

Análisis y cotejamiento de emisiones lingüísticas en idiomas inglés-español con el sistema AVIS+F para identificar hechos delictivos

Analysis and comparison of linguistic emissions in English-Spanish languages with the AVIS+F system to identify criminal acts

Carlos Francisco Osorio Vega¹

Roberth Smith Talavera Ayala²

Christian Fabián Carrazco Cali³

Recibido: 5 de marzo de 2021

Aceptado: 20 de mayo de 2021

Publicado: 28 de junio de 2021


Resumen


El análisis y cotejamiento de las emisiones lingüísticas es una de las experticias más importantes dentro del campo criminalístico ya que permite identificar de manera indubitable la identidad de una persona por medio de su voz (huella vocal). La presente investigación conlleva a la obtención de resultados fehacientes y técnicamente demostrables. Dicho análisis se realiza aplicando, primero, el método auditivo mediante el cual el perito escucha los archivos de audio. Posterior a esto, se aplica el método instrumental en la cual la técnica es la comparación espectrográfica de formantes. El trabajo desplegado en el presente proyecto, de índole técnico, amerita la utilización de equipos forenses especializados para la obtención de los resultados esperados. De este modo, se emite una conclusión determinante en lo que respecta a la identificación de una persona a través de su voz.


Palabras claves: análisis, comparación, cotejamiento, determinante, formantes, método, fonación

Abstract

The analysis and comparison of linguistic emissions is one of the most important expertise within the criminalistics field, since it allows us to undoubtedly identify the identity of a person through their voice (voice print). The present investigation in itself leads us to obtain reliable and technically demonstrable results. For the analysis and comparison of linguistic emissions, an objective methodology is applied, likewise the auditory method is used by which the expert listens to the audio files and after this the instrumental method is applied where the technique is the spectrographic comparison of formants. The

1 Teniente de la Policía Nacional del Ecuador, ingeniero en Informática y Ciencias de la Computación, carlos.osorio@policia.gob.ec
 <https://orcid.org/0000-0003-1945-6465>.

2 Sargento Primero de la Policía Nacional del Ecuador, tecnólogo en Criminalística, rstalavera@hotmail.com
 <https://orcid.org/0000-0002-0913-9258>.

3 Cabo Primero de la Policía Nacional del Ecuador, tecnólogo en Criminalística, christian.carrazco@hotmail.com
 <https://orcid.org/0000.0002-5406-9703>.

work carried out within this investigative project is technical, which warrants the use of specialized forensic teams to obtain the expected results and thus be able to issue a decisive conclusion regarding the identification of a person through their voice.

Keywords: Analysis, comparison, collation, determinant, formants, method, phonation

Introducción

Con el avance de las tecnologías digitales (teléfonos celulares, redes sociales) han proliferado las actividades delictivas que utilizan estos medios para perpetrar amenazas, extorsiones, secuestros y otros delitos. Para la administración de justicia es fundamental la identificación de las personas que realizan este tipo de actos atentatorios. No obstante, existen personas que cometen actos ilícitos empleando el idioma inglés para comunicarse, pero, al momento de ser investigados emplean el idioma español para así confundir a los administradores de justicia y evadir responsabilidades.

Dentro del trabajo técnico científico que realiza la Dirección Nacional de Investigación Técnica Científica de la Policía Nacional del Ecuador (Dinitec), no ha existido, aún, un antecedente de análisis y cotejamiento de voz de una misma persona que domina dos idiomas diferentes, inglés-español. Es por ello que, la presente investigación implementa el cotejamiento de voces en los idiomas inglés-español, aplicando el método auditivo y espectrográfico para la identificación de interlocutores y considerando ciertos factores importantes que intervienen en este tipo de cotejos. Se debe tener en cuenta que el análisis auditivo de un archivo de audio es imprescindible, ya que por medio del oído el perito detecta ciertas características importantes de la voz como el tono, el timbre, la intensidad, la rapidez y otros. Estos aspectos hacen que la voz de una persona sea única.

El análisis instrumental es el complemento del análisis auditivo ya que con la utilización de las herramientas del Sistema Biométrico

Multimodal de Identificación de Voz e Imagen Facial (AVIS+F) mediante el *software* SIS II, se realiza el análisis espectral del sonido, demostrando de una manera técnica la identidad de una persona mediante su voz. Para lograrlo, el perito debe seguir un proceso minucioso desde el momento que obtiene la evidencia y hasta el instante que expone sus conclusiones. Si omite algún paso los resultados obtenidos no tendrán validez técnica. Por tal razón, el experto debe obtener los doce espectrogramas resultantes de los fonemas vocálicos de los archivos dubitado e indubitado para tener la certeza de que se trata de la misma persona.

Las pericias de análisis y cotejamiento de voces en los últimos años han sido muy importantes para los administradores de justicia ya que les permite tener elementos de convicción para dictaminar sus sentencias. En consecuencia, los peritos requieren la actualización permanente para elevar sus conocimientos y que su aplicación sea satisfactoria ante cualquier evento delictivo. Es necesario que los peritos conozcan el alcance de los equipos forenses que se utilizan para este tipo de pericias.

Metodología

El objetivo de este estudio es realizar el análisis y cotejamiento de emisiones lingüísticas en los idiomas inglés-español con el sistema AVIS+F el cual se apoya en el método científico a partir de la experimentación controlada, la identificación de las características de la voz y la búsqueda de generalizaciones por medio de la técnica de la entrevista para obtener las muestras biométricas de voz (elementos indubitados). Esto se realiza en un ambiente controlado y con los equipos técnicos acordes que permitan receptar una muestra de voz óptima. Posteriormente, se aplica el método auditivo y de comparación espectrográfica de formantes. Abordados los principales parámetros de la metodología es necesario tener en cuenta las siguientes definiciones:

Sonido

Sensación producida en el órgano del oído por medio de los movimientos vibratorios de los

cuerpos, transmitido por un medio elástico como el aire.⁴

Frecuencia

Número de oscilaciones o variaciones de la presión por segundo. La frecuencia es una magnitud objetiva y mensurable referida a formas de onda periódicas. Tiene que ver con el número de ciclos por segundo que tiene que dar la onda e indicando la idea de rapidez con que se producen.⁵

Amplitud

El sonido se mueve a través del aire en forma de ondas. La amplitud es la distancia entre el punto más alto y el más bajo de una onda. La amplitud del sonido es la intensidad (lo que comúnmente llamamos “el volumen”).⁶

Tiempo

El tiempo es una dimensión que puede referirse a un momento determinado, una época o estación del año, incluso, al clima de un lugar en relación al estado atmosférico.⁷

Aparato fonador

Conjunto de órganos del cuerpo humano encargado de generar y ampliar el sonido que se produce al hablar. El aparato fonador está conformado por los órganos de respiración, fonación y articulación. Se debe tomar en cuenta que el sonido puede salir por dos vías tanto por vía oral o por vía nasal.⁸

Voz

Es el resultado de la función que realiza el conjunto de órganos del cuerpo humano, conocido como aparato fonador.⁹

Producción de la voz

La voz humana es producida en la laringe cuya parte esencial, la glotis, constituye el verdadero órgano de fonación humano. El aire procedente de los pulmones es forzado durante la espiración a través de la glotis haciendo vibrar los dos pares de cuerdas vocales las cuales se asemejan a dos lengüetas dobles membranáceas. Las cavidades de la cabeza, relacionadas con el sistema respiratorio y nasofaríngeo, actúan como resonadores.¹⁰

Fonología y fonética

La fonología es el estudio de los sonidos de la lengua en cuanto a su carácter simbólico o de representación mental. Por su parte, la fonética se refiere a los sonidos en el habla lo cual incluye su producción acústica y los procesos físicos y fisiológicos de emisión y articulación involucrados.¹¹

La transcripción del idioma inglés es la representación sistemática de una forma oral mediante signos escritos. Una transcripción es una adaptación mediante caracteres gráficos de una expresión de otra lengua o de un habla peculiar no estándar usando la ortografía y convenciones gráficas de una lengua o la lengua estándar. Es decir, la transcripción trata de representar fielmente los sonidos de una lengua con la ortografía de otra lengua.

La principal función del AVIS+F es la identificación de locutores a través de su voz la cual permite realizar el análisis y cotejamiento de emisiones lingüísticas mediante la utilización del software SIS II.

4 Significados, *Qué es el Sonido?*, <https://www.significados.com/sonido/>

5 P.W. López, *La frecuencia*. <https://es.slideshare.net/pattrilopezulloa/frecuencia-patt>

6 S. García Gago, *Amplitud*, (2018), <https://www.analfatecnicos.net/pregunta.php?id=5>

7 M. Porporatto, *El tiempo*, (2018), <https://quesignificado.com/tiempo/significados.com>

8 Diccionario Actual, *El aparato fonado*, (2018), <https://diccionarioactual.com/aparato-fonador/>

9 Acústica, *La voz*, (2003), <http://www.ehu.eus/acustica/espanol/musica/vohues/vohues.html>

10 *Ibid.*

11 Concepto Definición, *Fonología*, (2021), <https://concepto-definicion.de/fonologia/>

Transformada de Fourier

La transformada de Fourier es un análisis frecuencial de las señales. Es la suma de ondas simples mediante la aplicación de una fórmula matemática que descompone la frecuencia de un sonido en relación a la amplitud y el tiempo lo cual permite ubicar todas las ondas simples en un gráfico. Es decir, permite distribuir en función de la frecuencia la energía o la potencia de una señal dando por resultado el espectro.¹²

Cotejamiento de voces

Es el análisis técnico científico mediante el cual se compara una voz conocida con una voz desconocida, como una tentativa de determinar la identidad de la persona mediante su voz o, quizás, como medio para eliminar la voz de una persona como sospechoso de algún acto ilícito. Esto implica el estudio de la frecuencia, tiempo y amplitud de un espectrograma (la huella vocal).¹³

Método de análisis auditivo

En este método el experto escucha los archivos (al oído) para determinar ciertas características del archivo de audio como son el sexo del locutor, la altura de la voz, caracteres generales del sonido de la voz, fuerza, entonación, tiempo, ritmo, articulación, intermitencia del habla, maneras del lenguaje, pausas de vacilación, trastornos del lenguaje, respiración verbal y respiración del habla. Con el fin de familiarizarse con la voz y sus características es aconsejable realizar este análisis una y otra vez.

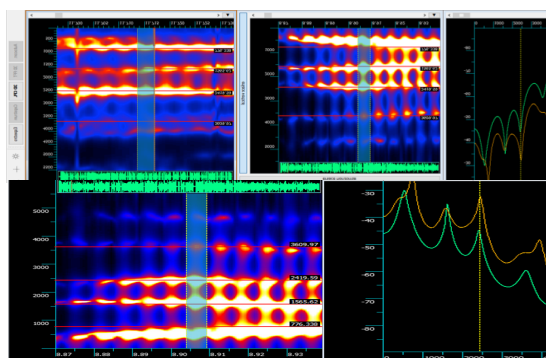
Método de comparación espectrográfica de formantes

El espectrograma es la representación de las variaciones de la frecuencia (eje vertical) y amplitud (nivel de grises) de la señal sonora a lo largo del

tiempo (eje horizontal). En este método el experto (perito) analiza y compara de manera visual los espectrogramas resultantes de las muestras dubitadas e indubitadas. Para ello se vale de cuatro formantes en las cuales las dos primeras van a representar el fonema vocálico que está siendo comparado, mientras que la tercera y cuarta formante indican de manera indubitable la identidad de las personas a partir de las características únicas de la voz. Se debe tener en cuenta que el *software* forense (SIS II) que permite visualizar los espectrogramas de manera dinámica se basa en la aplicación de la transformada de Fourier.

Gráfico 1

Espectrograma resultante obtenido del software SIS II perteneciente al sistema AVIS+F



Fuente: sistema AVIS+F, *software* SIS II

Elaboración: por los autores.

Espectrogramas resultantes de la comparación espectrográfica de formantes

Para demostrar la validez del análisis mencionado se realiza a continuación el cotejamiento de voces de los idiomas inglés/español aplicando el método auditivo y de comparación espectrográfica de formantes y utilizando el sistema AVIS+F y sus componentes.

12 Marco Martínez (12 de mayo de 2021). *¿Qué es la transformada de Fourier y para qué sirve?* Obtenido de www.nobbot.com: <https://www.nobbot.com/educacion/que-es-la-transformada-de-fourier-y-para-que-sirve/>

13 Cicres, J. (2014). *Cotejamiento de Voces*. Obtenido de scielo.conicyt.cl: <https://scielo.conicyt.cl/pdf/signos/v47n86/a02.pdf>

Tabla 1
Características de la voz

| Parámetro | Resultado |
|--|-----------------------------|
| Evidencia | Archivo dubitado 1 (inglés) |
| Archivo indubitado | Muestra de voz 1 (español) |
| Voz | Masculina |
| Intensidad | Fuerte |
| Tono o frecuencia de la voz | Grave |
| Respiración | Normal |
| Ritmo | Normal |
| Articulación, cómo el ser humano produce los sonidos del habla | Normal |

Fuente: análisis auditivo.

Elaboración: por los autores.

Carácter simbólico

Transcripción del archivo dubitado 1 (inglés)

Tiempo: 00:00:51

P1: Co mi Andrés Son yer ego ai disaired it was taim tu go bak tode si yes ai niret tu dere go of de erdly guorris dat brought mai enger and pein and bi en in de guarer uonsmor livin on de land vos estarting to teik to teik its to on mi der guer iven tai olmos jad to stop fairing songuan on de stret den laik meni men in de past ai uon egein turd toweird de si also der it te caind of magi in de werer and it can bi sin iven de works al reires artis ad not going to explor dis gie and de seint tat wil at o at som point aurselvs guosh aut to si (fin de la transcripction).

Método de análisis auditivo

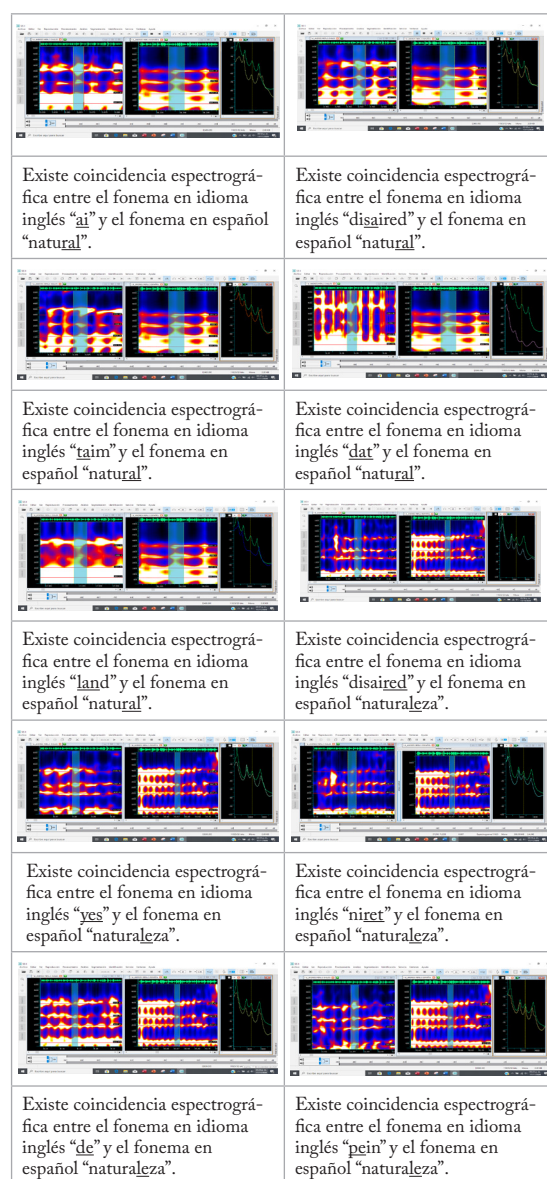
Luego de realizar el análisis auditivo (escuchar e identificar el contenido de las grabaciones de audio) se considera que existe similitud en la voz masculina signada como persona 1 (P1) en la transcripción del archivo dubitado 1 (inglés) con la muestra de voz 1 (español) que fue otorgada de manera libre y voluntaria.

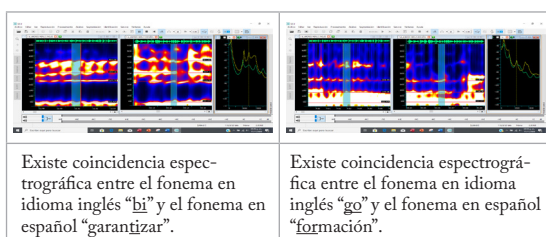
Método espectrográfico de formantes

El resultado obtenido del análisis visual de los espectrogramas resultantes de los archivos de audio es el siguiente: dubitado en el idioma inglés e indubitado en el idioma español.

Gráfico 2

Espectrogramas de los resultados obtenidos de un cotejamiento realizado en el software SIS II del sistema AVIS+F





Fuente: sistema AVIS+F, *software* SIS II.

Elaboración: por los autores.

Luego de haber realizado el análisis auditivo y de comparación espectrográfica de formantes se determina que la voz masculina signada como persona 1 (P1), en la transcripción del archivo dubitado 1 en el idioma inglés, corresponde auditiva y espectrográficamente con la muestra biométrica de voz 1 archivo indubitado en idioma español, otorgada de manera libre y voluntaria, es decir, se trata de la misma persona.

Discusión

En la Jefatura Zonal de Criminalística no existe un antecedente en el que se haya realizado un análisis y cotejamiento de emisiones lingüísticas en idiomas inglés-español con el sistema AVIS+F. Esto motivó a realizar la presente investigación considerando que el idioma inglés es una de las lenguas más habladas del mundo y que su ortografía no es una guía segura para la pronunciación del mismo. La pronunciación del idioma inglés refleja la forma en la que la palabra es hablada y, por tanto, la manera en que una persona escucha esa palabra. De allí tenemos que la transcripción es la representación sistemática de una forma oral mediante signos escritos. Por eso, cuando se realiza la transcripción se debe representar los sonidos con la ortografía de la otra lengua, es decir, se debe transcribir tal como se escucha en el idioma español.

El sistema AVIS+F, mediante la estación IKAR-Lab, y con la utilización del *software* SIS II, realiza la identificación de locutores, ya que este *software* posibilita la visualización de las señales del sonido para su análisis. El archivo de audio a ser cargado debe cumplir con ciertos parámetros técnicos: el *software* procesa los archivos

y realiza una identificación automática, aplicando un análisis estadístico de las características de la voz obtenidas por el sistema; para lo cual utiliza fórmulas matemáticas basadas en la transformada de Fourier.

El análisis y cotejamiento de voces está formado por la aplicación del método auditivo y este es complementado con la utilización del método de comparación espectrográfica de formantes, el cual consiste en un análisis visual y auditivo de las emisiones lingüísticas de una persona, específicamente de fonemas vocálicos. Luego de realizar dicho análisis, y de considerar lo expuesto, se demuestra que si es posible realizar un análisis y cotejamiento de emisiones lingüísticas en idiomas inglés-español con el sistema AVIS+F ya que el sistema permite la visualización de fonemas vocálicos, los cuales son procesados por el oído humano, al considerar la pronunciación de cada persona en un idioma determinado.

Conclusiones

Se realizó la toma de muestras biométricas de voz en un ambiente controlado, disminuyendo sonidos externos (personas hablando, música, naturaleza, tránsito, entre otros) y saturación de banda ancha (producidos por sistemas informáticos). Esto permitió la obtención de una muestra de voz adecuada para su análisis.

Mediante la utilización del sistema AVIS+F y aplicación del método auditivo y de comparación espectrográfica de formantes, se identificó de manera técnica a una persona a través de su voz, la cual domina dos idiomas diferentes como son el inglés y español.

La identificación de personas a través de la voz tiene un porcentaje del 98 % de efectividad, detrás de los análisis de ADN y de huellas dactilares, Pero, la voz de una persona es única y es considerada la huella vocal.

Referencias bibliográficas

- Asociación Española de Disfonia Espasmódica. *Aparato fonador*. Wikipedia.org. https://es.wikipedia.org/wiki/Aparato_fonador (consultado el 9/9/2021).
- Acústica. *La voz*. ehu.eus. <http://www.ehu.eus/acustica/espanol/musica/vohues/vohues.html> (consultado el 9/9/2021).
- Cicres, J. *Cotejamiento de voces*. scielo.conicyt.cl. <https://scielo.conicyt.cl/pdf/signos/v47n86/a02.pdf> (consultado el 9/9/2021).
- ConceptoDefinición. *Fonología*. conceptodefinition.de. <https://conceptodefinition.de/fonologia/> (consultado el 7/2/2021).
- Diccionario Actual. *El aparato fonador*. diccionarioactual.com. <https://diccionarioactual.com/aparato-fonador/> (consultado el 9/9/2021).
- García Gago, S. *Amplitud*. analfatecnicos.net. <https://www.analfatecnicos.net/pregunta.php?id=5> (consultado el 9/9/2021).
- López, P. W. *La frecuencia*. es.slideshare.net. <https://es.slideshare.net/pattrilopezulloa/frecuencia-patt> (consultado el 9/9/2021).
- Martínez, Marco. *¿Qué es la transformada de Fourier y para qué sirve?* nobbot.com. <https://www.nobbot.com/educacion/que-es-la-transformada-de-fourier-y-para-que-sirve/> (consultado el 12/5/2021).
- Porporatto, M. *El tiempo*. quesignificado.com. <https://quesignificado.com/tiempo/> (consultado el 9/9/2021).
- significados.com. *Qué es el Sonido?* significados.com. <https://www.significados.com/sonido/> (consultado el 9/9/2021).