulta ser una variable a tener en cuenta. En tercer lugar, según la lengua materna de los jueces (L1) encontramos:

Respuesta voz	Catalán	Español	Otro	Total
Voz 3	11 (15.50%)	41 (57.80%)	4 (5.60%)	56 78.90 %
Voz 9	2 (2.80%)	4 (5.70%)	3 (4.20%)	9 (12.70%)
Voz 6	0 (0%)	2 (2.80%)	1 (1.40%)	3 (4.20%)
Voz 2	1 (1.40%)	1 (1.40%)	0 (0%)	2 (2.80%)
Voz 1	0 (0%)	1 (1.40%)	0 (0%)	1 (1.40%)
Total	15 19.70%	48 69.10%	8 11.20%	71 100%

Tabla 13 Respuesta de voz según L1 de los jueces "Rueda de reconocimiento 1"

El p-valor de esta variable es de 0.03 y el F-valor de 3.64 por lo que la lengua materna influye en la percepción auditiva. La lengua materna en esta muestra es significativa. Además, la comparación múltiple del análisis estadístico mostró que la diferencia significativa era la que presentaban los jueces con L1 de catalán o castellano respecto a los jueces con L1 extranjera.

En cuarto lugar, en función de si los jueces sufrían o habían sufrido problemas auditivos los resultados fueron los siguientes:

Respuesta voz	Problemas auditivos: Sí	Problemas auditivos: No	Problemas auditivos: NS/NC	Total
Voz 3	4 (5.60%)	51 (71.90%)	1 (1.40%)	56 78.90 %
Voz 9	0 (0%)	9 (12.70%)	0 (0%)	9 12.70%
Voz 6	1 (1.40%)	2 (2.80%)	0 (0%)	3 4.20%
Voz 2	0 (0%)	2 (2.80%)	0 (0%)	2 2.80%
Voz 1	0 (0%)	1 (1.40%)	0 (0%)	1 1.40%
Total	5 7%	65 91.60%	1 1.40%	71 100%

Tabla 14 Respuestas de los jueces según problemas auditivos

La población de jueces con problemas auditivos es muy reducida, por lo tanto, consideraremos que nuestros jueces no tienen problemas auditivos. La prueba estadística, lógicamente, también dio negativa con un p-valor de 0.88 y un F-valor de 0.12. En quinto lugar, habían jueces con estudios musicales y sin, los resultados se observan en la tabla siguiente:

Respuesta voz	Estudios musicales: Sí	Estudios musicales: No	Total
Voz 3	21 (29.60%)	35 (49.30%)	56 78.90 %
Voz 9	1 (1.40%)	8 (11.30%)	9 12.70%
Voz 6	1 (1.40%)	2 (2.80%)	3 4.20%
Voz 2	1 (1.40%)	1 (1.40%)	2 2.80%
Voz 1	0 (0%)	1 (1.40%)	1 1.40%
Total	24 33.80%	47 66.20%	71 100%

Tabla 15 Respuestas jueces según estudios musicales

En este caso observamos que los jueces con estudios musicales (24 de 71) tienden a acertar más que no los jueces sin estudios musicales. Sin embargo, el p-valor de esta variable es de 0.17 y el F-valor de 1.85, por lo tanto, no es una variable significativa en el momento de identificar a un hablante. Entre los jueces existía también la división jueces con estudios universitarios” y “jueces sin estudios universitarios”. En esta diferencia se observa:

Respuesta voz	Estudios universitarios: Sí	Estudios universitarios: No	Total
Voz 3	49 (69%)	7 (9.90%)	56 78.90 %
Voz 9	7 (9.90%)	2 (2.80%)	9 12.70%
Voz 6	3 (4.20%)	0 (0%)	3 4.20%
Voz 2	2	0	2

	(2.80%)	(0%)	2.80%
Voz 1	1	0	1
	(1.40%)	(0%)	1.40%
Total	62	9	71
	87.30%	12.70%	100%

Tabla 16 Respuestas jueces según estudios universitarios

La variable de estudios universitarios tampoco resulta significativa si miramos la tabla, el análisis estadístico corroboró que es así puesto que tiene un p-valor de 0.69 y un F-valor de 0.15. Por último, este test estaba dividido entre jueces con conocimientos de fonética y jueces sin conocimientos de dicho ámbito. Los resultados fueron los siguientes:

Respuesta voz	Conocimientos fonética: Sí	Conocimientos fonética: No	Total
Voz 3	30	26	56
	(42.30%)	(36.60%)	78.90 %
Voz 9	7	2	9
	(9.90%)	(2.80%)	12.70%
Voz 6	1	2	3
	(1.40%)	(2.80%)	4.20%
Voz 2	0	2	2
	(0%)	(2.80%)	2.80%
Voz 1	0	1	1
	(0%)	(1.40%)	(1.40%)
Total	38	33	71
	53.50	46.50%	100%

Tabla 17 Respuestas jueces según conocimientos de fonética

En un primer momento podríamos considerar que, pese a que está muy equilibrado, los jueces con conocimientos fonéticos no cometían tantos errores, pero el análisis estadístico corroboró que tener estudios universitarios tampoco es una variable significativa puesto que obtuvo un p-valor de 0.84 y un F-valor de 0.0377.

Para observar otros factores más externos a la hora de reconocer una voz, los jueces debían especificar si realizaban el test con auriculares o con altavoces. De esa cuestión observamos los resultados en la tabla siguiente:

Respuesta voz	Altavoces	Auriculares	Total
Voz 3	38 (53.50%)	18 (25.40%)	56 78.90 %
Voz 9	6 (8.50%)	3 (4.20%)	9 12.70%
Voz 6	3 (4.20%)	0 (0%)	3 4.20%
Voz 2	1 (1.40%)	1 (1.40%)	2 2.80%
Voz 1	1 (1.40%)	0 (0%)	1 1.40%
Total	49 69%	22 31%	71 100%

Tabla 18 Respuesta jueces según uso de auriculares o altavoces "Rueda de reconocimiento 1"

El análisis estadístico corroboró que no existe una diferencia significativa puesto que el p-valor de esta variable es de 0.14 y el F-valor de 2.19.

La última cuestión que debían contestar los jueces era el motivo de la elección de aquella voz como la voz indubitada. Existían varias opciones entre las cuales solo podían escoger una: ritmo, tono medio de la voz, entonación, acento, si no sabían por qué u otros motivos que podían especificar. En la tabla 19 están todos los motivos que seleccionaron los diferentes jueces:

Motivo	Voz 3	Voz 9	Voz 6	Voz 2	Voz 1	Total
NS/NC	1 (1.40%)	0 (0%)	0	0	0	1 (1.40%)
Tono medio voz	30 (42.40%)	4 (5.60%)	1 (1.40%)	1 (1.40%)	1 (1.40%)	37 (52.20%)
Entonación	9 (12.70%)	3 (4.20%)	0 (0%)	1 (1.40%)	0 (0%)	13 (18.30%)
Acento	4 (5.70%)	1 (1.40%)	1 (1.40%)	0 (0%)	0 (0%)	6 (8.50%)
Todas las opciones	2 (2.80%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (2.80)
Ritmo	3 (4.20%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (4.20%)
Ot: cualidad de voz y pronunciación	0 (0%)	1 (1.40%)	0 (0%)	0 (0%)	0 (0%)	1 (1.40%)
Ot: Enunciación	1 (1.40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1.40%)
Ot: sibilantes, creaky voice	1 (1.40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1.40%)
Ot: Timbre	3 (4.20%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (4.20%)
Ot: Timbre y tono medio voz	1 (1.40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1.40%)
Ot: Cualidad de voz	1 (1.40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1.40%)
Ot: leve carraspera	0 (0%)	0 (0%)	1 (1.40%)	0 (0%)	0 (0%)	1 (1.40%)
Total	79%	12.60%	4.20%	2.80%	1.40%	100 %

Tabla 19 Motivos de las respuestas de los jueces

Como se observa, el principal motivo por el cual los jueces creen reconocer la voz es por su tono medio, sin embargo, también se observan otros jueces que consideraron aspectos distintos al tono, como el acento, el timbre o la entonación.

7.1.2. Test perceptivo número 2 "Rueda de reconocimiento 2"

De 51 jueces que participaron en el segundo test perceptivo tan sólo lo completaron 41, por lo tanto en este segundo test sólo se analizan las respuestas completadas de 41 jueces. La voz correcta en la rueda de reconocimiento 2 era la número 10 [NTL] –tal y como está marcada en verde en la tabla-, los resultados del segundo test perceptivo a nivel de respuesta fueron:

Respuesta voz	Frecuencias	Porcentajes
Voz 2	22	53.70%
Voz 10	4	9.80%
Voz 5	3	7.30%
Voz 6	6	14.60%
Voz 7	5	12.20%
Voz 8	1	2.40%
Total	41	100 %

Tabla 20 Frecuencia voces "Rueda de Reconocimiento 2"

Como se puede observar, en este segundo test perceptivo fue una minoría quien reconoció la voz correcta. Un 9.8% frente a un 51.20% que confundió la voz con la número 2 [ALB].

Para os resultados según el sexo de los jueces véase *Tabla 21*.

Respuesta voz	Hombre	Mujer	Total
Voz 2	5 (12.20%)	17 (41.46%)	22 53.66%
Voz 10	2 (4.88%)	2 (4.88%)	4 9.76%
Voz 5	2 (4.88%)	1 (2.44%)	3 7.32%
Voz 6	2 (4.88%)	4 (9.76%)	6 14.63%
Voz 7	3 (7.32%)	2 (4.88%)	5 12.20%
Voz 8	0 (0%)	1 (2.44%)	1 2.44%
Total	14 34.15%	27 65.85%	41 100%

Tabla 21 Respuestas según sexo jueces "Rueda de reconocimiento 2"

En este caso no hay ninguna diferencia significativa respecto al sexo en lo que respecta a acertar la voz. El p-valor de esta variable es de 0.15 y su F-valor de 2.11, lo que efectivamente, indica que no es una variable a tener en cuenta para el reconocimiento de hablantes.

Los resultados según la edad de la segunda rueda de reconocimiento fueron los siguientes:

Respuesta voz	G1	G2	Total
Voz 2	22 (53.66%)	0 (0%)	22 53.66%
Voz 10	3 (7.32%)	1 (2.44%)	4 9.76%
Voz 5	3	0	3

	(7.32%)	(0%)	7.32%
Voz 6	6	0	6
	(14.63%)	(0%)	14.63%
Voz 7	3	2	5
	(7.32%)	(4.88%)	12.20%
Voz 8	1	0	1
	(2.45%)	(0%)	2.44%
Total	38	3	41
	92.68%	7.32%	100%

Tabla 22 Respuesta de los jueces según edad "Rueda de reconocimiento 2"

Por desgracia no hay jueces suficientes entre las edades de 36 a 59 años para encontrar una diferencia significativa, ya que participaron más de un 90% de jueces entre 19 y 35 años. Su p-valor es de 0.40 y su F-valor de 1.08.

En este segundo test participaron castellanohablantes, catalanohablantes y los jueces con lengua extranjera fueron todos italianos:

Respuesta voz	Catalán	Italiano	Español	Total
Voz 2	8	2	12	22
	(19.51%)	(4.88%)	(29.27%)	53.66%
Voz 10	2	0	2	4
	(4.88%)	(0%)	(4.88%)	9.76%
Voz 5	2	0	1	3
	(4.88%)	(0%)	(2.44%)	7.32%
Voz 6	2	0	4	6
	(4.88%)	(0%)	(9.76%)	14.63%
Voz 7	1	2	2	5
	(2.44%)	(4.88%)	(4.88%)	12.20%
Voz 8	1	0	0	1
	(2.44%)	(0%)	(0%)	2.44%

Total	16	4	21	41
	39.02%	9.76%	51.22%	100%

Tabla 23 Respuesta de voz según L1 de los jueces "Rueda de reconocimiento 2"

En este caso, el análisis estableció que no existía un resultado significativo puesto que el p-valor era de 0.90 y el F-valor de 0.10, cosa que dista bastante del primer test. Seguramente esto ocurrirá porque el tamaño de la muestra es menor que en la primera rueda de reconocimiento.

La tabla siguiente muestra los resultados de los jueces con problemas auditivos y sin problemas auditivos, pero como se observa tan sólo existe un único juez con problemas auditivos que no acertó, por lo tanto, no podemos establecer que sea significativa esta característica a la hora de percibir mejor una voz:

Respuesta voz	Problemas auditivos:		Total
	No	Sí	
Voz 2	21 (51.22%)	1 (2.44%)	22 53.66%
Voz 10	4 (9.76%)	0 (0%)	4 9.76%
Voz 5	3 (7.32%)	0 (0%)	3 7.32%
Voz 6	6 (14.63%)	0 (0%)	6 14.63%
Voz 7	5 (12.20%)	0 (0%)	5 12.20%
Voz 8	1 (2.44%)	0 (0%)	1 2.44%
Total	40 97.56%	1 2.44%	41 100%

Tabla 24 Respuestas de los jueces según problemas auditivos "Rueda de reconocimiento 2"

Como en el test anterior, no se trata de un dato estadísticamente significativo, su p-valor es de 0.40 y el F-valor de 0.70. Por lo tanto, hay que descartar los problemas auditivos como una variable significativa en estas tres ruedas de reconocimiento porque no se dispone de jueces suficientes con problemas auditivos.

Las siguientes tablas muestran los resultados de los jueces con estudios musicales y sin estudios musicales, y también, los resultados de los jueces con estudios universitarios y sin estudios universitarios:

Respuesta voz	Estudios musicales: No	Estudios musicales: Sí	Total
Voz 2	21 (51.22%)	1 (2.44%)	22 53.66%
Voz 10	2 (4.88%)	2 (4.88%)	4 9.76%
Voz 5	2 (4.88%)	1 (2.44%)	3 7.32%
Voz 6	6 14.63%	0 (0%)	6 14.63%
Voz 7	4 (9.76%)	1 (2.44%)	5 12.20%
Voz 8	1 2.44%	0 (0%)	1 2.44%
Total	36 87.80%	5 12.20%	41 100%

Tabla 25 Respuestas jueces según estudios musicales "Rueda de reconocimiento 2"

Respuesta voz	Estudios universitarios: No	Estudios universitarios: Sí	Total
Voz 2	4 (9.76%)	18 (43.90%)	22 53.66%
Voz 10	1 (2.44%)	3 (7.32%)	4 9.76%
Voz 5	0 (0%)	3 (7.32%)	3 7.32%
Voz 6	2 (4.88%)	4 (9.76%)	6 14.63%
Voz 7	2 (4.88%)	3 (7.32%)	5 12.20%
Voz 8	0 (0%)	1 2.44%	1 2.44%
Total	9 21.95%	32 78.05%	41 100%

Tabla 26 Respuestas jueces según estudios universitarios "Rueda de reconocimiento 2"

Los jueces con estudios musicales no llegan a la mitad de la población del segundo test (5/41) pero, sin embargo, representa un dato representativo puesto que el p-valor es de 0.03 y el F-valor de 4.78. Parece relevante tener un "oído entrenado" como el de los músicos. Por el otro lado, los estudios universitarios siguen sin ser un dato relevante en este tipo de ruedas de reconocimiento. Para estudios universitarios el p-valor es de 0.51 y el F-valor de 0.43.

Observemos ahora si tener conocimientos de fonética ha influido en los resultados de la segunda rueda de reconocimiento:

Respuesta voz	Conocimientos fonéticos: No	Conocimientos fonéticos: Sí	Total
Voz 2	15 (36.59%)	7 (17.07%)	22 53.66%
Voz 10	4 (9.76%)	0 (0%)	4 9.76%
Voz 5	1 (2.44%)	2 (4.88%)	3 7.32%
Voz 6	6 14.63%	0 (0%)	6 14.63%
Voz 7	3 (7.32%)	2 (4.88%)	5 12.20%
Voz 8	0 (0%)	1 2.44%	1 2.44%
Total	29 70.73%	12 29.27%	41 100%

Tabla 27 Respuestas jueces según conocimientos fonéticos "Rueda de reconocimiento 2"

Los jueces con conocimiento de fonética que han participado en el segundo test no han acertado la voz correcta, lo que nos da a entender que no es una variable característica si además se observa un p-valor de 0.46 y un F-valor de 0.54. Pudiera ser que los conocimientos de fonética sirvan más de distractor que para acertar una identificación auditiva.

Observemos el factor externo del uso de auriculares o altavoces para intuir si será una diferencia en jueces significativa, o no:

Respuesta voz	Auriculares	Altavoces	Total
Voz 2	11 (26.83%)	11 (26.83%)	22 53.66%
Voz 10	3 (7.32%)	1 (2.44%)	4 9.76%
Voz 5	2 (4.88%)	1 (2.44%)	3 7.32%
Voz 6	2 (4.88%)	4 (9.76%)	6 14.63%
Voz 7	2 (4.88%)	3 (7.32%)	5 12.20%
Voz 8	1 (2.44%)	0 (0%)	1 2.44%
Total	21 51.22%	20 48.78%	41 100%

Tabla 28 Respuesta jueces según uso de auriculares o altavoces "Rueda de reconocimiento 2"

Los jueces de este test lo realizaron tanto con altavoces como auriculares y no presenta una diferencia demasiado grande como para tener en cuenta, además el p-valor corresponde a 0.46 y el F-valor a 0.54 lo que nos corrobora que no es significativo el uso de auriculares o altavoces.

Como se ha comentado anteriormente, la segunda rueda de reconocimiento y la tercera cuentan con una característica más y son aquellos jueces que han hecho el test anterior:

Respuesta voz	Test anterior: No	Test anterior: Sí	Total
Voz 2	17 (41.46%)	5 (12.20%)	22 53.66%
Voz 10	0 (0%)	4 (9.76%)	4 9.76%
Voz 5	2 (4.88%)	1 (2.44%)	3 7.32%
Voz 6	4 (9.76%)	2 (4.88%)	6 14.63%
Voz 7	5 (12.20%)	0 (0%)	5 12.20%
Voz 8	1 (2.44%)	0 (0%)	1 2.44%
Total	29 70.73%	12 29.27%	41 100%

Tabla 29 Respuestas jueces según realización test anterior "Rueda de reconocimiento 2"

Con la observación de la tabla anterior podría establecerse que hay una gran diferencia entre realizar el test anterior a este o no. Sin embargo, el análisis estadístico corroboró que se trata de un dato no significativo –aunque por poco- ya que su p-valor es de 0.06 y el F-valor de 3.53.

Por último, los motivos que llevaron a estos jueces a la elección de la voz seleccionada se ven en la *tabla 30*.

Motivo	Voz 2	Voz 10	Voz 5	Voz 6	Voz 7	Voz 8	Total
Tono medio voz	13 (31.71%)	2 (4.88%)	3 (7.32%)	4 (9.76%)	2 (4.88%)	0 (0%)	24 (58.54%)
Entonación	2 (4.88%)	1 (2.44%)	0 (0%)	1 (2.44%)	1 (2.44%)	0 (0%)	5 (12.20%)
Acento	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (2.44%)	1 (2.44%)	2 (4.88%)
Todas las opciones	7 (17.07%)	0 (0%)	0 (0%)	1 (2.44%)	1 (2.44%)	0 (0%)	9 (21.95%)
Ritmo	0 (0%)	1 (2.44%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (2.44%)
Total	22 53.66%	4 9.76%	3 7.32%	6 14.63%	5 12.20%	1 2.44%	41 100 %

Tabla 30 Respuesta jueces respecto a motivos "Rueda de reconocimiento 2"

Pese a que no se han obtenido buenos resultados de aciertos, observamos que el patrón mayoritario en las tres tablas es considerar que el tono medio de la voz es el principal motivo en el que se basan los jueces (31.71%), así mismo, observamos que un conjunto de todas las opciones también tiene una gran representación (17.07%).

7.1.3. Test perceptivo número 3 "Rueda de reconocimiento 3"

En el último test perceptivo participaron un total de 45 jueces, pero una vez más, sólo lo completaron 30. Por lo tanto, esta es la muestra más pequeña con 30 jueces. La voz correcta en esta rueda de reconocimiento era la número 8 –que como las anteriores se destaca en la tabla en verde-. Observemos las respuestas de los jueces en la *tabla 31*:

Respuesta voz	Frecuencias	Porcentajes
Voz 8	12	40%
Voz 9	6	20%
Voz 2	1	3.33%
Voz 3	4	13.33%
Voz 6	2	6.67%
Voz 10	1	3.33%
Voz 4	3	10%
Voz 1	1	3.33%
Total	30	100 %

Tabla 31 Respuestas voces "Rueda de reconocimiento 3"

Pese a que sólo un 40% de los jueces acertó la respuesta, podemos considerar que la segunda hipótesis de este trabajo se confirma puesto que es posible identificar una voz pese a que esté enmascarada con ruido.

Los resultados de los jueces según su sexo fueron los siguientes:

Respuesta voz	Hombre	Mujer	Total
Voz 8	5 (16.67%)	7 (23.33%)	12 40%
Voz 9	1 (3.33%)	5 (16.67%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	2 (6.67%)	2 (6.67%)	4 13.33%
Voz 6	1 (3.33%)	1 (3.33%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	1 3.33%
Voz 4	1	2	3

	(3.33%)	(6.67%)	10%
Voz 1	0	1	1
	(0%)	(3.33%)	3.33%
Total	11	19	30
	36.67%	63.33%	100 %

Tabla 32 Respuestas según sexo jueces "Rueda de reconocimiento 3"

El análisis estadístico determinó, una vez más, que el sexo no es una variable significativa. El p-valor obtenido fue de 0.65 y el F-valor de 0.20.

Los jueces que realizaron este último test tenían entre 20 y 58 años. En la siguiente tabla se muestran las respuestas del grupo 1 (20-35 años) y del grupo 2 (36-58 años).

Respuesta voz	G1 (20-35 años)	G2 (36-58 años)	Total
Voz 8	9	3	12
	(30%)	(10%)	40%
Voz 9	4	2	6
	(13.33%)	(6.67%)	20%
Voz 2	1	0	1
	(3.33%)	(0%)	3.33%
Voz 3	3	1	4
	(10%)	(3.33%)	13.33%
Voz 6	2	0	2
	(6.67%)	(0%)	6.67%
Voz 10	1	0	1
	(3.33%)	(0%)	3.33%
Voz 4	3	0	3
	(10%)	(0%)	10%
Voz 1	1	0	1
	(3.33%)	(0%)	3.33%

Total	24	6	30
	80%	20%	100 %

Tabla 33 Respuesta de los jueces según edad "Rueda de reconocimiento 3"

Por tanto, tan sólo contamos con 6 jueces que estén entre los 36 y los 58 años, algo que a simple vista ya no parece muy significativo. El p-valor es de 0.59 y el F-valor de 0.29, por lo tanto, no es una variable a tener en cuenta.

En la tabla siguiente se puede observar las respuestas de los jueces teniendo en cuenta su lengua materna (L1):

Respuesta voz	Español	Catalán	Otro	Total
Voz 8	9 (30%)	3 (10%)	0 (0%)	12 40%
Voz 9	2 (6.67%)	3 (10%)	1 (3.33%)	6 20%
Voz 2	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Voz 3	2 (6.67%)	1 (3.33%)	1 (3.33%)	4 13.33%
Voz 6	1 (3.33%)	0 (0%)	1 (3.33%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Voz 4	3 (10%)	0 (0%)	0 (0%)	3 10%
Voz 1	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Total	20 66.67	8 26.67	2 6.67	30 100 %

Tabla 34 Respuesta de voz según L1 de los jueces "Rueda de reconocimiento 3"

Si bien es cierto que en la primera rueda de reconocimiento la L1 era una variable significativa, a medida que la muestra es más pequeña, menos significativo lo considera el análisis estadístico (p-valor de 0.48, F-valor de 0.74).

Los jueces con problemas auditivos en esta tercera rueda de reconocimiento se reducen a 1, existe un juez que no sabe si tiene o ha tenido problemas auditivos por lo que se considera que los jueces no tienen problemas auditivos como en los test anteriores.

Respuesta voz	Problemas auditivos: No	Problemas auditivos: Sí	Problemas auditivos: NS/NC	Total
Voz 8	12 (40%)	0 (0%)	0 (0%)	12 40%
Voz 9	6 (20%)	0 (0%)	0 (0%)	6 20%
Voz 2	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Voz 3	3 (10%)	1 (3.33%)	0 (0%)	4 13.33%
Voz 6	2 (6.67%)	0 (0%)	0 (0%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Voz 4	2 (6.67%)	1 (3.33%)	0 (0%)	3 10%
Voz 1	1 (3.33%)	0 (0%)	0 (0%)	1 3.33%
Total	28 93.33%	1 3.33%	1 3.33%	30 100 %

Tabla 35 Respuestas de los jueces según problemas auditivos "Rueda de reconocimiento 3"

Observemos ahora los resultados de los jueces según si tienen estudios musicales o estudios universitarios:

Respuesta voz	Estudios musicales:		Total
	Sí	No	
Voz 8	2 (6.67%)	10 (33.33%)	12 40%
Voz 9	0 (0%)	6 (20%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	1 (3.33%)	3 (10%)	4 13.33%
Voz 6	0 (0%)	2 (6.67%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	1 3.33%
Voz 4	1 (3.33%)	2 (6.67%)	3 10%
Voz 1	0 (0%)	1 (3.33%)	1 3.33%
Total	5 16.67%	25 83.33%	30 100 %

Tabla 36 Respuestas jueces según estudios musicales "Rueda de reconocimiento 3"

En esta muestra los estudios musicales no son significativos, aparte de tener una población con estudios musicales de 5, el p-valor es de 1 y el F-valor de 0. Por lo tanto, podemos decir que se necesitaría una muestra equilibrada y con muchos jueces para poder demostrar si los conocimientos musicales son significativos o no.

Respuesta voz	Estudios universitarios: Sí	Estudios universitarios: No	Total
Voz 8	9 (30%)	3 (10%)	12 40%
Voz 9	5 (16.57%)	1 (3.33%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	2 (6.67%)	2 (6.67%)	4 13.33%
Voz 6	2 (6.67%)	0 (0%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	1 3.33%
Voz 4	2 (6.67%)	1 (3.33%)	3 10%
Voz 1	1 (3.33%)	0 (0%)	1 3.33%
Total	22 73.33%	8 26.67%	30 100 %

Tabla 37 Respuestas jueces según estudios universitarios "Rueda de reconocimiento 3"

Por otra parte, en ninguna de las muestras resultan ser significativos los estudios universitarios. En la tabla de los estudios universitarios, además, sólo tenemos 8 jueces sin estudios universitarios de los cuales acertaron 3.

En la siguiente tabla se muestran los resultados de los jueces según si tenían conocimientos fonéticos o no:

Respuesta voz	Conocimientos de fonética: Sí	Conocimientos de fonética: No	Total
Voz 8	4 (13.33%)	8 (26.67%)	12 40%
Voz 9	2 (6.67%)	4 (13.33%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	0 (0%)	4 (13.33%)	4 13.33%
Voz 6	0 (0%)	2 (6.67%)	2 6.67%
Voz 10	1 (3.33%)	0 (0%)	1 3.33%
Voz 4	2 (6.67%)	1 (3.33%)	3 10%
Voz 1	1 (3.33%)	0 (0%)	1 3.33%
Total	11 36.67%	19 63.33%	30 100 %

Tabla 38 Respuestas jueces según conocimientos fonéticos "Rueda de reconocimiento 3"

En este caso, los jueces con conocimientos fonéticos frente a los que no tenían conocimientos de fonética están más equilibrados que otras variables observadas, por lo tanto, la diferencia existente entre unos y otros es menor. Aun así, el resultado del análisis estadístico estableció esta variable, nuevamente, como no significativa (p-valor: 0.76, F-valor: 0.08).

Los resultados entre los jueces que utilizaron auriculares y los que utilizaron altavoces se muestran en la tabla a continuación:

Respuesta voz	Auriculares	Altavoces	Total
Voz 8	4 (13.33%)	8 (26.67%)	12 40%
Voz 9	3 (10%)	3 (10%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	3 (10%)	1 (3.33%)	4 13.33%
Voz 6	1 (3.33%)	1 (3.33%)	2 6.67%
Voz 10	0 (0%)	1 (3.33%)	1 3.33%
Voz 4	2 (6.67%)	1 (3.33%)	3 10%
Voz 1	0 (0%)	1 (3.33%)	1 3.33%
Total	12 40%	18 60%	30 100 %

Tabla 39 Respuesta jueces según uso de auriculares o altavoces "Rueda de reconocimiento 3"

Como se puede observar, también el uso de altavoces o auriculares es bastante equilibrado, pero como en los test anteriores el análisis estadístico tampoco ha mostrado que sea una variable significativa (p-valor 0.55, F-valor 0.35).

Respuesta voz	Test anterior: Sí	Test anterior: No	Total
Voz 8	9 (30%)	3 (10%)	12 40%
Voz 9	2 (6.67%)	4 (13.33%)	6 20%
Voz 2	0 (0%)	1 (3.33%)	1 3.33%
Voz 3	3 (10%)	1 (3.33%)	4 13.33%
Voz 6	0 (0%)	2 (6.67%)	2 6.67%
Voz 10	0 (0%)	1 (3.33%)	1 3.33%
Voz 4	2 (6.67%)	1 (3.33%)	3 10%
Voz 1	0 (0%)	1 (3.33%)	1 3.33%
Total	16 53.33%	14 46.67%	30 100 %

Tabla 40 Respuestas jueces según realización test anterior "Rueda de reconocimiento 3"

En la tabla anterior, que corresponde a los resultados según si los jueces eran los mismos que en las ruedas de reconocimiento anteriores, observamos que está equilibrado y que aciertan más los jueces que sí hicieron el test (o los tests) anteriores que no aquellos que lo realizaban por primera vez. Pero el análisis estadístico tampoco refleja que sea un dato significativo puesto que el p-valor es de 0.0549 y el F-valor de 4.02, son datos muy justos pero sigue siendo superior a 0.05, por lo tanto, no se considera una variable significativa.

Por último, se muestran los resultados de los jueces por los motivos que escogieron la voz que consideraban correcta:

Motivo	Voz 8	Voz 9	Voz 2	Voz 3	Voz 6	Voz 10	Voz 4	Voz 1	Total
Tono medio	7	3	1	3	1	0	2	1	17 56.67%
Entonación	3	1	0	1	1	0	1	0	7 23.33%
Acento	1	2	0	0	0	0	0	0	3 10%
Otro: breathy voice	0	0	0	0	0	1	0	0	1 3.33%
Ritmo	0	0	0	1	0	0	0	0	1 3.33%
Otro: por la "s"	1	0	0	0	0	0	0	0	1 3.33%
Total	12 40%	6 20%	1 3.33%	4 13.33%	2 6.67%	1 3.33%	3 10%	1 3.33%	30 100%

Tabla 41 Respuesta jueces respecto a motivos "Rueda de reconocimiento 3"

Los resultados de los jueces por lo que respecta al motivo, indican que en lo que más se fijan es en el tono medio de la voz y la entonación. Si está relacionado con acertar o no, se verá mejor en el análisis a continuación.

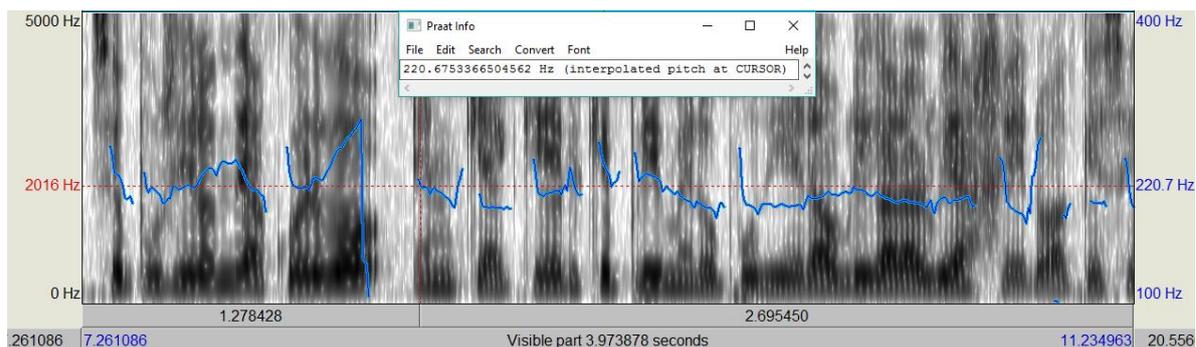
7.2. Análisis acústico de las voces

7.2.1. Rueda de reconocimiento 1:

En la primera rueda de reconocimiento la voz indubitada que se escuchaba primeramente en voz formal era la voz número 3 [EDL]. Los resultados que se han observado en el apartado anterior muestran como los jueces confunden la voz número tres mayormente con la voz número 9 [NTL] (12.70%) y la voz número 6 [MDM] (4.20%). Algún juez también consideró que se trataba de la voz número 2 [ABM] y otro juez pensó que era la voz número 1 [ALB]. En cambio, ningún juez consideró que se tratara ni de la voz número 4 [GSL], ni la 5 [LRD], ni la 7 [MR] ni la 8 [MRT] ni la 10 [SLV].

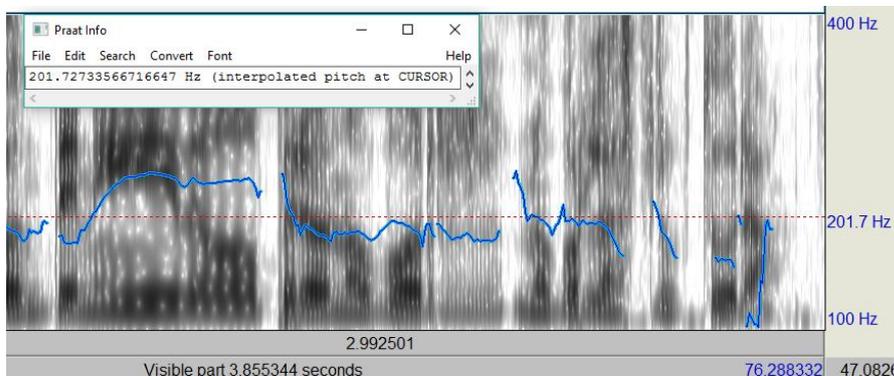
Observemos en este apartado qué características acústicas comparten las voces 3 [EDL], 9 [NTL] y 6 [MDM] y que otros rasgos las diferencian.

En primer lugar, si observamos el tono medio de la voz nos encontramos que la voz indubitada (voz 3) tiene un tono medio de 220 Hz.

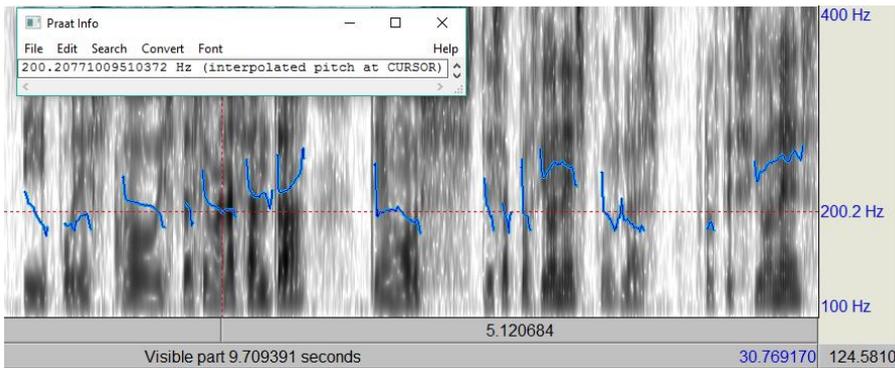


Voz número 3 EDL, 220Hz a

En cambio, la voz número 9 tiene un tono medio de 202Hz y la voz número 6 un tono medio de 200Hz.

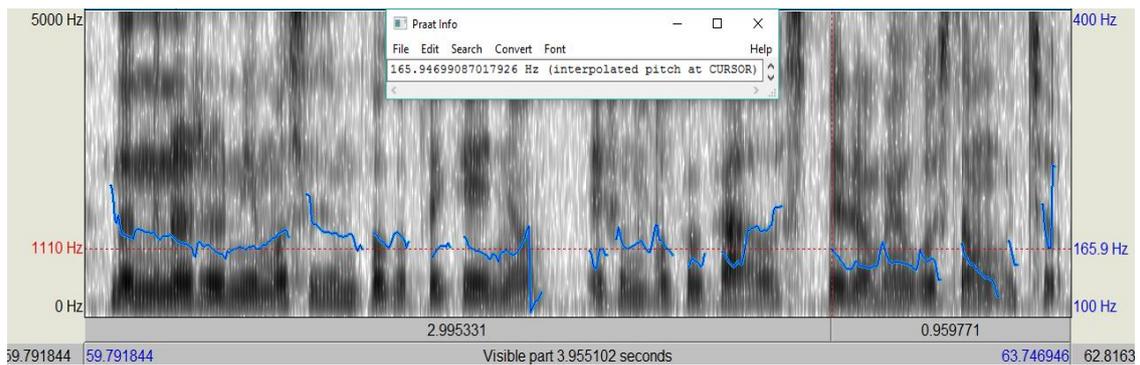


Voz número 9 NTL, 202Hz 1



Voz número 6 MDM, 200Hz 1

Por lo tanto, el tono medio de la voz no parece el rasgo más característico como para confundirlas ya que existe una diferencia de 20Hz entre la voz número 3 y las otras. Si analizamos el tono medio de las voces que no se han confundido, se observa que la voz número 5 tiene un tono de 186Hz y la voz número 8, 166Hz.

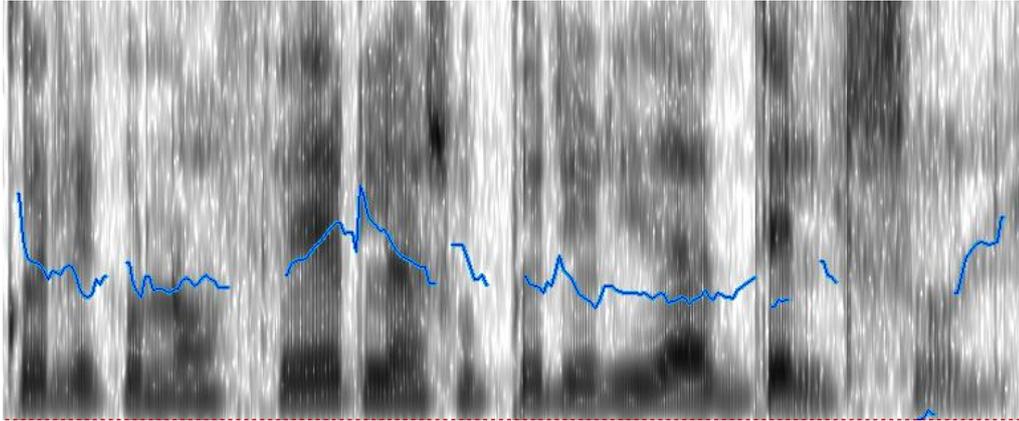


Voz número 8 MRT, 166Hz 1

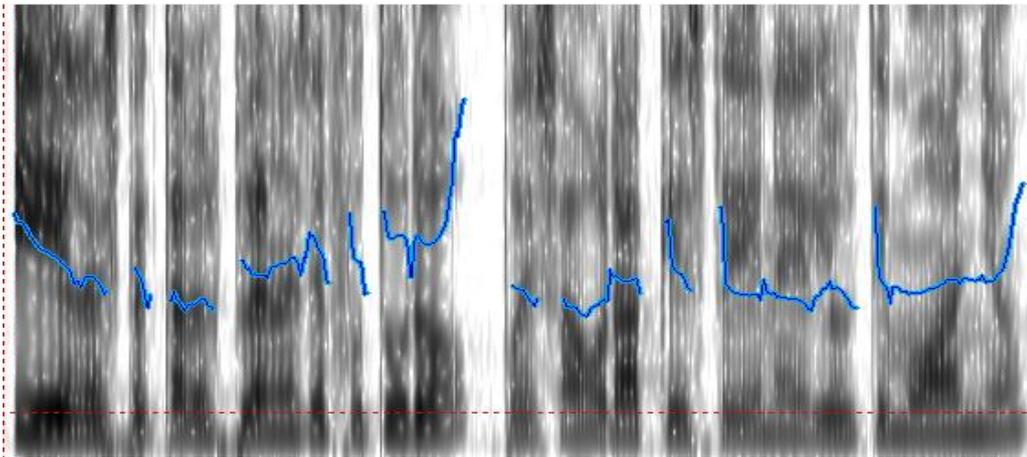
En estos casos sí que observamos que son voces mucho más graves que la indubitada y que, por lo tanto, es un posible rasgo para descartarlas.

En segundo lugar, si analizamos el ritmo –en los espectrogramas anteriores- tanto la voz indubitada como las que se confunden tienen un ritmo parecido bastante más rápido que las voces que no se confunden (voz 5 LRD y voz 8 NTL) puesto que su ritmo es más lento.

En tercer lugar, si analizamos el acento, observamos que EDL (voz 3) siempre tiene picos en la sílaba postónica, de la misma manera que NTL (voz 9) y MDM (voz 6), las tres voces presentan como característica que suele ser la última sílaba la que elevan:

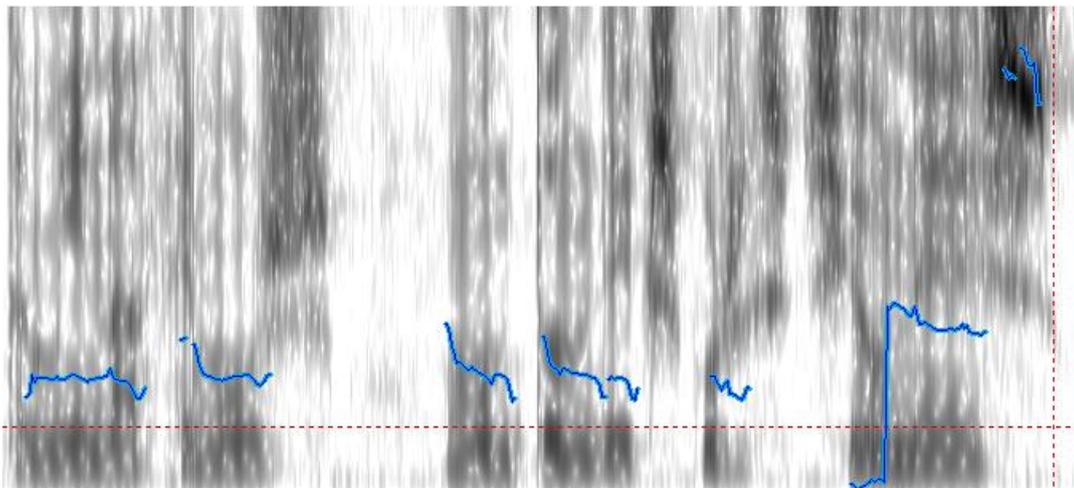


Voz número 3 EDL, 220Hz b"(re)conocer la superioridad del sol"



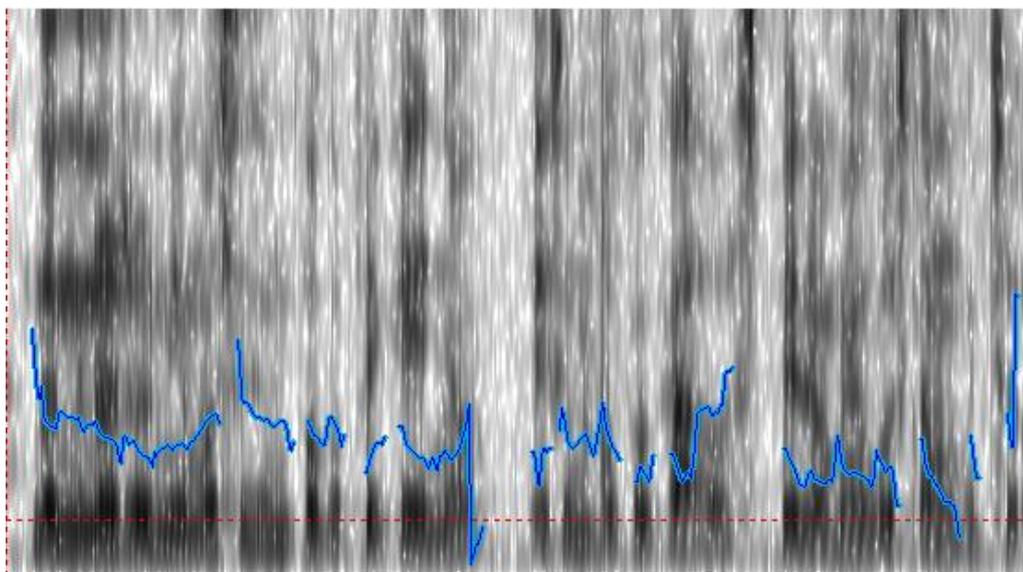
Voz número 9 NTL 1"y respecto a los planes de futuro ahora que estoy en el último año"

Por otra parte, las voces que no se confunden como la de LRD o MRT presentan rasgos distintos. En la voz número 5 LRD, se observa que las tónicas no están marcadas en exceso, tan sólo la sílaba final de oración están elevadas, en el espectrograma siguiente se puede observar un ascenso en la sílaba postónica final.



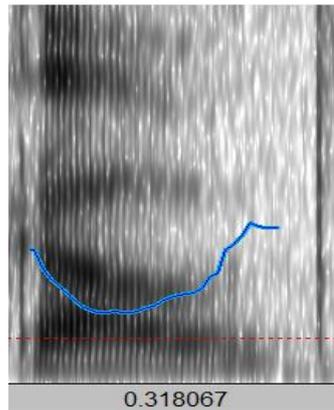
Voz número 5 LRD 1" También hacemos un montón de exposiciones"

En el caso de la voz número 8 MRT, tiene los mismos patrones que las voces confundidas y la voz indubitada. Los picos están situados en las postónicas como se puede observar en el siguiente sonograma:



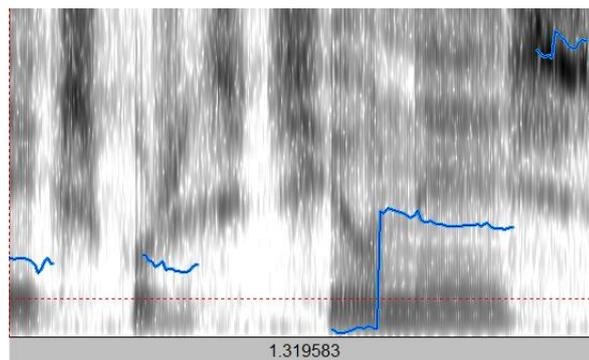
Voz número 8 MRT 1 "Donde nos reunimos es un poco pequeño, es un poco caos pero me gusta mucho"

Por lo que hace respecta a los elementos segmentales se observado que los más característicos son los sonidos fricativos y oclusivos sordos. Para empezar, la voz oída en primer lugar, número 3 EDL, es una voz bastante estándar, en algún punto del discurso parece que se nota una [t̚]. En el caso de la voz 9 NTL, ocurre lo mismo que con la 3 y es que es una voz muy estándar con una realización de fonemas de un español neutro. En cambio, es sorprendente que los jueces pudieran confundir la voz 3 con la voz 6 de MDM puesto que esta sí contiene realizaciones fonético-fonológicas por influencia del catalán. En la voz 6 encontramos [t̚], [s̺] y alguna fricativa alveolar sonora [z] en posición de la sorda [s] como en el caso de “proce[s]iones-proce[z]iones”, por influencia catalana, y también se encuentra otro rasgo característico que utiliza la fricativa palatal sorda en vez de la africada dental palatal sorda [t̚] por [t̚j̥]. Las voces que ningún juez ha considerado como la voz 8 MRT, presenta también una fuerte “contaminación” del catalán puesto que además de realizar [t̚], se encuentran oclusivas dentales sordas [t] en posición final de palabra :



1Voz 8 NTL, "escout"

En referencia a la voz 5 LRD, comentábamos en apartados anteriores que se trata de una voz distractora puesto que presenta rasgos fonético-fonológicos del italiano como la realización de una africada dental-alveolar sonora [tʃ] intervocálica donde corresponde una fricativa dental sorda [θ], por ejemplo, en vez de “exposi[θ]iones” ella realiza “exposi[tʃ]iones.



2Voz 5 LRD "esposiciones"

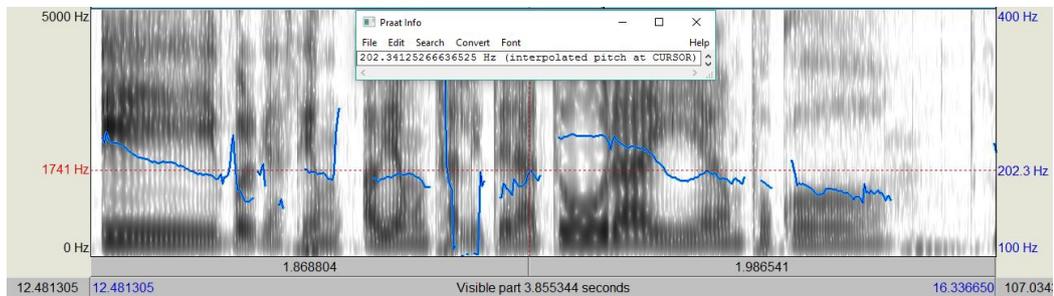
Por lo tanto, era de esperar que ningún juez confundiera esta voz con la voz indubitada, oída por primera vez.

7.2.2. Rueda de reconocimiento 2:

En la segunda rueda de reconocimiento la voz indubitada era la número 10 [NTL] en registro espontáneo. Sin embargo, recordemos que este test no tuvo demasiado éxito. Los jueces consideraron, casi por mayoría, que la voz correspondiente era la número 2 [ALB] (51.22%), también las voces número 6 [MR] y número 7 [SLV] obtuvieron más porcentaje que la voz indubitada (14.63% y 12.20% respectivamente). Además, aunque con un porcentaje más pequeño, la voz 10 [NTL] se confunde también con la voz 5 [EDL] tal y como ocurría en la primera

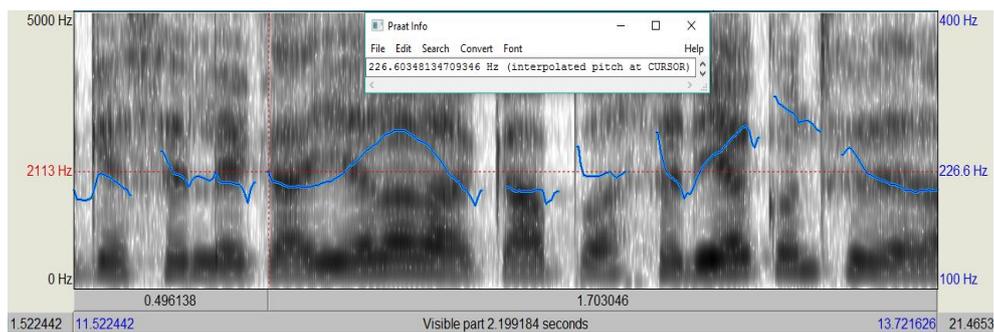
rueda de reconocimiento. Por otra parte, ningún juez consideró que se trataba ni de la voz número 1 [GSL], ni la 3 [LRD], ni la 4 [ABM] o la voz número 9 [MRT].

En primer lugar si se observan las características acústicas de estas voces, el tono medio de la voz 10 NTL es de 202Hz.

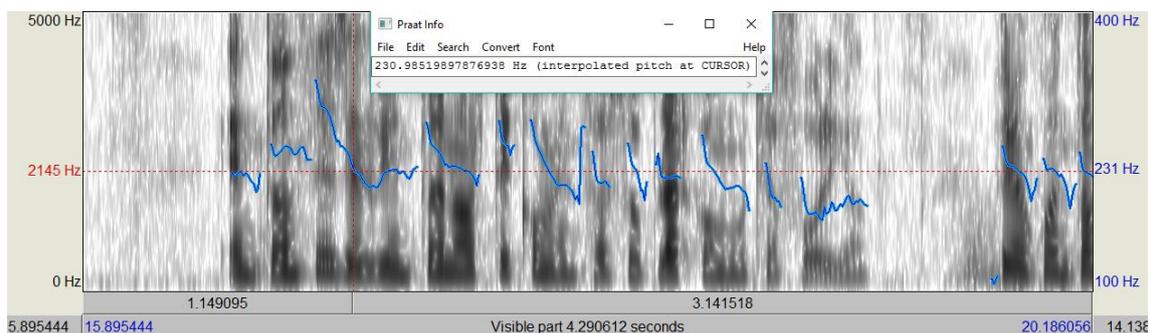


Voz 10 NTL, 202Hz 1

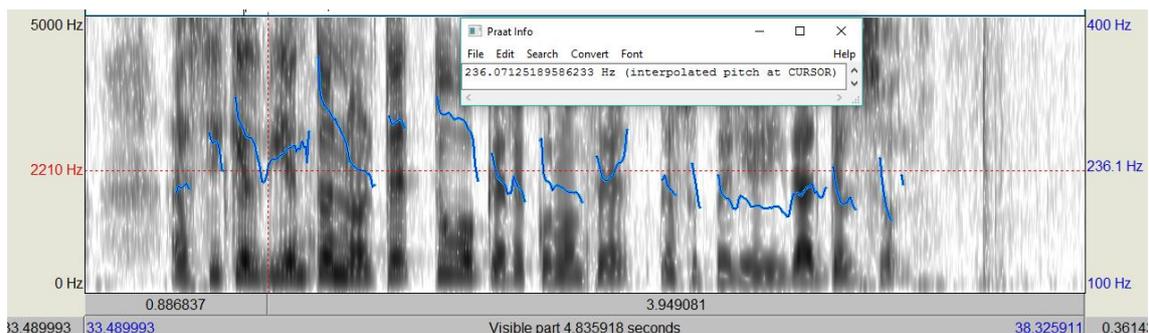
Mientras que el tono medio de las voces, en registro formal, es de 226Hz para la voz número 2 [ALB] y 230Hz para la voz 6 [MR] y 236.5Hz para la voz número 7 [SLV].



Voz 2, ALB 226Hz 1

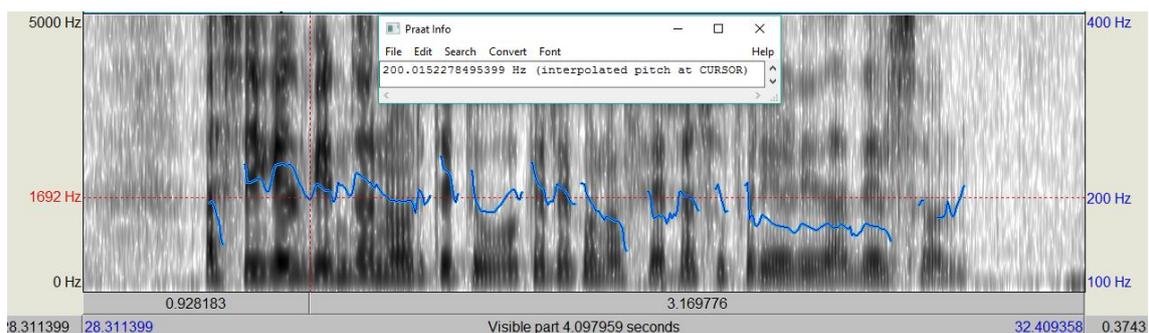


Voz 6 MR, 230Hz 1



Voz 7 SLV, 236Hz 1

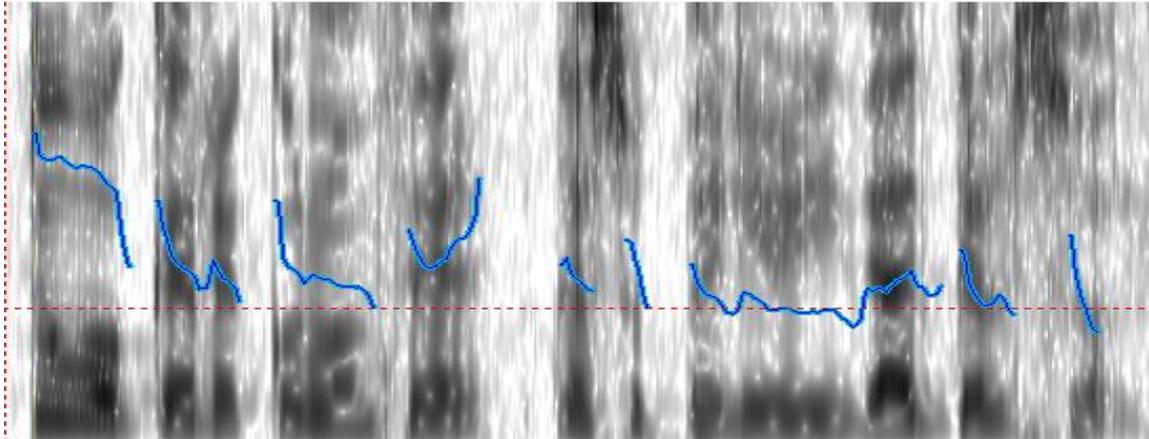
Una vez más los tonos de las voces confundidas tienen una diferencia de entre 20 y 30Hz más agudos que la voz indubitada. Si analizamos las voces que ningún juez confundió observamos que son más graves pero se acercan más a la voz indubitada ya que la voz número 3 [LRD] tiene un tono medio de 192Hz y de 200 Hz para la voz número 9 [MRT] –que es la que más parecida en cuanto a tono medio de voz es a la voz indubitada.



Voz 9 MRT, 200Hz 1

En segundo lugar, si analizamos el ritmo, se observa que ocurre lo mismo que en el caso anterior. Las voces que se confunden junto con la voz 9 [MRT] tienen un ritmo ágil y rápido –[ALB] una milésima más lenta, mientras que la voz 3 [LRD] es más lenta, por lo que podemos deducir que influye la lengua materna.

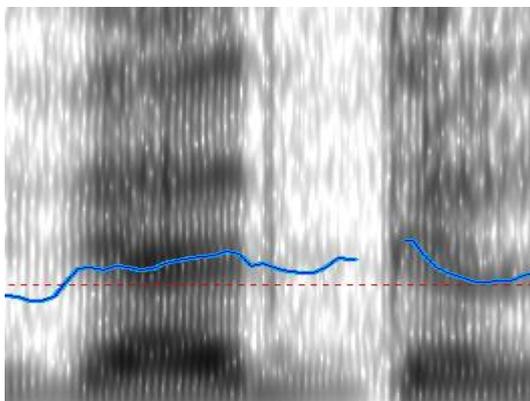
En tercer lugar, si analizamos el acento la voz 10 [NTL] siempre tiene picos en la postónica como ya analizamos en el apartado anterior. Las voces 2 [ALB] y 6 [MR] coinciden con este patrón y también realizan picos en las sílabas postónicas, en cambio, la voz 7 [SLV] que también es una de las confundidas, los picos tonales se encuentran en la sílaba tónica.



Voz 7 SLV, 236Hz 2"Tuvo que reconocer la superioridad del sol"

Lo mismo ocurre con la voz 3 [LRD] pero, en cambio, una de las voces que no se consideró por parte de los jueces, la 9 [MRT], tiene el mismo patrón que la voz 10 –los picos tonales se encuentran en las sílabas postónicas.

Los elementos segmentales que se han observado en esta rueda de reconocimiento muestran que la voz 10 [NTL], como se ha comentado en el análisis anterior, se trata de una voz muy neutra tanto en registro espontáneo como formal. Sin embargo, las voces que han sido confundidas con esta tienen rasgos dialectológicos. Por ejemplo, la voz 2 [ALB] que fue la más considerada es una voz forzada y un tanto nasalizada donde se encuentran [t̪] propias del catalán y una realización un poco exagerada de la vibrante [r]. La voz 6 [MR] es bastante más estándar que las demás y sólo en algún caso encontramos [t̪]. La voz 7 tiene rasgos fuertes en este sentido, puesto que además de las laterales velarizadas, las vocales son más cerradas y se encuentran oclusivas dentales sordas [t] a final de palabra como en el caso de “superioridad del sol” que lo realiza “superiorida[t d]el sol”



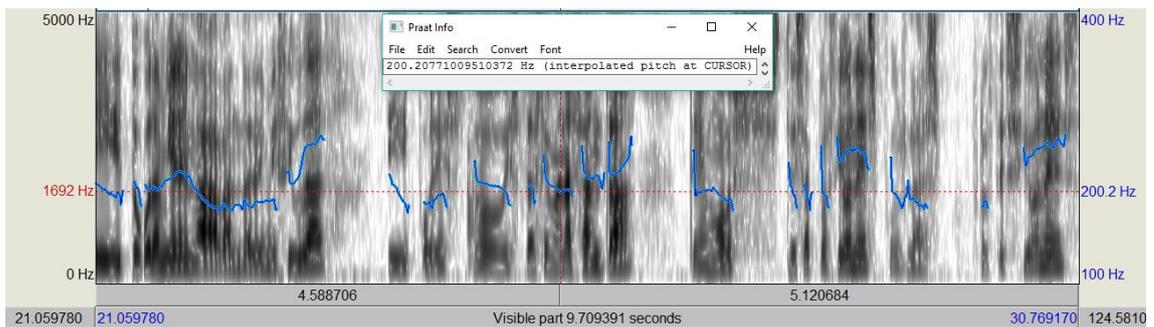
Voz 7 SLV, 236Hz 3"superiorita[t] del sol"

Una de las voces descartadas, diferente a las ya comentadas, es la voz 9 [MRT] que presenta una realización muy neutra como la voz indubitada de [NTL]. No presenta ningún rasgo fonético o fonológico característico.

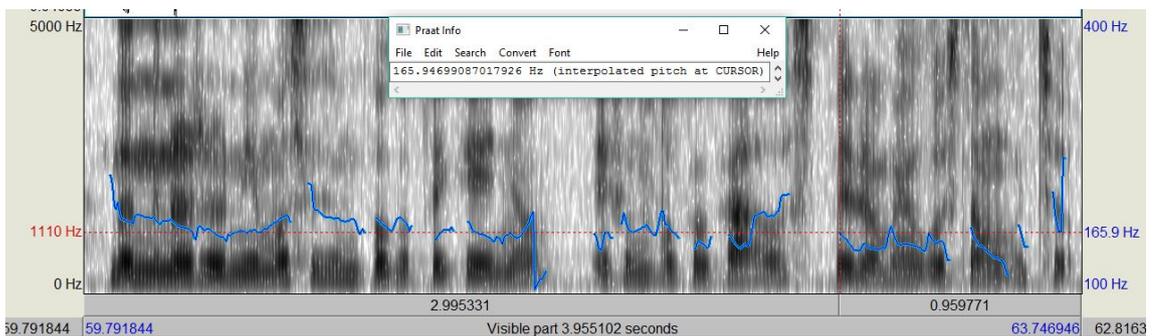
7.2.3. Rueda de reconocimiento 3:

En el tercer test perceptivo la voz indubitada era la número 8 [MDM] –enmascarada con ruido *party*-. El 40% de los jueces la identificó correctamente. Por otra parte, hubo jueces que la confundieron con la voz número 9 [MRT] (20%), también con la número 3 [LRD] (13.33%) y la voz número 4 [ABM] (10%). Sin embargo, ningún juez consideró que se tratase de la voz número 5 [EDL], ni la voz número 7 [SLV].

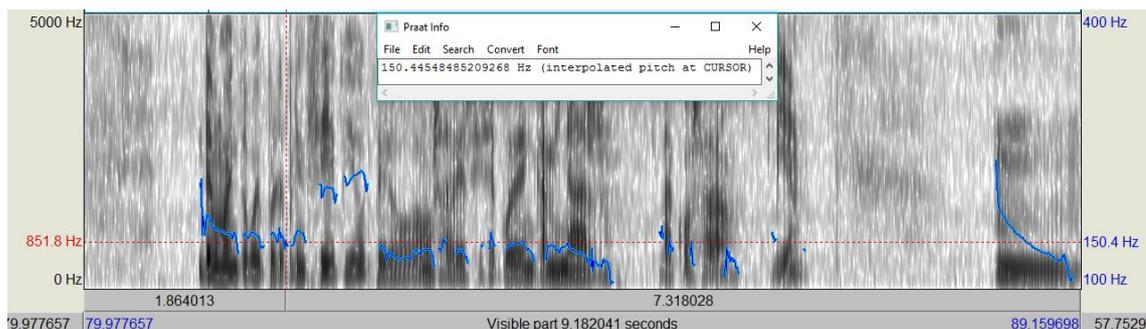
Si observamos las características acústicas de las voces de esta rueda de reconocimiento, la voz indubitada 8 [MDM] tiene un tono medio de 216Hz, las voces confundidas tienen un tono medio de 165Hz la voz 9 [MRT], 186Hz la voz 3 [LRD] y 150Hz la voz 4 [ABM].



Voz 8 MDM, 200Hz 1

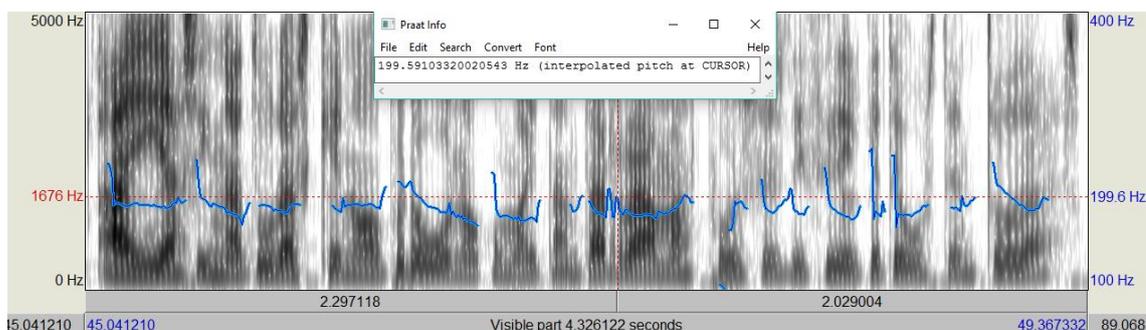


Voz 9 MRT, 166Hz 1

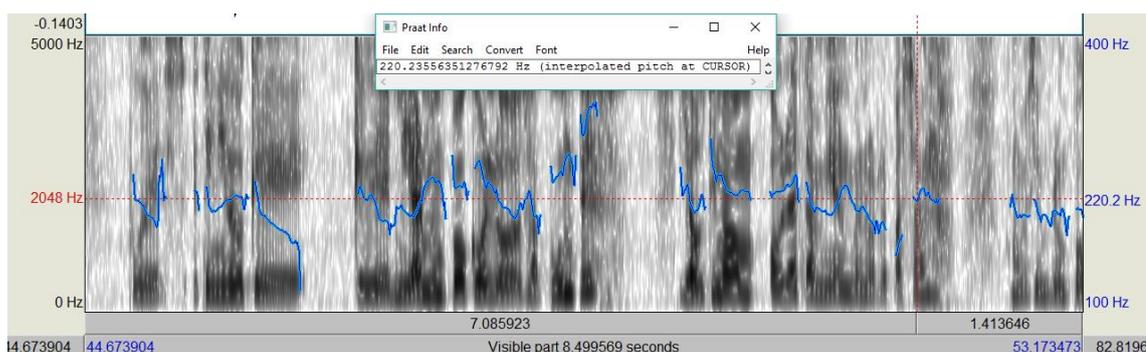


Voz 4 ABM, 150Hz 1

En cambio, las voces que no han sido considerada por los jueces como la número 5 [EDL] y la voz 7 [SLV] tienen un tono medio de 200Hz y 220Hz respectivamente. Como se observa la voz 5 [EDL] tiene el mismo tono medio que la voz 8, oída en primer lugar, pero ningún juez la ha considerado. El resto de voces tiene entre 50Hz más graves y 20Hz por encima en el caso de la voz 7 [SLV].



Voz 5 EDL, 200Hz 1

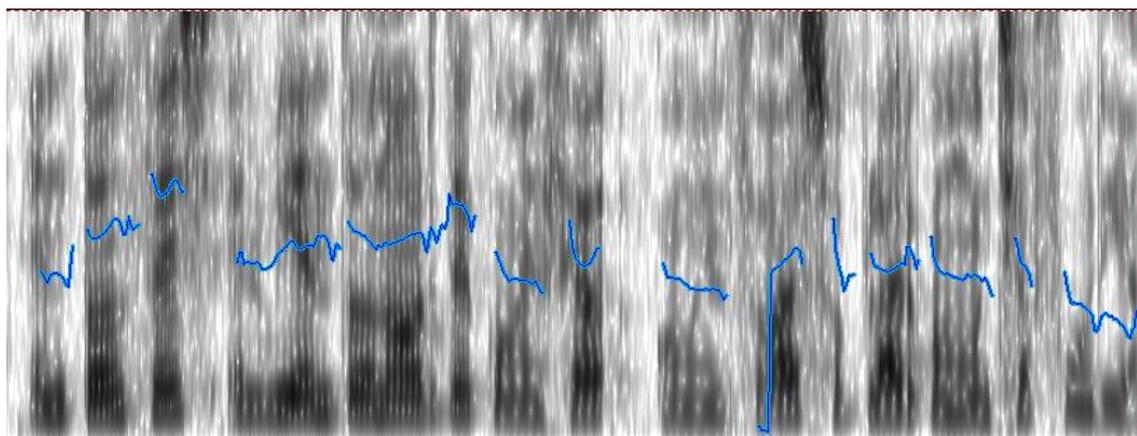


Voz 7 SLV, 220Hz 1

Una vez más se observa, que aunque el tono medio de la voz sea algo que condiciona mucho a los jueces en el momento de identificar una voz, no es el único rasgo característico puesto que voces muy dispares –en cuanto a tono medio- pueden ser confundidas y voces muy similares descartadas.

En segundo lugar, el ritmo en esta rueda de reconocimiento coincide en que todas las voces en registro espontáneo tienen un ritmo más lento excepto la voz 7 [SLV] que es más rápida.

En cuanto al acento, la voz indubitada, voz 8 [MDM] tiene el mismo patrón que las voces indubitadas de los otros test perceptivos. El pico tonal se encuentra en posición postónica.



Voz 8 MDM, 200Hz 2 "Entonces, el viento norte comenzó a soplar con mucha furia"

Por lo que respecta a las demás voces casi todas coinciden en tener el mismo patrón –los picos se encuentran en las sílabas postónicas- exceptuando la voz 3 [LRD] que aunque fue una de las que los jueces confundieron los picos tonales se encuentran en las sílabas tónicas pese a que no estén marcados en exceso.

Por último, por lo que respecta al análisis de los elementos segmentales encontramos que la voz 8 [MDM] es una voz un poco distractora puesto que presenta rasgos fonológicos fuertes del catalán. Como hemos comentado en el anterior test perceptivo, la voz de [MDM] es una voz que presenta realizaciones de fricativas apicales sonoras [ʒ] en vez de fricativas alveolares sordas, se encuentran oclusivas dentales sordas [t] a final de palabra, [t̚] ... Así pues, es normal que los jueces se confundieran con la voz 9 [MRT] puesto que presenta los mismos rasgos. En cambio, aquellos jueces que confundieron la voz 8 con la voz 3, podemos suponer que es a causa de las fricativas alveolares sonoras en posición intervocálica que realizan ambas informantes. La voz de [ABM] es bastante neutra pero como hemos observado en los resultados, hubo jueces que la confundieron con la oída primero. Cabe destacar que las voces que no fueron identificadas por ningún juez son voces casi neutras puesto que se

encuentran puntualmente [t] velarizadas tanto en la voz 5 [EDL] como en la voz 7 [SLV].

Por lo tanto, de estos resultados podemos deducir que uno de los mayores motivos por los que intuimos o creemos reconocer una voz es por su tono medio, además de valores añadidos como el ritmo o las realizaciones de algunos segmentos fonéticos.

8. DISCUSIÓN Y CONCLUSIÓN

Valorando los resultados del presente trabajo, se puede concluir que las dos hipótesis planteadas en el inicio del trabajo han sido corroboradas.

Por una parte, ha quedado demostrado que es más fácil reconocer e identificar una voz indubitada en registro formal entre voces en discurso espontáneo, que no al contrario, puesto que existe la tendencia de elevar el tono medio de la voz en registro formal y se estandariza más impidiendo reconocer todos los matices acústicos que existen cuando se habla de manera espontánea.

Por otra parte, se ha verificado que no es imposible reconocer una voz enmascarada con ruido entre un conjunto de voces sin enmascarar pero los resultados indican que hay una cierta dificultad –menor que reconocer una voz en elocución espontánea entre voces en registro formal-.

Si bien es cierto que el análisis estadístico ha mostrado que existen variables significativas, como por ejemplo, en el primer test la lengua materna, en el segundo test los estudios musicales y, en el último la realización de test anteriores, también se puede observar que ninguna variable es suficientemente estadísticamente relevante en nuestro estudio, puesto que seguramente con una muestra de jueces mayor podríamos establecer afirmaciones más fundamentadas.

Por último, este estudio muestra que el conjunto de motivos propuestos para la justificación de la respuesta de la voz es indispensable para identificar una voz oída en primer lugar. Es decir, que tanto la elección de los jueces del tono medio de la voz, como de su ritmo u otros aspectos como la realización específica de fonemas, por ejemplo, fricativos, caracterizan las voces que oímos.

9. REFERENCIAS BIBLIOGRÁFICAS

- Alcaraz, E., & Turell, M. T. (n.d.). • Lingüística general : – Modelos descriptivos • Lingüística forense : • Lenguaje Jurídico – Language of the Law • Lenguaje Judicial – Language of the Court & the Legal Process • Lenguaje Probatorio o Evidencial – Language as Evidence.
- Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glott International*
- Boersma, P. and D. Weenink (2001): «Praat, a system for doing phonetics by computer», *Glott International*, 5, pp. 341-345. 5:9/10, 341-345.
- Braz, A. (n.d.). La lingüística forense i l'interrogatori policial a Espanya: un enfocament lingüístic de les pràctiques actuals. *Llengua Societat I Comunicació*, (9), 5–11.
- Cerdà, R. (2008). Sobre alguns aspectes contraposats en fonètica forense. *Estudios de Fonética Experimental*, 17, 45–64.
- Cerdà, R. (2009). Camps de dispersió vocàlica en imitacions de veu: primers indicis d'un experiment sobre identificació de locutor. *Estudios de Fonética Experimental*, 18, 65–88. Retrieved from <http://www.raco.cat/index.php/EFE/article/view/140090>
- Cerdà, R. (2011). Creus que la teva veu és única? *Llengua Societat I Comunicació*, 9, 33–41. Retrieved from <http://www.ub.edu/cusc>
- Cicres, J. (2011). Transcripció i autenticació de gravacions en contextos judicials. *Llengua, Societat I Comunicació. Revista de Sociolingüística de La Universitat de Barcelona*, 9(9), 23–62. Retrieved from <http://revistes.ub.edu/index.php/LSC/article/view/3296>
- Colantoni, L. (2008). Variación micro y macro fonética en español. *Estudios de Fonética Experimental*, 17, 66–104. Retrieved from <http://www.raco.cat/index.php/EFE/article/view/140066/191245>
- Costell Jaime, M. (2014). Influencia del sexo del identificador en el reconocimiento de habla.
- Fernandez Pons, X. (2014). Anàlisi de la variació intra i inter parlant en la fricativa alveolar de la /s/ del castellà amb finalitat forense.
- Fernández, Planas, A. M. (2011). Introducció : lingüística forense. *Llengua Societat I Comunicació*, (9), 1–4.
- Hollien, H. (2002) *Forensic Voice Identification*. San Diego: Academic Press. Pp. 91 ss.
- Nolan, F. (2009). The phonetic bases of speaker recognition. *Cambridge Studies in Speech Science and Communication*. Cambridge University Press.
- Queralt, S., Spassova, M., & Turell, M. T. (2011). L'ús de les combinacions de seqüències de

categories gramaticals com a nova tècnica de comparació forense de textos escrits.
Llengua, Societat I Comunicació, (9), 60–68. Retrieved from
<http://revistes.ub.edu/index.php/LSC/article/view/3300>

Soler, Sabater, M. J.(1984). Experimentos sobre la percepci3n del acento. *Estudios de Fon3tica Experimental*, 1, 131–242. Retrieved from
<http://www.raco.cat/index.php/EFE/article/view/144192/195890>

Turell, M. T. (2004). La Tasca Del Lingüística Detectiu En Casos De Detecció De Plagi I Determinació D’Autoria De Textos Escrits , (1994), 69–85.

10. ANEXO

10.1. Tablas de Edad de los jueces.

Jueces Test perceptivo 1	Edad	Jueces test perceptivo 2	Edad	Jueces test perceptivo 3
11	22	13	22	5
7	23	6	23	5
5	25	2	25	3
0	20	0	20	1
4	21	4	21	0
4	26	1	26	1
0	27	2	27	2
4	40	0	40	0
3	29	0	29	2
2	28	1	28	0
2	32	0	32	0
2	33	0	33	0
2	38	0	38	0
2	44	0	44	0
2	46	1	46	0
2	52	0	52	0
2	54	0	54	0
2	47	0	47	1
1	19	0	19	0
1	24	6	24	2
1	30	0	30	1
1	31	0	31	0
1	34	0	34	2
1	35	0	35	0
1	36	0	36	1
1	37	0	37	0
1	39	0	39	0
1	42	0	42	0
1	43	0	43	0
1	45	0	45	0
1	50	0	50	1
0	52	1	52	1
1	55	0	55	0
1	57	0	57	1
0	58	1	58	1
0	16	1	16	0
0	59	1	59	0

Tabla 42Edades jueces participantes en los 3 test perceptivos.

10.2. Orden voces test perceptivos.

Orden rueda de reconocimiento 1:

Voz 1 [ALB]

Voz 2 [ABM]

Voz 3 [EDL]

Voz 4 [GSL]

Voz 5 [LRD]

Voz 6 [MDM]

Voz 7 [MR]

Voz 8 [MRT]

Voz 9 [NTL]

Voz 10 [SLV]

Orden ruedas de reconocimiento 2 y 3:

Voz 1 [GSL]

Voz 2 [ALB]

Voz 3 [LRD]

Voz 4 [ABM]

Voz 5 [EDL]

Voz 6 [MR]

Voz 7 [SLV]

Voz 8 [MDM]

Voz 9 [MRT]

Voz 10 [NTL]

10.3. Test 1

Rueda de reconocimiento

PASO 1.

Antes de empezar permíteme saber un poco más de ti

1. Sexo: *

- Mujer
- Hombre
- Otro

2. Edad: *

3. Lengua materna: *

- Español
- Catalán
- Castellano y catalán
- Catalán y castellano
- Otros- especifica, por favor,

4. ¿Tienes estudios musicales? *

Sí

No

5. ¿Sufres o has sufrido problemas auditivos? *

Sí

No

NS/NC

6. ¿Has estudiado o estudias en la universidad? *

Sí

No

7. Por último, ¿estás utilizando auriculares o altavoces? *

Auriculares

Altavoces

8. ¿Cursas o has cursado alguna asignatura de fonética? *

Sí

No

Next

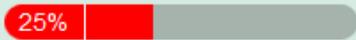
Escucha atentamente la siguiente voz.

No es importante el contenido (es decir, qué dice).



Back

Next



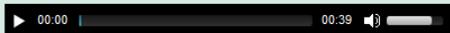
¿Ya la has escuchado bien?

De acuerdo, pues a continuación vas a encontrar 10 archivos de voces femeninas recitando el mismo texto. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

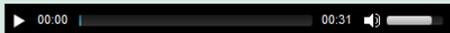
Voz 1



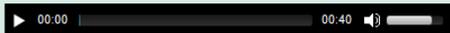
Voz 2



Voz 3



Voz 4



Voz 5



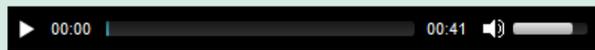
Voz 6



Voz 7



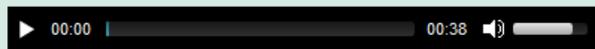
Voz 8



Voz 9



Voz 10



Back Next



9. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

- Voz 1
- Voz 2
- Voz 3
- Voz 4
- Voz 5
- Voz 6
- Voz 7
- Voz 8
- Voz 9
- Voz 10

10. ¿En qué te has basado? *

- Ritmo
- Tono medio de la voz
- Entonación
- Acento
- Otro - especifica: *

Otro - especifica: *

NS/NC

Back

Submit

75%

Rueda de reconocimiento

¡Muchas gracias por tu participación!

iiiiiiFIIIIIN!!!!!!

Thank you for taking our survey. Your response is very important to us.

100% 

Rueda de reconocimiento 2

PASO 1.

Antes de empezar permíteme saber un poco más de ti

1. Sexo: *

- Mujer
- Hombre
- Otro

2. Edad: *

3. Lengua materna: *

- Español
- Catalán
- Castellano y catalán
- Catalán y castellano
- Otros- específica, por favor,

4. ¿Tienes estudios musicales? *

- Sí
- No

5. ¿Sufres o has sufrido problemas auditivos? *

- Sí
- No
- NS/NC

6. ¿Has estudiado o estudias en la universidad? *

- Sí
- No

7. Por último, ¿estás utilizando auriculares o altavoces? *

Auriculares

Altavoces

8. ¿Cursas o has cursado alguna asignatura de fonética? *

Sí

No

9. ¿Hiciste el test anterior a este? *

Sí

No

Next

0%

Rueda de reconocimiento 2

Escucha atentamente la siguiente voz.

No es importante el contenido (es decir, qué dice).



Back

Next

25%

Rueda de reconocimiento 2

¿Ya la has escuchado bien?

De acuerdo, pues a continuación vas a encontrar 10 archivos de voces femeninas recitando el mismo texto. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

Voz 1



Voz 2



Voz 3



Voz 4



Voz 5



Voz 6



Voz 7



Voz 8



Voz 9



Voz 10



Back

Next

25%

Rueda de reconocimiento 2

10. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

- Voz 1
- Voz 2
- Voz 3
- Voz 4
- Voz 5
- Voz 6
- Voz 7
- Voz 8
- Voz 9
- Voz 10

11. ¿En qué te has basado? *

- Ritmo
- Tono medio de la voz
- Entonación
- Acento
- Todas las anteriores
- Otro - especifica: *
- NS/NC

Back

Submit

75%

Rueda de reconocimiento 2

¡Muchas gracias por tu participación!

iiiiiiFIIIIIN!!!!!!

Thank you for taking our survey. Your response is very important to us.

100%

Rueda de reconocimiento 3

PASO 1.

Antes de empezar permíteme saber un poco más de ti

1. Sexo: *

- Mujer
- Hombre
- Otro

2. Edad: *

3. Lengua materna: *

- Español
- Catalán
- Castellano y catalán
- Catalán y castellano

4. ¿Tienes estudios musicales? *

Sí

No

5. ¿Sufres o has sufrido problemas auditivos? *

Sí

No

NS/NC

6. ¿Has estudiado o estudias en la universidad? *

Sí

No

7. Por último, ¿estás utilizando auriculares o altavoces? *

Auriculares

Altavoces

8. ¿Cursas o has cursado alguna asignatura de fonética? *

- Sí
- No

9. ¿Has hecho o hiciste los tests anteriores a este?

- Sí
- No

Next

0%

Rueda de reconocimiento 3

Escucha atentamente la siguiente voz.

No es importante el contenido (es decir, qué dice).

.



Back

Next

25%

Rueda de reconocimiento 3

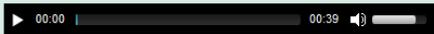
¿Ya la has escuchado bien?

De acuerdo, pues a continuación vas a encontrar 10 archivos de voces femeninas recitando el mismo texto. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

Voz 1



Voz 2



Voz 3



Voz 4



Voz 5



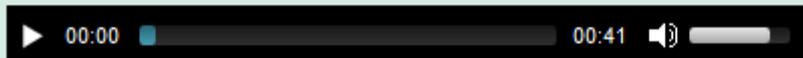
Voz 6



Voz 7



Voz 8



Voz 9

Voz 8



Voz 9



Voz 10



Back

Next

50%

Rueda de reconocimiento 3

10. ¿Sabrías decir cuál de esas 10 voces es la que acabas de escuchar?

- Voz 1
- Voz 2
- Voz 3
- Voz 4
- Voz 5
- Voz 6
- Voz 7
- Voz 8
- Voz 9
- Voz 10

11. ¿En qué te has basado? *

Ritmo

Tono medio de la voz

Entonación

Acento

Otro - especifica:

*

NS/NC

Back

Submit

75%

Rueda de reconocimiento 3

¡Muchas gracias por tu participación!

iiiiiiFIIIIIN!!!!!!

Thank you for taking our survey. Your response is very important to us.

100%