

**PONTIFICIA UNIVERSIDAD
CATÓLICA DEL PERÚ**

Escuela de Posgrado



**Identificación de locutor en el marco de la
fonética forense en el Perú**

Tesis para optar el grado académico de
Maestro en Lingüística que presenta:

Fernando Aarón Torres Castillo

Asesor:

Roberto Daniel Zariquiey Biondi

Lima, 2023


Informe de Similitud

Yo, Roberto Daniel Zariquiey Biondi, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor de la tesis de investigación titulado *Identificación de locutor en el marco de la fonética forense en el Perú*, del autor Fernando Aarón Torres Castillo dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 12%. Así lo consigna el reporte de similitud emitido por el software Turnitin el 20/09/2023.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, 20 de setiembre del 2023

Apellidos y nombres del asesor / de la asesora: Roberto Daniel Zariquiey Biondi	
DNI: 40203566	Firma: 
ORCID: 0000-0002-1421-1314	

Resumen

Esta tesis describe y compara las características fonéticas de *muestras indubitadas* (muestras en las que no hay dudas acerca de la identidad del locutor pues se obtienen a partir de audiencias, declaraciones y tomas de muestra de voz) y *muestras dubitadas* (muestras de las que se tiene dudas acerca de la identidad del locutor pues se obtienen de audios de llamadas telefónicas o grabaciones donde el locutor no se identifique). Los audios analizados pertenecen a un caso de corrupción de funcionarios en una región del sur del Perú, y se optó por mantener a los implicados en el anonimato, son siete (07) los efectivos policiales implicados, pero solo se consideraron tres (03). Las muestras de voz dubitadas han sido proporcionadas por el Ministerio Público – Fiscalía de la Nación. Dos de las muestras indubitadas se recogieron a través de diligencias de toma de muestra de voz; y la tercera corresponde al archivo informático de una audiencia solicitada al Poder Judicial que corresponde a otro imputado.

En esta investigación, las muestras se organizan e interpretan en el software *SplitsTree4* con el objetivo de verificar si las voces de los locutores de las muestras indubitadas se encuentran también en los audios de las muestras dubitadas. Asimismo, se utiliza el programa *SIS II* (del paquete *IKAR LAB del Speech Technology Center*) para corroborar la correspondencia de voces a partir del análisis biométrico de voz.

En primer lugar, todos los audios fueron segmentados y anotados con rasgos fonéticos particulares, cuya presencia y/o ausencia fue codificada de manera binaria y sistemática para cada voz participante. Luego, esta información fue procesada por el programa *SplitsTree4* con la finalidad de reordenar los rasgos conforme a la cantidad de voces participantes y dilucidar la coincidencia entre las voces de las muestras intervinientes. La presente tesis obtuvo resultados positivos con respecto del método implementado; en relación al uso de rasgos fonéticos para la identificación forense de voz. Esto establece un aporte significativo concerniente a la aplicación de tecnología forense a la justicia en el Perú y establece un precedente metodológico que puede ser replicado por otros países.

Abstract

This thesis describes and compares the phonetic characteristics of unquestionable-speaker samples (samples in which there is no doubt about the identity of the speaker as they are obtained from hearings, statements and sampled voices) and questionable-speaker samples (samples in which there is doubt about the identity of the speaker as they are obtained from audios of telephone calls or recordings where the speaker does not identify him/herself). The audios analyzed belong to a case of corruption of officials in a region of southern Peru, and it was decided to keep those involved anonymous, and the decision was made to keep those involved anonymous, there are seven (07) police officers implicated, but only three (03) were considered. The questionable-speaker samples have been provided by the Public Prosecutor's Office (Ministerio Público - Fiscalía de la Nación). Two of the unquestionable-speaker samples were collected through voice sampling diligences, and the third sample was obtained from the digital files of a hearing requested to the Judiciary that corresponds to another defendant.

In this research, the samples are organized and interpreted in the SplitsTree4 software in order to verify if the voices of the speakers of the unquestionable-speaker samples are also found in the audios of the questionable-speaker samples. In addition, the SIS II software (from the IKAR LAB package of the Speech Technology Center) is used to corroborate the voice correspondence from the biometric voice analysis.

First, all the audios were segmented and annotated with particular phonetic features, whose presence and/or absence was coded in a binary and systematic form for each participating voice. Then, this information was processed by the SplitsTree4 software in order to reorder the features according to the number of participating voices and to elucidate the coincidence between the voices of the intervening samples. The present thesis obtained positive results with respect to the implemented method; in relation to the use of phonetic features for forensic voice identification. This establishes a significant contribution concerning the application of forensic technology to justice in Peru and establishes a precedent method that can be replicated by other countries.

Dedico esta tesis...

A mis padres, por el apoyo constante.

A mi familia, por estar siempre en los momentos más difíciles.

A Akuma, por su compañía y atención a todas las actividades que realizo en casa.

A mis abuelitos que están en el cielo.

A Lupita Mendoza



Agradecimientos

En primer lugar, agradezco a mis profesores de la Maestría en Lingüística de la Pontificia Universidad Católica del Perú por sus enseñanzas en el transcurso del programa. Segundo, agradezco a mi asesor, Roberto, por su constante ayuda desde mi inicio en la maestría, también agradezco su guía y paciencia para seguir adelante con la presente tesis. Tercero, agradezco a la profesora Verónica Lazo por generarme interés en el ámbito de la lingüística forense, así como su apoyo constante en los primeros años en el trabajo, y al profesor Lovón por sus comentarios en la tesis. Cuarto, agradezco al Dr. Danny Humpire por dar la oportunidad a la Lingüística como carrera necesaria para combatir la criminalidad a partir de la ciencia, además de haber inculcado en la Oficina de Peritajes la idea de aportar a través de nuestros conocimientos a la criminalística. Quinto, agradezco a mis amigos del Área de Fonética y Acústica Forense por crecer juntos día a día en el ámbito laboral y en el aspecto científico, además de admirar la templanza de superar con seguridad y probidad los casos que se presentan en la Oficina. Sexto, agradezco a mis amigos lingüistas por sus aportes y críticas que me ayudaron a mejorar la presente investigación, a Dzon y Osgar por trabajar en conjunto en temas de lingüística forense, a Smis, por su honestidad, a Erike y Charles, por los momentos de risas y *ayni*, a Ruru por su ayuda a entender el funcionamiento de los programas de biometría de voz y su innovación respecto al uso del programa Praat, a Yuyu, por los gratos momentos en la oficina, a Chistris, por dar momentos amenos en un ambiente tétrico, al equipo transcripción, Clau, Andrea, Milagros, Diana y Edwin; agradezco también a Prikila y al Dixiri, por ser parte de mi vida, y a Luz, por su constante ayuda y amistad desde que nos conocimos. Al Ministerio Público por permitir realizar investigaciones científicas para combatir la criminalidad.

Por último, y no menos importante, agradezco al Juficiente, por su compañía en los momentos más difíciles desde que nos conocimos, también a Kuromi, por su constante ayuda y en el diseño de la infografía.

Índice

Resumen.....	3
Agradecimientos.....	6
Índice	7
Lista de Figuras	11
Lista de Tablas	12
CAPÍTULO I	13
INTRODUCCIÓN	13
1.1. Datos generales	13
1.2. Lingüística forense	16
1.3. Fonética forense.....	18
1.4. Sobre el funcionamiento del sistema pericial en el Perú (OPERIT).....	20
1.4.1. Casos en el Perú.....	23
1.4.1.1. Caso Solsiret	23
1.4.1.2. Caso Mil voces	25
1.4.1.3. Caso Vieira.....	26
1.5. Antecedentes	28
1.6. Problema de investigación.....	29
1.7. Objetivos	31
1.8. Metodología.....	31
1.8.1. Alcance, enfoque y diseño de la investigación	32
1.8.1.1. Alcance.....	32
1.8.1.2. Enfoque.....	33
1.8.1.3. Diseño	33
1.8.2. Población	33
1.8.2.1. Muestra	34
1.8.2.2. Variables	34
1.8.2.2.1. Operacionalización de variables	35
1.8.3. Metodología combinada	36
1.8.3.1. Método clásico.....	36
1.8.3.2. Método del sistema automático	37
1.8.3.3. Método de la red filogenética.....	37
1.8.4. Herramientas metodológicas.....	38

1.8.4.1. Sobre el software Praat y el análisis lingüístico-espectral	38
1.8.4.2. Sobre el software SIS II (STC) y el análisis biométrico	39
1.8.4.3. Sobre el software SplitsTree 4 y el análisis filogenético	44
1.8.4.4. Locutores indubitados.....	50
1.8.4.5. Locutores dubitados	51
1.8.4.5.1. Metadatos de la muestra dubitada	51
1.8.4.5.2. Asignación de rótulo y referencia a la muestra dubitada	52
1.9. Justificación	53
1.9.1. Justificación científica	53
1.9.2. Justificación social	53
1.9.3. Limitaciones	54
CAPÍTULO II	55
MARCO CONCEPTUAL.....	55
2.1. Fonemas y alófonos	55
2.2. Sonidos del español	56
2.2.1. Vocales.....	56
2.2.1.1. Formantes de vocales	56
2.2.1.2. Carta de formantes	57
2.2.2. Consonantes	58
2.2.2.1. Modo de articulación.....	59
2.2.2.2. Punto de articulación	61
2.2.2.3. Función de las cuerdas vocales.....	62
2.2.2.4. Clasificación de los sonidos fonéticos consonánticos del español	62
2.2.2.5. Descripción acústica de las consonantes	63
2.3. Estudio de la percepción auditiva	67
2.3.1. Elementos para el análisis fonético-fonológico y lingüístico	68
2.3.2. Tabulación del análisis lingüístico	69
2.4. Procesos en vocales y consonantes	74
2.4.1. Procesos que afectan directamente en vocales	74
2.4.1.1. Disfonía vocálica	74
2.4.1.2. Laringalización vocálica.....	75
2.4.1.3. Alargamiento vocálico.....	76
2.4.1.4. Aspiración vocálica	77
2.4.1.5. Ensondecimiento vocálico	78
2.4.2. Procesos que cambian la vocal.....	78
2.4.2.1. Cambio vocálico o sustitución vocálica	78

2.4.2.2. Monoptongación	80
2.4.2.3. Elisión vocálica	80
2.4.2.4. Sinalefa	81
2.4.3. Procesos que afectan directamente a la consonante	81
2.4.3.1. Sonorización consonántica	81
2.4.3.2. Aspiración de consonantes oclusivas	83
2.4.3.3. Alargamiento de la vibrante simple	84
2.4.3.4. Alargamiento de la vibrante múltiple	84
2.4.3.5. Simplificación de la vibrante múltiple	85
2.4.4. Procesos característicos en las consonantes	86
2.4.4.1. Producción de la /s/ en frecuencias bajas del espectro	86
2.4.4.2. Producción del <i>chiwchi sound</i>	87
2.4.4.3. Producción estridente de la fricativa velar /x/	87
2.4.4.4. Producción aproximante de la vibrante múltiple	88
2.4.4.5. Producción aproximante de la vibrante simple	89
2.4.5. Procesos que cambian a la consonante	90
2.4.5.1. Elisión consonántica	90
2.4.5.2. Rotacismo por sustitución de consonantes	90
2.4.5.3. Africación de la fricativa /s/	91
2.4.5.4. Palatalización de la fricativa /s/	92
2.4.5.5. Aspiración de la fricativa /s/	92
2.4.5.6. Lateralización	93
2.4.5.7. Supresión consonántica	94
2.5. Elementos paralingüísticos	94
2.5.1. Las risas	94
2.5.2. Acento	95
2.6. Estudio de la comparación biométrica de voces	96
2.6.1. Métodos automáticos por voz	96
2.6.2. <i>Software</i> SIS II	96
2.7. <i>Software</i> SplitsTree 4	102
2.8. Sobre los métodos combinados en identificación de locutores	105
2.8.1. Niveles de conclusión metodológica	105
2.9. Identificación de locutor	105
2.10. Verificación de locutor	106
CAPÍTULO III.....	107
ANÁLISIS	107

3.1. Muestras indubitadas.....	107
3.1.1. Análisis de la calidad de las muestras indubitadas.....	107
3.2. Muestras dubitadas	107
3.2.1. Análisis de la calidad de las muestras dubitadas.....	108
3.3. Análisis de comparación automática a partir del software SIS II	108
3.4. Análisis fonético-acústico entre las muestras indubitadas y dubitadas	110
3.4.1. Análisis de las características lingüísticas de la muestra indubitada LI_01.....	110
3.4.2. Análisis de las características lingüísticas de la muestra indubitada LI_02.....	113
3.4.3. Análisis de las características lingüísticas de la muestra indubitada LI_03.....	120
3.5. Uso del software SplitsTree 4	122
3.5.1. Análisis de los rasgos fonético-acústicos con el <i>software</i> SplitsTree 4.....	123
3.5.1.1. Tabulación de la matriz de rasgos fonético-acústicos	124
CAPÍTULO IV	126
RESULTADOS	126
4.1. Resultados del método de ordenamiento con el software SplitsTree 4	128
4.2. Balance sobre el método filogenético	129
CONCLUSIONES	132
RECOMENDACIONES.....	134
REFERENCIAS BIBLIOGRÁFICAS.....	135
ANEXOS.....	142
Anexo 01: Modelo de Protocolo de toma de muestra de voz	142
Anexo 02: Segmentación de la participación de las voces de las muestras dubitadas ..	148
Anexo 03: Capturas de pantalla de los resultados de análisis de la calidad de la señal de las muestras indubitadas y dubitadas	169
Anexo 04: Capturas de pantallas de los resultados de la comparación automática con el programa SIS II	172
Anexo 05: Alfabeto Fonético Internacional	177
Anexo 06: Script para generar redes filogenéticas con <i>Simple Data</i>	178
Anexo 07: Transcripción de las muestras dubitadas.....	179
Anexo 08: Transcripción de las muestras indubitadas	192

Lista de Figuras

Figura 1.....	16
Figura 2.....	21
Figura 3.....	26
Figura 4.....	27
Figura 5.....	30
Figura 6.....	30
Figura 7.....	39
Figura 8.....	40
Figura 9.....	41
Figura 10.....	41
Figura 11.....	42
Figura 12.....	43
Figura 13.....	44
Figura 14.....	45
Figura 15.....	46
Figura 16.....	49
Figura 17.....	52
Figura 18.....	57
Figura 19.....	58
Figura 20.....	59
Figura 21.....	63
Figura 22.....	64
Figura 23.....	65
Figura 24.....	66
Figura 25.....	67
Figura 26.....	68
Figura 27.....	75
Figura 28.....	76
Figura 29.....	77
Figura 30.....	77
Figura 31.....	78
Figura 32.....	79
Figura 33.....	79
Figura 34.....	80
Figura 35.....	81
Figura 36.....	82
Figura 37.....	82
Figura 38.....	83
Figura 39.....	83
Figura 40.....	84
Figura 41.....	85
Figura 42.....	85
Figura 43.....	86
Figura 44.....	87
Figura 45.....	88
Figura 46.....	89

Figura 47.....	89
Figura 48.....	90
Figura 49.....	91
Figura 50.....	91
Figura 51.....	92
Figura 52.....	93
Figura 53.....	93
Figura 54.....	94
Figura 55.....	95
Figura 56.....	96
Figura 57.....	98
Figura 58.....	106
Figura 59.....	106
Figura 60.....	128

Lista de Tablas

Tabla 1.....	35
Tabla 2.....	51
Tabla 3.....	51
Tabla 4.....	52
Tabla 5.....	53
Tabla 6.....	56
Tabla 7.....	62
Tabla 8.....	69
Tabla 9.....	73
Tabla 10.....	99
Tabla 11.....	107
Tabla 12.....	108
Tabla 13.....	109
Tabla 14.....	109
Tabla 15.....	111
Tabla 16.....	114
Tabla 17.....	121
Tabla 18.....	123
Tabla 19.....	124

CAPÍTULO I

INTRODUCCIÓN

1.1. Datos generales

Uno de los primeros estudios sobre la zonificación de las variedades dialectales peruanas es el de Escobar (1978), en el cual utiliza un número limitado de isoglosas, o como las llama el autor “fronteras entre rasgos sistemáticos, permanentes y de considerable abstracción”. Escobar llamó al español andino como español de Tipo I, al español ribereño (conformado por el de la costa y la Amazonía) como español Tipo II; distinguía fonético-fonológicamente estas dos variedades a partir de dos sonidos, la yod /j/ y la lateral palatal /ʎ/, la última presente solo en la variedad de Tipo I. Caravedo (1992) propone una distinción entre las variedades que se hablan en la costa y la Amazonía, y en ese sentido se basa en las regiones tradicionalmente conocidas por la geografía (costa, sierra y selva). En esa línea, Pérez (2004) también señala que el panorama lingüístico del español en el Perú es diverso; en diferentes partes del país se manifiestan distintas variedades dialectales de español, que pueden ser agrupadas en grandes conglomerados dialectales definidos geográficamente como español amazónico, español andino y español costeño.

Una investigación reciente de Andrade (2023) propone seis zonas principales de variedades dialectales, en primer lugar se tiene la variedad de (1) los andes norteños, la cual comprende las zonas del sur de Cajamarca, extremo norte de Áncash la sierra de La Libertad, de Piura, de Lambayeque y posiblemente Chachapoyas; en esta variedad de los andes norteños no solo han intervenido lenguas como el quechua y aymara, sino también lenguas extintas como el culle, presente hasta inicios de la mitad del siglo XX. Estas lenguas dejaron su legado tanto en el plano fonético-fonológico como en el morfológico

La segunda variedad que propone Andrade es la de los (2) andes centrales, surcentrales y sureños. Los andes centrales comprenden las regiones de Ancash, Lima, Huánuco, Pasco y Junín; los andes surcentrales, Huancavelica, Ayacucho y los andes occidentales de Apurímac, y los andes sureños están conformados por el oriente de Apurímac, de Arequipa, de Cusco, y de Puno. Esta variedad ha tenido un contacto histórico mayoritario con lenguas como el quechua y el aymara en Puno y otras regiones aledañas. Los rasgos más característicos de esta variedad son los que

conforman el componente fonético fonológico, tales como el cambio vocálico entre los pares vocálicos /i, e/ o /u, o/.

Asimismo, la producción de las vibrantes es llamativo, ya que, tal como señala Andrade (2023, p. 336) en sus inicios se solía describir como netamente asibilada, sin embargo, se identifican otros la retroflexión, el ensordecimiento o el debilitamiento. Otras características resultantes del contacto entre lenguas resaltan en la morfología, ejemplos como el doble posesivo, uso de los clíticos de objeto, entre otros.

Ahora bien, para el caso de los (3) andes occidentales sureños y costa sur, en primera instancia, Andrade señala que, en las primeras investigaciones sobre las regiones de Moquegua y Tacna, estas no fueron clasificadas como parte del “español altiplánico” o también llamado español andino, esto por algunas características fonético-fonológicas, por ejemplo, la producción debilitada de la fricativa velar en posición coda y presidida de las vocales /i, e/. Sin embargo, Cerrón-Palomino (2020) indica que la presencia de lenguas andinas como el quechua, aimara y puquina en el territorio sureño hasta fines del siglo XIX abordaría al español de esta zona y las clasificaría como parte del español andino.

El español de la costa se divide en dos, y el punto de separación entre ambas variedades es el extremo norte de Áncash; en ese sentido, la (4) costa norte, inicia desde Tumbes, hasta el departamento de La Libertad, uno de los rasgos más representativos es la flexión de la primera persona plural en el imperfecto del modo indicativo, en vez de decir *estábamos*, dirían *estábanos*.

Para la (5) costa central se dividió desde el extremo norte de Ancash hasta el departamento de Ica. Algunos investigadores como Caravedo (1983) o O'Rourke (2008) hicieron estudios sobre las variedades de Lima, y actualmente se ha observado la influencia de rasgos andinos por la migración.

Por último, se tiene al español de la (6) Amazonía, o también llamado español amazónico, que resalta por aquellos rasgos fonético-fonológicos y morfosintácticos concurrentes en esta variedad, uno de los rasgos particulares más conocidos es la presencia del sonido africado sonoro [d̪ʒ], es una realización fonética del fonema yod. Según Valenzuela y Jara (2021) otro rasgo particular es la neutralización de la fricativa labial /f/ y la fricativa velar /x/. Andrade (2023) añade que las variedades dialectales propuestas deben ser estudiadas con más detalle, y en todos los planos lingüísticos y extralingüísticos existentes.

Lo mismo indica Vásquez (2017) en su tesis, ya que el propósito de su investigación consistió en exponer las particularidades entonacionales específicas del enfoque estrecho en oraciones afirmativas dentro del dialecto español de Iquitos. Aunque las características entonacionales de esta variante del español muestran una convergencia significativa con el español de Pucallpa, también se identifican algunas diferencias. Estos hallazgos sugieren que, tal como se planteó al inicio y se ha mantenido en diversos ámbitos académicos, investigaciones más detalladas sobre el español hablado en las diferentes regiones de la Amazonía peruana podrían revelar un nivel de diversidad mayor al descrito habitualmente en la literatura. En resumen, esta tesis buscó analizar las particularidades de la entonación del foco estrecho en oraciones declarativas dentro del español de Iquitos. Aunque se encontraron similitudes significativas con el español pucallpino, también se observaron diferencias entre ambas variantes. Estos resultados respaldan la idea de que se debe realizar un estudio más exhaustivo del español hablado en las diversas regiones de la Amazonía peruana, ya que es probable que se descubra un nivel de heterogeneidad lingüística mayor al que se ha descrito hasta ahora en la literatura existente.

Según Hudson (1980), es difícil afirmar que dos personas hablan la misma lengua debido a que su experiencia lingüística difiere. Por lo tanto, las conclusiones obtenidas del análisis del habla de un individuo carecerían de relevancia en estas áreas, ya que no sería posible generalizarlas. Como resultado, surge la necesidad científica de agrupar a las personas en categorías superiores, aunque estos grupos no sean completamente homogéneos debido a las particularidades individuales, tal como menciona López (1989).

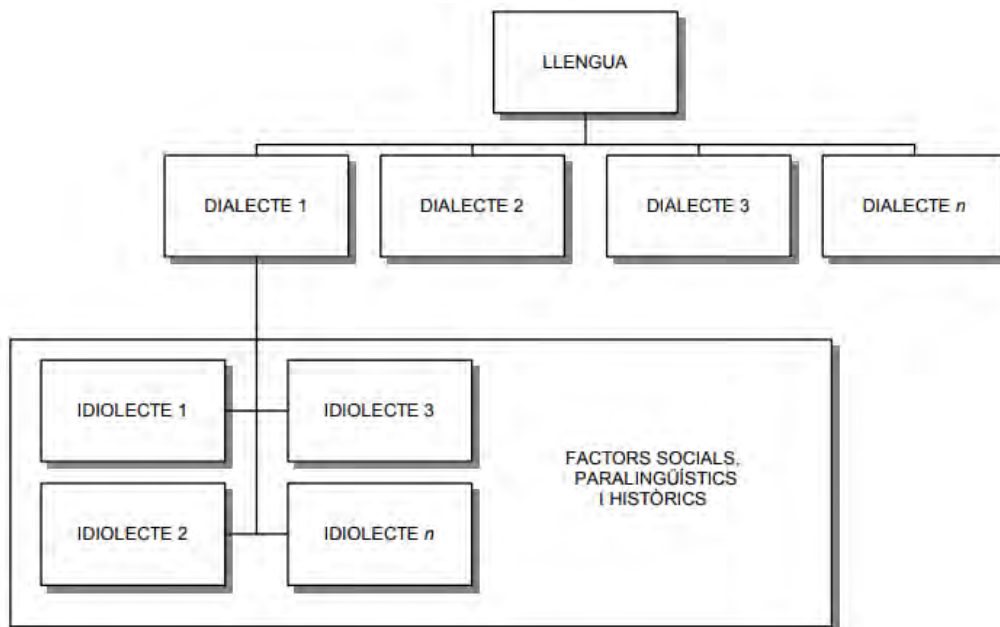
Es en ese sentido que, para la lingüística forense, la idea de idiolecto es sumamente relevante porque se debe entender que una variedad lingüística es un conjunto de idiolectos, y cada uno de ellos debe ser entendido como el uso individual que establece un hablante y que diversos factores como los culturales, económicos, educativos, sociales, de género u profesión se manifiestan en estos idiolectos, tal como lo describe Cicres (2007) en la Figura 1.

Cabe destacar que también existe el concepto de *interlecto*, entendido y descrito por Escobar (1989, 147) como «el español hablado como segunda lengua por personas con una de las dos lenguas maternas u originarias de mayor número, sea el quechua o el aimara». La noción de idiolecto nos ayuda a definir y a referirnos a la forma o manera de expresarse de un individuo en específico. Existen particularidades que se

manifiestan en los niveles del lenguaje: estos pueden ser en el fonético-fonológico, morfológico, sintáctico y tanto el léxico como el semántico (Sánchez-Iglesias, 2005; Turell, 2010). El factor individual tiene mayor ponderación al realizar estudios de comparación forense (Cicres, 2011) y, por ende, en este estudio, **nos centraremos en los sonidos o fonos que son particularmente identificables en el habla de una persona.**

Figura 1

Esquema de lenguas, formadas por dialectos y estos por idiolectos



Nota. Datos sobre el comportamiento de los idiolectos, de Cicres, 2007, p. 50

1.2. Lingüística forense

¿Qué impresión se tiene al momento de escuchar la frase **lingüística forense**? Una de las primeras ideas estaría relacionada con los textos escritos y a la voz humana, pero, ¿por qué forense?, la palabra *forense* refiere al hecho de que el material lingüístico es analizado en los tribunales para luego entrar a debate en público o *foro*, siendo este *foro* directamente vinculado con la justicia. Entonces, una ciencia, como lo es la lingüística, estaría vinculada al servicio de la ley cuando se refiere a una ciencia forense. Toda ciencia que sirve al procedimiento judicial es forense. La lingüística forense se describe como una ciencia que relaciona aspectos de la lingüística con ámbitos del derecho. El lenguaje sirve, en este contexto, como una prueba ante un crimen. Entonces, ¿qué elementos se puede analizar con ayuda de la

lingüística? Como medios probatorios se tienen textos escritos y orales, documentos que se utilizan para cometer crímenes o quedan registros de ellos, o la voz de un criminal, también es materia de análisis. Como consenso de lo que significa lingüística forense, existe una definición de la Asociación Internacional de Lingüística Forense (AILF), la cual la conceptualiza como una disciplina situada entre la lengua y el derecho. (Garayzábal *et al.*, 2019)

Es así que, en palabras de Coulthard *et al.*, (2017), define a la lingüística forense como un subcampo de la lingüística, y está comprometido con la interacción profesional e institucional de textos (orales y escrito) legales. La lingüística consta de disciplinas lingüísticas: tales como la fonética, fonología, morfología, léxico, sintaxis, semántica y pragmática, de las cuales todas son esenciales para describir lo que se observa en un caso que involucre al lenguaje como materia de análisis.

Algunos subcampos que se involucran en el aspecto legal son la sociolingüística, el análisis de la conversación, lingüística del corpus, entre otros. La sociolingüística se ocupa del lenguaje en la sociedad, y estudia la variación del lenguaje según los locutores y los usos que le dan a determinadas áreas del lenguaje. Entre ellas destacan algunas variables como género, edad, sexo, lugar de nacimiento, educación, profesión, entre otros; estas variables interactúan entre sí, y se plasma en la cotidianidad. Los asuntos legales se conectan con el entorno social, es por esa razón que la sociolingüística es importante para la lingüística forense, el lenguaje no solo se difiere a grupos, sino también a individuos, un concepto importante para los lingüistas forenses, en relación a determinar la autoría de algo, como textos escritos u orales. (Delgado, 2001)

Todos los oyentes son capaces de reconocer las voces familiares que están en su entorno, a través de un audio, una llamada o un video; sin embargo, la información auditiva no siempre es certera, por lo que podría producirse una confusión o un resultado equívoco. Es en ese sentido que surge la necesidad de la intervención de un experto, con el fin de esclarecer las dudas en la investigación científica, para ello nos valdremos de la **identificación forense de locutores de voz**.

Existen muchas dificultades a las que se enfrenta el experto o “perito” al analizar las muestras de voz, ya que son característicamente complejas y variadas. Tal como lo presentamos en un inicio de esta tesis, el llamado castellano peruano es en realidad un conjunto entramado de variedades lingüísticas geográficas, sociales e individuales, que representan un desafío para la lingüística forense. Más aún, existe un plano

emocional que tiene un efecto en la producción del habla; no todas las elocuciones de un individuo son iguales y, por ejemplo, este hablará distinto en un ámbito formal (como una audiencia o diligencia fiscal) y en un ámbito informal (como llamadas telefónicas) (Lazo, 2023). Ello no quiere decir, sin embargo, que no se tengan las herramientas y metodologías necesarias para analizar de la manera conclusiva la identidad de una muestra de voz (Delgado, 2001).

El primer caso conocido de lingüística forense se remonta a 1960, cuando el lingüista forense Roger Shuy fue contratado para analizar una serie de grabaciones de voz en un caso de extorsión en Estados Unidos. En ese caso, un individuo desconocido había estado realizando llamadas telefónicas amenazantes y exigía un rescate a cambio de no hacer daño a la víctima, Shuy utilizó su experiencia en lingüística para analizar las grabaciones de voz y determinar si el autor de las llamadas era la misma persona en diferentes ocasiones. A través del análisis del habla, pudo identificar ciertos patrones y características distintivas en la pronunciación, la entonación y el uso del lenguaje, que indicaban que se trataba de la misma persona. Sus análisis fueron presentados como evidencia en el juicio, y el acusado fue finalmente condenado. Este caso sentó las bases para el uso de la lingüística forense como una disciplina científica para la resolución de casos judiciales y la identificación de locutores a través del análisis del habla. (Shuy, 1998)

1.3. Fonética forense

La fonética forense es una rama de la lingüística forense que se enfoca en el análisis y la interpretación científica de la voz y los sonidos del habla para propósitos legales. Su objetivo principal es utilizar métodos y técnicas fonéticas para determinar la identidad de una persona o para proporcionar evidencia en investigaciones criminales y procedimientos judiciales. La fonética forense se basa en el principio de que cada persona tiene una forma única de hablar y producir sonidos vocales, lo que se conoce como la "huella vocal". Sin embargo, de acuerdo con Olsson (2008), el método de la huella vocal, también conocido como *voiceprint*, generó cuestionamientos debido a su variabilidad inherente, ya que la voz de una persona puede fluctuar debido a factores diversos como la edad, la salud y las emociones, lo que puede conducir a errores en la identificación. Además, la precisión de esta tecnología puede ser cuestionable, dado que las similitudes entre las características vocales de diferentes individuos pueden generar identificaciones incorrectas. Asimismo, indica que la idea

de una “huella vocal” es un concepto poderoso, y atrae a diferentes entidades a hacer cumplir la ley, pero hay escasas pruebas sólidas que respaldan la idea de una “huella vocal” para identificar a un locutor.

Los expertos en fonética forense utilizan una variedad de herramientas y técnicas para examinar y comparar muestras de voz, como grabaciones de llamadas telefónicas, testimonios grabados o cualquier otro material de audio relacionado con un caso. Estas técnicas incluyen el análisis acústico, que estudia las características físicas de la señal de sonido, como la frecuencia, la intensidad y la duración de los sonidos.

También se utiliza el análisis espectrográfico, que representa visualmente la energía de diferentes frecuencias en el habla. Además, se emplean métodos de análisis fonético y fonológico para estudiar los rasgos distintivos del habla de una persona, como el acento, la entonación, la pronunciación de sonidos específicos y otros aspectos de la producción vocal.

Asimismo, Aguirrezabala (2015) realiza un estudio sobre verificación biométrica de voz y evalúa la posibilidad de autenticar locutores a través de la biometría vocal. Se extraerán características vocales clave, como los coeficientes MFCC, de una base de datos con 10 grabaciones por locutor. Se desarrollarán clasificadores y se realizarán pruebas para determinar si un locutor es el buscado. Se utilizarán Máquinas de Soporte Vectorial (SVM) para la autenticación. Los resultados demuestran la capacidad óptima del sistema para verificar la identidad del locutor en comparación con otros en la base de datos.

La fonética forense puede ser utilizada para resolver casos de identificación de locutores, donde se busca determinar si una persona en particular es la responsable de una grabación de voz. También se puede usar para el análisis de testimonios grabados y su autenticidad, así como para la detección y el análisis de alteraciones en grabaciones de voz con el objetivo de identificar manipulaciones o ediciones fraudulentas. Es importante destacar que la fonética forense se basa en métodos científicos y su aplicación requiere de expertos altamente capacitados y calificados en el campo de la fonética y la lingüística forense. Su trabajo y análisis pueden proporcionar pruebas y evidencias valiosas en casos judiciales y contribuir a la justicia y la resolución de crímenes. (Delgado, 2011)

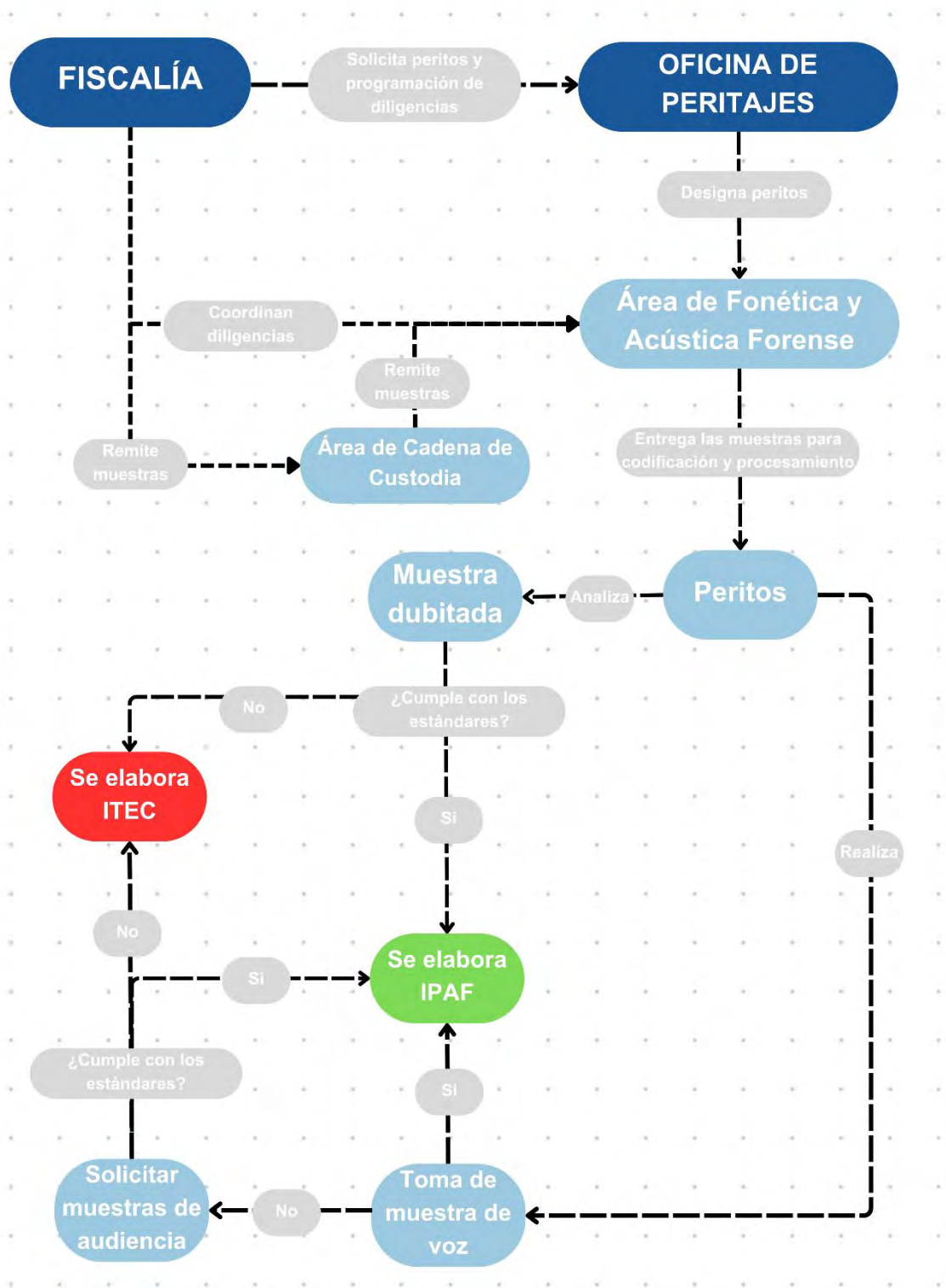
1.4. Sobre el funcionamiento del sistema pericial en el Perú (OPERIT)

La Oficina de Peritajes (desde ahora OPERIT) del Ministerio Público – Fiscalía de la Nación nace de la separación de un grupo de peritos, aproximadamente veinte (20), que formaban parte del Instituto de Medicina legal, legalizándose como tal en julio de 2019 como un órgano de apoyo a la Gerencia General del Ministerio Público. Entre sus actividades esenciales está la designación de peritos para encargarse de investigaciones de las carpetas fiscales a nivel nacional, según sea la competencia del perito, asimismo, los peritos y el personal de cada área tienen la tarea de participar en diligencias, emitir informes periciales y/o técnicos, entre otros como la investigación y capacitación. La OPERIT está conformada por un grupo de peritos separados por especialidades o áreas, entre ellos están lingüistas, traductores, ingenieros, físicos, antropólogos, especialistas en balística, grafotécnicos, especialistas en escena de crimen, y otros. A través de los años, los peritos han participado en casos polémicos, tales como el caso Cuellos Blancos, los asesinatos de Inti y Bryan, en el gobierno de Manuel Merino, entre otros. Hasta diciembre de 2022, la OPERIT contaba con ciento catorce (114) trabajadores CAS y doscientos (200) trabajadores locadores de servicios. De los cuales, actualmente solo quedan menos de cien (100), entre peritos y administrativos. A pesar de las múltiples carencias administrativas y presupuestales que existen en la OPERIT, la oficina cumple eficientemente con las investigaciones a su cargo (Ojo Público, 2023).

La OPERIT cuenta con un área especializada en homologación de voz/voces, en la cual labora un grupo multidisciplinario conformado por lingüistas, ingenieros electrónicos, físicos y traductores. Para poder abordar el desarrollo de elaboración de un informe pericial, se detalla a continuación, en la Figura 2 el flujograma de llegada y salida de documentos, pericias y muestras por Carpeta Fiscal.

Figura 2

Flujograma de atención de pedidos de pericias de homologación o identificación de voz/voces



Toda investigación en la OPERIT inicia desde la solicitud de alguna entidad del estado, en su mayoría despachos fiscales, algunas veces desde el mismo Poder Judicial o la Policía Nacional del Perú, presidida por un fiscal, pero todos ellos articulan de la misma forma.

Desde la dependencia de inicio, se solicita a la OPERIT la designación de un perito para encargarse de la pericia que ayudará al fiscal a sustentar en juicio con pruebas científicas, a su vez, juramenta, dando su palabra de acabar con la investigación en un plazo determinado de tiempo, de acuerdo a la naturaleza de las muestras y cantidad de investigados. En este caso, se designa a un perito del Área de Fonética y Acústica Forense, en la cual se realizan, actualmente los servicios de homologación o identificación de voz/voces, análisis del discurso y de traducción e interpretación¹.

Después de la juramentación, la Fiscalía entrega las muestras dubitadas (por ejemplo, voces de las que se tiene dudas de autoría o pertenencia) y acompañados obligatoriamente de documentos para la investigación, como actas de transcripción de los audios dubitados, al Área de Cadena de Custodia, la cual cumple con la función de salvaguardar las muestras hasta que el perito las solicite para poder peritarlas.

Una vez entregada las muestras al Área de Cadena de Custodia, el perito designado del caso solicita las muestras para que se remitan al, en este caso, Área de Fonética y Acústica Forense, el perito se encarga de codificar internamente las muestras y con los documentos a cargo, deslaca el contenido en el que se encuentran las muestras, estos pueden ser dispositivos como discos, USBs, memorias externas, entre otros, que contengan los archivos informáticos como audios, estos audios dubitados pueden ser grabaciones de llamadas, grabaciones de audio y video, mensajes de voz o interceptaciones telefónicas del sistema Constelación (Ramírez, 2015), Constelación es un sistema que intercepta comunicaciones telefónicas, opera desde el 2009 y se ha utilizado en casos emblemáticos como “Los Cuellos Blancos del Puerto”, “Los Intocables Ediles”, entre otros.

A la par, se puede coordinar las diligencias pertinentes para realizar la pericia solicitada, en este caso, la Fiscalía coordina con el perito designado las fechas y horas para las diligencias de toma de muestra de voz (ver 1.8.4.4. Locutores indubitados).

¹ Antes del recorte del personal y reestructuración de la Oficina de Peritajes, por pedido expreso de la Fiscalía de la Nación, se ofrecían adicionalmente los servicios de transcripción lingüística (Caso Cuellos Blancos) y pericias semánticas.

Tal sea el caso, el investigado, testigo o denunciante logre dar su voz, esta muestra se encripta con un código de aseguramiento informático para ser enviado a la base de datos interna del Área de Fonética y Acústica Forense, si la persona o personas citadas no entregan voluntariamente su voz, la fiscalía deberá solicitar muestras de declaraciones o audiencias públicas en la que logre identificarse la persona objetivo; de no existir aquellas muestras, no se podrá realizar la pericia solicitada.

Al tener la muestra dubitada y la muestra indubitada, se procede a realizar el análisis de los estándares de calidad, de acuerdo a ello se tomará en cuenta que muestras utilizar o no. Si sucede el caso de que alguno de ellos no cumpla con los estándares de calidad conforme el software de biometría de voz SISII lo permita, se emitirá un informe técnico indicando que se remitan más muestras y que no se pueden trabajar con las que ya se tiene en custodia; si logran superar los estándares de calidad, se procederá a realizar el o los informes periciales solicitados por la Fiscalía u otras dependencias.

Una vez terminada la o las pericias solicitadas, la OPERIT emitirá un oficio de remisión que acompañará a la pericia y cumplirá con la trazabilidad de informes hasta llegar a la dependencia solicitante, a través del Área de Cadena de Custodia, cabe resaltar que a este informe pericial se le adjunta las muestras remitidas por la dependencia al inicio de la investigación, con el formato de cadena de custodia respectivo y las actas de transcripción de las muestras dubitadas.

Cabe resaltar que el estudio de Holgado (2020, 41) señala que tan importante es la valoración de la pericia fonética en los juicios orales en delitos relacionados a crímenes organizados cumplen con «[...] los estándares de valoración establecidos, tales como la estimación, legitimidad, veracidad, aprobación técnica y científica.»

1.4.1. Casos en el Perú

En el presente subcapítulo se presentan algunos casos fiscales públicos que han sido concluidos y en la que han participado peritos lingüistas del Área de Fonética y Acústica Forense.

1.4.1.1. Caso Solsiret

Este acápite se adaptó de la investigación periodística de Zuloeta (2022, p. 61 - 94). Solsiret Rodríguez Aybar estudió en la Universidad Nacional Federico Villareal, en Lima, era madre de dos niños, feminista y realizó actividades por la lucha de los

derechos de las personas, en especial aquellas que sufrían violencia de género y abusos sexuales contra mujeres. Los padres de Solsiret la vieron por última vez un domingo 21 de agosto de 2016.

El 23 del mismo año, su pareja, Bryan Villanueva, llamó a los padres de Solsiret para comunicarles que se había ido, y que prácticamente abandonaba a su familia. Los padres de Solsiret llamaron con insistencia, pero no hubo respuesta alguna, así que al día siguiente acudieron a la Dirincri del Callao a asentar la denuncia por desaparición, sin embargo, el personal policial no quiso recibirla, quienes utilizaron excusas machistas como que era mayor de edad y quizá “se habrá ido con otro”.

Luego, Solsiret fue denunciada por abandono de hogar por su pareja. La buscaron exhaustivamente y escucharon sobre ella hasta principios del 2017, fecha en la que se rastreó el celular de Solsiret, lo cual trajo muchas suspicacias de qué había sucedido realmente con ella. La llamada que se efectuó desde el celular de Solsiret, provenía de su casa en el Callao que pertenece a la familia de Bryan Villanueva.

Hasta julio del 2017 no hubo ningún avance de investigación de la desaparición de Solsiret; persistió la búsqueda por la obligación social a las autoridades. La investigación se dejaría en manos de la Dirección de Investigación Criminal de Lima. Lograron geolocalizar la línea telefónica de Solsiret, cuya señal indicaba el condominio donde vivía con Bryan Villanueva, en la Alameda Colonial en el Callao, como posible ubicación. Posteriormente, se solicitó a la compañía telefónica Entel el levantamiento de las comunicaciones del celular de Solsiret, demoraron aproximadamente seis meses para entregar la información requerida por el fiscal encargado del caso. En febrero de 2018, citaron a la familia de Bryan Villanueva, de los cuales estuvieron Kevin Villanueva y Andrea Aguirre. La demora de la compañía telefónica en entregar los audios tuvo como consecuencia el archivamiento del caso temporal.

A inicios del 2019, el fiscal Jimmy Mansilla Castañeda solicita la ampliación de información de llamadas que le pertenecían a los investigados, la cual demoró aproximadamente nueve meses más. A finales del 2019, llamaron a los padres de Solsiret con la finalidad de escuchar un audio de un reporte de un robo del celular de su hija, pero, los padres de Solsiret mencionaron que esa no era su voz, por lo que enviarían el audio a la Oficina de Peritajes para peritar.

En ese momento, el caso pasó a manos del Área de Acústica y Fonética Forense, de la Oficina de Peritajes (OPERIT). Los peritos lingüistas concluyeron que la voz no era

compatible con la de Solsiret, por lo que, en aquel entonces, el gerente de la OPERIT, Dr. Danny Humpire, se movilizó con agentes de la Dirincri de Lima, e indicaron el chip del celular fue insertado al equipo de Andrea Aguirre, según la empresa telefónica; por lo que se citó a Andrea para recabar su voz en la OPERIT.

Andrea acudió a las instalaciones de toma de muestra de voz de la OPERIT, en la entrevista, la lingüista Giuliana Tito observó y analizó particularidades de habla en Andrea que eran compatibles con la voz de la llamada telefónica, el gerente, al tener esa información, intuía que ella podía ser la culpable. En febrero del 2020 se obtuvo el resultado de la pericia de homologación de voz, ante esa situación, Andrea confesó dónde estaba el cuerpo de Solsiret.

El 14 de febrero, Andrea y su pareja, Kevin Villanueva, fueron detenidos por agentes policiales. Andrea manejaba una versión de la historia, pero su confesión fue cambiando conforme pasaba el tiempo, lograron encontrar partes del cuerpo de Solsiret, la cabeza de Solsiret estuvo en la ducha de su habitación, asimismo se encontraron algunos huesos.

Pronto los padres de Solsiret sabrían la verdad del paradero de su hija. Finalmente, pese a todo lo sucedido, el 3 de junio del 2022, sentenciaron a Andrea a 28 años de prisión por homicidio calificado, mientras que, a Kevin, 30 años por los delitos de feminicidio, asimismo, reconoció el acoso sexual que propinaba a Solsiret. Los cómplices de encubrimiento recibieron aproximadamente 3 años.

1.4.1.2. Caso Mil voces

Este acápite resume los informes periodísticos sobre el caso “El mil voces”, elaborado por América Noticias y Perú21 (2018). Jean Jhonatan Rodríguez Espinoza, o también conocido en el mundo del crimen como “El mil voces”, está recluido en el centro penitenciario de Lurigancho desde el 2012, en el distrito de San Juan de Lurigancho; los delitos que le imputaron en ese momento estaban relacionados a estafas a jubilados y a esposas de efectivos policiales.

En varias oportunidades, desde el penal, el criminal lograba llamar e imitar las voces de funcionarios públicos, tales como la voz del viceministro del interior: Carlos León Romero, con el propósito de escarmentar a personal policial y solicitar la entrega de dinero o algún bien como licores u otros.

A partir del peritaje lingüístico (ver Figura 3), lograron concluir que las voces de la muestra dubitada y de la muestra indubitada de “El mil voces”, presentaban

coincidencias lingüísticas fuertemente compatibles en los niveles acústico, gramatical y discursivo. Para el análisis lingüístico compararon el contorno tonal de enunciados de las cinco muestras dubitadas e indubitadas, indicaron que el hablante utiliza un tono extremadamente alto (65%) para marcar preguntas absolutas y un rasgo de pausa entre enunciados; a nivel discursivo se encontraron coincidencias entre el uso no convencional del locativo “donde” con valor ilativo, así como la coincidencia de la formulación del saludo, la identificación del interlocutor y la presentación con el nombre completo y el cargo que suplantaba. Cabe resaltar que, en una entrevista posterior, el imputado confiesa y detalla su *modus operandi*, desde el penal de Lurigancho.

Figura 3

Captura de pantalla del peritaje del caso "Mil voces"



Nota. Muestra de segmentación en el programa Praat del caso “Mil voces”, de Latina Noticias, 2018.

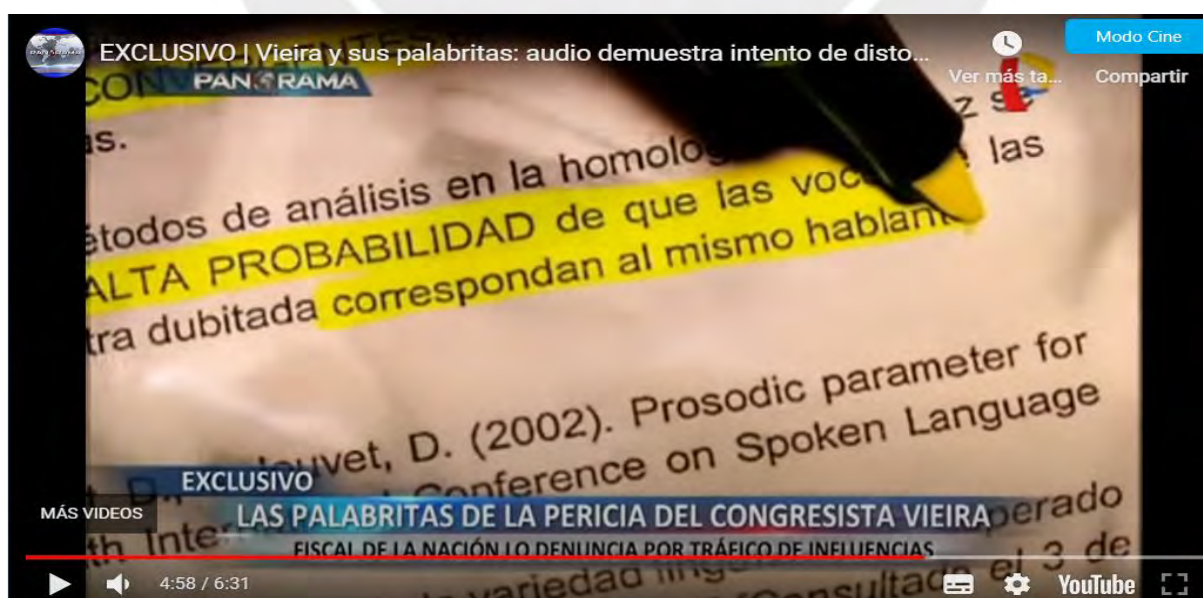
1.4.1.3. Caso Vieira

Este acápite se resume a partir de las noticias sobre el ex congresista Roberto Vieira, por Panorama; Redacción RPP; Macronorte, 2019. El caso del ex congresista no agrupado Roberto Vieira estuvo en la mira por los medios de comunicación respecto a su actitud poco colaborativa en una diligencia de extracción de voz en la OPERIT. Vieira fue investigado por presuntamente solicitar dinero a su primo, una cantidad exuberante para favorecerlo en unos trámites. Su propio familiar terminaría denunciándolo, por lo que se solicitó en ese momento una pericia de homologación

de voz. La fiscalía responsable del caso lo citó en más de una oportunidad, dos frustradas por su inasistencia, en la tercera, el 16 de abril de 2019, Vieira acudió a las instalaciones de la Oficina de Peritajes con el fin de someterse a la entrevista por parte de los peritos del Área de Fonética y Acústica Forense, los peritos Jhon Jimenez y Verónica Lazo iniciaron con la entrevista, al percatarse de la actitud poco colaborativa, las respuestas monosilábicas por parte de Vieira, intentos de distorsionar su voz, y su tan pausada respuesta ante las preguntas, se le indicó en varias ocasiones que hablara con normalidad, esto con el fin de obtener datos de habla espontánea, sin embargo, hizo caso omiso, y siguió respondiendo de la misma forma, tanto en la entrevista como en la lectura del protocolo de toma de muestra de voz del Área de Fonética y Acústica Forense. A pesar de la forma de hablar que empleaba el investigado no cumplía con las características de una interacción comunicativa “natural” o “fluida”, la fiscal de la nación, quien en ese entonces fue Zoraida Ávalos, solicitó muestras públicas de la voz del ex congresista Vieira para complementar a la muestra recabada por los peritos en la Oficina de Peritajes. Finalmente, se concluyó que la voz objetivo de la muestra dubitada, o sea de la voz que presuntamente le pertenece a Vieira eran compatibles con las del mencionado ex congresista (ver Figura 4), en la pericia se indica que existe similitud lingüística y alta probabilidad de que sean las mismas voces.

Figura 4

Captura de pantalla del reportaje Vieira y sus palabritas



Nota. Captura de pantalla del resultado de la pericia de homologación de voz del caso Vieira, de Panorama, 2019.

1.5. Antecedentes

Como antecedente directo para la presente tesis se tiene el artículo «*Identificación de locutor a partir de la fonética forense: aplicación del software SplitsTree4 para una organización esquemática de los datos lingüísticos*» (2022), de Jimenez *et al.*; en el cual, los autores utilizan este software para esquematizar y organizar las características fonéticas de locutores en su investigación. Los investigadores utilizan los métodos complementarios y el software SplitsTree 4 para esclarecer un caso de la fiscalía, en la cual se presumía que quien realizó ciertas llamadas para pedir un soborno a la familia de una persona agraviada era uno (01) de seis (06) efectivos policiales objetivo. Con este método, se concluyó que uno de ellos era el autor de las llamadas delictivas.

Otro estudio relacionado a la identificación de locutor es el caso de Lusi Tralcal, Rosas *et al.*, (2011) compararon muestras de habla de Luis Tralcal, un investigado por terrorismo, de acuerdo al Ministerio Público de Chile. La muestra indubitada de Tralcal no se obtuvo mediante una toma de muestra de voz, sino de una declaración de audiencia, la cual cumplía con los estándares para ser analizada, y la muestra de comparación, o sea la muestra dubitada, era la grabación de una reunión entre dirigentes mapuche y un funcionario de la CONADI, para lo cual, los investigadores utilizaron el cálculo del cociente de probabilidad, también llamado por sus siglas en inglés como LR. Compararon el segundo formante (F2) del sonido vocálico /a/ de toda la muestra, que consistía en veinticinco mapuches como población de referencia, cuatro locutores del CONADI y la voz de Tralcal, los autores concluyen que el F2 comparado de las muestras indican un cociente menor a la unidad, lo cual se interpreta como locutores diferentes.

Asimismo, en la fonética forense se utilizan diferentes parámetros para analizar, por ejemplo, Cicres (2014) utiliza las pausas llenas de muestras dubitadas para la comparación forense de voz. En su estudio corrobora si el parámetro de las pausas llenas es discriminante para identificar voces.

En el ámbito de la lingüística ha sido utilizado para el ordenamiento de datos con corpus más amplios y de mayores categorías gramaticales, con el propósito de elaborar redes filogenéticas, por ejemplo, Zariquiey *et al.* (2017) utiliza el programa SplitsTree 4 para realizar una clasificación con redes filogenéticas de las lenguas

pertenecientes a la familia Pano (Purús, Ucayali). Ellos consideran un corpus compuesto por rasgos léxicos y gramaticales.

Otro estudio en particular es la documentación que realiza Heggarty (2005) para la creación de redes y esquemas arbóreos para lenguas andinas peruanas de las familias uru-chipaya, quechua y aimara. En total utilizó ciento cincuenta mil conceptos para crear cuantificadores de semejanza, los mismos que fueron representados con la red filogenética de NeighborNet, algoritmo ofrecido dentro del paquete del software SplitsTree4².

1.6. Problema de investigación

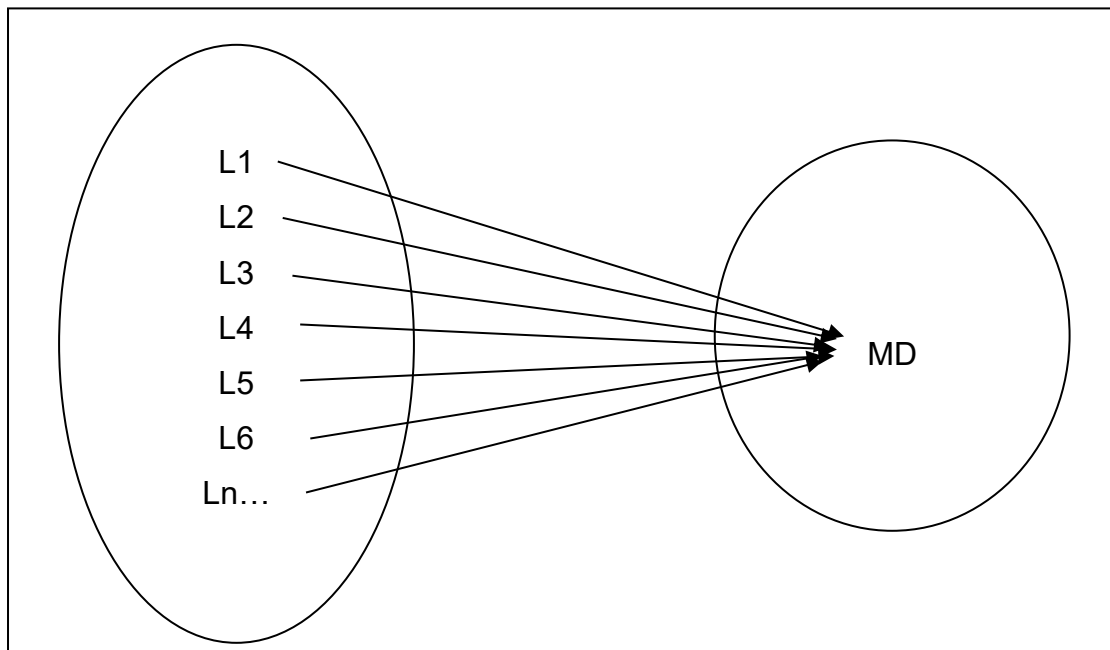
En esta investigación se tiene varios implicados en el caso, la hipótesis fiscal no se limita a una sola correspondencia de voz, por lo que se optó por realizar una pericia de identificación de voz. Al existir una cantidad de datos exorbitantes, se obtuvo un resultado del análisis fonético-acústico extenso cuya conclusión no podría interpretarse correctamente.

De acuerdo con el trabajo de Jimenez *et al.* (2022), una de las alternativas para explicar óptimamente un caso de identificación en el Perú es utilizar el programa SplitsTree4. Pero, el caso fiscal que ellos evaluaron consta de seis locutores (L01 a L06) y una sola voz correspondiente a dos audios de la muestra dubitada (MD_01 y MD_03), por lo que no hubo complicaciones si la muestra dubitada le correspondería a más de una persona. La forma en la que trabajaron Jimenez *et al.* correspondería a la Figura 5.

² La Facultad de Ciencias, de la Universidad Tubingen, trabaja constantemente en actualizar el software SplitsTree, actualmente existe la versión 6, llamado *SplitsTreeCE*, la cual indica que el programa aún sigue en desarrollo activo (D. H. Huson and D. Bryant, 2006).

Figura 5

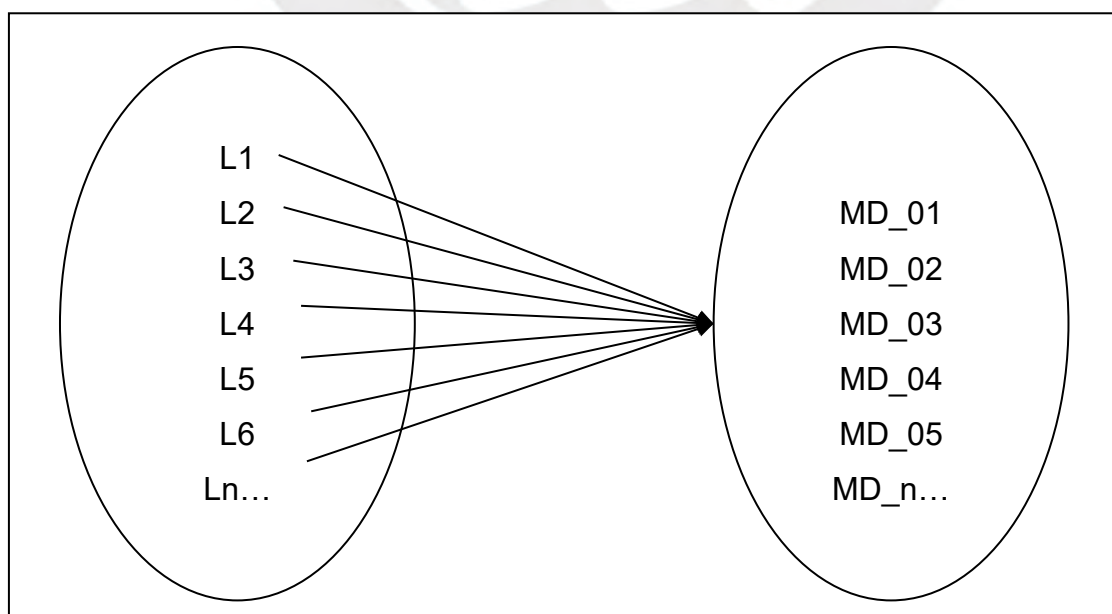
Tipo de análisis ejecutado en Jimenez et al. (2022)



La comparación de voces se ejecuta en una sola dirección, todos los locutores indubitados versus la muestra dubitada. En contraste a la presente investigación, la muestra dubitada no correspondería a una sola voz, sino a distintas voces, de acuerdo a las actas de transcripción remitidas por la fiscalía, el análisis se realizaría como se muestra en la Figura 6.

Figura 6

Tipo de análisis de la tesis



A partir de lo expuesto, la Fiscalía solicitó comparar la voz de las personas implicadas, en total tres (03), en un caso de corrupción de funcionarios y así determinar si las voces de las muestras dubitadas (MD_01 a MD_08) pertenecen a uno o a varios de ellos. Por ello, en el presente estudio se plantea como interrogante lo siguiente: **¿el software SplitsTree4 permite organizar adecuadamente muestras de habla para la identificación de locutores con más de una voz dubitada?**

1.7. Objetivos

El objetivo central de la tesis es corroborar si el *software* SplitsTree4 sirve para ordenar rasgos fonéticos extraídos de una cantidad determinada de muestras indubitadas y voces que pertenecen a muestras dubitadas diferentes, y dilucidar si con el algoritmo de ordenamiento del software la compatibilidad entre muestras indubitadas con muestras dubitadas.

1.8. Metodología

La Fiscalía remitió a la Oficina de Peritajes dos discos: uno contiene muestras de audio para su análisis y el otro contiene grabaciones de audiencias relacionadas con un individuo investigado. Esta investigación de la Fiscalía está relacionada con la presunta comisión del delito de Cohecho Pasivo Propio en el ejercicio de la función policial, tal como se establece en el artículo 395-A del Código Penal, y afecta a la Policía Nacional del Perú (PNP). La Fiscalía solicitó la presencia de siete (07) individuos investigados para llevar a cabo la toma de muestras de voz, pero solo cuatro (04) de ellos se presentaron. Además, la Fiscalía ha adjuntado las transcripciones correspondientes a la muestra dubitada.

Las muestras dubitadas ya estaban configuradas en formato *wav* y en un solo canal de audio, o sea, *mono*, por lo que no fue necesario acondicionarlos para la posterior lectura con los software biométrico y análisis fonético-acústico. Ya con los audios segmentados, concatenados y consignados a un rótulo, se procedió a ejecutar el *software* SIS II para verificar los estándares de calidad de la señal, esto en base a la relación entre la señal-ruido o también llamado (SNR), también se observa el porcentaje de ruido tonal, reverberación, dispersión, saturación y la actividad vocal o de voz (VAD), ver 1.8.4.2. Sobre el software SIS II (STC). De esa forma, se evidencia si todos los audios o solo algunos de ellos son óptimos para ser utilizados en los análisis

biométricos de comparación automática y el fonético-acústico. Para el primer análisis, el sistema de reconocimiento automático nos brinda, a partir de la aplicación de cálculos y la clasificación de distancias de similitud entre vectores de caracterización matemática del habla, la razón de verosimilitud entre muestras indubitadas y muestras dubitadas, es decir, si se conoce la voz en la que el locutor la reconoce como suya (indubitada) y las voces de las que no se sabe la autoría (dubitadas) son en alta o baja probabilidad el mismo o diferente hablante.

Para las diligencias de tomas de muestra de voz se utilizó las siguientes herramientas:

- Una grabadora profesional de marca Zoom H5n Pro,
- Un micrófono de tipo diadema modelo Shure WH20,
- Y con lo que respecta al lugar físico del análisis del caso, se usaron una computadora con procesador Core i7, de 8gb de RAM y un audífono AKG.

El procedimiento a seguir en las diligencias de toma de muestra de voz consiste en realizar entrevistas libres, en las que el lingüista forense a cargo pregunte a los imputados uno por uno sobre temas ajenos a la investigación fiscal, tales como sus pasatiempos, gustos, metas, entre otros; adicional a ello se les indica a los imputados a leer el protocolo de toma de muestra de voz sugerido en la *Guía de elaboración de pericias fonético-acústicas de homologación de voz*, perteneciente al Ministerio Público (2020). Esta guía protocolar consta de un número limitado de palabras y frases, también consta del ejercicio de pronunciación de las vocales de la lengua materna del, en este caso, imputado, la lectura de un trabalenguas y un cuento (Ver Anexo 01: Modelo de Protocolo de toma de muestra de voz»)

Por otra parte, las muestras dubitadas fueron segmentadas y agrupadas según lo solicitado por la Fiscalía, por lo que se obtuvieron distintas voces intervinientes, a las cuales se les asignó un rótulo (ver Tabla 5. Lo mismo para las muestras indubitadas, fueron segmentadas y concatenadas para obtener solo las voces de los investigados y así poder utilizarlas al momento del análisis.

1.8.1. Alcance, enfoque y diseño de la investigación

1.8.1.1. Alcance

El alcance de la presente investigación es descriptivo porque, según Hernandez *et al.* (2014, 92), «se busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos [...] o cualquier otro fenómeno que se

someta a un análisis.» En ese sentido, el estudio se concentra en la recopilación de datos fonético-acústicos de muestras dubitadas e indubitadas que servirán como variables para corroborar si alguna de las voces con compatibles entre ellas.

1.8.1.2. Enfoque

Hernandez *et al.* (2014, 7) indica que el enfoque de los estudios cualitativos «pueden desarrollar preguntas e hipótesis antes, durante o después de la recolección y el análisis de los datos». Asimismo, los autores señalan que los estudios de corte cualitativo se «basan más en una lógica y proceso inductivo [...]. Van de lo particular a lo general. Por ejemplo, [...] el investigador entrevista a una persona, analiza los datos que obtuvo y saca conclusiones [...] y analiza más entrevistas para comprender el fenómeno que estudia.» (p. 8)

Por tanto, en este estudio se adopta un enfoque cualitativo, ya que se emplean métodos de recolección de datos que no siguen un estándar previamente establecido en su totalidad. El objetivo principal es obtener las perspectivas y opiniones de los participantes a través de preguntas abiertas y la captura de datos a partir del lenguaje hablado. Se centra en examinar los fenómenos y características fonético-acústicas presentes en las muestras, tanto las dubitadas como las indubitadas. Además, se incorpora un programa biométrico para analizar los registros de voz, y los rasgos fonético-acústicos se codifican de manera binaria para su posterior procesamiento mediante el programa SplitsTree 4.

1.8.1.3. Diseño

El tipo de diseño es un estudio de caso cualitativo, ya que la «materia prima de los datos la constituyen narrativas verbales, escritas y audiovisuales, [...] tienen como objetivo documentar una experiencia [...] en profundidad o entender un fenómeno desde la perspectiva de quienes lo vivieron. El estudio de caso cualitativo no persigue ninguna clase de generalización.» (Hernandez *et al.*, 2014, 17 - 18), como parte de las funciones de los estudios de caso están el identificar patrones de un fenómeno en su ambiente natural y proveer datos para evaluar procesos, programas e individuos.

1.8.2. Población

Una población abarca a todos los individuos o elementos que comparten ciertas características en común, representando en su totalidad a los miembros de la unidad de análisis. Esta concepción de población guarda similitud con la idea de un conjunto,

y su alcance es definido por el investigador de acuerdo con los criterios que considere apropiados. En consecuencia, el tamaño de la población puede fluctuar según la definición que el investigador establezca, pudiendo ser más extenso o más limitado. (Mejía, 2015) La población está constituida por los hablantes de español peruano a nivel nacional.

1.8.2.1. Muestra

Una muestra se define como un grupo más pequeño que forma parte de la población total. Para que un conjunto de individuos de la población sea considerado como muestra, es esencial que todos los elementos que la componen sean miembros de la población en su totalidad. Es por ello que se enfatiza que una muestra debe ser representativa de la población, es decir, debe reflejar de manera precisa las características generales que se encuentran en la población completa. (Mejía, 2005) La muestra está constituida por los tres (03) locutores implicados que poseen muestra indubitada, sus edades oscilan entre los 27 a 55 años, los tres son hablantes de español y son de distintas partes del Perú. El corpus de la investigación se obtuvo mediante diligencias de toma de muestra de voz y en el caso de un locutor, mediante dos audios de audiencia.

1.8.2.2. Variables

Una variable de estudio es una característica, atributo, propiedad o concepto que se investiga, mide o analiza en un estudio de investigación con el fin de comprender mejor un fenómeno o responder a una pregunta de investigación. Estas variables pueden ser observables y medibles, como la edad, el género, la altura, el nivel de ingresos, la temperatura, etc., o pueden ser conceptos abstractos que se definen y operacionalizan de manera específica para fines de la investigación, como la satisfacción del cliente, la calidad de vida, la actitud hacia un producto, etc. (Hernandez *et al.*, 2014, 105)

La elección de variables de estudio en investigaciones forenses de la voz es diversa, los parámetros y características de la voz son muchos, y la mayoría de ellos sirven para la comparación, verificación y/o identificación de voces. De acuerdo a los parámetros que escoja el especialista, se debe guiar la investigación a cumplir requisitos mínimos. Colls (2016, 11) señala algunos de ellos, tales como presentar «un número elevado de apariciones en la muestra de habla [...]; soportar bien las

distorsiones que el canal de comunicación pueda generar; ser relativamente fáciles de extraer y de medir; [...] tener un grado pequeño de variabilidad intra-hablante pero un elevado grado de variabilidad inter-hablante y, por último, ser independientes entre sí.»

La variable más importante para el presente estudio son el *rasgo fonético-fonológico*, el cual se manifiesta de diferentes formas de acuerdo al análisis del habla espontánea.

1.8.2.2.1. Operacionalización de variables

La variable *rasgo fonético-fonológico* se define como particularidades individuales con alto valor identificativo o discriminatorio diferenciadas del valor estándar. (Garayzábal, 2019) En la Tabla 1 se presentan los correlatos fonéticos de los indicadores.

Tabla 1

Operacionalización de las variables

Variable	Definición conceptual	Dimensión	Indicadores	Ítems
Rasgo fonético-fonológico	De acuerdo con Garayzábal (2019) es una particularidad individualizante del habla de un locutor comparado con el lenguaje estándar.	Procesos que afectan al segmento		¿Cómo se realizan los fonemas a nivel superficial de los individuos?
		Procesos que cambian al segmento	Modo y lugar de articulación	¿Es frecuente la aparición del rasgo fonético-fonológico?
		Procesos característicos del segmento		¿En qué se diferencia de la realización estándar?

1.8.3. Metodología combinada

Ante la solicitud de la Fiscalía para realizar la homologación de voz de los investigados y conocer si es que alguno de ellos ha participado en la conversación de las muestra dubitadas (muestras de voz de las cuales no se sabe su pertenencia), se decidió utilizar dos enfoques complementarios en la elaboración del informe pericial fonético forense: el primero es el método clásico o manual, que implica el análisis perceptivo-auditivo del habla de un hablante, y el segundo método es la comparación automática mediante algoritmos que analizan las características biométricas de las muestras de voz proporcionadas al programa. (Garayzábal *et al.*, 2019).

1.8.3.1. Método clásico

De acuerdo con Garayzábal *et al.* (2019), el método clásico se divide en tres: a) análisis perceptivo-auditivo, este análisis se desarrolla mediante la percepción auditiva de las grabaciones dubitadas e indubitadas para ser comparadas a nivel alofónico, por grupos fónicos y estructuras discursivas de rango superior como la entonación, el acento, las dudas, entre otros, asimismo es de carácter complejo, debido a que se debe realizar una escucha activa de los registros de voz con atención a las particularidades del sociolecto, y a nivel suprasegmental como el timbre de voz, las pausas o los patrones entonativos; el b) análisis fonoarticulatorio implica dos enfoques: uno centrado en el estudio de la fonética y la articulación y otro centrado en aspectos lingüísticos. En el primero, se busca identificar características específicas del habla de un locutor, comparándolas con estándares de pronunciación del español. Esto implica utilizar la percepción auditiva y representaciones gráficas de la señal de voz para detectar las diferencias individuales que son altamente identificativas. En el segundo enfoque, se abordan los análisis desde diferentes niveles lingüísticos, que incluyen aspectos fonéticos, morfosintácticos y léxico-semánticos; por último el c) análisis acústico-espectrográfico que evalúa las resonancias y formantes, la frecuencia fundamental o F0, la desviación estándar, la estructura acústica, los rasgos de sonoridad y tonalidad, entre otros parámetros acústicos relacionados a la forma producida de las consonantes y vocales de, en este caso, el español, estos rasgos son también llamados índices acústicos.

En ese sentido, se emplea este método desde un enfoque cualitativo mediante la comparación de rasgos lingüísticos, los cuales representan los rasgos fonético-acústicos más comunes presentes en las voces de cada sospechoso, con las voces

de las muestras dubitadas. Este análisis permite identificar similitudes y diferencias visuales en las características de las ondas sonoras, mediante el programa Praat.

1.8.3.2. Método del sistema automático

Se pueden encontrar programas que llevan a cabo comparaciones automáticas de voz utilizando tecnología biométrica, conocidos como Sistemas Automáticos de Reconocimiento de Locutores (SARL). Estos programas realizan cálculos y clasificaciones basados en las características biométricas de la voz. Determinan distancias de similitud entre vectores matemáticos que caracterizan el habla, con el propósito de verificar o identificar según la hipótesis planteada. El proceso de identificación de locutores logra clasificar la señal de habla de un desconocido como concerniente a otro dentro de un conjunto de locutores determinados; este proceso toma diferentes parámetros tales como el sexo, la lengua, el canal por el que se transmite, entre otros, por su parte, el proceso de verificación de locutores se caracteriza por la comparación de dos señales distintas de habla, en la que tiene dos opciones, o la de rechazar o aceptar que ambas muestras pertenezcan a un mismo locutor, esta decisión se le atribuye a la razón de verosimilitud (Likelihood Ratio o LR). Este estudio comparativo de muestras tiene un proceso previo, el cual es constatar si las señales vertidas en el programa poseen la calidad suficiente para determinar si son aptas o no para las conclusiones científicas dentro de un dictamen pericial. Si la calidad de las grabaciones es suficiente, se iniciará la comparación de las muestras. Algunos de los programas que se utilizan para el estudio biométrico de voz son el Glottex, extrae y evalúa los parámetros fisiológicos de una fuente glotal, Identivox, programa de reconocimiento utilizado por la Guardia Civil de España, Batvox, utilizado por las fuerzas de seguridad a nivel mundial y la policía científica española. Estos programas de biometría de voz miden parámetros en las vocales, tales como los formantes, la frecuencia fundamental o también llamada F0, y utiliza como base de datos diferentes hablantes, diferentes circunstancias o canales, distintos tipos de habla. (Garayzábal *et al.*, 2019).

1.8.3.3. Método de la red filogenética

Posterior al método combinado, se pasa a una etapa cuantitativa en la que los rasgos fonético-acústicos identificados son codificados en códigos binarios para su posterior procesamiento (ver 1.8.4.3. Sobre el software SplitsTree 4 y el análisis filogenético). Para

este fin, se utiliza el programa Splitstree 4, que facilita el análisis y la representación gráfica de relaciones filogenéticas entre las muestras de voces. Dicha herramienta permite visualizar los datos codificados y construir árboles filogenéticos basados en similitudes y discrepancias en los rasgos acústicos identificados. En ese sentido, este estudio combina un enfoque cualitativo, mediante el análisis de espectrogramas para comparar las características fonético-acústicas, y un enfoque cuantitativo, donde los rasgos identificados son codificados y procesados utilizando el software Splitstree 4.

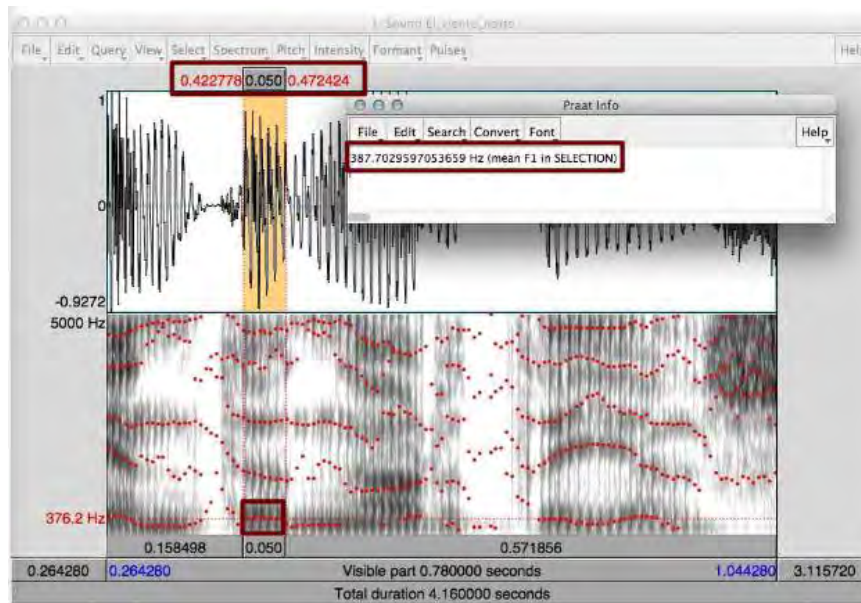
1.8.4. Herramientas metodológicas

1.8.4.1. Sobre el software Praat y el análisis lingüístico-espectral

En la actualidad, para un análisis de carácter fonético, se suele emplear el *software* Praat (Paul & David, 2022), en este se puede observar diversas formas de representar el sonido, una de ellas es el espectrograma, en el cual se pueden introducir ediciones y conducir análisis, además de otros elementos y representaciones acústicas. En la ventana de Praat (ver Figura 7) se puede añadir la visualización de los formantes, esto con la finalidad de medirlos, o lo que compete de acuerdo al análisis particular del investigador, se puede seleccionar un fragmento de la señal sonora y se pueden obtener los valores como la cantidad de Hertz de un grupo formántico (Listerri, 2022). El espectrograma sirve como una herramienta de representación para el análisis de una señal acústica, que es procesada eléctricamente. «En un espectrograma se puede visualizar la energía del contenido frecuencial de la señal a través de la duración de tiempo. Esta representación se realiza en tres dimensiones: temporal, frecuencial y de amplitud.» (Paz *et al.*, 2020, 35) Gracias a la representación del espectrograma, el investigador puede seleccionar un segmento de muestras por medio de una ventana (de algún software), en el tamaño que se requiera, según el tipo de análisis a realizar (Machuca *et al.*, 2014).

Figura 7

Extracción de valores formánticos de la señal sonora



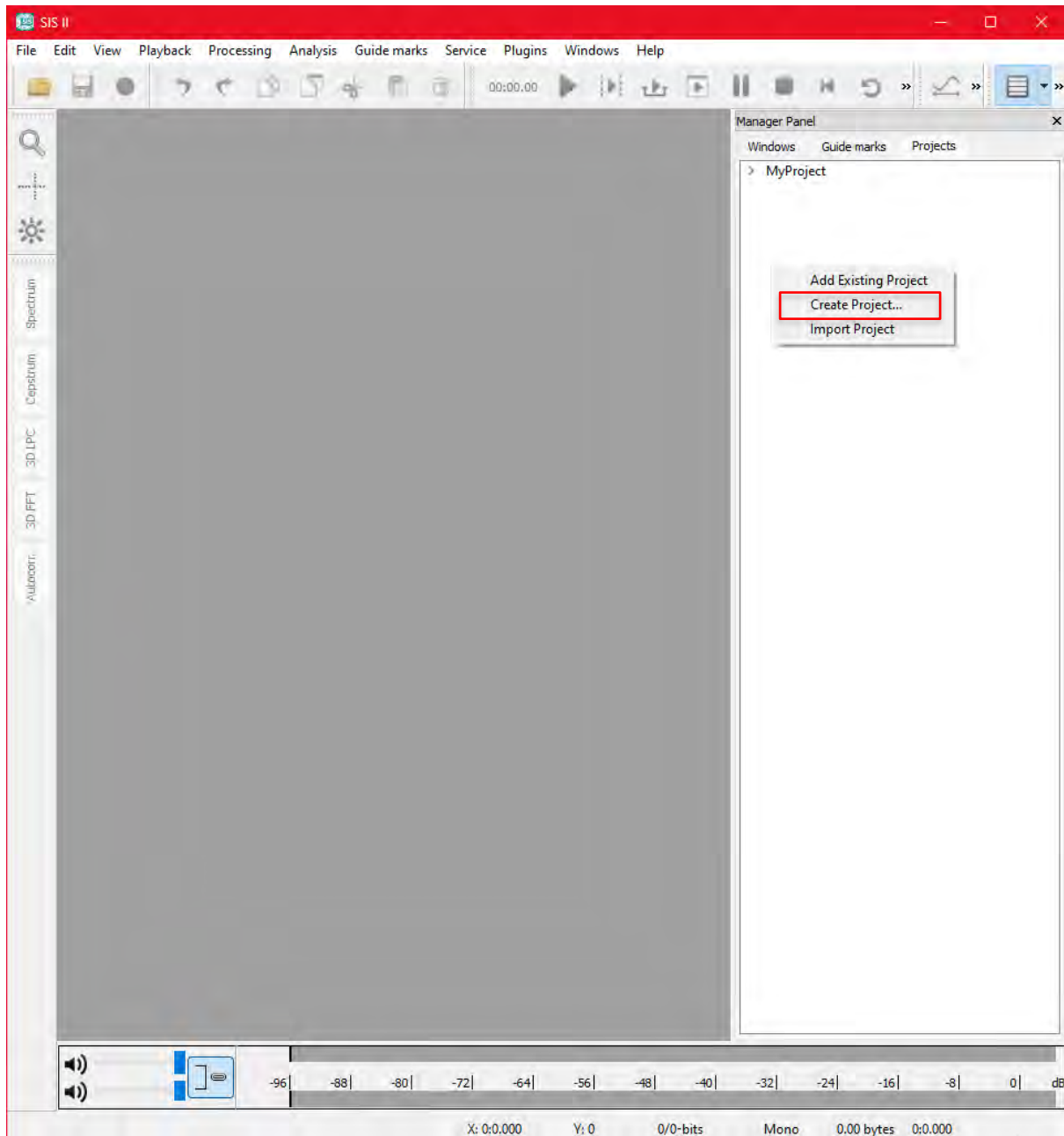
Nota. Captura de pantalla de la ventana del programa Praat en la que se visualiza el sonido mediante un espectrograma, en él se observan los formantes, de Listerri, 2022.

1.8.4.2. Sobre el software SIS II (STC) y el análisis biométrico

En este acápite se describe y explica los pasos a seguir para comparar dos muestras de voz en formato *wav*. Después de segmentar y concatenar los audios de las muestras dubitada e indubitada, se procede a crear un proyecto en el panel administrativo del programa (ver Figura 8), dar un clic derecho y seleccionar *Create Project*.

Figura 8

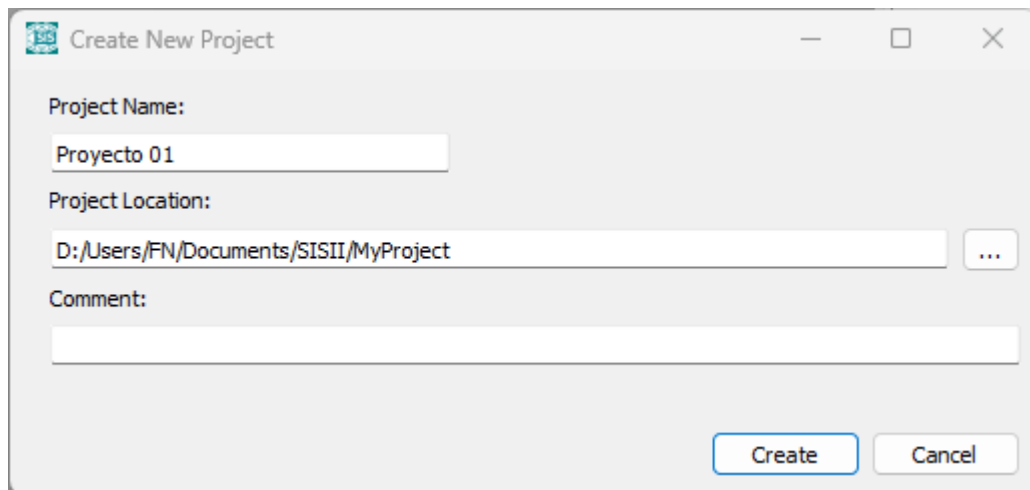
Captura de la pantalla principal del programa SIS II



Esto permite que cuando el proyecto se cree, se pueda volver a abrir los archivos que se hayan guardado. Después de seleccionar la opción *Create Project*, se abre una ventana emergente para colocar los datos del proyecto, como el nombre, la ubicación del guardado del proyecto y comentario del proyecto (ver Figura 9).

Figura 9

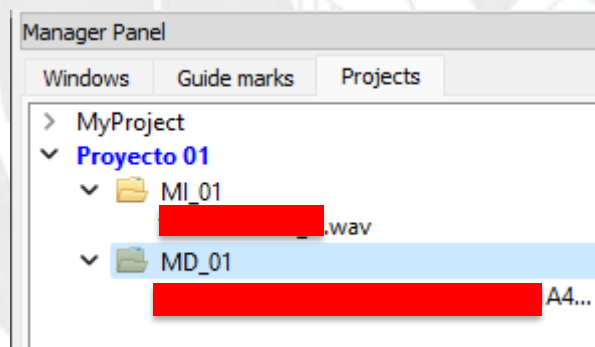
Captura de la pantalla de la ventana emergente al crear un proyecto en SIS II



Al crear un proyecto, el programa permite generar *Folders* en las que irán los audios, en este caso se tiene dos muestras como ejemplo, una muestra indubitada y una muestra dubitada (ver Figura 10).

Figura 10

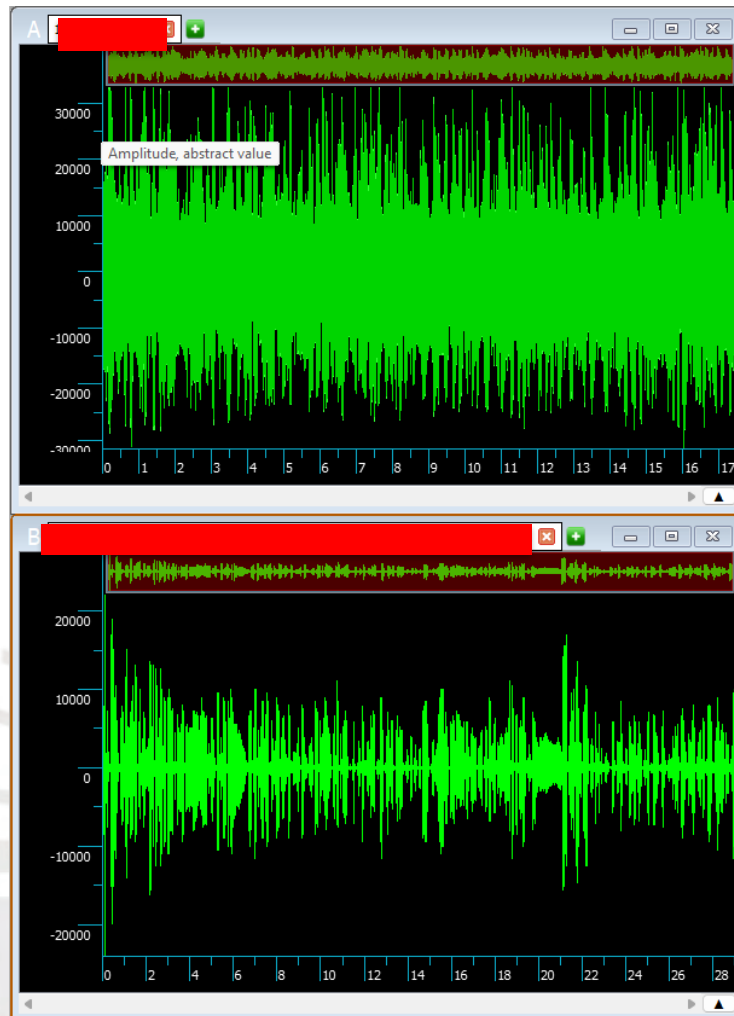
Captura de la pantalla de los audios importados a SIS II



Al seleccionar estos audios, instantáneamente aparecen en la pantalla principal, tla como se muestra en la Figura 11, pueden reproducirse, seleccionar intervalos determinados de tiempo y realizar distintos análisis, en este caso, se realizará un análisis de calidad a cada uno.

Figura 11

Captura de la pantalla de los audios importados en la pantalla principal de SIS II



La opción de análisis de calidad de los audios a comparar se encuentra en la opción de *Plugins*, en la barra de herramientas superior. Se debe seleccionar el audio a analizar y presionar la opción de *Análisis de calidad*. En este se verificará si es aceptado, parcialmente aceptado o no aceptado, según corresponda (ver Figura 12).

Figura 12

Captura de la pantalla del análisis de la calidad de un audio en SIS II

The screenshot displays the audio quality analysis interface in SIS II. It is organized into several sections, each with a help icon (question mark):

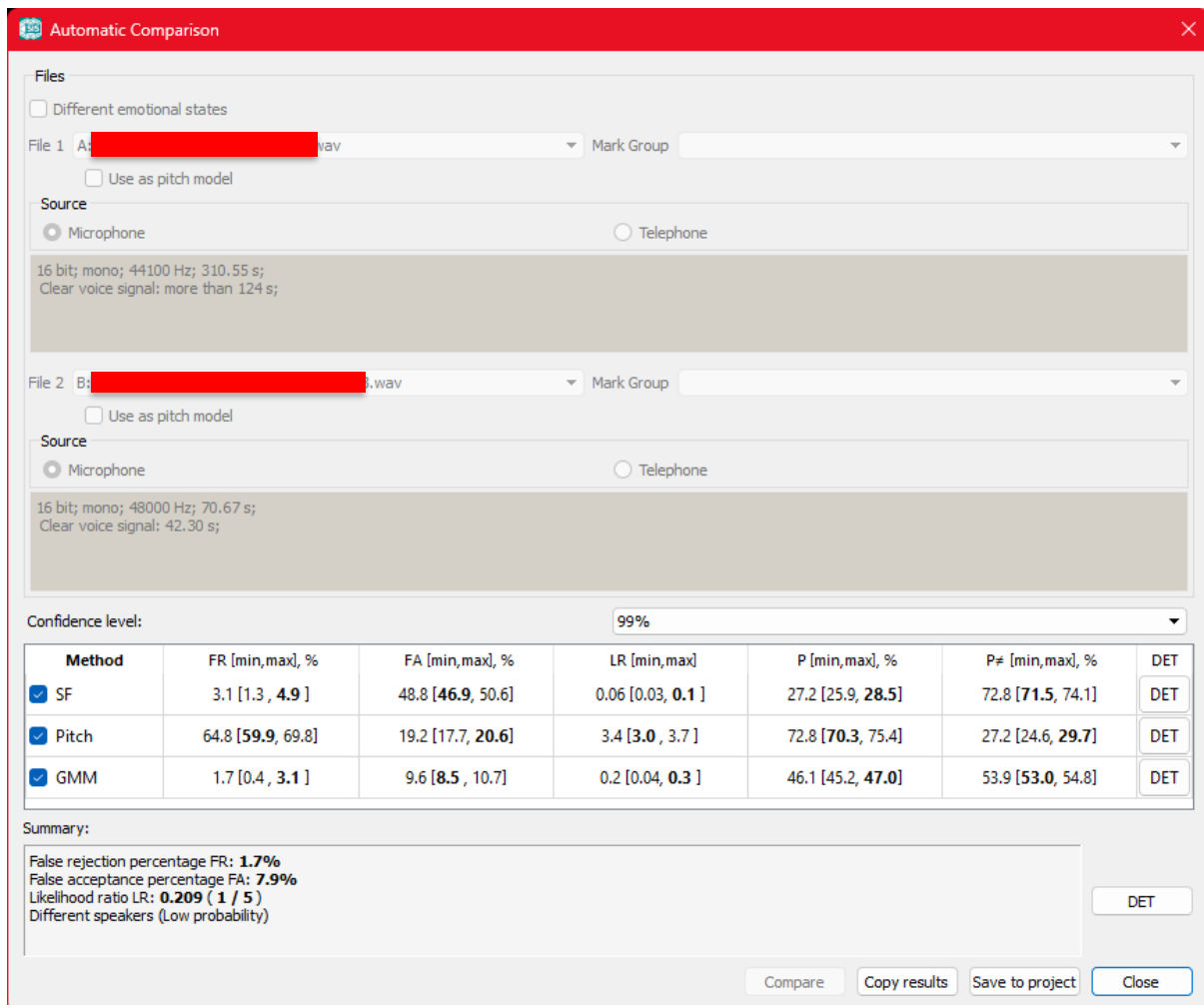
- Emotional State:** A dropdown menu set to "Regular".
- Spectrum Characteristics:** Includes "Frequency response" (188 --- 1547 Hz) and "SNR" (100 --- 3300 Hz, 24 dB). Both have "Plot a chart" buttons.
- Tonal Noise:** States "No stationary tonal noise detected".
- Reverberation Time:** Shows "Average Time" (241 ms) and "Dispersion" (10).
- Clipping:** States "No clipping detected".
- Voice:** Shows "Total" (14.95 s) and "Total" (51.48 %). Below this, a summary message reads: **The recording can be partially accepted for voice ID analysis**.
- Process:** Two radio buttons: "Entire Signal" (selected) with a value of "29.03 seconds", and "Selected Area" with a value of "0.0_ seconds".
- Source:** Two radio buttons: "Microphone" (selected) and "Telephone".

At the bottom, there are four buttons: "Save to project", "Copy", "Calculate", and "Close".

De acuerdo a lo que se indica en la Tabla 10, el audio resultara óptimo o no para el proceso de comparación de voz. Finalmente, si los audios son aprovechables para el proceso de comparación de voz, se comparan con el *plugin Automatic Comparison*, este emite un resultado de acuerdo a los parámetros de la estructura formántica, frecuencia fundamental y el modelo mixto gaussiano. Si el resultado es menor a 1, el programa lo califica como diferente hablante, pero si es igual o mayor a uno, lo califica como mismo hablante, por ejemplo, en la Figura 13 se muestra que al comparar las dos muestras de voz, el programa calcula que son diferentes hablantes con baja probabilidad.

Figura 13

Captura de la pantalla del análisis de la calidad de un audio en SIS II



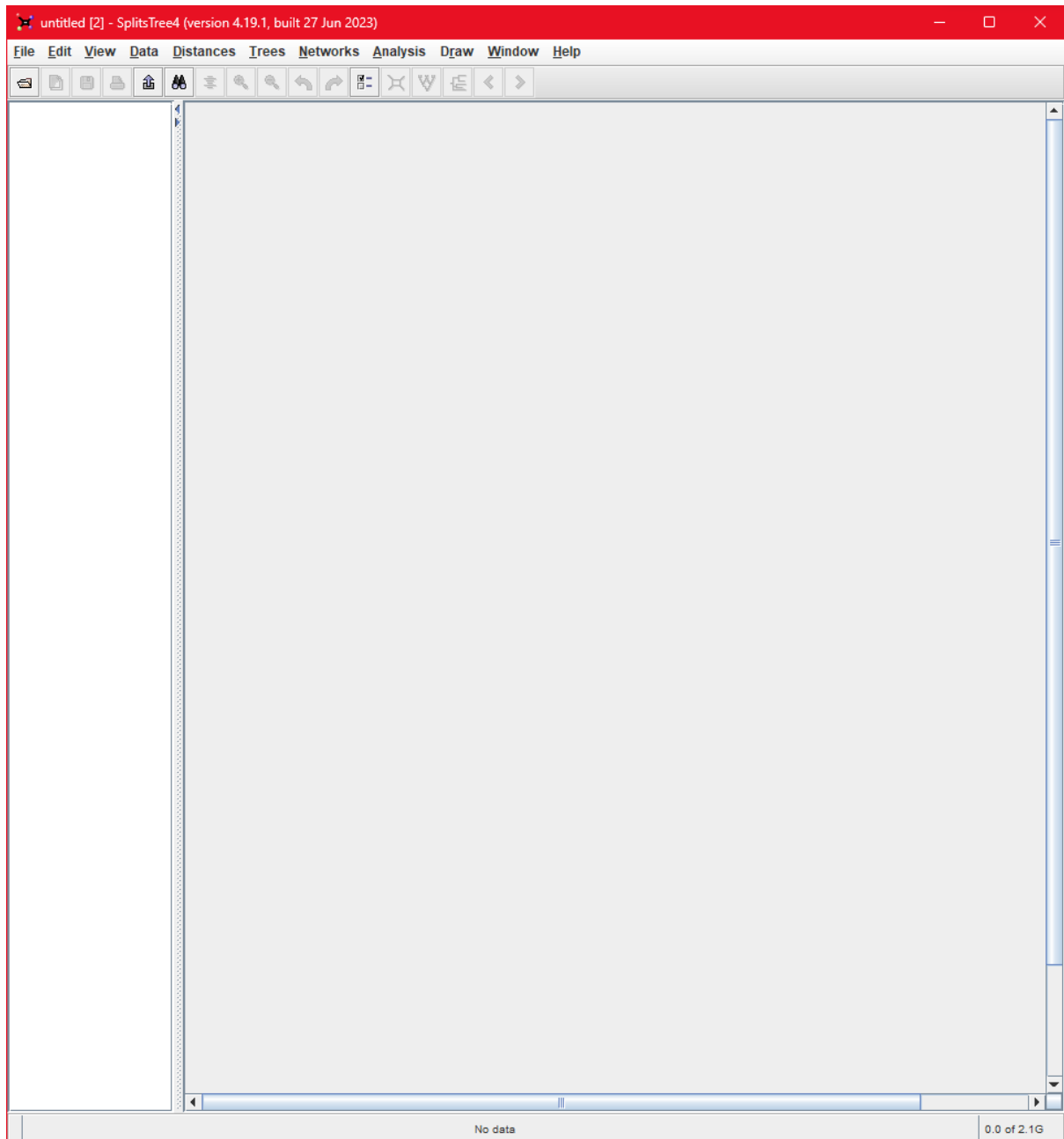
1.8.4.3. Sobre el software SplitsTree 4 y el análisis filogenético

El programa SplitsTree 4 es una herramienta de análisis de datos simples o complejos, de acuerdo a la naturaleza de los datos y a su sintaxis, interpretará mediante diferentes redes filogenéticas la mejor opción de ordenamiento y clasificación. Por ello, en este acápite se describe y detalla los pasos a seguir para introducir los datos que interpretará el programa, para eso se tiene como guía el estudio de Peetz (2014).

En primer lugar, se tiene la ventana principal del programa, en la que se interpretarán los datos que se vayan a introducir (ver Figura 14), por motivos de ejemplificación se trabajará con datos ficticios binarios, en total 45 rasgos de nueve locutores del mismo sexo y edades entre 30 a 40 años, y una voz que es compatible con alguno de los nueve locutores, en total sumarán diez voces.

Figura 14

Captura de pantalla de la ventana principal del programa SplitsTree4

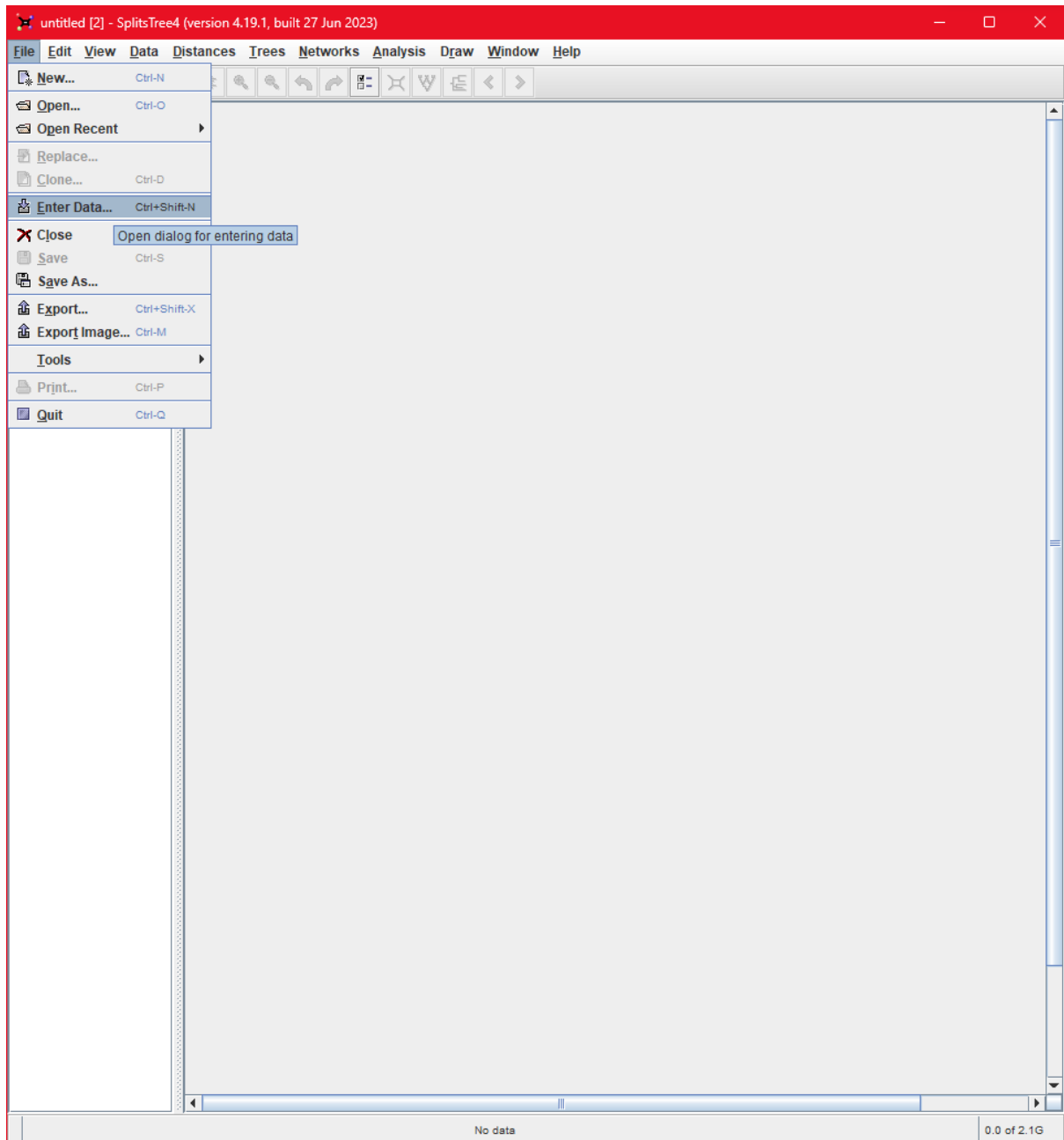


Cabe resaltar que el programa SplitsTree 4 fue creado para el uso de la bioinformática, pero eso no quiere decir que no se pueda usar para otras ciencias como la lingüística, es en ese sentido que el programa también es usado para crear redes filogenéticas entre lenguas de una misma familia (Maguire *et al.*, 2011; McMahon *et al.*, 2013). Sin embargo, para este caso, observaremos un ejemplo hipotético de cómo ordenar datos lingüísticos para corroborar la compatibilidad entre locutores.

En la ventana principal del programa se selecciona la opción de *File* y luego *Enter data*, tal como se muestra en la Figura 15. Luego aparecerá otra ventana para escribir el código con la información numérica de cada locutor.

Figura 15

Captura de pantalla del menú File del programa SplitsTree4



Ahora bien, aparecerá una ventana emergente para escribir el código, en él se escribe lo siguiente:

```
#nexus
```

Es el código inicial que permite interpretar la data posterior al programa, en este caso es un archivo *nexus*.

BEGIN Taxa;

Con esta sección del código el programa puede leer cuántos elementos o *taxones* se van a comparar.

DIMENSIONS ntax=10;

En esta parte del código, se coloca la cantidad específica de elementos a comparar, se debe reemplazar el número por la cantidad real de elementos.

TAXLABELS

Con este comando se le dice al programa cuáles son los nombres asignados a los elementos a comparar. En este caso se tienen diez (10), así que habrán diez líneas de código que comparará el programa.

[1] 'L1'

[2] 'L2'

[3] 'L3'

[4] 'L4'

[5] 'L5'

[6] 'L6'

[7] 'L7'

[8] 'L8'

[9] 'L9'

[10] 'LD'

Después de haber escrito los nombres de acuerdo a la cantidad de elementos (*ntax*), se debe asegurar que los nombres estén en comillas simples y evitando espacios.

;

END; [Taxa]

Aquí termina la parte inicial de la sintaxis, se debe copiar tal y como aparece. Sino al ejecutar el programa, aparecerá error en la codificación.

BEGIN Unaligned;

Esta codificación le dice al programa que debe iniciar a leer el siguiente contenido.

DIMENSIONS ntax=10;

Esta parte del código sólo es la repetición del número de elementos o *taxas*.

FORMAT

datatype=STANDARD

Aquí se describe la naturaleza de los datos.

missing=?

Si algo falla en la codificación de los elementos, el programa utilizará en la matriz el símbolo escogido.

symbols="0 1 2 3 4 5 6 7 8 9"

Esto indica qué símbolos se utilizan para diferenciar los elementos de la matriz.

labels=left

;

MATRIX

'L1' 011000011010101001001100000111110010110001101,
'L2' 111001010000101101110011100110100100110111011,
'L3' 100111010011110010000110001011100111000000011,
'L4' 001000111100101010011101010010011101111010011,
'L5' 000110111101001100101110001011100101001000110,
'L6' 010010001011101111001111010001111010010110010,

'L7' 111101110011010010111001101110011101011001001,
 'L8' 101111000010101100000110001101101000001111101,
 'L9' 011101000010011000010111110001011110101101100,
 'LD' 111001010000101101110011000110100100110111011;

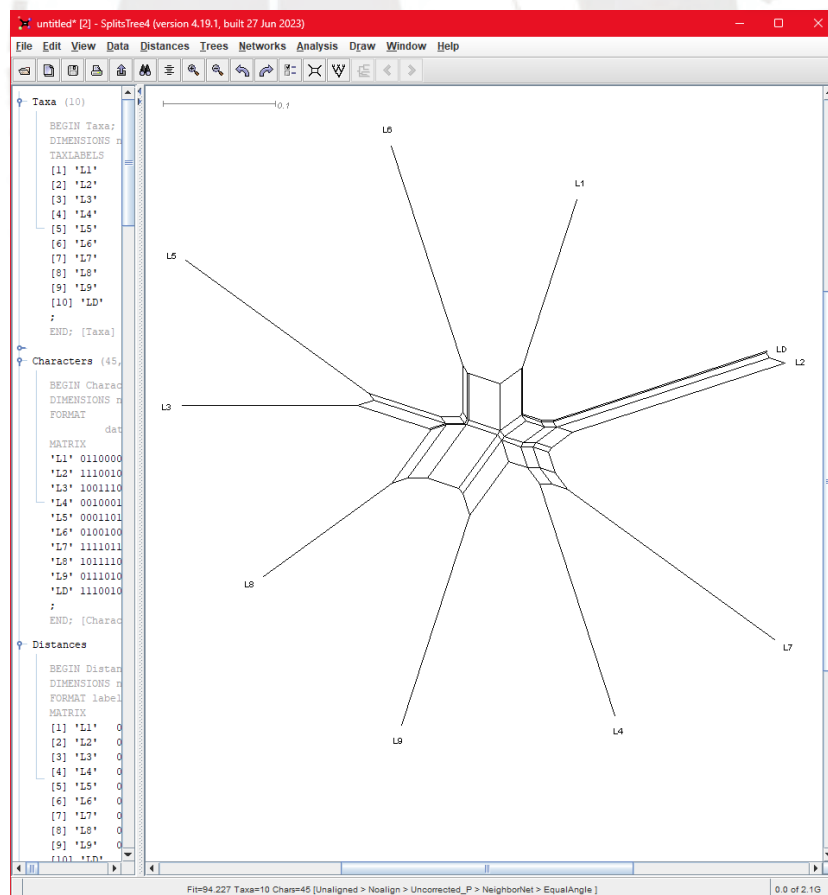
Debido a que la data de la matriz es una codificación binaria, el programa solo computó los números “0” y “1”. La matriz se obtiene de acuerdo a los procesos o fenómenos fonético-fonológicos, lingüísticos y paralingüísticos encontrados en cada hablante.

END; [Unaligned]

Finalizar la codificación con este parámetro. Luego presionar el botón *Execute*. El resultado de esta codificación es el que se muestra en la Figura 16. Se observa que la muestra L2 y LD son las menos distantes, por ende, son las más compatibles entre todas.

Figura 16

Captura de pantalla del menú File del programa SplitsTree4



1.8.4.4. Locutores indubitados

Lazo (2014) indica que la muestra indubitada se recoge a través de diligencias de toma de muestra de voz, en algunas ocasiones, el encargado de hacer la pericia de comparación de voz no es quien toma la muestra, por ello, el personal asignado a aquella tarea debe tener conocimientos sobre el habla y el discurso oral. Asimismo, propone tres aspectos a tener en cuenta en la diligencia de toma de muestra de voz:

a) El aspecto administrativo se relaciona con la Guía de Elaboración de pericias fonético-acústicas de homologación de voz (2020), en ella se señala que la muestra indubitada debe ser obtenida por una diligencia de toma de muestra de voz y utilizando un equipo profesional, asimismo, la muestra indubitada también puede ser remitida por el fiscal encargado del caso, siempre y cuando el locutor niegue a dar su voz. Algunos requisitos que debe tener la muestra indubitada remitida por el fiscal es que debe ser óptima, sin ruidos, y con un tiempo determinado de habla mayor a un minuto. Las muestras deben de tener un código de aseguramiento, en la Oficina de Peritajes se utiliza el algoritmo Hash SHA.

b) En el aspecto lingüístico se considera el código de ética del encargado de tomar la muestra, en este caso, al recolectar la muestra de la persona a peritar se debe considerar que el lugar y la situación sea lo suficientemente cómoda para obtener datos de habla espontánea y que estén al mismo nivel que los discursos de la muestra dubitada, así como las preguntas y el cuestionario sean claras. Es necesario que todo lo que suceda al inicio de la diligencia quede consignado en un acta elaborado por el personal fiscal, esto con el fin de respetar el derecho del locutor.

Ahora bien, de los siete (07) investigados del caso fiscal, solo asistieron tres (03) a las diligencias de tomas de muestra de voz, mientras que para un imputado se remitió mediante un disco sellado y lacrado dos (02) audios de audiencias que contienen sus participaciones orales en los procesos. De acuerdo a lo solicitado por la fiscalía y mediante la hipótesis formulada por el despacho fiscal, se tiene en la Tabla 2, las iniciales de los investigados con sus respectivos códigos, esto con el fin de mantener el anonimato de los imputados. Los números de muestras 3 y 4 corresponden a un solo investigado. Estas muestras se segmentaron y concatenaron según las voces que aparecen en el audio, y posteriormente se transcribieron (Ver Anexo 08: Transcripción de las muestras indubitadas).

Tabla 2*Iniciales de los investigados con sus respectivos códigos de referencia*

Número de muestra	Iniciales	Referencia
1	O.C.M.	LI_01
2	M.E.M.V.	LI_02
3	E.A.C.M.	LI_03
4	E.A.C.M.	LI_04

En la Tabla 3 se recopila la información sobre las características lingüísticas y socioculturales de los locutores implicados para su posterior comparación. Se optó por no indicar su lugar de residencia por motivos de seguridad.

Tabla 3*Características sociolingüísticas de los locutores indubitados*

Referencia	Edad	Género	Nivel de educación	Lengua materna	Lugar de origen
LI_01	27	masculino	Técnico superior	español	Puno
LI_02	55	masculino	Técnico superior	español	Madre de Dios
LI_03	35	masculino	Técnico superior	español	Tacna
LI_04					

1.8.4.5. Locutores dubitados**1.8.4.5.1. Metadatos de la muestra dubitada**

La muestra dubitada fue remitida por la Fiscalía en un disco DVD color blanco, marca Princo, con su respectiva cadena de custodia, tal como se muestra en la Figura 17. Se optó por la confidencialidad del rótulo del disco de la muestra dubitada.

Figura 17

Disco de la muestra dubitada



El disco contiene dos audios en formato wav que corresponderían a la muestra dubitada para peritar, en la Tabla 4 se muestran las propiedades de los archivos de audio.

Tabla 4

Propiedades de los archivos de audio de la muestra dubitada

Nombre	Tipo	Tamaño	Duración (min)
BVR_2019_07_26_21_54_09_01	WAV	773,824 KB	34:23
BVR_2019_07_26_22_30_26_01	WAV	266,141 KB	11:49

1.8.4.5.2. Asignación de rótulo y referencia a la muestra dubitada

En este apartado se detalla las voces que intervienen en las muestras dubitadas, según la hipótesis fiscal. Como muestras dubitadas se tienen dos (02) archivos de audio en formato wav, por lo que no fue necesario su adecuación para el proceso de identificación de voz, a estos dos archivos se les asigna, tal como se detalla en la Tabla 5, los siguientes rótulos.

Tabla 5

Rótulos de las voces intervinientes en las muestras dubitadas

Número de muestra	Rótulo de la voz según acta de transcripción	Referencia
1	BVR_2019_07_26_21_54_09_01-C_voz2	MD_01
2	BVR_2019_07_26_21_54_09_01-C_voz1	MD_02
3	BVR_2019_07_26_21_54_09_01-C_voz4	MD_03
4	BVR_2019_07_26_21_54_09_01-C_voz3	MD_04
5	BVR_2019_07_26_22_30_26_01-C_voz2	MD_05
6	BVR_2019_07_26_22_30_26_01-C_voz1	MD_06
7	BVR_2019_07_26_22_30_26_01-C_voz3	MD_07
8	BVR_2019_07_26_22_30_26_01-C_voz4	MD_08

Estas muestras se segmentaron y concatenaron según las voces que aparecen en el audio, y posteriormente se transcribieron (Ver Anexo 07: Transcripción de las muestras dubitadas).

1.9. Justificación

1.9.1. Justificación científica

El estudio busca impulsar el desarrollo científico de la lingüística forense en el Perú, una rama poco explorada a nivel nacional que está empezando a ganar importancia en el país. Como miembro del equipo de la Oficina de Peritajes, es fundamental buscar formas innovadoras de abordar científicamente la criminalidad.

Por ello, el estudio se justifica en el marco metodológico, utilizando una nueva herramienta para el campo de la lingüística forense, la relevancia recae en la cantidad de muestras que el software logra ordenar. En la presente investigación, a diferencia de Jimenez *et al.* (2022) que analizan y comparan una voz dubitada con diferentes voces indubitadas, existen más voces dubitadas.

1.9.2. Justificación social

El Perú enfrenta diversos desafíos en materia de criminalidad y corrupción, delitos que la sociedad peruana vive día a día. La aplicación de técnicas y metodologías de lingüística forense puede brindar herramientas adicionales para investigar y combatir

delitos como el fraude, el lavado de dinero, el crimen organizado y otros actos ilícitos que afectan la sociedad peruana. Además, el Perú es un país multilingüe y multicultural, con una gran diversidad de lenguas y dialectos. La lingüística forense puede desempeñar un papel importante en el análisis de comunicaciones en diferentes idiomas y en la identificación de características lingüísticas que ayuden a establecer la autoría o la autenticidad de documentos, grabaciones o testimonios en procesos judiciales.

Asimismo, el desarrollo de la lingüística forense en el Perú contribuiría a fortalecer el sistema de justicia y a garantizar la igualdad de acceso a la justicia para todos los ciudadanos. Esto implica la formación de profesionales especializados en lingüística forense, el establecimiento de protocolos y estándares de investigación, y la implementación de laboratorios especializados para el análisis de evidencias lingüísticas.

Por ello, realizar investigaciones en lingüística forense en el Perú es esencial para mejorar la capacidad de investigación y persecución de delitos, garantizar la justicia y promover la transparencia en el sistema judicial. Además, contribuiría al desarrollo de la ciencia y la tecnología en el país, posicionando al Perú como un referente en el campo de la lingüística forense a nivel nacional e internacional.

1.9.3. Limitaciones

Una de las limitaciones del uso del software es la cantidad de datos que tiene que interpretar para establecer distancias a partir de las similitudes y semejanzas de los rasgos fonético-fonológicos. Se entiende que una muestra dubitada puede ser una llamada telefónica o la grabación de una conversación, y muchas veces no dura el tiempo esperado. Sería propicio utilizar el software con más datos distintos a los fonético-fonológicos e incluir datos morfológicos, sintácticos, léxicos y semánticos, siempre y cuando se obtenga aquella información.

CAPÍTULO II

MARCO CONCEPTUAL

En este capítulo se presentan de forma sistemática los conceptos y herramientas de análisis de la fonética y la fonología que resultan relevantes para el desarrollo de investigaciones en lingüística forense en el contexto peruano, es decir con foco en el español hablado en nuestro país. Los ejemplos e imágenes presentadas a lo largo de esta exposición provienen de nuestra propia investigación a partir de las voces de diferentes hablantes anónimos que participaron en diligencias de toma de muestra de voz a nivel nacional.

De esta forma, en este capítulo, se detallan las definiciones de fonema y alófono; y la descripción de los sonidos que forman parte del sistema español en su variedad estándar (con referencias a variedades específicas de esta lengua), todo esto desde una perspectiva articulatoria y acústica. Por otra parte, se propone una lista de procesos fonéticos, lingüísticos y paralingüísticos que afectan a la producción del habla.

2.1. Fonemas y alófonos

El fonema es una unidad lingüística abstracta que manifiesta realizaciones concretas, llamadas alófonos, que por lo general suelen ser predecibles a partir del contexto. Para distinguir entre un fonema y un alófono, se establece un conjunto de reglas relacionadas a la permutación de elementos fónicos en relación a otros estratos del lenguaje (Trubetzkoy, 1973). Para la fonología generativa, la idea de fonema mantiene el rasgo distintivo, en esta operan reglas fonológicas que intervienen del nivel subyacente al nivel superficial, Chomsky y Halle (1968) propusieron un inventario de veintiún (21) rasgos en base a categorías articulatorias, esto permitió caracterizar a los fonemas del inglés. Un fonema no siempre se realiza de la misma forma, la producción de un sonido está supeditado a otros sonidos que están en contacto, a la forma de habla, la velocidad de locución, entre otros; por ejemplo, las personas de habla hispana no son conscientes de que producen la consonante /d/ de dos formas distintas, es probable que sea una oclusiva [d] o una aproximante [ð], esto dependiendo del contexto, son dos pronunciaciones diferentes pero no existe ningún contraste significativo entre ellas (Hualde, 2014).

Con el fin de diferenciar de manera gráfica los fonemas de los alófonos, los fonemas se escriben dentro de barras oblicuas /i/, y los alófonos entre corchetes [i]. Esto es útil para la transcripción fonética del habla. (Esquivel *et al.*, 2016).

2.2. Sonidos del español

Nos referimos al sistema de sonidos del español a aquellos fonemas que pertenecen al inventario fonológico del español, para definir los conceptos de los siguientes acápite se utiliza principalmente como base a Hualde (2014).

2.2.1. Vocales

Las vocales son sonidos producidos cuando las cuerdas vocales vibran sin encontrar ninguna obstrucción o constricción entre la laringe y las aberturas oral o nasal. Esta vibración se origina a partir de una fuente de energía generada por los pulmones, en forma de un flujo de aire (Miyara, 2004). Para el caso del español, se tiene cinco vocales y se pueden clasificar por su articulación, tal como se indica en la Tabla 6

Tabla 6

Clasificación de las vocales del castellano según la posición de la lengua

Posición vertical	Tipo de vocal	Posición horizontal (avance)		
		Anterior	Central	Posterior
Alta	Cerrada	i		u
Media	Media	e		o
Baja	Abierta		a	

Nota. Vocales del castellano, de Miyara, 2004.

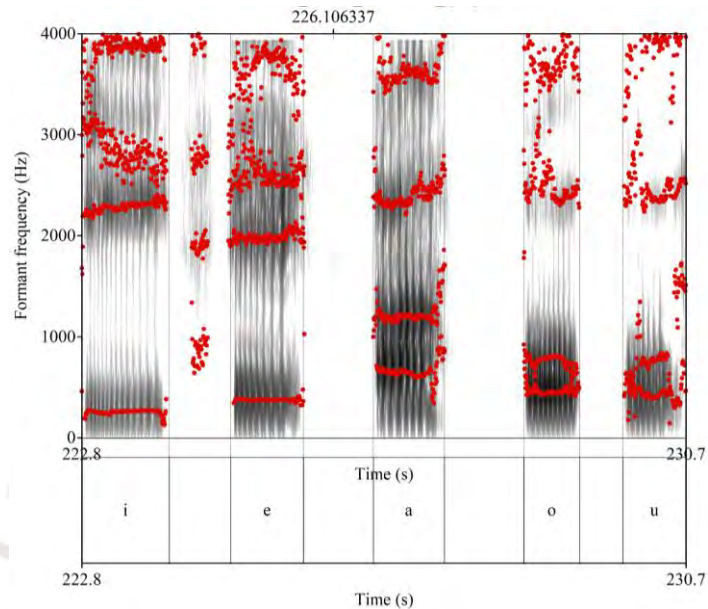
2.2.1.1. Formantes de vocales

La fonética acústica se enfoca en el estudio de las ondas sonoras generadas durante el proceso del habla. Estas ondas son producidas por los órganos articulatorios del aparato fonador, siendo las cuerdas vocales el componente principal. Estas ondas se denominan formantes y se clasifican en diferentes grupos, pudiendo llegar hasta cinco. Los formantes vocálicos más importantes son el F1 (formante 1) y el F2 (formante 2). Para representar de manera visual estos formantes, se utilizan espectrogramas, como se muestra en la Figura 18. En dicho espectrograma, que corresponde a las cinco vocales del español pronunciadas por un hablante masculino

monolingüe, los puntos de color rojo indican la ubicación de los formantes. Se pueden apreciar agrupaciones horizontales, donde los dos primeros (de abajo hacia arriba) representan los formantes F1, F2, F3 y F4.

Figura 18

Espectrograma de las cinco vocales del español de un hablante

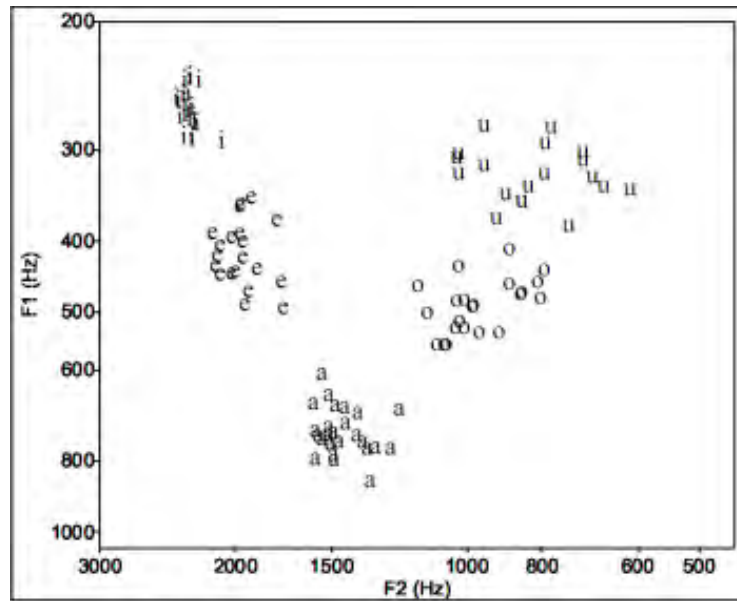


2.2.1.2. Carta de formantes

La carta de formantes es una herramienta que nos permite observar las realizaciones de sonidos que poseen formantes en los ejes F1 y F2, en este caso se ejemplificarían de una forma más precisa las vocales, estas no tienen una realización única, ninguna vocal se va a producir con la misma cantidad de formantes que otra de la misma naturaleza fonológica; tal y como señalan Pérez *et al.* (2008), las vocales en una carta de formantes se reflejan como islas o campos de dispersión (ver Figura 19).

Figura 19

Ejemplo de carta de formantes



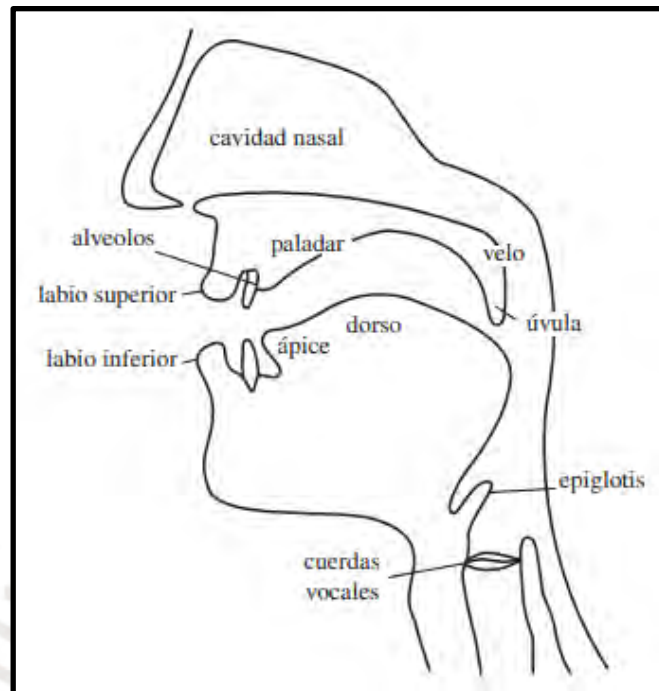
Nota. Carta de formantes del español limeño, de Pérez *et al.*, 2008.

2.2.2. Consonantes

La producción de sonidos consonánticos se distingue por la presencia de obstrucciones en el aparato fonador. En este proceso participan un articulador activo y un articulador pasivo (ver Figura 20). El articulador activo se mueve para establecer contacto u obstrucción con otro órgano, que en este caso sería el articulador pasivo. La clasificación de las consonantes se basa en tres parámetros: su modo de articulación, su punto de articulación y la participación de las cuerdas vocales, que determina el grado de intervención de las mismas.

Figura 20

Principales articuladores del aparato fonador



Nota. Articuladores del aparato fonador visto desde perfil, de Hualde, 2014.

2.2.2.1. Modo de articulación

El modo de articulación se refiere al tipo de obstrucción presente en la articulación de una consonante. Hay tres posibilidades distintas en términos de formación del obstáculo. En primer lugar, se encuentra la obstrucción total, que impide por completo la salida del aire, lo que resulta en una breve interrupción. En segundo lugar, está la obstrucción parcial, en la que el aire se escapa con una ligera obstrucción, lo que genera turbulencia característica. En tercer lugar, se encuentra la obstrucción aproximante, donde el articulador pasivo está en posición de aproximación, permitiendo que el aire salga por un canal más amplio y sin turbulencia. Estas tres formas de articulación se combinan para formar otros modos de articulación. Actualmente, se reconocen siete (07) modos de articulación, que se presentan a continuación (Hualde, 2014).

- 1) Oclusivas. Son consonantes que tienen obstrucción total entre los articuladores, lo que interrumpe la salida del aire. Estas consonantes se distinguen por ser plosivas, lo que significa que al separarse los articuladores que estaban en posición fija, se produce una pequeña explosión que permite

- la liberación de aire concentrado en el aparato fonador. Los fonemas oclusivos que pertenecen a este grupo son [p, t, k] y su correspondiente sonora [b, d, g].
- 2) Fricativas. Estos sonidos se caracterizan por la salida turbulenta del aire permitida por la ligera constricción entre los articuladores. Los fonemas que se encuentran en esta categoría son /f, s, x/, en ciertos contextos, la /s/ se manifiesta con sus alófonos sonoro [z] y aspirado [h].
 - 3) Aproximantes. La constricción característica de este tipo de segmentos es que no llega a haber contacto, de tal modo que no es una fricativa, sino que la apertura es un poco más amplia. Esta consonante se manifiesta en español en contextos intervocálicos, por ejemplo, en la palabra “coda” [ko.da], se produce generalmente el fonema /d/ como una aproximante [ð]. Los fonemas /b, d, g/ son los que tienen aproximantes como alófonos [β, ð, γ].
 - 4) Africadas. Estas consonantes se describen como sonidos con dos fases, la primera es la oclusión total y la segunda es la salida del aire como una consonante fricativa. En el español, solo existe el fonema /tʃ/, como en la palabra “chupete” /tʃu.pe.te/.
 - 5) Nasaes. Los sonidos nasales se caracterizan por la emisión de aire a través de la cavidad nasal y la obstrucción en la cavidad oral; el aire se libera al descender el velo. El español posee tres fonemas nasales: /m/, /n/ y /ɲ/, un ejemplo de cada uno de ellos sería en las palabras “mamá” /ma.ma/, “cana” /ka.na/ y “piña” /pi.ɲa/, respectivamente.
 - 6) Laterales. Estos sonidos se distinguen por el contacto entre los articuladores en la región central de la cavidad oral, lo que permite que el aire salga por uno o ambos lados laterales. En español, se encuentra el fonema lateral /l/, como en la palabra “pala” /pa.la/.
 - 7) Vibrantes. Se pueden identificar dos fonemas vibrantes: la vibrante múltiple y la vibrante simple, representadas como /pe.ro/ y /pe.ro/, respectivamente. Estos sonidos se generan mediante un contacto rápido entre los articuladores. La vibrante múltiple implica dos o más contactos, mientras que la vibrante simple implica solo un contacto.

2.2.2.2. Punto de articulación

Este es un parámetro que se utiliza para categorizar a las consonantes por los articuladores que ejecutan la constricción. Para ello, se describen nueve puntos de articulación relacionados al español: (Hualde, 2014)

1. Bilabial. Una consonante bilabial se forma mediante el contacto entre los dos labios, como se observa en la palabra "pepa" [pe.pa]. En esta palabra, se presenta la consonante oclusiva [p], su contraparte sonora [b] y la aproximante [β], que también pertenecen a este grupo de sonidos. La consonante nasal [m], presente en la palabra "mamá" [ma.ma], comparte el mismo punto de articulación.
2. Labiodental. Estas consonantes se generan mediante el contacto del labio inferior con los dientes superiores. Un ejemplo de este grupo es el sonido [f], presente en palabras como "foca" [fo.ka] o "café" [ka.fe].
3. Interdental. Este tipo de sonidos no existe como fonema en el español de latinoamérica, la grafía que lo representa se pronuncia en esa variedad como un sonido fricativo alveolar [s], sin embargo, en el español peninsular si tiene valor fonémico, la consonante interdental aparece en palabras como "César" [θe.sar] o "cemento" [θe.men.to].
4. Alveolar. El articulador pasivo de esta zona es la región alveolar, estos sonidos los conforman [t], [d], [s], [n], [l], [r] y [r̄]. son el grupo más numeroso respecto al punto de articulación.
5. Prepalatal. O también llamados sonidos post-alveolares, en español existe el fonema africado [tʃ] como en la palabra "choclo" [tʃo.klo], otros sonidos como la fricativa [ʃ] y su respectiva sonora [ʒ] no forman parte del español.
6. Palatal. Dentro del español, se encuentran los sonidos [j] y [ɲ], los cuales se articulan con el dorso de la lengua en contacto con el paladar. En algunos dialectos del español, también se incluye el sonido [ʎ] en este grupo.
7. Velar. Las oclusivas que conforman este grupo son las consonantes [k] y [g], y se producen con la parte posterior del dorso de la lengua en contacto con la última sección del paladar. La forma aproximante de la oclusiva velar sonora es [ɣ] y se produce entre dos vocales, regularmente. También forma parte de este punto de articulación el sonido fricativo velar [x] como en la palabra "caja" [ka.xa].

2.2.2.3. Función de las cuerdas vocales

Las cuerdas vocales cumplen una función muy importante en la clasificación de los sonidos, estas tienen la tarea de generar sonoridad mediante el junte o la cerrazón de los pliegues vocales. Estas se encuentran en la glotis, lugar donde se ejecuta la vibración por la salida del aire. Hualde (2014, 38) señala que «si nos colocamos los dedos índice y pulgar sobre la garganta, en el lugar donde se sitúa la llamada nuez de Adán, y alternamos entre la producción de [ssss] y de [zzzz] (como en el inglés zoo /zu/ 'zoológico'), notaremos la vibración en la pronunciación de [z], pero no en la de [s]».

2.2.2.4. Clasificación de los sonidos fonéticos consonánticos del español

A partir de lo descrito anteriormente, es posible clasificar y distinguir los sonidos del español con los tres parámetros desarrollados, para el presente estudio se muestra la Tabla 7 de los fonemas consonánticos que describe Hualde (2014). Algunos de los sonidos que están entre paréntesis no son considerados fonemas distintivos en todas las variedades del español. Por ejemplo, el sonido interdental /θ/ se encuentra únicamente en el español peninsular, la lateral palatal /ʎ/ está presente en algunas variedades peninsulares y sudamericanas, la fricativa postalveolar o prepalatal /ʒ/ solo se encuentra en el español de Argentina, y el sonido /j/ tiene un estatus fonémico que es objeto de debate.

Tabla 7

Fonemas consonánticos del español

	Bilabial	Labiodental	Interdental	Dental	Alveolar	Prepalatal	Palatal	Velar	Glotal
Oclusivas	p b			t d				k g	
Fricativas		f	(θ)		s		(j)	x	
Africadas						ʃ			
Nasales		m			n		ɲ		
Laterales					l		(ʎ)		
Vibr. simple					r				
Vibr. múltiple					̄r				

Nota. Consonantes fonológicas del español, de Hualde, 2014, p. 39.

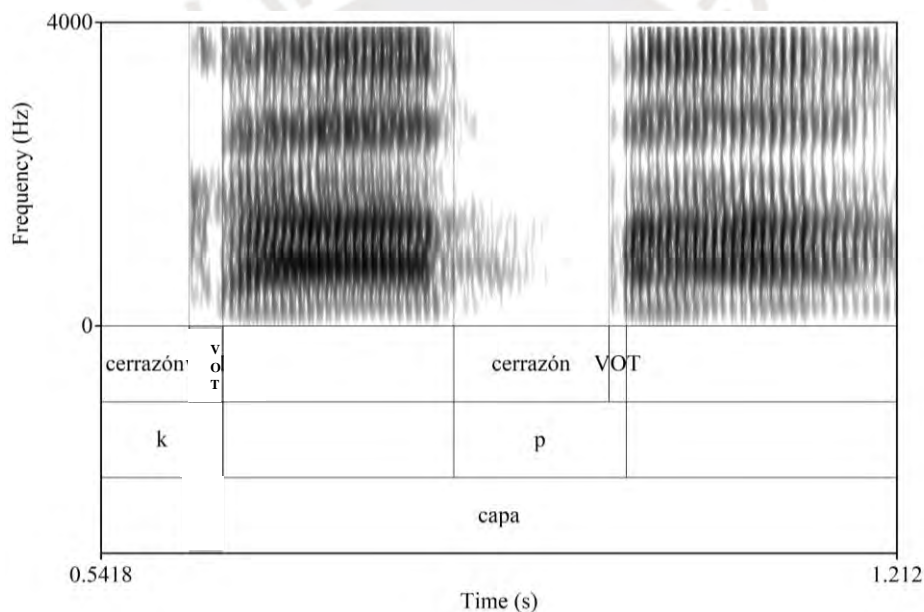
2.2.2.5. Descripción acústica de las consonantes

Oclusivas

Las oclusivas se articulan con una constricción total entre los órganos articulatorios (activo y pasivo), lo que causa la interrupción de la salida del aire y luego este es liberado de forma abrupta. Las oclusivas presentan tres fases, la implosión, la cerrazón y la explosión. La cerrazón y la explosión son las dos fases que tienen su correlato acústico, estos se manifiestan por el silencio y la barra de plosión, respectivamente, en el espectrograma. (Jiménez, 2018) En la Figura 21 se muestra un ejemplo de sonido oclusivo con las dos fases acústicas delimitadas, se tiene la palabra “capa”, las oclusivas [k] y [p].

Figura 21

Características acústicas de un sonido oclusivo

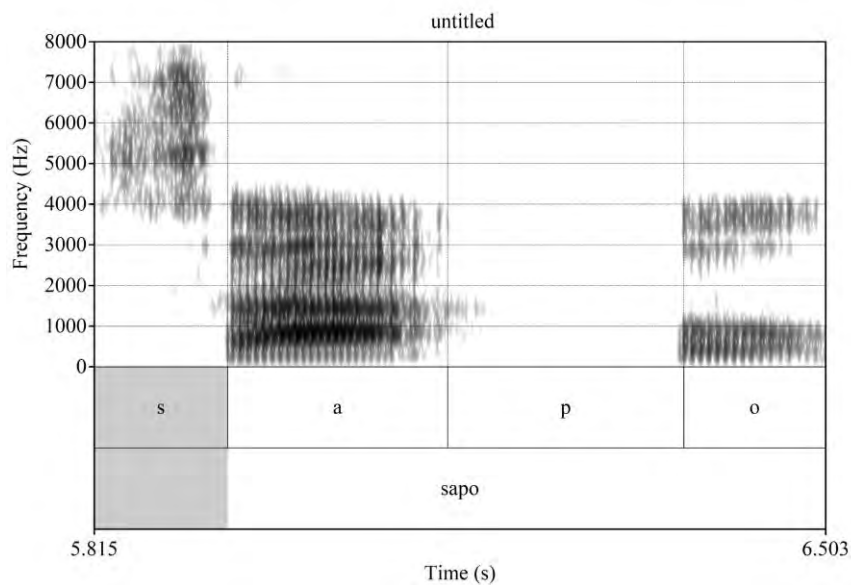


Fricativas

Los sonidos fricativos se caracterizan acústicamente por la turbulencia a lo largo de la duración del sonido, la energía que se distribuye es dispersa y, a diferencia de las vocales, posee ciclos aperiódicos. En la Figura 22, se tiene un ejemplo de la consonante fricativa alveolar [s], al ser una fricativa sibilante, su energía comienza a distribuirse a partir de los 4000 Hz a 8000 Hz. A pesar de ser un sonido que mantiene su energía dispersa, su posición en el espectrograma logra regularidad de acuerdo a su naturaleza articulatoria.

Figura 22

Características acústicas de un sonido fricativo

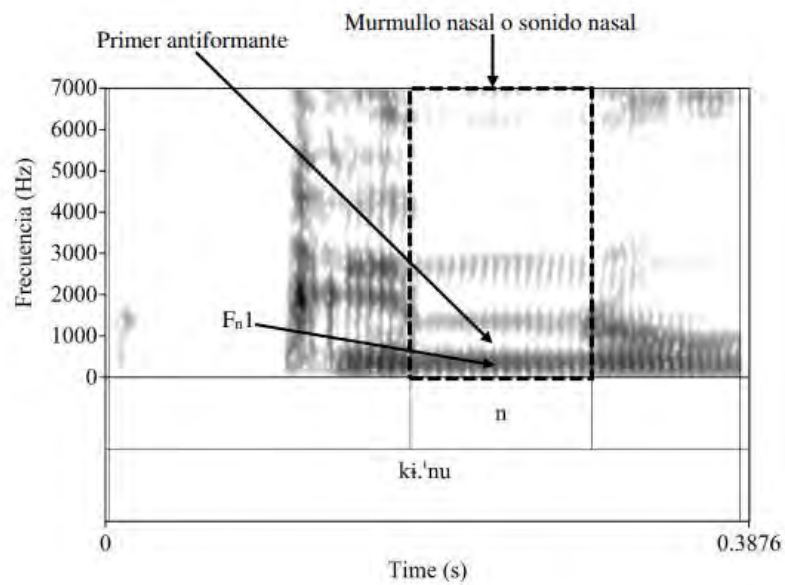


Nasales

Según Elías-Ulloa (2011), una señal para reconocer los sonidos nasales es la resonancia nasal, producida por el cierre prolongado de la cavidad oral, permitiendo que el flujo de aire pase a través de la cavidad nasal. Las consonantes nasales, al igual que las vocales nasales, presentan formantes, pero en menor intensidad que ellas. Este formante F1 o también llamado formante nasal, se manifiesta por debajo de los 1000 Hz. Asimismo, 1000 Hz sería también el límite para la aparición de antiformantes, entendidos como una zona de frecuencia donde la energía disminuye drásticamente. En la Figura 23 se puede evidenciar los tres parámetros para identificar a los sonidos nasales: murmullo nasal, formantes y antiformantes.

Figura 23

Características acústicas de un sonido nasal



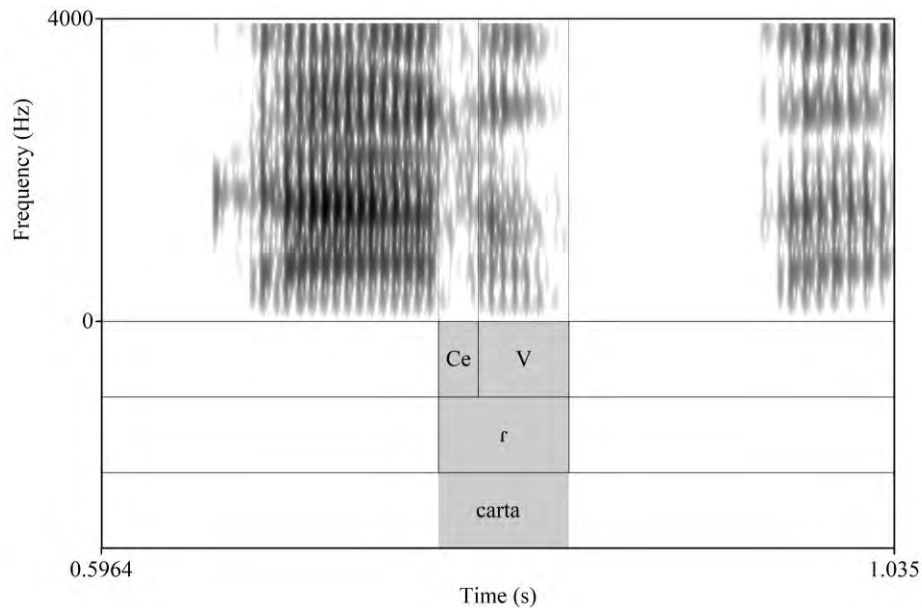
Nota. Consonante nasal en arabela, de Jiménez, 2018, p. 77.

Vibrantes

Las vibrantes se caracterizan por presentar dos fases, la primera es la fase cerrada y la segunda se describe como un elemento vocálico, este tipo de sonido puede presentar fricción al momento de la oclusión. En la Figura 24, se observa a la vibrante simple en la palabra “carta”, con su fase de cerrazón y posteriormente el momento vocálico.

Figura 24

Características acústicas de un sonido vibrante

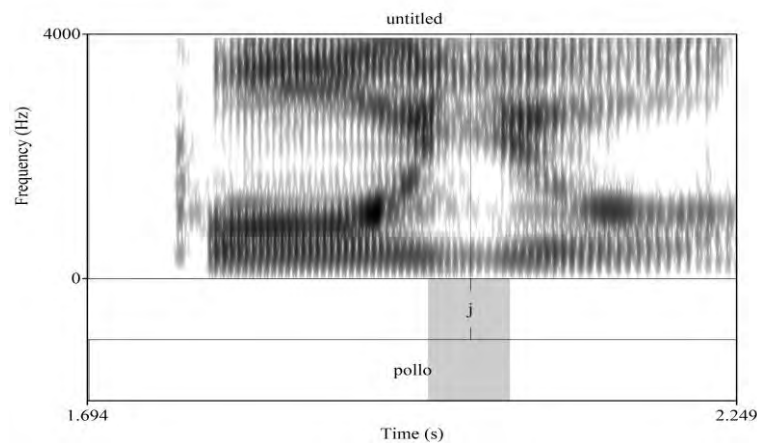


Glides

Las glides son un tipo de sonidos que se producen cuando los articuladores activo y pasivo se acercan sin generar turbulencia, a diferencia de los sonidos fricativos. Estos elementos también se conocen como transitorios, ya que se producen durante la transición entre una articulación y otra distinta. La posición de las glides es difícil de determinar, ya que no tienen un inicio fijo y un final estático, presentan formantes al igual que las vocales, pero no pueden ser núcleos de sílaba. En la Figura 25, se muestra un ejemplo de glide, en este caso la yod en la palabra “pollo” [po.jo].

Figura 25

Características acústicas de un sonido glide



2.3. Estudio de la percepción auditiva

El experto observa la muestra dubitada y limita lo que puede comparar con la muestra indubitada, su conocimiento científico permitirá «determinar los sonidos que se pueden analizar, atendiendo a distintos factores, entre otros, las condiciones grabación, las características acústicas de los sonidos o la relevancia de determinados parámetros acústicos para la identificación de locutor» (Machuca *et al.*, 2014, 97).

Además, Machuca *et al.* (2014, 100) señala que «el análisis auditivo, complementado con un análisis acústico detallado, permite establecer los fenómenos que se han de contemplar en el informe pericial sobre las grabaciones dubitadas. Estos fenómenos serían los que se deberían obtener en las grabaciones indubitada». El análisis auditivo tradicional debe ser complementado con un análisis acústico porque por si solo resulta subjetivo, con el enfoque acústico-auditivo se puede medir y demostrar las propiedades acústicas del discurso oral (Lazo, 2023).

El análisis de la percepción auditiva y acústica se compone de dos fases: a) en la primera fase, se escuchan las grabaciones de voz objetivo con el propósito de caracterizar las diferentes realizaciones fonético-articulatorias y lingüísticas emitidas por el hablante; b) en la segunda fase, se determinan y evalúan estas realizaciones con el fin de establecer su relación con aspectos dialectales, sociolectales e idiolectales, con el objetivo de determinar la compatibilidad entre las voces registradas. Para ello, definiremos en qué consiste un análisis lingüístico y cómo se

enfoca a nivel descriptivo, se describirán diversos procesos fonético- acústicos que se identifican en el habla, esto no quiere decir que sean todos, cabe la posibilidad de que existan más.

2.3.1. Elementos para el análisis fonético-fonológico y lingüístico

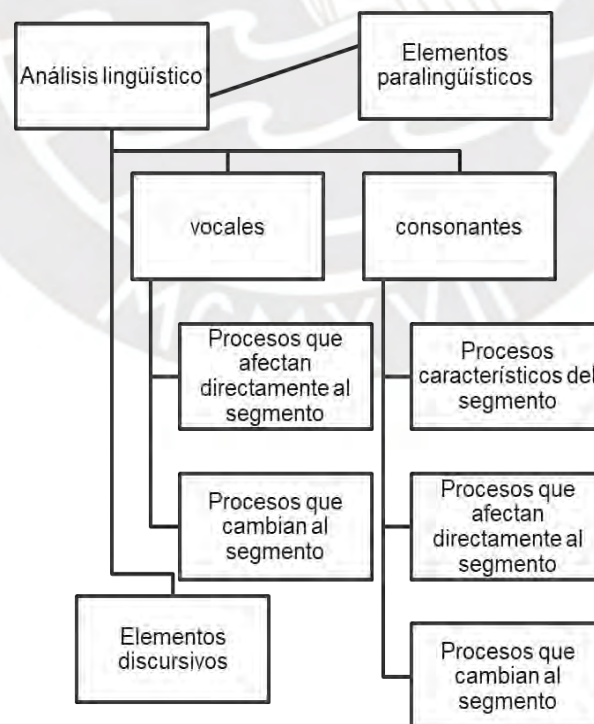
Ahora bien, dentro de un análisis lingüístico existen componentes de análisis, los cuales se encuentran en diferentes niveles, está el fonético-acústico o articulatorio, el morfológico, sintáctico, léxico - semántico, discursivo y paralingüístico. De los cuales solo nos enfocaremos en el fonético-acústico. Para ello, es necesario una estructuración del análisis lingüístico enfocado a los datos fonético-acústicos.

La extracción de datos fonético-acústicos para la comparación forense de voz ha sido la base principal de sustento pericial para la sección del método manual o perceptual, esto debido a que en la producción del habla se encuentran diversos fenómenos recurrentes y/o particulares del habla (ver Figura 26).

Asimismo, conforme se han realizado investigaciones, se han añadido otros dos parámetros complementarios (elementos discursivos y paralingüísticos).

Figura 26

Reestructuración del análisis lingüístico



Dentro del análisis lingüístico intervienen cuatro (04) elementos principales, los sonidos de la lengua, en este caso el español, vocales y consonantes, en los cuales

existen procesos que afectan al segmento, eso quiere decir que el segmento permanece con los mismos rasgos (p. ej. la laringalización interviene como un chirrido que se puede observar en el espectro, pero no afecta a los rasgos principales del sonido, una vocal laringalizada sigue siendo la misma vocal) y procesos que cambian al segmento (p. ej. una vocal /i/ se convierte en [e, eso quiere decir que sus rasgos principales son sustituidos por otros). Los procesos que son característicos del segmento se refieren a la producción de un sonido en particular y que no cumpla con los estándares frecuentes (p. ej. un sonido fricativo alveolar que tenga menor cantidad de Hertz que lo regular). En el análisis lingüístico también intervienen elementos discursivos y paralingüísticos, el primero de ellos se consideran muletillas, interjecciones, entre otros, en el segundo intervienen elementos como las risas, los llantos, pausas llenas, entre otros. Estos son los elementos más frecuentes dentro del análisis lingüístico, sin embargo, en ocasiones suelen presentarse procesos compatibles relaciones a la morfosintaxis.

2.3.2. Tabulación del análisis lingüístico

Este paso es muy importante al momento de presentar el análisis lingüístico dentro del informe pericial, al identificar los procesos que se encuentran en el habla espontánea de un locutor, este puede constar de elementos fonético-fonológicos, lingüísticos y paralingüísticos; estos procesos pueden ser tabulados de acuerdo a la cantidad de muestras que se necesite comparar junto con la voz del locutor a investigar, en otras palabras, la tabulación de rasgos lingüísticos nos permite corroborar la existencia de acuerdo a la frecuencia de rasgos dentro del habla de un locutor. Para esta investigación, se tiene en cuenta lo planteado por Celis (2009), ya que considera y organiza a los rasgos fonético-fonológicos, lingüísticos y paralingüísticos en variaciones relacionadas y no relacionadas al campo expresivo, ver Tabla 8.

Tabla 8

Variaciones relacionadas y no relacionadas con el plano expresivo del habla

Variaciones no relacionadas con el plano expresivo	Variaciones relacionadas con el plano expresivo
<i>Contemporaneidad</i>	<i>Modificaciones de rangos fonatorios y articulaciones estándar</i>

<p>Este factor describe el cambio en la voz a medida que una persona envejece. Se mencionan estudios que han investigado el reconocimiento auditivo a largo plazo y el descenso en la frecuencia fundamental de la voz a lo largo del tiempo. Se señala que la madurez vocal se alcanza alrededor de los 15 años y que en los hombres hay un descenso gradual hasta los 40 años, seguido de un aumento entre los 60 y 80 años. En las mujeres, el decrecimiento ocurre principalmente antes de los 15 años, con un 43% después de esa edad. Se mencionan factores como la menstruación y la menopausia que pueden afectar la frecuencia fundamental en las mujeres. Se concluye que estos cambios no tienen un impacto significativo en la identificación vocal, pero se reconoce un cambio gradual en los componentes del habla a lo largo del tiempo.</p>	<p>Asociadas a cambios notables en aspectos esenciales como el tono o la intensidad, niveles de tensión y relajación en la articulación, grado de nasalización y oralidad, sonoridad y atenuación, apertura y cierre, velarización y palatalización, fricativización, suavización, entre otros.</p>
<p><i>Cambios en los procesos y órganos de la fonación</i></p> <p>El autor describe tres grupos, estos son: Anatómico (Modificaciones en la estructura dental, alteraciones en la pronunciación debido a prótesis bucales, presencia de tumores y otras condiciones relacionadas.), fisiológico (Resfriados, inflamaciones, periodo menstrual, menopausia, irritaciones y</p>	<p><i>Alteraciones elocutivas de elementos fonéticos simples</i></p> <p>Donde se incluirían las diversas variantes fonéticas de cada sonido o grupo de sonidos, así como sus efectos relacionados de inicio, finalización, transición y coarticulación.</p>

<p>otros eventos similares.) y psicológico y neurológico (Movimientos involuntarios temporales, trastornos de la voz de origen psicogénico, alteraciones emocionales, efecto Lombard y otras condiciones relacionadas.)</p>	
<p><i>Cambios producidos por químicos exógenos</i></p> <p>Estos cambios pueden ser ocasionados por el consumo de medicamentos, tabaco, alcohol, drogas y otras sustancias. Estas sustancias pueden afectar de diferentes maneras, como se mencionó previamente. Por ejemplo, en el caso de las mujeres, el uso de anticonceptivos aumenta los niveles de progesterona en el cuerpo, lo cual resulta en un aumento en el tamaño de los pliegues vocales. Esto, a su vez, produce un descenso en el tono de la voz, llevándola hacia tonalidades más graves.</p>	<p><i>Alteraciones elocutivas relativas al tiempo y carácter suprasegmental o melódico expresivo</i></p> <p>Aquí mencionaremos todas las variaciones relacionadas con la entonación (ascendente, descendente, suspensiva), la acentuación, la velocidad (relacionada con las sílabas o la articulación de las emisiones), la fluidez, el ritmo (normal, lento, rápido), las pausas (articulatorias, respiratorias, de duda, para seleccionar unidades léxicas, para estructurar la sintaxis, para examinar o preparar la información comunicativa, para ceder el turno en una conversación, etc.).</p> <p><i>Variaciones de construcción lingüística y de códigos de relación comunicativa</i></p> <p>Incluye estructuras morfosintácticas, estilo, recursos retóricos, elementos paralingüísticos o extralingüísticos (como chasquidos con la lengua, etc.), vocabulario, idiomas, dialectos, variaciones verticales de dialectos o jergas, proxémica (organización cultural de microespacios de comunicación), y variaciones diafásicas (alternancia de</p>

	registros de expresión, como el paso del coloquial al técnico, por ejemplo).
	<p><i>Alteraciones transitorias de la cualidad de la voz</i></p> <p>Aparte de las mencionadas que involucran limitaciones en la fonación o en la articulación según los estándares, se incluyen las variaciones no intencionales como voz quebrada, áspera, ronca, cavernosa, fañosa, entre otras, así como las variaciones voluntarias como voz susurrada, imitada, fingida o disimulada, entre otras.</p>
	<p><i>Variaciones de los componentes de construcción emocional o comunicativa del discurso</i></p> <p>Relacionados con los niveles de emoción, estabilidad, sorpresa, tristeza, miedo, amenaza, ansiedad, ira, felicidad, persuasión, y otros similares.</p>

Nota. Variaciones relacionadas y no relacionadas con el plano expresivo, de Celis, 2009, p. 30-31.

Por ejemplo, si se tuviera un “Locutor n”, y este constara de quince (15) rasgos lingüísticos verificados en su muestra indubitada, se estaría creando un “perfil de habla del Locutor n”, y estos van a ser comparados con una, dos o más muestras dubitadas, según lo requiera la solicitud de la Fiscalía, en general. Entonces, para cada muestra, sea indubitada o dubitada, se crean perfiles de rasgos lingüísticos (esto está sujeto a diferentes factores como duración de la muestra, tipo de recojo de la muestra, entre otros).

En una tabla se colocan los rótulos que representan a las muestras por columnas (p. ej. MI_00 o MD_00), y en las filas los rasgos más frecuentes identificados, según lo considere el especialista. Esta tabla permite observar el *match* entre rasgos, de tal

forma que se concluye que las muestras son o no son compatibles, por ejemplo, en la Tabla 9, se muestra una representación hipotética de cómo se vería una tabla con muestras que hagan *match* en la mayoría de los rasgos de LI_00 y MC_00, mientras que es menos compatible con la muestra MC_01. Para el presente estudio se utiliza la siguiente codificación de los procesos “palabra / [representación fonética] / tiempo en segundos”, de acuerdo al rasgo identificado.

Tabla 9

Ejemplo de tabulación de muestras

RASGOS LINGÜÍSTICOS	LI_00	MC_00	MC_01...
R1	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00
R2	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	
R3	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00
R4	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	
R5	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	
RN	Palabra [Representación fonética] 00.00 – 00.00	Palabra [Representación fonética] 00.00 – 00.00	

Ahora bien, siguiendo el método manual, esta tabulación permite observar eficazmente la compatibilidad en las muestras comparadas con un solo investigado, para la presente investigación, la tabulación es el primer paso para proponer al SplitsTree4 como solución a la cantidad de tablas que se pueden crear por tener más de un investigado en un caso criminal.

2.4. Procesos en vocales y consonantes

En este apartado se describen diversos procesos que afectan, cambian y que son característicos en vocales y consonantes, cada proceso presenta como mínimo un espectrograma como ejemplo, cabe recalcar que estos ejemplos pertenecen a audios de la OPERIT, eso quiere decir que las muestras pertenecen a hablantes de todo el territorio peruano, de los cuales, los nombres permanecerán en anonimato. Esta categorización de procesos tiene la finalidad de servir como un banco de procesos que pueden ocurrir en el habla natural, por lo que su interpretación requiere de un especialista en el área. Asimismo, todos los gráficos exportados de Praat se han realizado con el *plugin* TgDraw (versión 0.3), elaborado por Muñoz (2020).

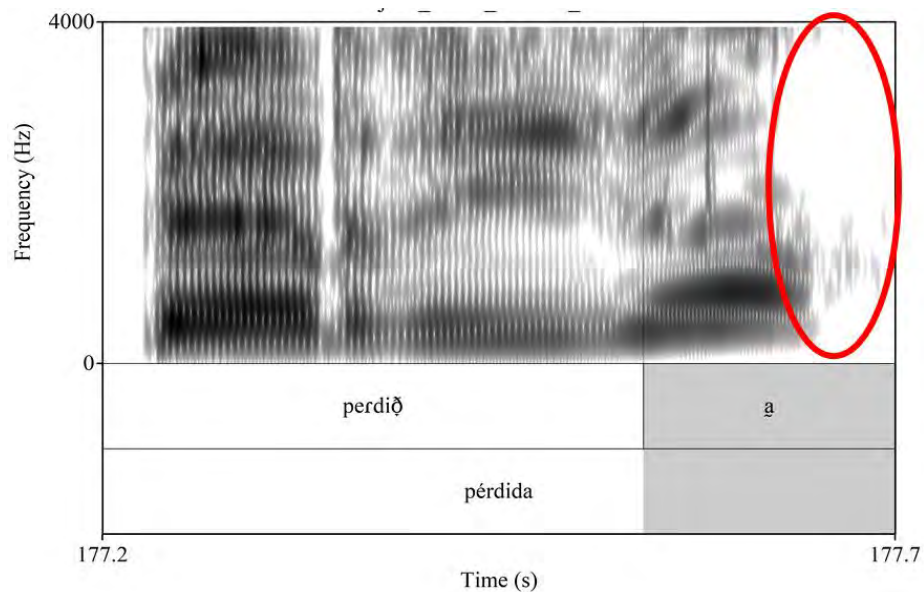
2.4.1. Procesos que afectan directamente en vocales

2.4.1.1. Disfonía vocálica

Se describe a la disfonía vocálica como una patología de la voz, y se manifiesta con alteraciones al momento de producir o ejercer la voz, el tono o el volumen en una comunicación, esto se produce porque hay múltiples afecciones en la laringe o por el abuso al esforzar la voz (Saavedra-Mendoza y Akaki-Caballero, 2014). En la Figura 27 se observa un ejemplo de disfonía vocálica al final de la palabra, lo que nos indica que en la vocal “a” no se produce de forma regular al final del segmento, el espectrograma de la vocal se va difuminando conforme va terminando su producción.

Figura 27

Ejemplo de disfonía vocálica

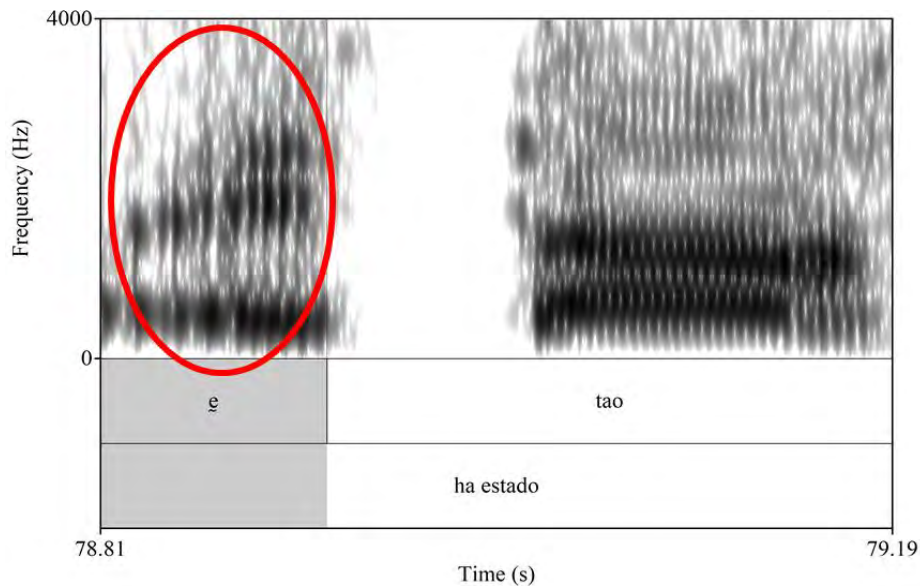


2.4.1.2. Laringalización vocálica

La laringalización o también llamada *creaky voice* se manifiesta a partir de pulsos glotales irregulares y más largos, esto puede afectar a todo el segmento vocálico o a una parte de él (Elías-Ulloa, 2019; Ladefoged, 2004) y no necesariamente aparece al final de la palabra o sílaba, en el espectrograma (ver Figura 28) se estaría evidenciando como estrías más amplias e incluso menos nítidas.

Figura 28

Ejemplo de laringalización vocálica

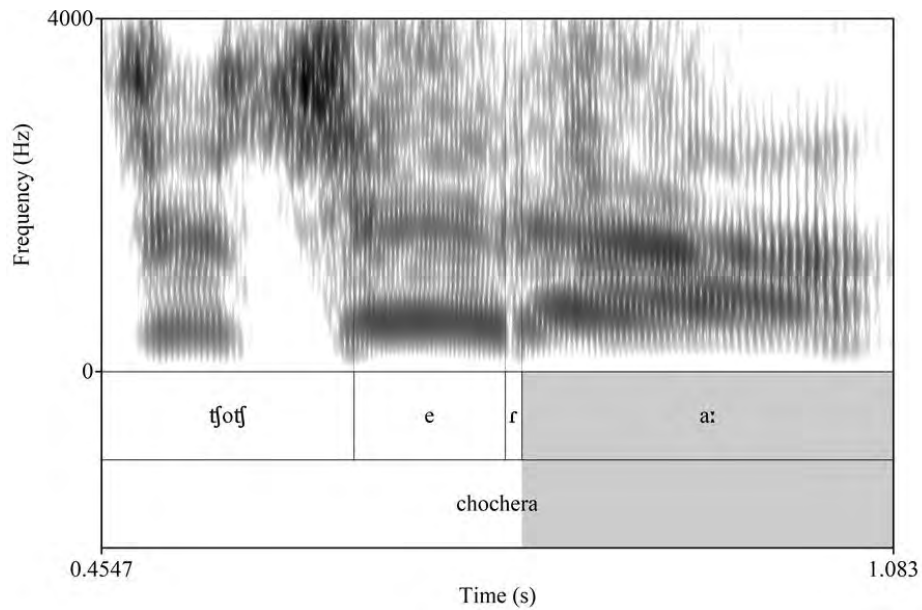


2.4.1.3. Alargamiento vocálico

Monroy-Casas (1980) lleva a cabo una investigación en la que entrevista a colaboradores de Zaragoza, Murcia, Madrid, León y Huelva. Los resultados de este estudio indican que la vocal /a/ es la que tiene una duración más prolongada, seguida de la vocal /e/, la vocal /i/, la vocal /o/ y, por último, la vocal /u/. Por otro lado, Stensrud (2019) realiza un estudio sobre la duración de las vocales en el español de Costa Rica. Este estudio señala que la duración de las vocales está relacionada con su apertura vocálica, lo que significa que las vocales abiertas tienden a ser más largas que las vocales cerradas. El alargamiento es un proceso en el cual un segmento tiene una duración mayor en comparación con otros segmentos. En este caso, la duración de una vocal alargada es mayor que la de una vocal breve. En conclusión, es más común que se observe una vocal alargada en segmentos con una mayor apertura vocálica, ver Figura 29.

Figura 29

Ejemplo de alargamiento vocálico

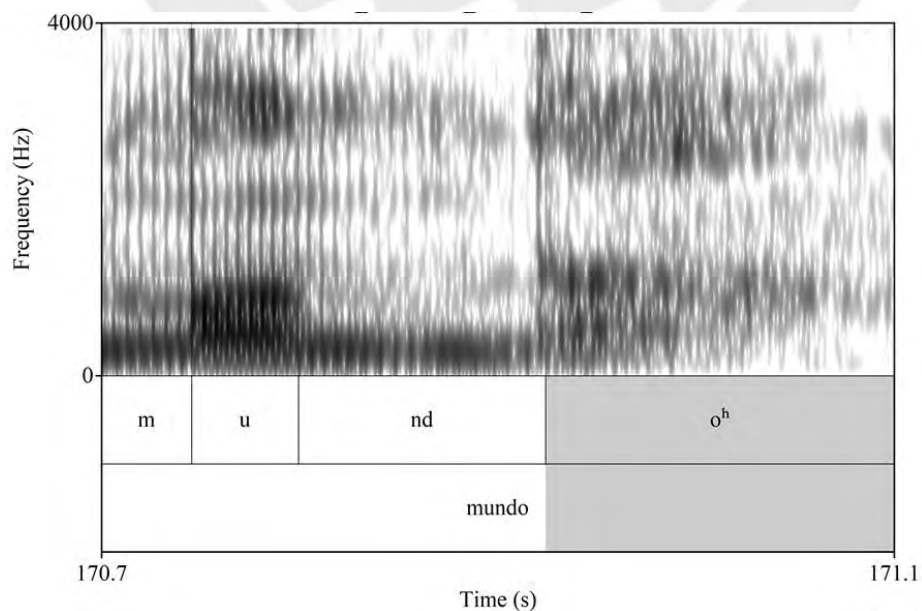


2.4.1.4. Aspiración vocálica

Se puede describir al proceso de aspiración vocálica como aquel segmento vocálico que posee dentro de su producción un grado de aspiración, la cual se representa como V^h , esta se genera por el cierre glótico completo (Darby, 1981). Por ejemplo, se tiene la Figura 30.

Figura 30

Ejemplo de aspiración vocálica

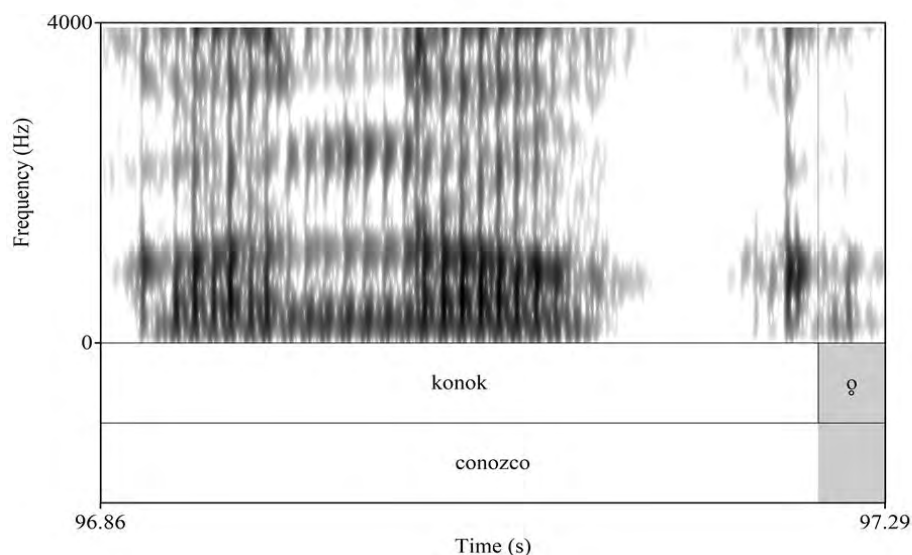


2.4.1.5. Ensordecimiento vocálico

Este proceso es común cuando se produce una vocal a final de palabra y cuando está entre dos consonantes sordas, o en contacto con alguna consonante sorda. La energía de la vocal en el espectro es difusa y menor a comparación de una vocal “sonora”, por ejemplo, se tiene la Figura 31, en la que la vocal “o” está ensordecida, además que se encuentra posterior a una consonante velar sorda oclusiva y a final de palabra.

Figura 31

Ejemplo de ensordecimiento vocálico



2.4.2. Procesos que cambian la vocal

2.4.2.1. Cambio vocálico o sustitución vocálica

El cambio o sustitución vocálica se define como cualquier cambio en el nivel fonético o fonológico del habla, este cambio es puramente fonético, cuando el sistema fonológico no interviene, o sea, algunos fonemas reciben una articulación distinta a la usual antes de la sustitución. Por ejemplo, se tiene en la Figura 32 el cambio vocálico de /e/ por [i], sin embargo, el significado no se altera, esto se debe a cuestiones dialectales, tal como se describe al inicio del estudio. Una herramienta para representar la posición vocálica a partir de los formantes F1 y F2 es la carta de formantes, en la cual se pueden comparar la posición de una vocal con otra para tener la certeza de que ha habido algún cambio respecto a la altura o posterioridad de la pronunciación de una vocal, es en ese sentido que se tiene la palabra “viceversa” en

la cual se puede identificar a las vocales [i] en la segunda sílaba “-ce-” y [e] en “-ver-”, lo cual indicaría, según la carta de formantes, que ha habido un cambio vocálico por la posición de la altura del que debería ser un segmento vocálico [e] (ver Figura 33).

Figura 32

Ejemplo de cambio o sustitución vocálica

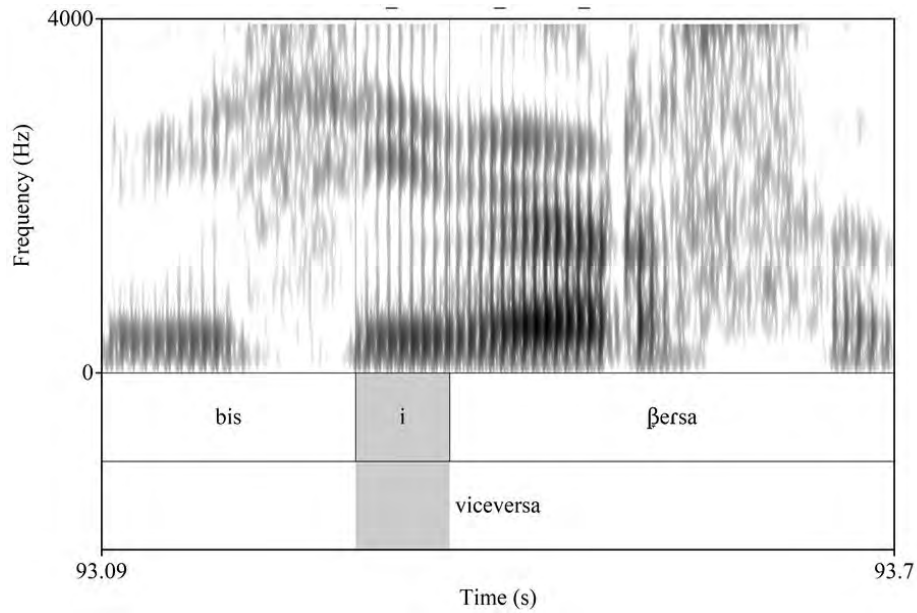
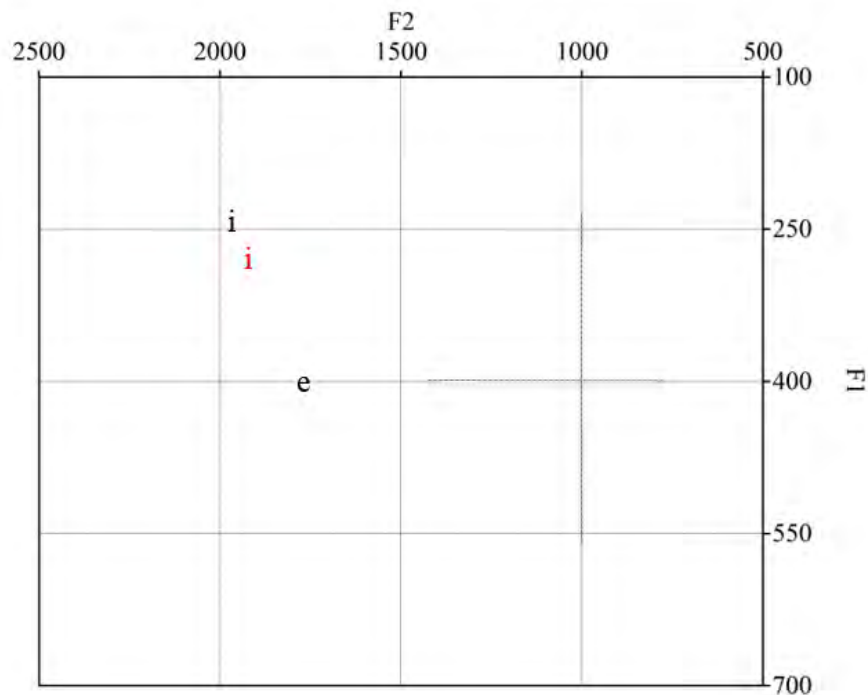


Figura 33

Carta de formantes de las vocales de la palabra “viceversa”

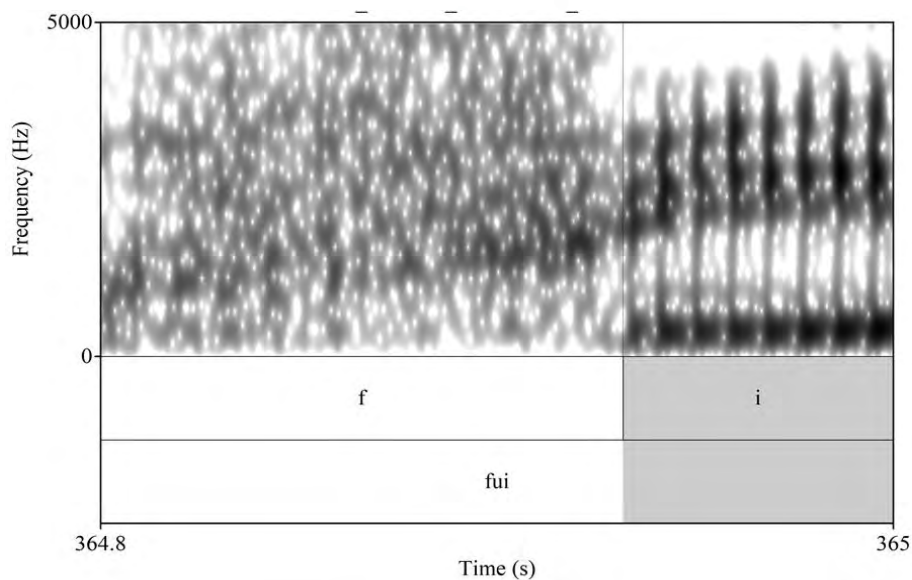


2.4.2.2. Monoptongación

El proceso de monoptongación permite a un diptongo convertirse en una sola vocal, como en la palabra /pie/, este diptongo se convierte en un solo segmento, resultando en [pe] (Ciccía *et al.*, 2009) o, como se muestra en la Figura 34, la palabra /fui/ se convierte en [fi].

Figura 34

Ejemplo de monoptongación

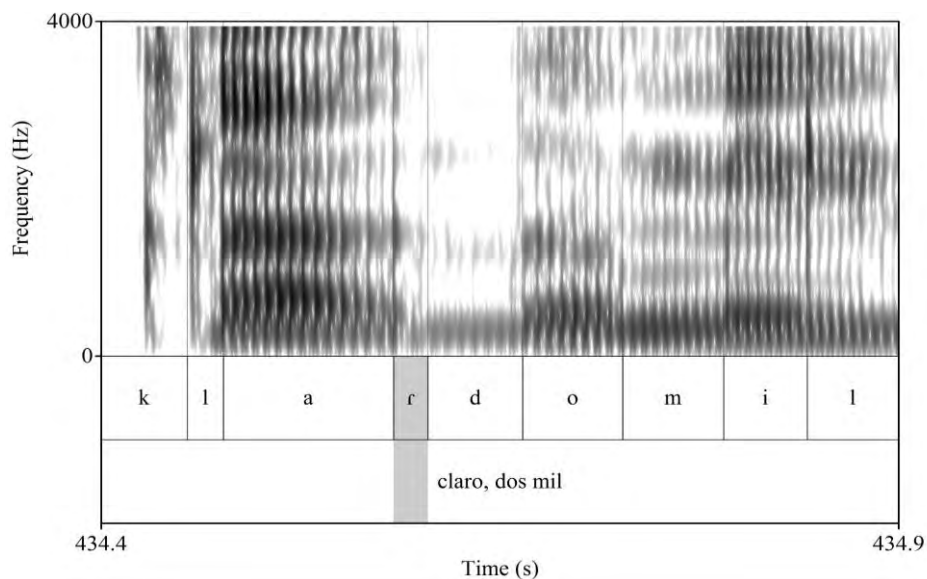


2.4.2.3. Elisión vocálica

La elisión vocálica es un proceso que elimina completamente a la vocal en cuestión dentro de la palabra, o sea no se manifiesta en el plano superficial. Esto se observa en la Figura 35, en la cual, se tiene la frase “claro, dos mil”, la vocal [o] se elide.

Figura 35

Ejemplo de elisión vocálica



2.4.2.4. Sinalefa

La sinalefa es el resultado del contacto entre palabras de contracción análoga, también al grupo de vocales formados por la unión de palabras o enunciados en una sola sílaba, pueden juntarse en grupos de dos hasta de cinco vocales, por ejemplo en la secuencia *ae* “amad(a e)sposa” para el grupo de dos vocales, y la secuencia *ioaeu* “envid(io a Eu)sebio” para el grupo de cinco vocales. (Navarro, 1971)

2.4.3. Procesos que afectan directamente a la consonante

2.4.3.1. Sonorización consonántica

La sonorización consonántica se caracteriza por la presencia de la sonoridad en segmentos que fonológicamente son sordas y que pasan a ser sonoras en la representación superficial, este proceso es muy característico cuando la consonante sorda se encuentra entre vocales. Por ejemplo, se tiene la palabra “paquete”, el segmento consonántico velar oclusivo sordo /k/ se encuentra en un contexto intervocálico, lo cual produce la sonorización del segmento, sin embargo, el segmento oclusivo bilabial sordo no cumple con las condiciones para una sonorización espontánea, puede producirse aislada y sin ningún segmentos sonoro previo, si se sonoriza, sería un caso particular de sonorización de dos segmentos en diferentes

contextos, tal como se observa en la Figura 36; o también como describe Cueva (2018) para el español limeño, muestra un contexto donde el segmento en cuestión se sonoriza por estar entre una vocal y una consonante oclusiva sonora, en la Figura 37.

Figura 36

Primer ejemplo de sonorización consonántica

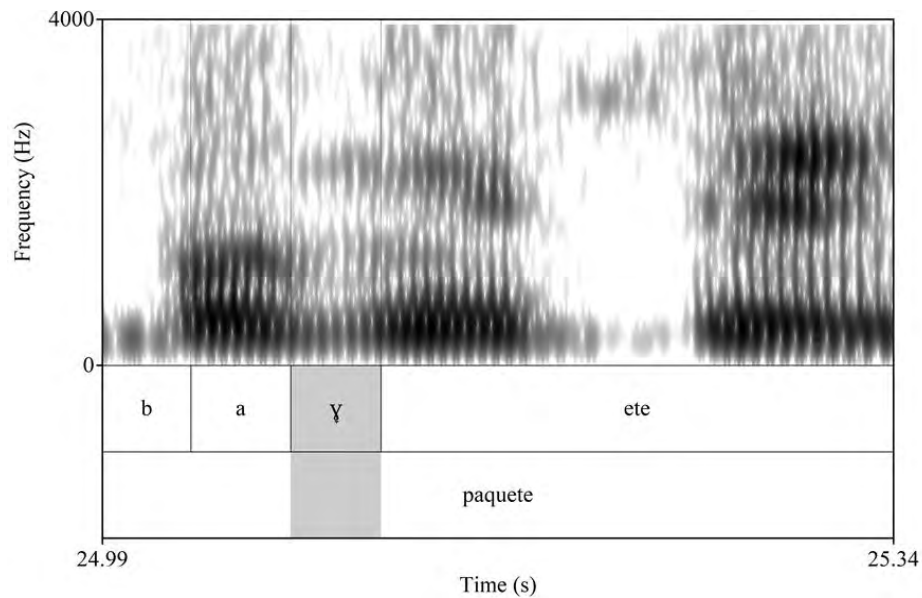
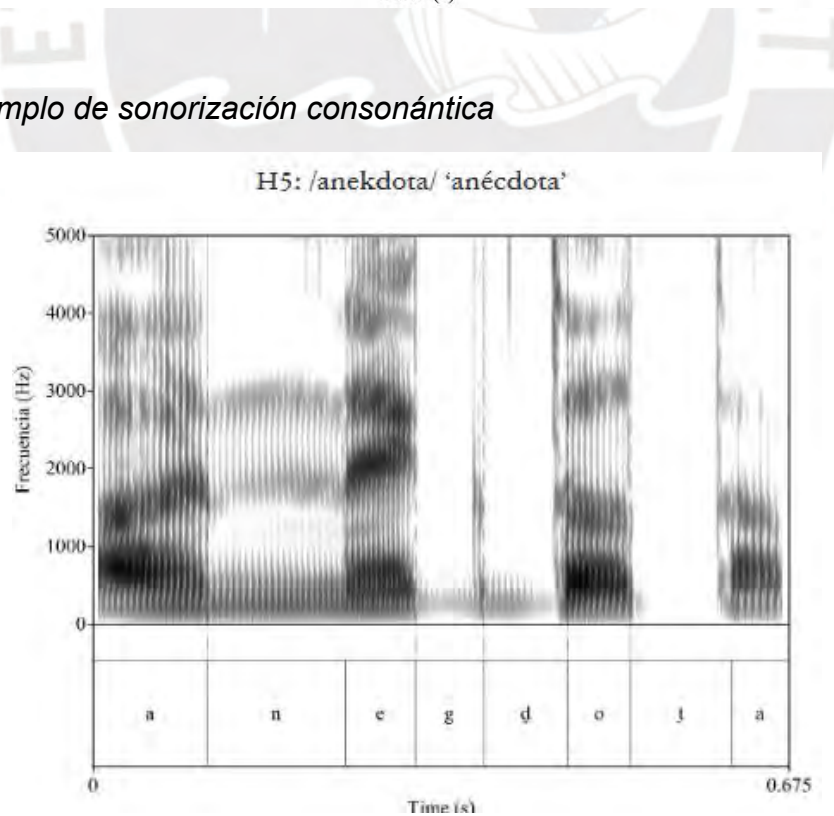


Figura 37

Segundo ejemplo de sonorización consonántica



Nota. Observar la barra de sonoridad en las consonantes sonorizadas a menos de 500 Hz. De Cueva, 2018, p. 168.

2.4.3.2. Aspiración de consonantes oclusivas

Este proceso fonético se manifiesta por la aspiración en la producción de una consonante oclusiva sorda, eso quiere decir que además de la explosión al momento de la apertura de los articuladores, se eyecta un flujo de aire. En la Figura 38, podemos observar la producción de la oclusiva velar [k] con una aspiración que se representa con una h en superíndice, y en la Figura 39 la producción de la oclusiva velar sin aspiración consonántica.

Figura 38

Ejemplo de aspiración consonántica

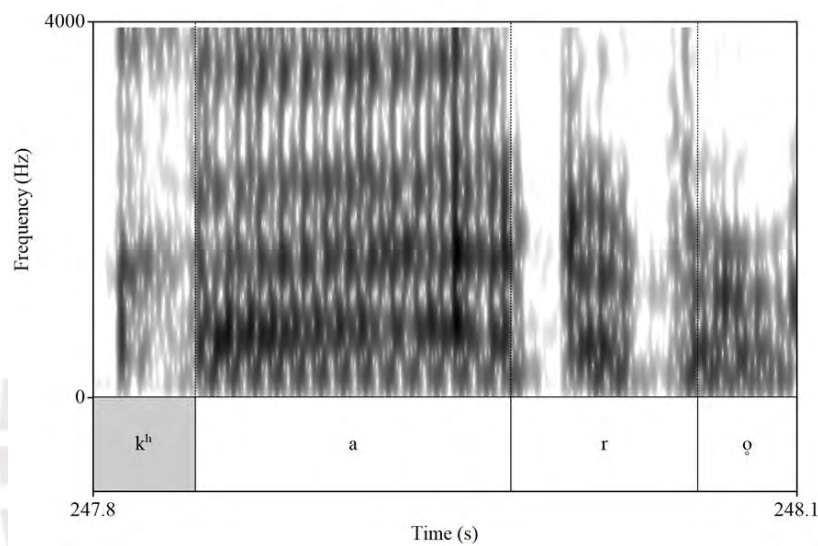
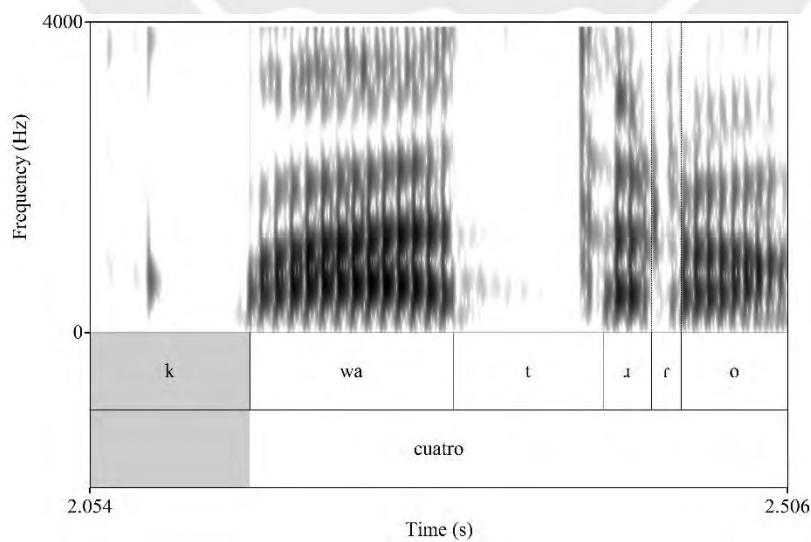


Figura 39

Ejemplo de producción de la [k] sin aspiración

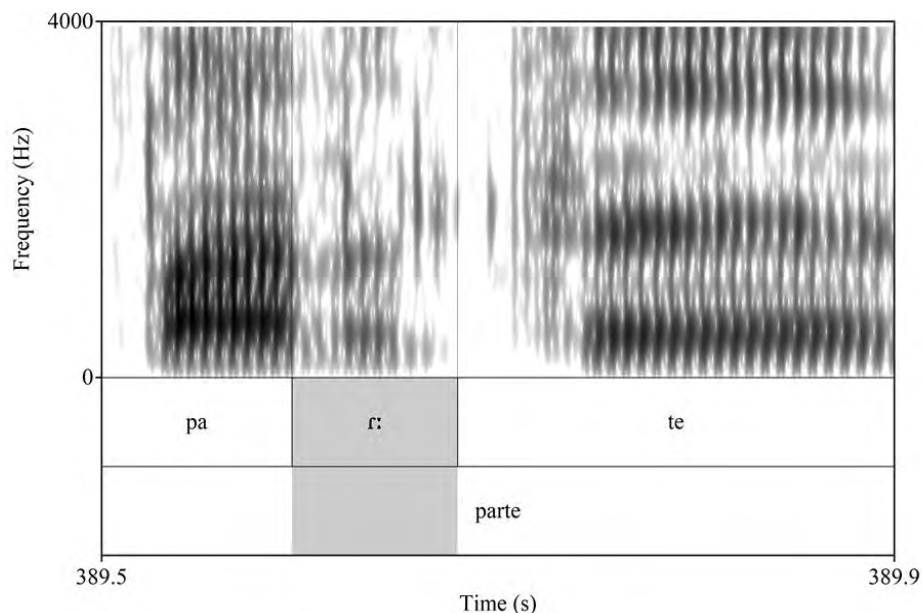


2.4.3.3. Alargamiento de la vibrante simple

El alargamiento de la vibrante simple consiste en la producción de más de una constricción, lo que hace que el sonido tenga mayor duración que una vibrante simple regular, por ejemplo, en la palabra “parte” se puede observar la producción alargada de la vibrante simple por tener más de una obstrucción (ver Figura 40).

Figura 40

Ejemplo de alargamiento de la vibrante simple

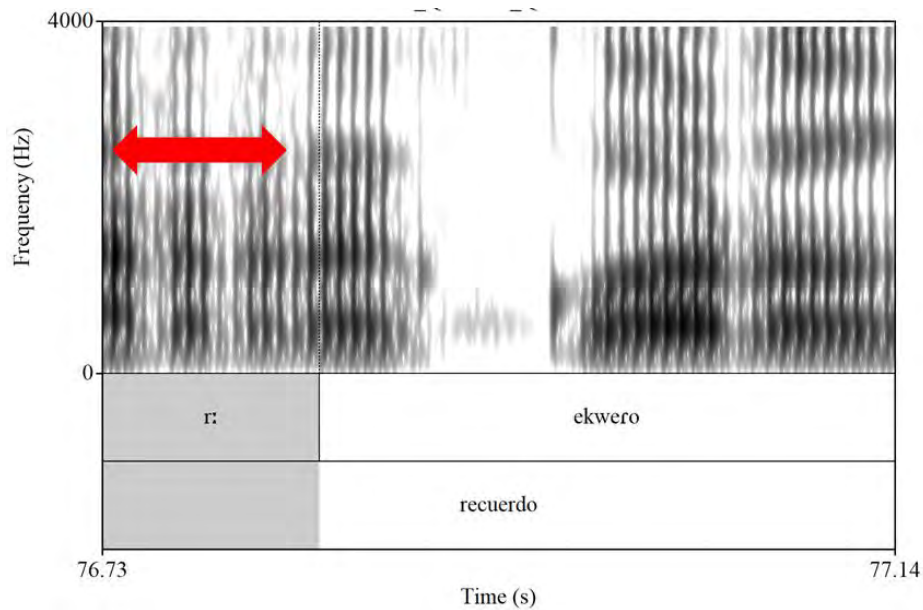


2.4.3.4. Alargamiento de la vibrante múltiple

Al igual que en el caso del proceso anterior, el alargamiento de la vibrante múltiple consiste en la producción de más constricciones de lo regular, ocasionando que la vibrante múltiple se prolongue y adquiera mayor duración que las vibrantes regulares. Esto se puede observar en la Figura 41.

Figura 41

Ejemplo de alargamiento de la vibrante múltiple

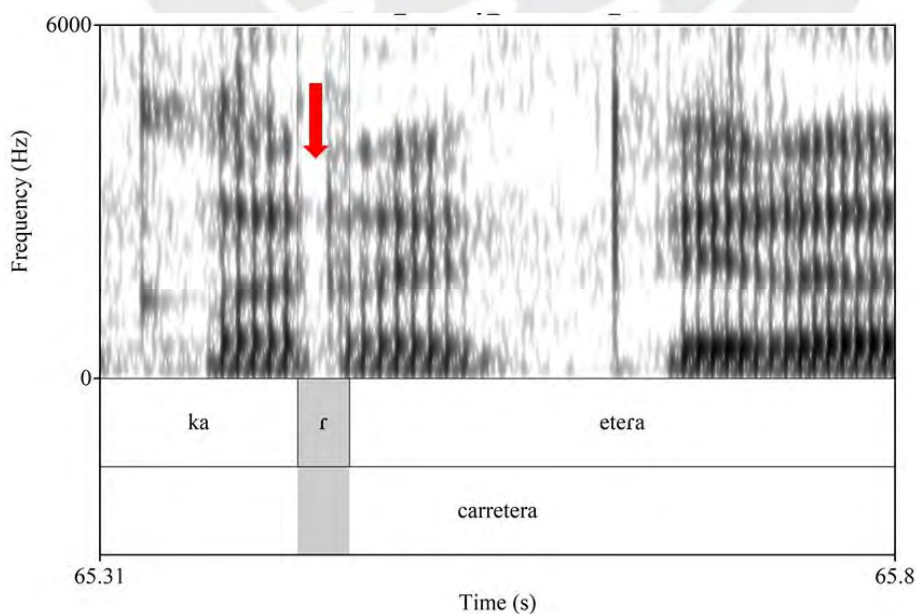


2.4.3.5. Simplificación de la vibrante múltiple

Esta vibrante se caracteriza por poseer de dos constricciones a más, por lo que, en este proceso de simplificación, se logra observar solamente una obstrucción, generando la disminución de duración del sonido, tal como se muestra en la Figura 42.

Figura 42

Ejemplo de simplificación de la vibrante múltiple



2.4.4. Procesos característicos en las consonantes

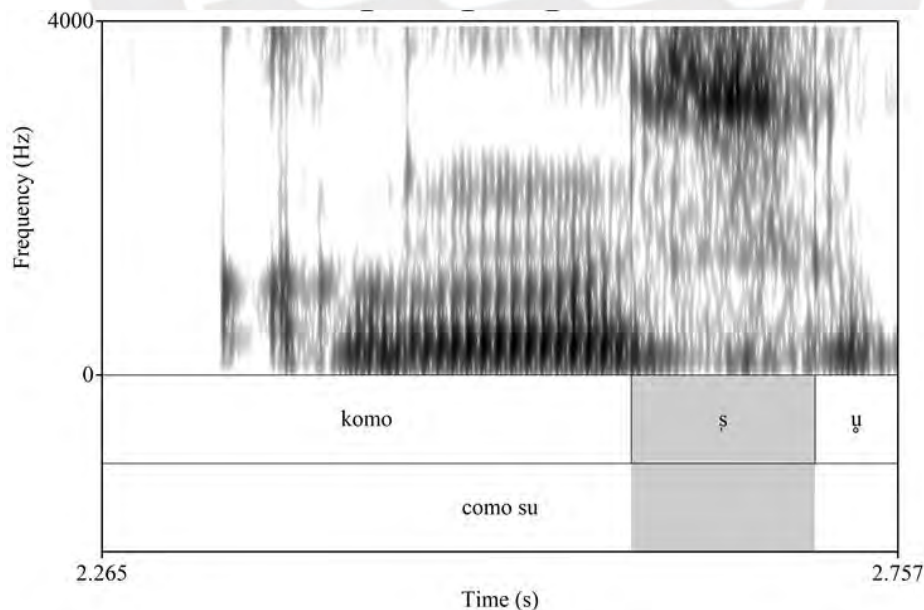
2.4.4.1. Producción de la /s/ en frecuencias bajas del espectro

Como se señaló anteriormente, los sonidos fricativos no poseen una obstrucción completa, por lo que el aire que se expulsa de los pulmones no se encuentra con algún obstáculo que impide totalmente su salida, asimismo, los sonidos fricativos poseen turbulencia, y se dividen en sibilantes y no sibilantes, donde la primera de ellas se caracteriza por concentrar mayor energía en frecuencias altas. Elías-Ulloa (2011, 104) presenta un promedio de las frecuencias de los «picos más altos de los sonidos fricativos del castellano: a) fricativa alveolar [s]: 7000-8000 Hz, b) fricativa postalveolar [ʃ]: 4000-5000 Hz y c) fricativa velar [x]: 1000-2000 Hz. Esto nos brinda una idea más clara de la posición más regular de las fricativas del castellano».

Ahora bien, para este caso en particular, la fricativa /s/ debería concentrar la energía en la cantidad de Hz que se indican en a), sin embargo, esto no sucede en algunos casos, como es la muestra de un individuo (ver Figura 43), la fricción se concentra en niveles más bajos, esto es entre 3500-4000, lo cual nos estaría indicando la producción particular del segmento /s/ respecto a su concentración de energía.

Figura 43

Ejemplo de producción de la /s/ en bajas frecuencias del espectro

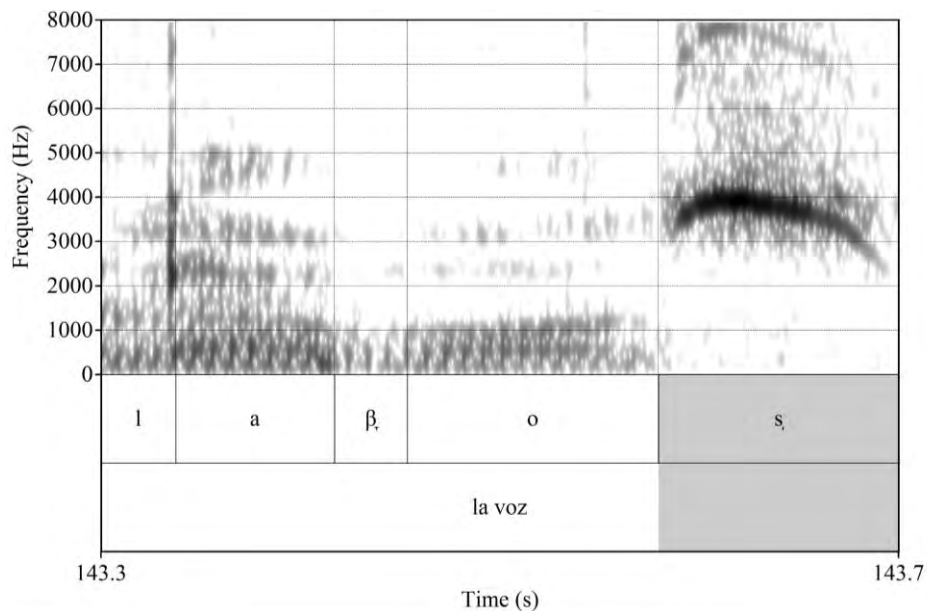


2.4.4.2. Producción del *chiwchi* sound

Este proceso es sumamente interesante, debido a que en la producción del segmento fricativo [s] se concentra la energía en una limitada área, perceptualmente asemejándose al silbido de un ave, en la Figura 44 se puede apreciar aquella concentración de energía que oscila entre los 3000 y 4000 Hz, característico de la producción de la [s] en las frecuencias bajas del espectro.

Figura 44

Ejemplo de producción del chiwchi sound

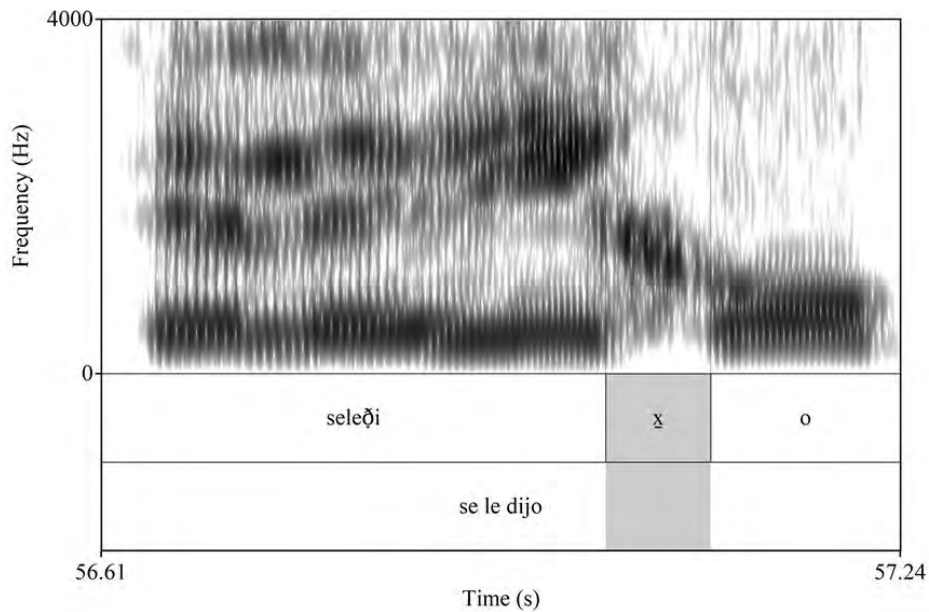


2.4.4.3. Producción estridente de la fricativa velar /x/

La producción estridente de un segmento fricativo se manifiesta por la intensidad de energía que concentra un segmento, sea sibilante o no, para el caso de la fricativa velar, no es un sonido sibilante, por lo que, de manera general, es un sonido que no presenta características de turbulencia concentrada en un nivel, sin embargo, cabe la posibilidad que estos segmentos se produzcan con mayor fuerza en la salida del aire, lo que generaría un sonido estridente. Por ejemplo, en la Figura 45, se tiene la palabra "dijo" en la cual el sonido fricativo velar [x] se encuentra oscilando entre los 1000 a 2000 Hz, además de la concentración de energía notoria en el espectrograma.

Figura 45

Ejemplo de producción estridente de la fricativa velar /x/

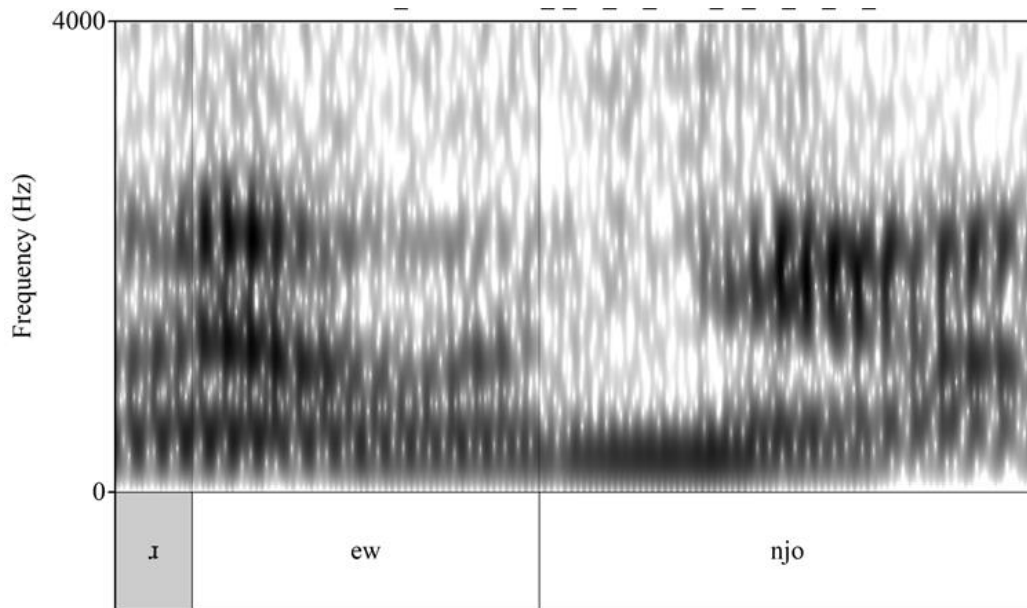


2.4.4.4. Producción aproximante de la vibrante múltiple

En el habla natural, sucede en algunos casos la producción de la vibrante múltiple como una aproximante, su modo de articulación se produce por el acercamiento del tracto vocal al lugar de la articulación, que en este caso es alveolar, asimismo, es sonora y se extiende por toda la duración del sonido. La vibrante aproximante no forma parte del inventario fonológico del español, pero si está presente como alófono, por ejemplo, en la Figura 46, se tiene la palabra “reunión”, y la vibrante múltiple /r/ se encuentra a inicio de palabra, sin embargo, en el plano superficial se manifiesta como una vibrante aproximante. Más ejemplos, en el español peruano del departamento de Puno, de variaciones de la vibrante múltiple se pueden encontrar en el trabajo realizado por Cueva (2021).

Figura 46

Ejemplo de producción aproximante de la vibrante múltiple

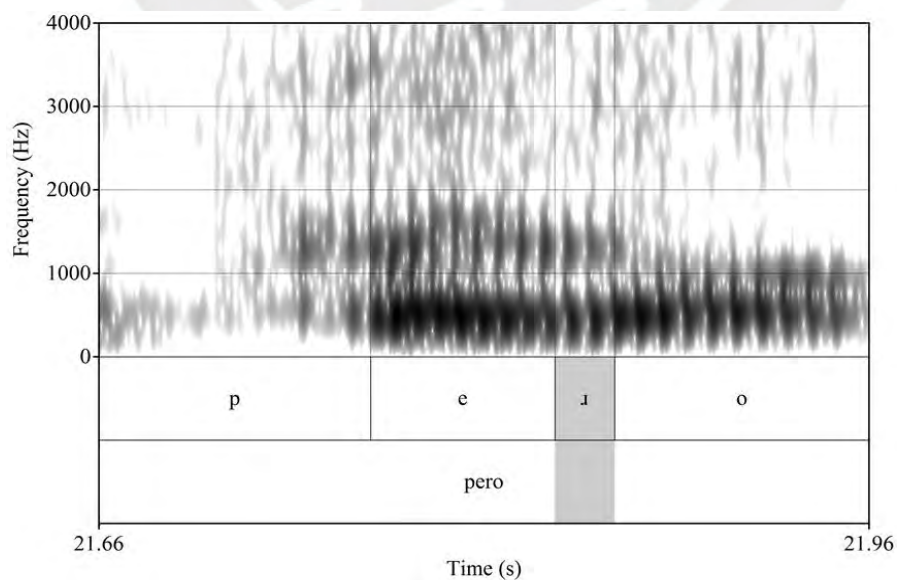


2.4.4.5. Producción aproximante de la vibrante simple

La vibrante simple también se puede producir como una vibrante aproximante, tal y como señala Cueva (2021) en su investigación sobre las vibrantes de Puno. La constricción no logra concretarse, y eso se logra observar en el espectrograma de la Figura 47.

Figura 47

Ejemplo de producción aproximante de la vibrante simple



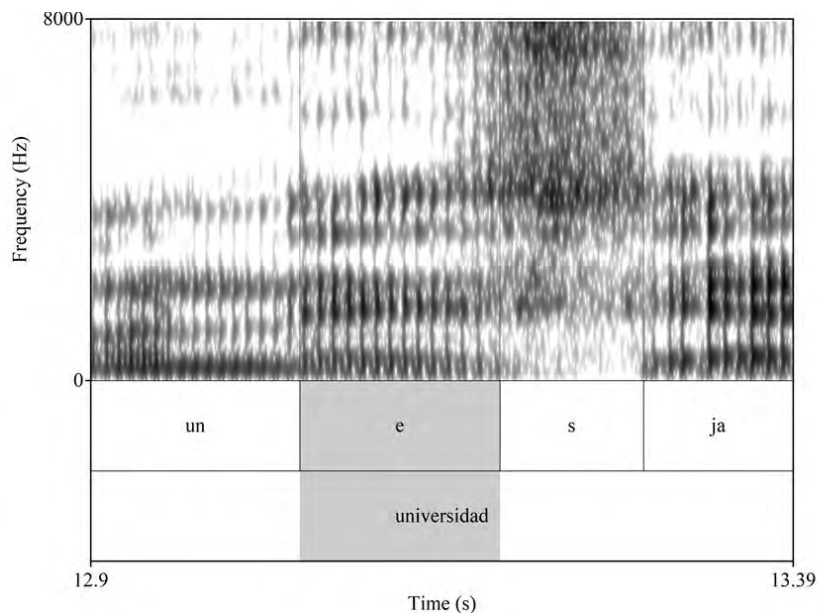
2.4.5. Procesos que cambian a la consonante

2.4.5.1. Elisión consonántica

En la descripción de los procesos que cambian a las vocales se observó que la elisión del segmento es total, no aparece en el espectrograma, lo mismo puede suceder con las consonantes, por ejemplo, se observa en la Figura 48 la palabra “universidad”, y esta es reducida a [u.ne.sja].

Figura 48

Ejemplo de elisión consonántica

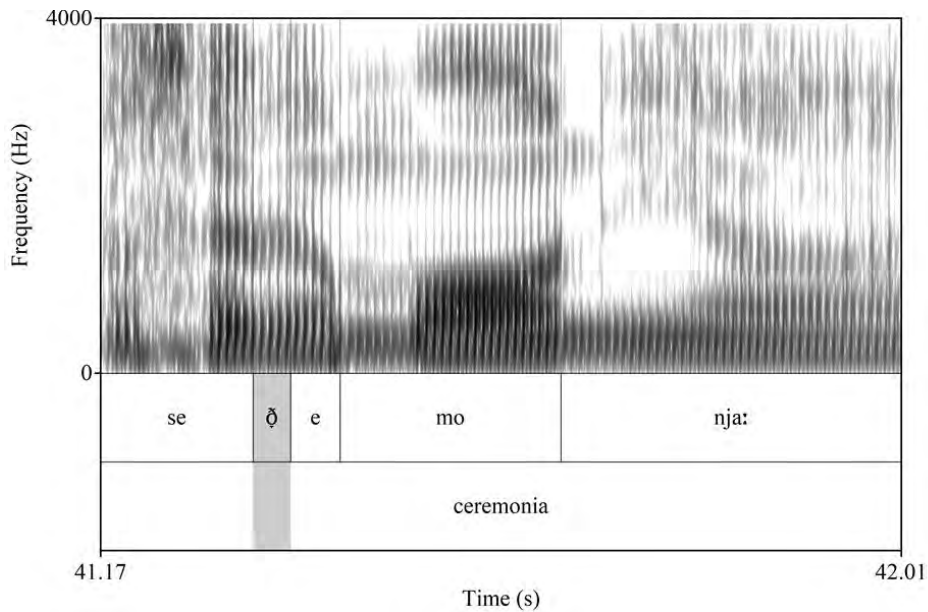


2.4.5.2. Rotacismo por sustitución de consonantes

El rotacismo es una alteración del fonema vibrante, y suele suceder porque son los sonidos más difíciles de pronunciar, existen casos en la que estos sonidos son sustituidos por otras consonantes, pueden ser por [d] o [t] (Uzco, 2018). En este caso, se tiene en la Figura 49, la palabra “ceremonia”, en la cual la vibrante simple es sustituida por el sonido consonántico aproximante alveolar.

Figura 49

Ejemplo de rotacismo por sustitución

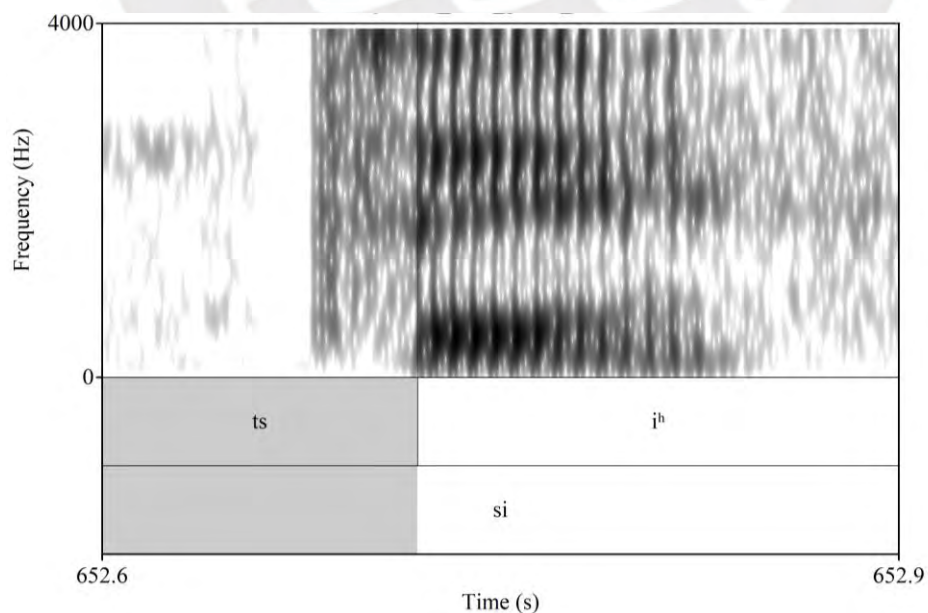


2.4.5.3. Africación de la fricativa /s/

Este proceso de /s/ → [ts] cambia totalmente a la consonante fricativa alveolar, convirtiéndolo en un segmento africado en el plano superficial, se adiciona un segmento oclusivo para formar un sínfone, esto se demuestra en la Figura 50.

Figura 50

Ejemplo de africación de la fricativa /s/

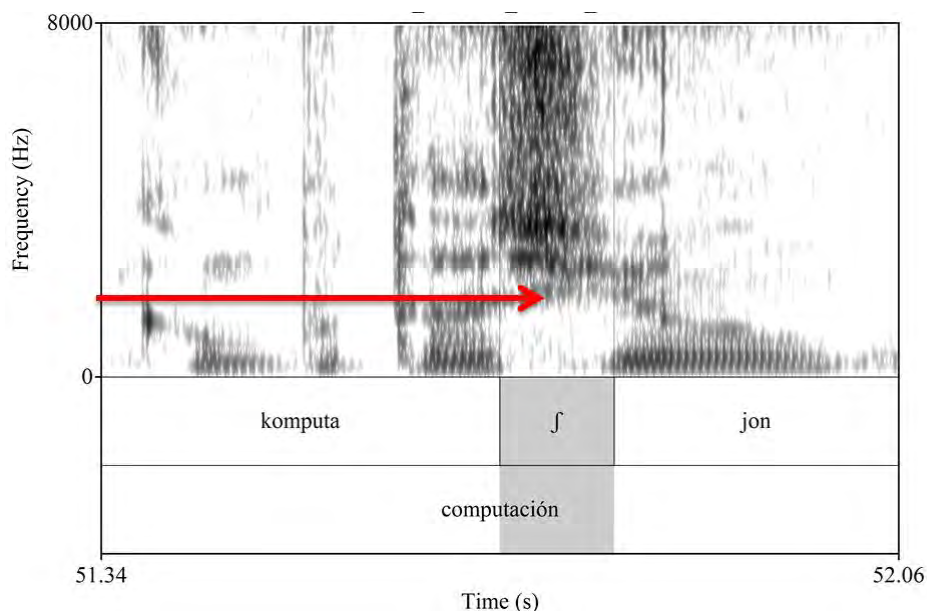


2.4.5.4. Palatalización de la fricativa /s/

Este fenómeno se describe como la producción de la fricativa palatal [ʃ] en vez de la fricativa alveolar [s], cabe la posibilidad de que el segmento cambie debido al contacto con un segmento de la misma naturaleza palatal, por ejemplo, en la Figura 51, el caso de la jod, en la palabra “computación”, que le precede el segmento /s/, y en la producción del habla se manifiesta como la fricativa palatal [ʃ].

Figura 51

Ejemplo de palatalización de la fricativa /s/

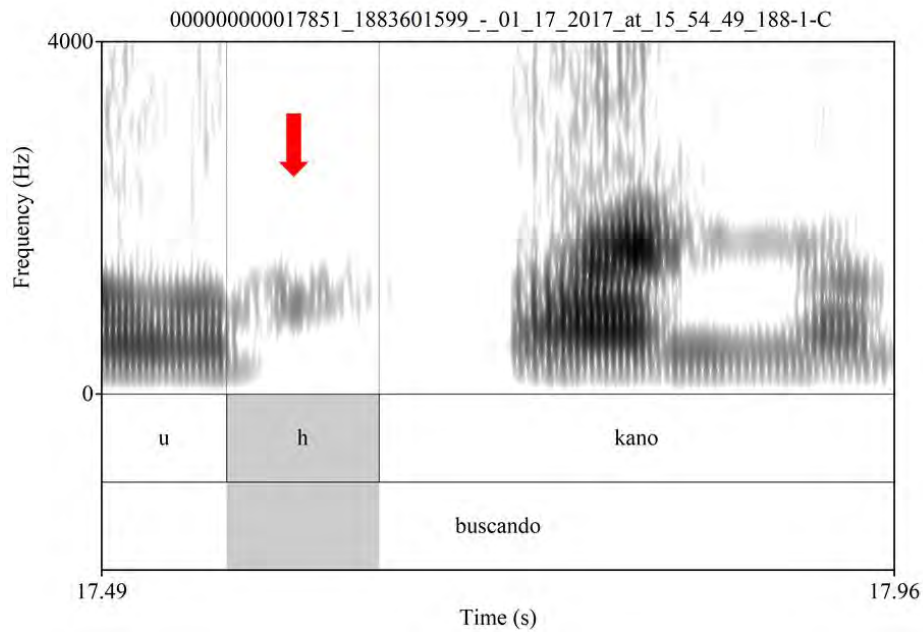


2.4.5.5. Aspiración de la fricativa /s/

La aspiración de la fricativa alveolar /s/ es un proceso, en la cual el segmento cambia naturaleza, y se convierte en un sonido glotal [h], el sonido glotal no forma parte de los sonidos reconocidos en la lengua española, pero se puede encontrar en la pronunciación de diferentes variantes del español. (Hualde, 2014). Un ejemplo de ello es en la palabra “buscando”, en la cual, /s/ se realiza como un sonido glotal, ver Figura 52.

Figura 52

Ejemplo de aspiración de la fricativa /s/

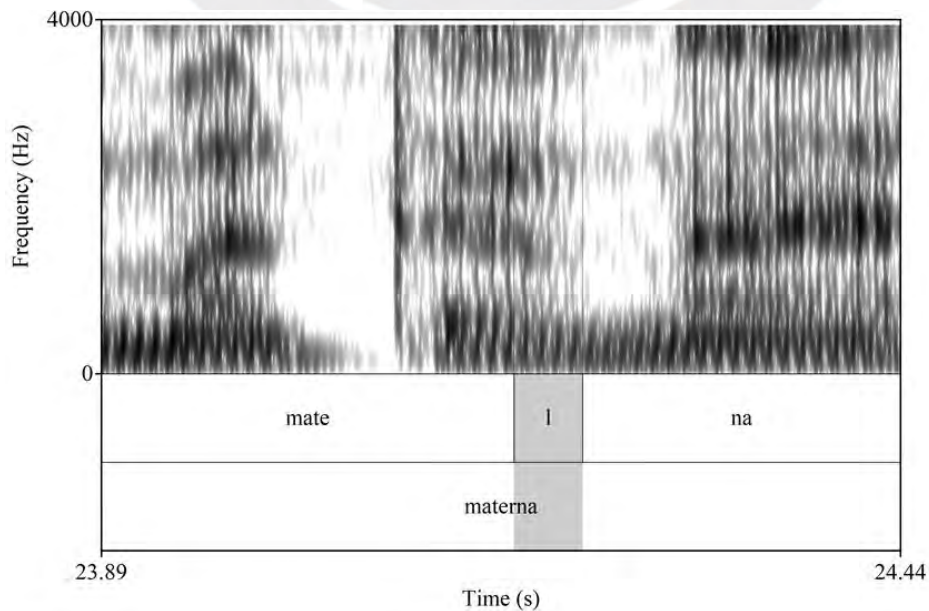


2.4.5.6. Lateralización

Este fenómeno fonético se presenta cuando un segmento, que en muchos casos es la vibrante simple, se convierte en un sonido líquido lateral, esto se evidencia en la Figura 53.

Figura 53

Lateralización

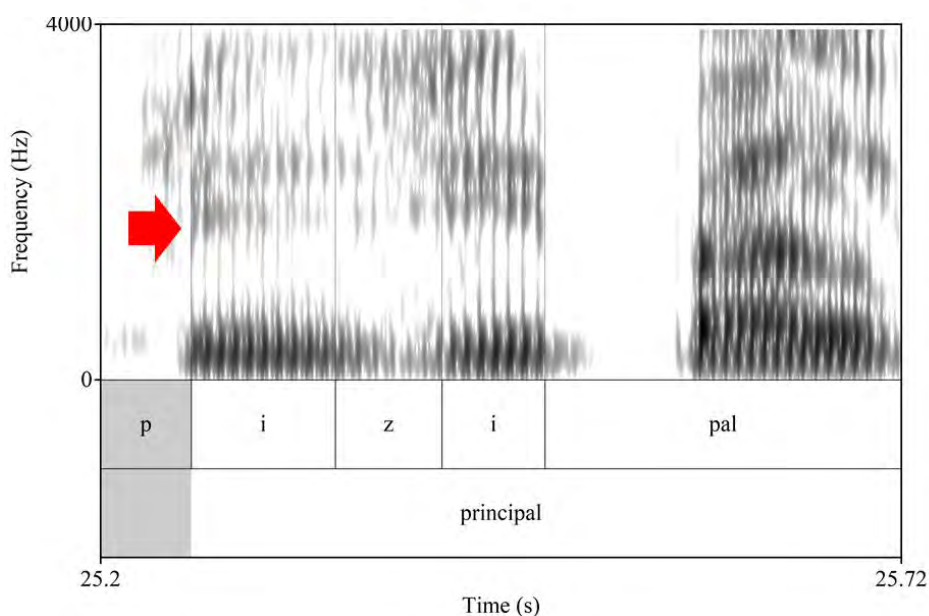


2.4.5.7. Supresión consonántica

Un grupo consonántico o también llamado sílfone son consonantes trabadas que están formadas por la combinación de una consonante obstruyente y una consonante líquida (Gutierrez. s./f.). La supresión consonántica consiste en la elisión de un elemento de una sílfone, eso quiere decir que del grupo consonántico, solo uno de ellos se va a manifestar en el plano superficial del habla. Por ejemplo, se tiene en la Figura 54 la palabra “principal”, y que del grupo /pr/ solo se expresa la consonante [p].

Figura 54

Ejemplo de supresión consonántica



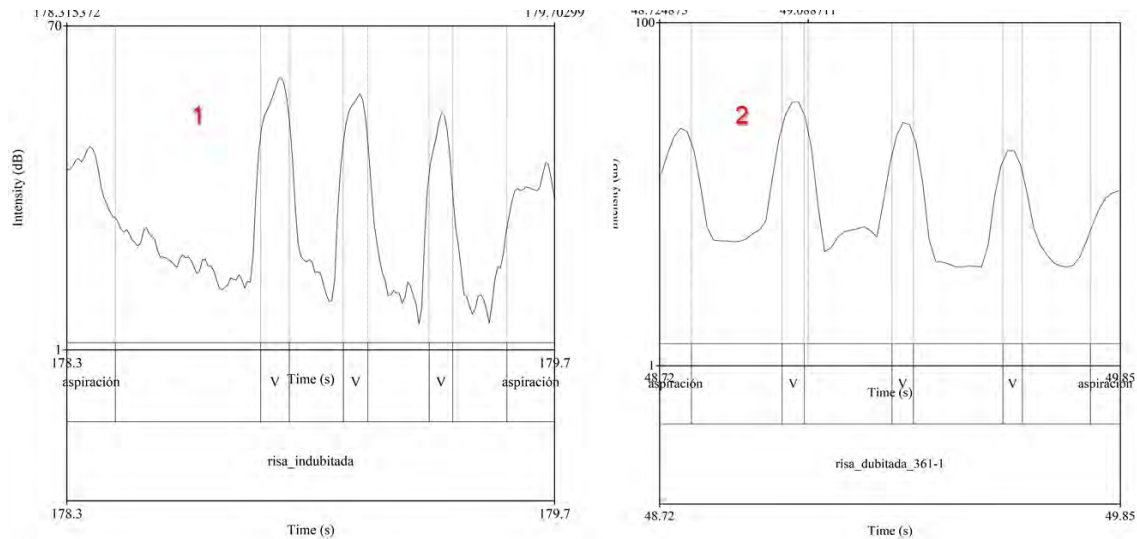
2.5. Elementos paralingüísticos

2.5.1. Las risas

Las risas conforman un elemento de comparación único, con ellas es posible determinar, desde el plano paralingüístico, si muestras de comparación de voz son compatibles o no. Es posible con el programa Praat observar el espectrograma de este fenómeno y verificar las similitudes y/o diferencias entre muestras. Un ejemplo de similitud entre dos muestras se observa en la Figura 55, donde al inicio y al final de las risas de ambas muestras se manifiesta una aspiración, junto con tres picos vocálicos. Esto puede variar conforme a las muestras que se analicen.

Figura 55

Ejemplo de comparación de risas en dos muestras

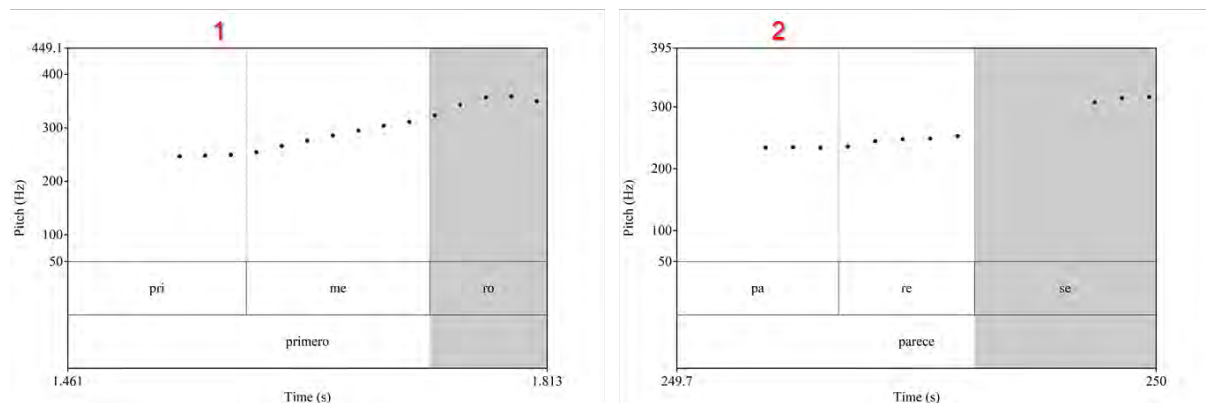


2.5.2. Acento

Otro factor importante de identificación de correspondencia de voces es el análisis del acento particular de un individuo, por ejemplo, en la Figura 56, se tiene que en dos palabras de naturaleza de acentuación grave, se producen como agudas; en (1) se observa la palabra “primero” [pri.me.'ro], la cual es una palabra que fonológicamente posee una acentuación en la segunda sílaba /pri.'me.ro/, esta sección del audio pertenece a una muestra indubitada de un hablante de la región del norte del Perú, en (2) que corresponde a una muestra dubitada, de la cual se tiene sospecha que pertenece a la voz de la muestra indubitada (1), se observa el mismo proceso, existe un cambio en la acentuación: en la palabra “parece” [pa.re.'se] hay un cambio en la posición del acento, a pesar de ser una palabra grave fonológicamente /pa.'re.se/, sin embargo, esto no sucede. Mediante el programa Praat, en la opción de *show Pitch*, se muestra el recorrido del pitch a través de líneas o, en este caso, puntos.

Figura 56

Ejemplo de comparación del acento en dos muestras



2.6. Estudio de la comparación biométrica de voces

2.6.1. Métodos automáticos por voz

En el caso peruano, existe una propuesta de dictamen pericial acústico forense, en el cual no se consigna el uso de programas de biometría de voz, ni la consideración de la evaluación de otros parámetros concernientes a la lengua como medio probatorio, solo se indica en (Gutierrez, 2015) el análisis de los formantes F1 y F2. Por otra parte, específicamente en la Oficina de Peritajes, autorizado por la Guía del Ministerio Público (2020), se incluye al *software* SIS II para ser utilizado por el personal del Área de Acústica y Fonética Forense con la finalidad de incluir en las conclusiones los resultados del programa biométrico y poder darles mayor solidez a los informes periciales.

2.6.2. Software SIS II

Existen pocas investigaciones respecto al uso del software biométrico de voz de comparación automática «SIS II, perteneciente al paquete IKAR Lab3, creado por la empresa rusa *Speech Technology Center* (STC); sin embargo, esto no demerita su óptima funcionalidad, ya que es utilizado en diferentes países, y un caso de ellos es el peruano» (Jimenez *et al.* 2022, 434). Asimismo, este software posee *plugins* para la homologación de voz, que permiten su funcionamiento a cabalidad (Pop-Dimitrijoska *et al.*, 2012)

En suma, los resultados de una investigación realizada por Morrison *et al.* (2016), en López-Escobedo & Solórzano-Soto (2016), indican que, primero, las organizaciones

que combaten el crimen en el mundo utilizan en su mayoría los sistemas de reconocimiento automático de voz son el BATVOX, consecutivo al IKAR LAB, segundo, que ambos programas de comparación biométrica de voces tienen la capacidad de clasificar las muestras de voz relacionándolas con determinados atributos.

El *software* ruso SIS II pertenece al paquete IKAR Lab3 del *Speech technology Center*, con código STC-S521, este contiene herramientas de análisis de señal de la voz, visualización del habla mediante espectrogramas, permite segmentar y reconocer el habla, asimismo, cuenta con *plugins* de identificación automática y semiautomática, entre otros (STC, 2001).

En la *Guía del Usuario*, se presenta la funcionalidad de los *plugins*: a) identificación rápida, b) características de la señal, c) marcación de locutores, d) detección de voz, e) comparación automática, f) comparación de formantes, g) comparación de tono, h) comparación de las características auditivas, y por último, la i) conclusión final.

Cada *plugin* cumple una función única, y para el presente estudio sólo nos ceñimos a mostrar el análisis de la calidad de la señal y la comparación automática de voces.

Características de la señal

El *software* SIS II nos permite calcular las características de la señal a partir de los siguientes factores: respuesta en frecuencia, relación señal/ruido, ruidos tonales, tiempo de reverberación medio, dispersión, clipping y duración del fragmento de la voz, asimismo, el cálculo permite seleccionar si es de canal stereo o mono, así como la fuente de grabación, si se ha obtenido por micrófono o por teléfono.

Los siguientes factores se definen según se indica en el documento *SIS II: Suite de plugins de identificación automática* del programa SIS II (v2.6.357), *software* utilizado para la evaluación de calidad de los archivos de audio.

a) SNR (*Signal-to-Noise Ratio*): Este parámetro, denominado relación señal-ruido, es el encargado de medir cuánto ruido existe en una señal de habla.

b) Ruido tonal (*Stationary Tonal Noise*): Es un parámetro que determina cuánto es la afección de ruidos indeseables provocados por campos magnéticos, electromagnéticos de las tensiones/corrientes eléctricas. Los ruidos tonales son «unas señales periódicas de frecuencia del período fundamental desde las decenas de hercios hasta frecuencias más altas» (SIS II).

c) Tiempo de reverberación (*Reverberation Time*): Es un parámetro que determina la disminución gradual de la intensidad ocasionada por las repeticiones de la misma señal. La dispersión es subparámetro ligado a la reverberación.

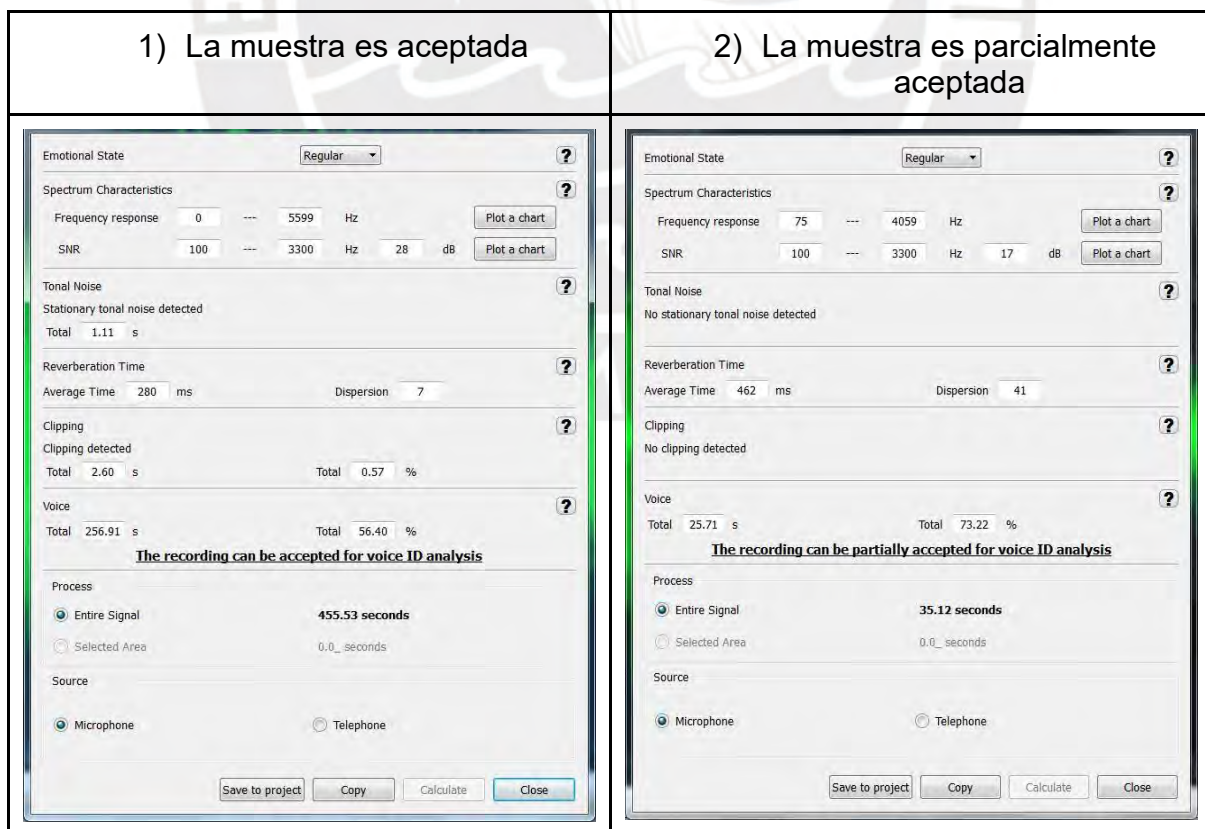
d) Saturación (*Clipping*): Es una clase de alteración de la señal cuando existe una sobrecarga de un amplificador de manera que supera el límite de la tensión de la alimentación. Estos se perciben como “chirridos” en la señal.

e) VAD (Voice Activity Detection): Es el instrumento de programa para separar el habla activa del ruido de fondo o del silencio.

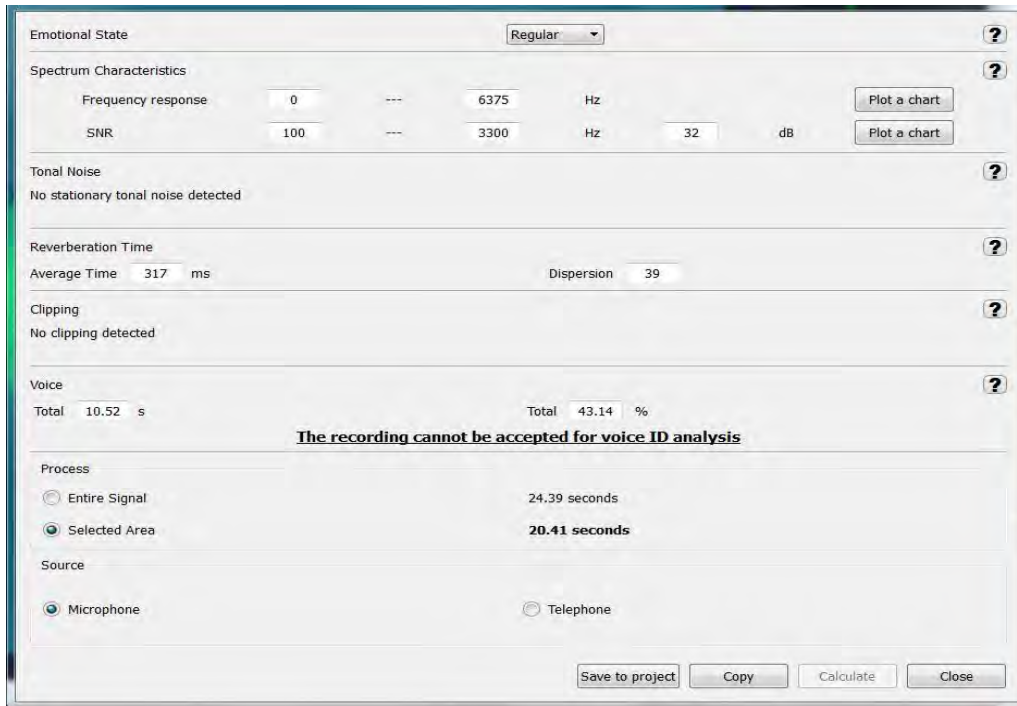
En la Figura 57, se muestran tres imágenes, en las cuales describe el análisis de tres archivos de audio, donde en 1) se observa un análisis en la que la muestra es aceptada para el proceso de comparación de voces, en 2) la muestra es parcialmente aceptada para la comparación de voces y 3) donde la muestra no es aceptada para la comparación de voces.

Figura 57

Tipos de resultados en el análisis de la calidad de la señal



3) La muestra no es aceptada



Según la *Speech Technology Center*, el programa SIS II posee un estándar que determina si las muestras son o no aprovechables para el análisis de comparación de voz, esto se detalla en la Tabla 10:

Tabla 10

Estándar de resultados según el análisis de calidad de la señal del software SIS II

Clasificación de las grabaciones de audio apropiadas o no apropiadas para el análisis	Parámetro		
	Relación señal/ruido	Tiempo de reverberación	Duración de la grabación de voz
La grabación de audio es apropiada para el análisis	Más de 20 dB	Menos de 500 ms	Más de 30 s
La grabación de audio es moderadamente apropiada para el análisis	De 10 dB a 20 dB	De 500 ms a 700 ms	11 s a 30 s
La grabación de audio no es apropiada para el análisis	Menos de 10 dB	Más de 700 ms	Menos de 11 s

Si alguno de los valores es menor a lo resaltado en la tabla por el color amarillo, el *software* forense concluirá que no es aceptado para la comparación automática de voces.

De ser el caso que se utilice alguna herramienta para el mejoramiento del audio, si el juez solicita las transcripciones de un audio, se puede proceder a mejorar el audio y quitar el ruido que intersecta la voz de la o las personas a transcribir, pero si el juez, solicita la comparación de muestras de habla, el perito no debe manipular la señal. (Cicres, 2011)

Comparación automática de voces

La comparación automática de voces se basa en tres métodos: a) método de estadísticas del tono, b) método de espectro de formantes y c) método de variabilidad tonal. Según la Guía de usuario del *Speech Technology Center*, los métodos del SIS II (2001) actúan de la siguiente forma:

- a) El análisis del tono basado en un enfoque estadístico utiliza dieciséis características relacionadas con el tono fundamental de la voz. Estas características abarcan el promedio, los valores máximo y mínimo, la mediana, el porcentaje de intervalos con tono ascendente, la dispersión logarítmica del tono, la asimetría logarítmica del tono, el exceso logarítmico del tono y otros parámetros. Este método tiene la ventaja de evaluar el estado emocional del hablante al momento de la pronunciación, permitiéndolo utilizar como un complemento adicional. La extracción del tono fundamental se lleva a cabo mediante el análisis espectral de la señal de audio grabada. El fundamento de este método se basa en un algoritmo que analiza los armónicos del tono presentes en el espectro de la señal. El objetivo del algoritmo es superar los problemas derivados de la supresión de cierta parte del espectro en un canal. Durante la etapa de conclusión e identificación, se comparan las características estadísticas del tono, se calculan los errores de tipo I y tipo II (FA y FR) para cada valor, y se obtienen los valores correspondientes. Estos últimos se determinan mediante métricas estadísticas del tono que representan la suma ponderada de todas las desviaciones relativas de las características estadísticas del tono fundamental en comparación entre las dos grabaciones. Se emplea el método de bagging según la duración de la señal de voz (los valores de FR/FA varían en función de la duración de la señal de voz utilizada para calcular cada modelo de MDN).

- b) El método de análisis del espectro de formantes se sustenta en la premisa de que cada persona tiene una estructura vocal única y que la voz presenta diversas características en su espectro. Estas características se evidencian principalmente en la distribución de frecuencias y en la posición relativa de los formantes. Además, este método se basa en la extracción y comparación de las posiciones y dinámicas de al menos tres formantes. El enfoque del espectro de formantes ofrece cierta resistencia al ruido de amplio espectro, como el ruido aleatorio o de fondo, así como a las variaciones en las señales debidas a diferencias entre canales.
- c) El enfoque de variabilidad total se basa en el uso de los coeficientes MFCC (Mel Fourier Cepstrum Coefficients, coeficientes de cepstrum en la escala Mel) como características que describen la representación espectral de una señal de voz en momentos específicos. Estos coeficientes son modelados mediante una combinación de distribuciones normales (MDN) para representar la densidad de su distribución. En la etapa de comparación, los i-vectores son sometidos a un proceso de clasificación utilizando el clasificador SVM, un método probado y efectivo que permite obtener una identificación rápida y confiable. Los errores de tipo I y tipo II (FR y FA) se calculan en función de la distancia del SVM. Además, se aplica la técnica de bagging teniendo en cuenta la duración de la señal de voz, lo cual influye en los valores de FR/FA al calcular cada modelo de MDN.

De acuerdo con SIS II (2001), la conclusión de los métodos, que puede realizarse con uno, dos o tres métodos indicados anteriormente, se realiza a partir del método de la pseudoprobabilidad P, el valor que muestra el porcentaje de probabilidad del hecho que los locutores de la grabación 1 y la grabación 2 son la misma persona; 0% significa que los locutores son positivamente diferentes, mientras que 100% muestra que los locutores cuya voz está grabada en los dos archivos audio son en efecto la misma persona. Esto se puede resumir de la siguiente manera:

- Alta probabilidad, mismo locutor (100 a <)
- Media probabilidad, mismo locutor (10 a 99)
- Baja probabilidad, mismo locutor (1 a 9)
- Baja probabilidad, diferente locutor (0.9 a 0.1)
- Media probabilidad, diferente locutor (0.09 a 0.001)
- Alta probabilidad, diferente locutor (0.0009 a >)

2.7. Software SplitsTree 4

El *software* SplitsTree (Huson & Bryant, 2006), desarrollado por la Facultad de Ciencias de la Universidad de Tübingen, pertenece a la categoría de algoritmos utilizados en bioinformática. Esta herramienta permite generar redes filogenéticas no enraizadas a partir de datos de secuencias moleculares. Utilizando una alineación de secuencias como entrada, el programa genera un árbol filogenético o una red utilizando diferentes métodos, dependiendo del tipo de análisis, ya sea por hibridación o por datos simples. El software SplitsTree4 es un programa cuya aplicación ha sido utilizada para calcular redes filogenéticas, a partir de datos de secuencias moleculares. Según la alineación de secuencias, se crea una matriz de distancias o un conjunto de esquemas arbóreos, el programa utiliza métodos de descomposición dividida, redes vecinas, redes de consenso, métodos de superredes o métodos para computar redes de hibridación o recombinación simple (Huson y Bryant, 2006).

Actualmente, se encuentra disponible la versión SplitsTree Community Edition (CE), que es la sexta versión del software. Sin embargo, para este estudio se utiliza la versión 4, debido a los siguientes factores:

1. Estabilidad y madurez: Splitstree 4 ha sido ampliamente utilizado y probado a lo largo del tiempo. Como resultado, es un programa estable y confiable que ha demostrado su eficacia en numerosos estudios filogenéticos. En contraste, Splitstree CE es una versión beta temprana que aún está en desarrollo activo, lo que implica que podría haber errores y características faltantes.
2. Funcionalidades disponibles: Splitstree 4 ofrece una amplia gama de funcionalidades y herramientas para la construcción y visualización de árboles filogenéticos, incluyendo opciones como árboles enraizados y no enraizados, redes filogenéticas, tanglegrams y más. Por otro lado, dado que Splitstree CE es una versión en desarrollo, es posible que algunas funcionalidades estén ausentes o no estén completamente implementadas.
3. Soporte y documentación: Splitstree 4 ha existido durante bastante tiempo, lo que ha permitido que se genere una cantidad significativa de documentación, tutoriales y comunidades de usuarios que pueden brindar apoyo en caso de preguntas o problemas. Splitstree CE, al ser una versión más nueva, podría tener una disponibilidad limitada de recursos de apoyo y documentación.

Además, los creadores del programa en cuestión, indican que el programa al ser una versión *beta*, puede presentar fallas en los cálculos y que los formatos de los archivos pueden cambiar.

Splitstree4 también puede ser utilizado en el campo de la lingüística para analizar y visualizar datos relacionados con la evolución y las relaciones entre las lenguas. En lingüística histórica y comparativa, por ejemplo, se pueden utilizar datos lingüísticos, como vocabulario o características gramaticales, para construir redes filogenéticas y explorar las relaciones entre diferentes lenguas. Este software permite representar visualmente las similitudes y diferencias entre las lenguas, identificar posibles relaciones de parentesco lingüístico, y detectar fenómenos de reticulación o hibridación lingüística. Esto puede ayudar a los lingüistas a comprender mejor la historia y la evolución de las lenguas, así como a reconstruir posibles proto-lenguas y rutas migratorias.

Además, Splitstree4 también puede ser utilizado en el análisis de dialectos y variaciones regionales de una lengua. Al construir redes filogenéticas basadas en características lingüísticas específicas, los lingüistas pueden visualizar las relaciones entre los diferentes dialectos y determinar patrones de cambio lingüístico a lo largo del tiempo.

En resumen, Splitstree4 ofrece a los lingüistas una herramienta visual y analítica para explorar y analizar datos lingüísticos y comprender las relaciones evolutivas entre las lenguas y sus variantes. Su capacidad para representar redes filogenéticas complejas lo convierte en una herramienta valiosa en el campo de la lingüística histórica, comparativa y dialectológica.

Un ejemplo de utilidad para la lingüística está relacionado con la lingüística histórica, como el de Heggarty *et al.* (2010), en el que trabajan con lenguas germánicas y en especial, el inglés, observan cambios fonéticos e innovaciones que permiten delimitar mediante divergencias las distancias entre las lenguas germánicas. Utilizan redes filogenéticas para representar la divergencia entre lenguas.

McMahon *et al.* (2007) llevaron a cabo un estudio sobre las diferentes variantes del idioma inglés, donde se describen las similitudes fonéticas entre las muestras relacionadas con el inglés, especialmente las de origen británico. Además, realizaron un análisis de los cambios sincrónicos que ha experimentado el idioma a lo largo del tiempo. Por otro lado, se han realizado también estudios en lenguas peruanas, como el de Heggarty (2005), en el cual se comparan veinte (20) variantes de lenguas

andinas pertenecientes a las familias quechua, aimara y uru-chipaya. El objetivo de este estudio fue cuantificar la similitud entre las diferentes variantes lingüísticas de estas familias mencionadas, y se presentaron datos fonéticos, léxicos y semánticos. Estos datos resultan útiles para contrastar la propuesta inicial de Torero con la información recopilada por Heggarty.

Un estudio reciente realizado utilizando el software SplitsTree 4 se centró en las lenguas indígenas peruanas, específicamente en el caso de las lenguas y dialectos del Purús (Zariquiey, *et al.* 2017). En esta investigación se empleó un enfoque filogenético para clasificar las diversas variedades lingüísticas presentes en la provincia de Purús, ubicada en Ucayali. Los investigadores utilizaron datos léxicos de nueve (09) variedades lingüísticas de la familia pano, así como información gramatical de ocho variedades. Mediante el uso del programa SplitsTree 4, lograron generar diagramas Neighbor-nets que ayudaron a explicar la distribución de las distintas variedades lingüísticas del Purús.

El algoritmo Neighbor-net utiliza varias fórmulas matemáticas para calcular las distancias y construir la red de fragmentación. Estas fórmulas pueden variar dependiendo de la implementación específica del algoritmo y las características de los datos que se están analizando. A continuación, se mencionan algunas de las fórmulas comunes utilizadas en el algoritmo Neighbor-net:

1. Fórmula de distancia: Se utiliza una fórmula específica para calcular las distancias entre los elementos o muestras que se están analizando. Esta fórmula puede variar según el tipo de datos, como distancias léxicas, distancias fonéticas, etc.
2. Fórmula de disimilitud: Para convertir las distancias en medidas de disimilitud, se pueden utilizar diversas fórmulas, como la resta de las distancias del máximo valor o la inversión de las distancias.
3. Fórmula de ajuste de disimilitud: Para ajustar las medidas de disimilitud y tener en cuenta factores como la varianza o la escala de los datos, se pueden aplicar fórmulas de ajuste, como la transformación logarítmica o la normalización de las disimilitudes.
4. Fórmula de ponderación: En algunos casos, se pueden aplicar fórmulas de ponderación para asignar diferentes pesos a las distancias o disimilitudes según su importancia relativa en el análisis.

2.8. Sobre los métodos combinados en identificación de locutores

El método combinado en la identificación forense del habla se refiere a la integración de diversas perspectivas básicas, como el análisis fonético-fonológico, lingüístico y biométrico. Este enfoque metodológico se considera el más confiable. Estas diferentes aproximaciones representan dos ideas fundamentales: a) cada análisis por separado no es suficientemente preciso para determinar si la voz de una persona es compatible con otra, ya que los valores de precisión no son exactos, y b) al combinar los análisis, se logra una eficacia máxima en la técnica de identificación forense. Para estudiar un objeto científico variable, es necesario abordarlo desde múltiples perspectivas con el fin de lograr un grado óptimo de objetividad y llegar a conclusiones precisas. Si todos los enfoques convergen en la misma dirección, se puede obtener una conclusión más clara.

2.8.1. Niveles de conclusión metodológica

De acuerdo a las comparaciones efectuadas de las muestras con la metodología combinada, se puede establecer criterios de conclusión, estos están relacionados a los resultados de los métodos que se han utilizado en el análisis de identificación forense de voz. Delgado (2001, 244) propone siete niveles de conclusión, son los siguientes: «nivel de identificación, nivel de alta probabilidad, nivel medio-alto de probabilidad, nivel inconclusivo, nivel medio-bajo de probabilidad, nivel de baja probabilidad, nivel de eliminación»

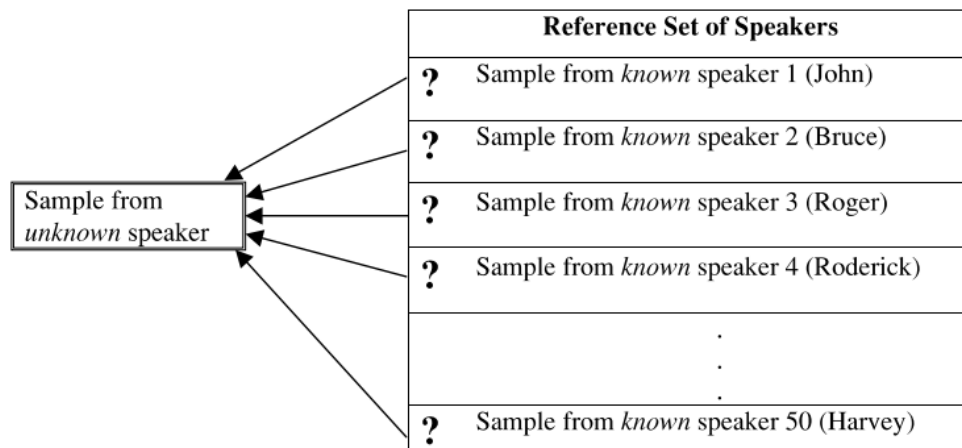
Es importante entender que los niveles de identificación y eliminación en la identificación forense de voz no deben interpretarse como un porcentaje absoluto del 100%. En cambio, deben considerarse como el grado máximo de certeza alcanzable dentro de este campo

2.9. Identificación de locutor

Rose (2002) define a la identificación de locutor como la comparación de una muestra dubitada y muestras indubitadas. Tiene como objetivo determinar si alguna de las muestras indubitadas le pertenece la voz de la muestra dubitada. En la Figura 58, Rose esquematiza el concepto de identificación de locutor.

Figura 58

Esquematzación del concepto de identificación de locutor



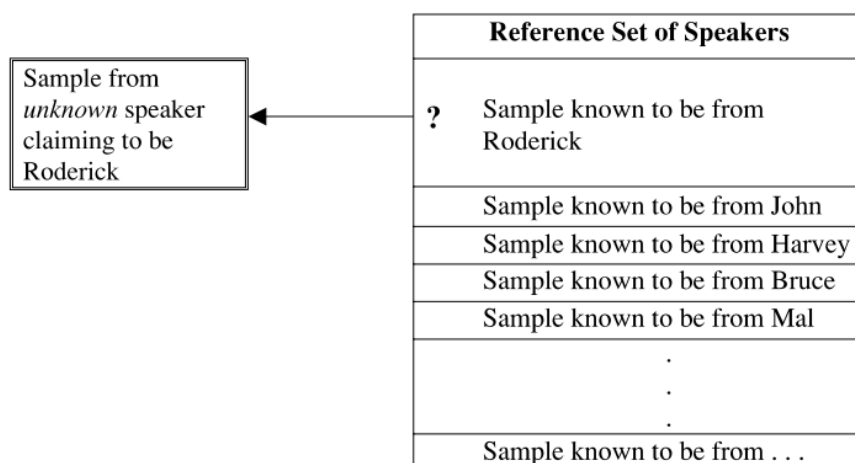
Nota. Interacción de las muestras indubitadas y muestra dubitada para la identificación de locutor, de Rose, 2002, p. 94

2.10. Verificación de locutor

Según Rose (2002), la verificación de locutor consiste en el reconocimiento de hablantes de acuerdo a la hipótesis de afirmación o negación del locutor. Por ejemplo, en llamadas telefónicas, queda almacenado el número y nombre a quien pertenece la llamada, por ello, Rose ejemplifica en la Figura 59, que la interrogante solo existiría para verificar si es o no compatible la voz de un locutor indubitado con la voz de una muestra dubitada.

Figura 59

Esquematzación del concepto de identificación de locutor



Nota. Interacción de las muestras indubitadas y muestra dubitada para la verificación de locutor, de Rose, 2002, p. 94

CAPÍTULO III

ANÁLISIS

Las muestras indubitadas y dubitadas previamente han sido segmentadas, etiquetadas y concatenadas para delimitar la participación de las voces identificadas, esto de acuerdo a las actas de transcripción otorgadas por la Fiscalía. En el Anexo 02: Segmentación de la participación de las voces de las muestras dubitadas se muestran las segmentaciones realizadas en el programa Praat, a partir de tablas, en las cuales se indica el tiempo de inicio y final de la segmentación.

3.1. Muestras indubitadas

En este apartado se describe la duración y las características de la calidad de la señal de las muestras indubitadas, según el software SIS II.

3.1.1. Análisis de la calidad de las muestras indubitadas

Al aplicar el *plugin* de análisis de la señal de la calidad de los audios con el programa SIS II, se obtuvieron los resultados que se muestran en la Tabla 11.

Tabla 11

Análisis de la calidad de señal de los audios de las muestras indubitadas

Referencia	SNR	Ruido tonal	Reverberación (ms)	Saturación	VAD	Duración	Resultado
LI_01	28	9.66	308	0.43	176.22	310.55	Aceptado
LI_02	29	23.22	299	0.09	152.61	289.85	Aceptado
LI_03	46	0	239	0	46.59	78.24	Aceptado
LI_04	46	0.19	281	0	10.27	16.52	No aceptado

3.2. Muestras dubitadas

En este apartado se describe la duración y las características de la calidad de la señal de las muestras dubitadas, según el software SIS II.

3.2.1. Análisis de la calidad de las muestras dubitadas

Al aplicar el plugin de análisis de la señal de la calidad de los audios con el programa SIS II, se obtuvo los resultados que se muestran en la Tabla 4, las capturas de pantalla del resultado individual se encuentran en el Anexo 03: Capturas de pantalla de los resultados de análisis de la calidad de la señal de las muestras indubitadas y dubitadas .

Tabla 12

Análisis de la calidad de señal de los audios de las muestras dubitadas

Referencia	SNR	Ruido tonal	Reverberación (ms)	Saturación	VAD	Duración	Resultado
MD_01	19	0	378	0	152.77	312.00	Parcialmente aceptado
MD_02	19	0.19	419	0	129.08	256.18	Parcialmente aceptado
MD_03	15	0.56	503	0	59.08	100.36	Parcialmente aceptado
MD_04	12	0	435	0	42.30	70.67	Parcialmente aceptado
MD_05	17	0	387	0	112.79	188.53	Parcialmente aceptado
MD_06	29	0	418	0	23.41	39.37	Parcialmente aceptado
MD_07	13	0	789	0	13.92	19.65	No aceptado
MD_08	16	0	0	0	2.01	2.42	No aceptado

Como se puede observar, seis (06) de las muestras son parcialmente apropiadas para el proceso de identificación de voz, sin embargo, dos de ellas no, dado que no satisfacen los criterios requeridos para ser tomados en cuenta en el análisis. Por ello, estos audios se excluyen de las conclusiones, eso no quiere decir que por esta razón esos audios no le puedan pertenecer a alguno de los investigados del caso.

3.3. Análisis de comparación automática a partir del software SIS II

Siguiendo la hipótesis fiscal, se procedió a comparar las voces de las muestras indubitadas con las muestras dubitadas, el pedido expreso de la fiscalía es el que se muestra en la Tabla 13.

Tabla 13*Pedido de homologación de voz según la fiscalía*

Muestras indubitadas	Muestras dubitadas
LI_01	MD_03, MD_04, MD_07 y MD_08
LI_02	MD_01, MD_02, MD_03, MD_04, MD_05, MD_06, MD_07 y MD_08
LI_03	MD_03, MD_04, MD_07 y MD_08

En ese sentido, se realizó la comparación automática de las voces de las muestras indubitadas con las muestras dubitadas, cabe resaltar que no se consideran las muestras dubitadas MD_07 y MD_08, debido a que no fueron óptimas según el análisis de la calidad de la señal. Los resultados de la comparación automática indican lo siguiente, ver Tabla 14.

Tabla 14*Resultados de la comparación automática con el software SIS II*

Referencia 1	Referencia 2	Método	FR	FA	LR	Resultado	Probabilidad
LI_01	MD_03	SF, PITCH, GMM	0.17	50.8	0.003	Diferente hablante	Alta
LI_01	MD_04	SF, PITCH, GMM	88.6	0.1	1023.313	Mismo hablante	Alta
LI_02	MD_01	SF, PITCH, GMM	0.16	65.4	0.0025	Diferente hablante	Alta
LI_02	MD_02	SF, PITCH, GMM	0.6	66.8	0.009	Diferente hablante	Alta
LI_02	MD_03	SF, PITCH, GMM	0.4	57.4	0.008	Diferente hablante	Alta
LI_02	MD_04	SF, PITCH, GMM	0.2	47.7	0.004	Diferente hablante	Alta
LI_02	MD_05	SF, PITCH, GMM	0.16	56.4	0.003	Diferente hablante	Alta
LI_02	MD_06	SF, PITCH, GMM	0.6	63.6	0.009	Diferente hablante	Alta
LI_03	MD_03	SF, PITCH, GMM	0.7	79.6	0.008	Diferente hablante	Alta

LI_03	MD_04	SF, PITCH, GMM	0.3	84.3	0.004	Diferente hablante	Alta
-------	-------	-------------------	-----	------	-------	-----------------------	------

A partir del método de comparación automática se puede concluir que la muestra indubitada LI_01 presenta compatibilidad con la muestra dubitada MD_04, mientras que las otras muestras indubitadas no son compatibles con alguna de las muestras dubitadas. Eso quiere decir que, respecto a este método, solo la voz de LI_01 si participa en la conversación en el universo de las muestras dubitadas.

3.4. Análisis fonético-acústico entre las muestras indubitadas y dubitadas

Las características fonético-acústicas descritas en el marco conceptual permite crear un perfil lingüístico de habla particular a cada locutor, a lo que nos estaríamos refiriendo como idiolecto. En este análisis, se procede a segmentar y etiquetar aquellas características particulares de habla de cada muestra de voz, a partir de la percepción auditiva y como referencia los segmentos con información acústica espectral, se transcribe de acuerdo al alfabeto fonético internacional (AFI, 1999) los fenómenos que se logren identificar (ver Anexo 05: Alfabeto Fonético Internacional). Para ello, es necesario el uso del software Praat (Paul y David, 2022), programa que se emplea para análisis lingüísticos.

El análisis de las características lingüísticas se divide en los siguientes acápite a partir de las muestras indubitadas de los imputados (LI_01, LI_02 y LI_03), solo se trabajó con seis (06) muestras dubitadas, porque fueron óptimos según el análisis de la señal.

3.4.1. Análisis de las características lingüísticas de la muestra indubitada LI_01

En primera instancia, se tiene la Tabla 15, que corresponde a la comparación de características fonético-acústicas entre las muestras LI_01 y las muestras dubitadas MD_03 y MD_04, los espacios vacíos indican la ausencia de aquellas características. De acuerdo a ello, solo se encontró compatibilidad entre cuatro (04) procesos lingüísticos de doce (12) entre la LI_01 y la MD_03, mientras que con la MD_04 existe coincidencia entre nueve (09) procesos de doce.

Tabla 15

Perfiles lingüísticos entre las muestras LI_01 con MD_03 y MD_04

Características Lingüísticas	Muestra Indubitada	Muestra Dubitada Voz 4	Muestra Dubitada Voz3
	LI_01	MD_03	MD_04
Aproximación de vibrante	Departamento [depa.ɽamento] 13.49-14.16 Martes [ma.ɽtes] 188.61-189.67 Viernes [bie.ɽnes] 191.97-192.99 Ratón [ɽa.ɽon] 265.20-265.78	Internas [inte.ɽnas] 36.75-37.05 Comprarte: [komp.ɽa.ɽte] 67.36-67.83	Revienten [ɽe.βien.ɽen] 20.13-20.54 Reventón [ɽe.βen.ɽon] 20.81-21.26 Arriba: [a.ɽi.βa] 46.04-46.29
Sonorización de oclusivas	Departamento [depa.ɽamento] 13.49-14.16 Cumplí: [ɽumpli] 39.23-39.53 Lo que pasa [lokeae] 50.29-50.70 Parte [pad] 57.47-57.65 Pues [βwe] 101.28-101.45	Compare [ɽompa] 51.45-51.68	Quédate [ɽea.ɽe] 15.61-15.87 Rato [ra.ɽo] 39.97-40.23
Alargamiento de vibrantes	Materno [ma.ɽer.no] 22.35-22.78 Recién: [re.ɽien] 38.22-38.60 Deporte [depor.te] 125.30-125.65 Sierra [siera] 137.52-137.98 Perro [per.o] 164.08-164.64 Rason [ra.ɽon] 239.41-239.91		Rato [ra.ɽo] 39.97-40.23 Arriba: [a.ɽi.βa] 42.25-42.52 Arregla: [a.ɽre.ɽla] 43.88-44.18

<p>Sinalefa</p>	<p>Mi papá : [miapa] 23.58-23.96</p> <p>Lo que pasa es que [lokeae] 50.29-50.70</p> <p>Toda mi familia [tomifamilia] 50.98-51.41</p> <p>mi mamá: [mjama]59.02-59.27</p>		<p>Ustedes se van a ir [t̥eseβair] 12.45-13.08</p>
<p>Elision de vocales <i>/i/ /o/ /a/</i></p>	<p>Mil novecientos sesenta y ocho [milnoβesientosesentaot̥] 26.45-27.70</p>		<p>Ahorita [oj̥ta] 8.03-8.33 / 10.56-10.83</p>
<p>Elisión de /s/</p>	<p>Nos venimos [noβenimos] 36.13-37.15</p> <p>Pues [βwe] 101.28-101.45</p>	<p>No más:[noma] 22.80 - 22.96</p>	<p>Dos días [doḷja] 2.12-2.36</p> <p>Nosotros [nosot̥o] 17.60-17.85</p> <p>Nosotros [nosotro] 26.62-26.99</p> <p>Vamos [βamos] 32.10-32.27</p> <p>No más [noma:] 61.87-62.27</p>
<p>Laringalización</p>	<p>e: [] 42.07-42.46</p>		<p>Vaina [βajna] 4.78-5.38</p>

Apócope	Toda: [to] 50.98-51.41 Parte: [pad] 57.47-57.66		Para probarte [paproβarte] 1.22-1.71 Para [pa] 14.07-14.19 Simplemente: [simple] 53.82-54.12 Nada: [na] 63.30-63.51
Aféresis	Estaba [taβa] 55.77-56.00 Está: [ta] 88.44-88.58	Estamos: [tamos] 0.59-086 Está: [ta] 31.75-31.84	Estamos: [tamos] 17.87-18.05
Monoptongación		Pues: [pes] 0.35-0.62	
Alargamiento de vocal		Dale [da:le] 39.75-40.05 / 48.78-49.13 Vas [ba:s] 55.53-55.84 Hablo [a:blo] 84.95-85.26	
Marcador discursivo		Ya pe [jape] 69.01 - 69.28	

3.4.2. Análisis de las características lingüísticas de la muestra indubitada LI_02

En la Tabla 16, se observan las características lingüísticas entre la muestra indubitada LI_02 y las muestras dubitadas de la MD_01 a la MD_06. No se encontraron rasgos coincidentes entre ninguna muestra.

Tabla 16

Perfiles lingüísticos entre las muestras LI_02 con MD_01 a MD_06

Características Lingüísticas	Muestra Indubitada	Muestra dubitada voz 2	Muestra dubitada voz 1	Muestra dubitada voz 4	Muestra dubitada voz 3	Muestra dubitada voz 2	Muestra dubitada voz 1
	LI_02	MD_01	MD_02	MD_03	MD_04	MD_05	MD_06
Producción de la vibrante simple como dos componentes	93.42 – 93.76 porque [por.ɣe]						
Elisión de la lateral	92.82 – 93.38 la catedral [lakatedra]						
Ensondecimiento vocálico	56.67 – 57.23 Arequipa [arekip]						
Producción estridente de la fricativa velar	100.08 – 100.08 lejos [lexos] 145.72 – 146.22 ejemplo [exemplo]						

<p>Aproximación de vibrante</p>				<p>Internas [inteɾnas] 36.75-37.05</p> <p>Comprate: [kompɾate] 67.36-67.83</p>	<p>Departamento [depaɾtamento] 13.49-14.16</p> <p>Martes [maɾtes] 188.61-189.67</p> <p>Viernes [bieɾnes] 191.97-192.99</p> <p>Ratón [ɾaɾon] 265.20-265.78</p>		
<p>Sonorización de oclusivas</p>			<p>Cuatro [ɾwaɾo] 54.04 - 54.23</p> <p>Lucas [luɾas] 76.32 - 76.62</p> <p>Que [ɾe] 81.29 - 81.38</p>	<p>Compare [ɾompa] 51.45-51.68</p>	<p>Departamento [depaɾtamento] 13.49-14.16</p> <p>Cumplí: [ɾumpli] 39.23-39.53</p> <p>Lo que pasa [lokeae] 50.29-50.70</p> <p>Parte [paɾ] 57.47-57.65</p> <p>Pues [βwe] 101.28-101.45</p>	<p>5.63 – 6.25</p> <p>me ha dicho que [miɾfoɾe]</p>	
<p>Alargamiento de vibrantes</p>					<p>Materno [maɾe:ɾno] 22.35-22.78</p> <p>Recién: [ɾiesien] 38.22-38.60</p> <p>Deporte [depo:ɾte] 125.30-125.65</p> <p>Sierra [siera] 137.52-137.98</p> <p>Perro [pe:ɾo] 164.08-164.64</p> <p>Rason [ɾa:son] 239.41-239.91</p>		

Sinalefa					<p>Mi papá : [miapa] 23.58-23.96</p> <p>Lo que pasa es que [lokeae] 50.29-50.70</p> <p>Toda mi familia [tomifamilia] 50.98- 51.41</p> <p>mi mamá: [mjama]59.02-59.27</p>		
Elision de vocales /i/ /o/ /a/					<p>Mil novecientos sesenta y ocho [milnoβesient̃osesen ʔaot̃] 26.45-27.70</p>		
Elisión de /s/			<p>Fiscal [f̃ika] 4.75 - 5.19</p> <p>Has visto [aβĩto] 27.03 - 27.29</p> <p>Cosas [kosa] 59.05 - 59.31</p>	<p>No más:[no ma] 22.80 - 22.96</p>	<p>Nos venimos [noβenimos] 36.13- 37.15</p> <p>Pues [βwe] 101.28- 101.45</p>	<p>27.46 – 27.94</p> <p>diez días [ðjeʰðja]</p>	
Laringalización					<p>e: [] 42.07-42.46</p>		

<p>Apócope</p>			<p>Arranca do [aranka] 1.60 - 1.97</p> <p>Toda la [tola] 3.45 - 3.61</p> <p>Sacado [saka] 28.99 - 29.25</p> <p>Revisad o [reβisa] 30.08 - 30.45</p> <p>Para [pa] 71.35 - 71.48</p>		<p>Toda:[to] 50.98- 51.41</p> <p>Parte: [pad] 57.47- 57.66</p>	<p>1.96 – 2.31</p> <p>compad re [kompa h]</p>	
<p>Aféresis</p>			<p>Está [ta] 28.33 - 28.49</p> <p>Estabas [ta] 60.55 - 60.70</p>	<p>Estamos: [tamos] 0.59-086</p> <p>Está: [ta] 31.75- 31.84</p>	<p>Estaba [taβa] 55.77- 56.00</p> <p>Está: [ta] 88.44- 88.58</p>	<p>6.72 – 6.72</p> <p>ya están [jatan]</p>	
<p>Monoptongac ión</p>		<p>14.60 – 14.86</p> <p>quiero [keðo]</p>		<p>Pues: [pes] 0.35-0.62</p>			

Alargamiento de vocal				Dale [da:le] 39.75- 40.05 / 48.78- 49.13 Vas [ba:s] 55.53- 55.84 Hablo [a:blo] 84.95- 85.26			
Marcador discursivo				Ya pe [jape] 69.01 - 69.28			
Elisión de /d/		2.59 – 2.98 malograd o [malorao]	Quedan [kean] 22.18 - 22.41 Huevad as [weβas] 34.51 - 34.96			24.63 – 25.09 vas a dejar [βasaex a]	3.69 – 4.14 acercado [asekaw]
Producción de la /s/ en bajas frecuencias del espectro							

<p>Supresión consonántica</p>		<p>2.59 – 2.98</p> <p>malograd o [malorao]</p> <p>5.47 – 5.60</p> <p>creo [ʎeo]</p> <p>20.01 – 20.74</p> <p>yo ofrezco [jo:fesko]</p>				<p>2.77 – 2.77</p> <p>no entra [noenta]</p> <p>26.13 – 26.13</p> <p>tranquilo [taŋkilo]</p>
<p>Elisión de la yod</p>		<p>3.71 – 4.54</p> <p>¿Cómo se llama?- komosea ma</p>				<p>19.15 – 19.15</p> <p>diciendo [disendo]</p>
<p>Rotacismo por sustitución</p>		<p>14.60 – 14.86</p> <p>quiero [keʝo]</p>				<p>17.28 – 17.95</p> <p>ya estaríamos s [jastaʝjamos]</p>

<p>Elisión de la /r/</p>			<p>Ubicar [uβjka] 20.76- 21.02</p> <p>Cuatro [waɾo] 54.04 - 54.23</p> <p>Negocia r [neɾoʒe] 75.29 - 75.67</p>			<p>24.63 – 25.09</p> <p>vas a dejar [βasaex a</p>	
<p>Síncopa</p>			<p>Ahorita [ojta] 19.85 - 20.10</p> <p>Ahorita [ajta] 45.31 - 45.50</p>			<p>7.84 – 8.15</p> <p>subido [sujo]</p>	<p>1.24 – 1.90</p> <p>subido [sujo]</p>

3.4.3. Análisis de las características lingüísticas de la muestra indubitada LI_03

Por último, se observan en la Tabla 17, las características lingüísticas de la LI_03 con las muestras dubitadas MD_03 y MD_04. Entre la LI_03 y la MD_03 solo se encontró la compatibilidad entre cuatro procesos de trece, mientras que la LI_03 y la MD_04 solo de dos.

Tabla 17

Perfiles lingüísticos entre las muestras LI_03 con MD_03 y MD_04

Características Lingüísticas	Muestra Indubitada	Muestra Dubitada Voz 4	Muestra Dubitada Voz 3
	LI_03	MD_03	MD_04
Elisión de /d/	Suspendido [suspɛ̃ɲio] 25.64 - 26.15 Cercado: [serkao] 54.94- 55.42 Abogado: [aβoɣao] 59.54-60.04		
Aproximación de vibrante	Conocer [konoseɻ] 26.47-20.93 Poder [poðeɻ] 71.76 - 72.40	Internas [in̄ɛ̃nas] 36.75-37.05 Comprarte: [kompɻaɽe] 67.36-67.83	Revienten [ɻeβɛ̃ɲɛ̃n] 20.13- 20.54 Reventón [ɻeβɛ̃ɲɔ̃n] 20.81- 21.26 Arriba: [aɻiβa] 46.04-46.29
Sonorización de oclusivas		Compare [ɣɔmpa] 51.45-51.68	Quédate [ɻeaɽe] 15.61-15.87 Rato [raðo] 39.97-40.23
Alargamiento de vibrantes			Rato [r:aðo] 39.97-40.23 Arriba: [ar:iβa] 42.25-42.52 Arregla: [ar:egla] 43.88- 44.18
Sinalefa			Ustedes se van a ir [ɽeseβair] 12.45-13.08
Elisión de vocales /i/ /o/ /a/			Ahorita [oɽta] 8.03-8.33 / 10.56-10.83

Elisión de /s/		No más:[noma] 22.80 - 22.96	Dos días [doðja] 2.12-2.36 Nosotros [nosoɽo] 17.60-17.85 Nosotros [nosotro] 26.62-26.99 Vamos [βamos] 32.10-32.27 No más [noma:] 61.87-62.27
Laringalización			Vaina [βajna] 4.78-5.38
Apócope			Para probarte [paproβarɽe] 1.22-1.71 Para [pa] 14.07-14.19 Simplemente: [simple] 53.82-54.12 Nada: [na] 63.30-63.51
Aféresis	Todavía: [toβja] 58.28-58.62	Estamos: [ɽamos] 0.59-0.86 Está: [ɽa] 31.75-31.84	Estamos:[ɽamos] 17.87-18.05
Monoptongación	Diecinueve [disinueβe] 52.44-52.83	Pues: [pes] 0.35-0.62	
Alargamiento de vocal	Número [n mero] 5.59-6.06 O sea [osea] 51.66-52.16 Poder: [poð r] 71.76-72.40	Dale [da:le] 39.75-40.05 / 48.78-49.13 Vas [ba:s] 55.53-55.84 Hablo [a:blo] 84.95-85.26	
Marcador discursivo		Ya pe [jape] 69.01 - 69.28	

3.5. Uso del software SplitsTree 4

Durante el análisis de las características lingüísticas de los sujetos investigados junto con las muestras dubitadas, se identificaron rasgos que eran consistentes con una

muestra indubitada, mientras que eran diferentes en relación a las otras muestras. Sin embargo, este hallazgo no proporcionaba una confirmación definitiva de la presencia o ausencia de los rasgos encontrados. Por lo tanto, se utilizó el software SplitsTree 4 para ordenar y organizar las características fonético-fonológicas y lingüísticas identificadas.

Se construyó una matriz que contenía todos los rasgos encontrados, y se asignó una codificación binaria para indicar la presencia o ausencia de cada rasgo en cada hablante. Esto se hizo con el objetivo de evitar interpretaciones erróneas durante el proceso judicial y presentar una evidencia clara y comprensible de la interpretación de los datos ante el tribunal. (Jiménez *et al.*, 2022).

Con el fin de lograrlo, se asignaron valores numéricos binarios de "0" y "1" para indicar la ausencia y presencia más frecuente del rasgo específico en cada muestra, respectivamente. Por lo tanto, se muestra en la Tabla 18 una matriz hipotética de rasgos binarios, donde cada rasgo fonético-articulatorio está identificado con la referencia "R" seguida de su número correspondiente, que puede ser "0" o "1".

Tabla 18

Matriz hipotética de rasgos fonético-acústicos de las muestras

	LI_01	LI_02	LI_03	MD_01	MD_02	MD_03	MD_04	MD_06	MD_07
R1	0	1	1	1	0	1	1	1	0
R2	1	1	0	1	1	1	0	1	1
R3	0	1	0	1	0	0	1	1	0
R4	1	1	1	1	1	1	1	1	1
Rn...	0	1	0	0	0	1	0	0	0

Por último, esta matriz debe ser incluida dentro del software SplitsTree4 como parte del *enter data*, para poder generar con los algoritmos del programa construcciones diferentes de redes según la estructuración de los rasgos y muestras.

3.5.1. Análisis de los rasgos fonético-acústicos con el software SplitsTree 4

En esta sección, se presenta información detallada sobre la tabulación de los rasgos fonético-acústicos utilizando una codificación binaria. Se describe la creación de una matriz de rasgos para cada voz en las muestras indubitadas y dubitadas, así como la codificación de las categorías taxonómicas de cada voz en función de los rasgos

fonético-acústicos identificados. (Zariquiey, R., Vásquez, A., & Tello, G., 2017; Jimenez Peña, J., Torres Castillo, F. A., & Cueva Sanchez, O. E., 2022)

3.5.1.1. Tabulación de la matriz de rasgos fonético-acústicos

Se revisó e identificó cada rasgo fonético interviniente en las muestras (indubitadas y dubitadas) para generar la matriz de rasgos (ver Tabla 19), y a estos se les agregó la codificación binaria de “0” y “1”, según corresponda.

Tabla 19

Matriz de rasgos fonético-acústicos de las muestras indubitadas y dubitadas

	LI_01	LI_02	LI_03	MD_01	MD_02	MD_03	MD_04	MD_05	MD_06
Producción de la vibrante simple como dos componentes	0	1	0	0	0	0	0	0	0
Elisión de la lateral	0	1	0	0	0	0	0	0	0
Ensondecimiento vocálico	0	1	0	0	0	0	0	0	0
Producción estridente de la fricativa velar	0	1	0	0	0	0	0	0	0
Aproximación de vibrante	1	0	1	0	0	1	1	0	0
Sonorización de oclusivas	1	0	0	0	1	1	1	1	0
Alargamiento de vibrantes	1	0	0	0	0	0	1	0	0
Sinalefa	1	0	0	0	0	0	1	0	0
Elisión de vocales /i/ /o/ /a/	1	0	0	0	0	0	1	0	0
Elisión de /s/	1	0	0	0	1	1	1	1	0
Laringalización	1	0	0	0	0	0	1	0	0
Apócope	1	0	0	0	1	0	1	1	0
Aféresis	1	0	1	0	1	1	1	1	0
Monoptongación	0	0	1	1	0	1	0	0	0
Alargamiento de vocal	0	0	1	0	0	1	0	0	0
Marcador discursivo	0	0	1	0	0	1	0	0	0
Elisión de /d/	0	0	1	1	1	0	0	1	1
Producción de la /s/ en bajas frecuencias del espectro	0	0	0	0	0	0	0	0	0
Supresión consonántica	0	0	0	1	0	0	0	0	1
Elisión de la yod	0	0	0	1	0	0	0	0	1
Rotacismo por sustitución	0	0	0	1	0	0	0	0	1
Elisión de la r	0	0	0	0	1	0	0	1	0
Síncopa	0	0	0	0	1	0	0	1	1

Esta lista de codificación binaria, se consigna en el software SplitsTree 4, el llenado de datos consiste en una lista de asignaciones numéricas para su interpretación y esquematización. Para ello, se utilizó un script de *simple data* (Ver Anexo 06: Script para generar redes filogenéticas con *Simple Data*), en la cual se ingresó el "código genético" de cada muestra taxonómica (voz) en la categoría "Enter Data Dialog". Los comandos correspondientes al formato nexus fueron codificados en esta sección, mientras que el comando "Matrix" se refiere a la matriz de rasgos previamente mencionados



CAPÍTULO IV

RESULTADOS

En este apartado se muestra y explica el gráfico elaborado por el *software* SplitsTree 4 a partir de los datos introducidos en el programa, para este caso se generó un esquema con el algoritmo *Neighbor-net*, este algoritmo fue desarrollado por los científicos Daniel Bryant y Vincent Moulton y fue presentado por primera vez en un artículo de nombre "Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks" en el 2002. Desde entonces, Neighbor-net ha sido ampliamente utilizado en el campo de la filogenética para construir redes de fragmentación y analizar relaciones evolutivas entre especies o lenguas. (Dan & Lior, 2011)

El algoritmo Neighbor-net y la lingüística están relacionadas en el campo de la filogenética lingüística, que es el estudio de la relación evolutiva entre las lenguas. Este algoritmo es utilizado en la construcción de árboles filogenéticos a partir de datos lingüísticos, como las similitudes léxicas o las características gramaticales de diferentes lenguas. En la filogenética lingüística, se utilizan métodos computacionales para analizar y modelar las similitudes y diferencias entre las lenguas, con el objetivo de reconstruir su historia evolutiva y establecer relaciones de parentesco entre ellas. Es una herramienta útil en este sentido, ya que permite visualizar y representar gráficamente las relaciones entre las lenguas en forma de una red. Este algoritmo puede ayudar a identificar agrupaciones lingüísticas, determinar el grado de parentesco entre lenguas, y analizar la evolución y dispersión de características lingüísticas a lo largo del tiempo. Al representar las relaciones entre las lenguas como una red, es posible examinar patrones de similitud y divergencia, así como identificar posibles influencias o contactos entre diferentes grupos de lenguas. En resumen, el algoritmo Neighbor-net es una herramienta valiosa en la investigación filogenética lingüística, ya que contribuye a comprender las relaciones entre las lenguas y proporciona una base para el estudio de la evolución y diversidad lingüística. El algoritmo Neighbor-net también puede utilizarse para analizar y comparar rasgos lingüísticos entre diferentes variedades de una lengua, incluyendo los idiolectos. Como se describió al principio de la investigación, un idiolecto se refiere al conjunto único de rasgos lingüísticos que caracterizan a un hablante individual. Cada persona

tiene su propio idiolecto, que se desarrolla a lo largo de su vida y puede estar influenciado por diversos factores, como el entorno social, la educación, la experiencia personal y los contactos lingüísticos.

Por ello, cuando se aplican otras técnicas metodológicas al análisis lingüístico, como el algoritmo Neighbor-net, a los idiolectos, se pueden identificar similitudes y diferencias entre ellos. Esto permite examinar la variación individual en el uso de rasgos lingüísticos y comprender cómo se desarrollan y se distribuyen estos rasgos dentro de una comunidad lingüística. El algoritmo Neighbor-net puede ayudar a visualizar las relaciones entre los idiolectos y determinar patrones de similitud o agrupaciones basadas en los rasgos lingüísticos analizados. En un esquema Neighbor-net, las distancias se refieren a las medidas de disimilitud o diferencias entre las variedades o elementos que se están analizando. Estas distancias se utilizan para construir una red de fragmentación que representa las relaciones entre las diferentes variedades o elementos en términos de similitud o divergencia. En el contexto del algoritmo Neighbor-net, se pueden utilizar diferentes métodos para calcular las distancias entre los elementos. Estos métodos pueden variar dependiendo del tipo de datos que se esté analizando. Por ejemplo, en el análisis de rasgos lingüísticos, las distancias pueden calcularse utilizando medidas de similitud léxica, como la comparación de vocabulario compartido o el cálculo de la distancia léxica entre palabras. También pueden utilizarse medidas de similitud fonética, morfológica o sintáctica, dependiendo de los aspectos lingüísticos que se estén estudiando. Para este caso, nos estamos centrando en la similitud fonético-fonológica.

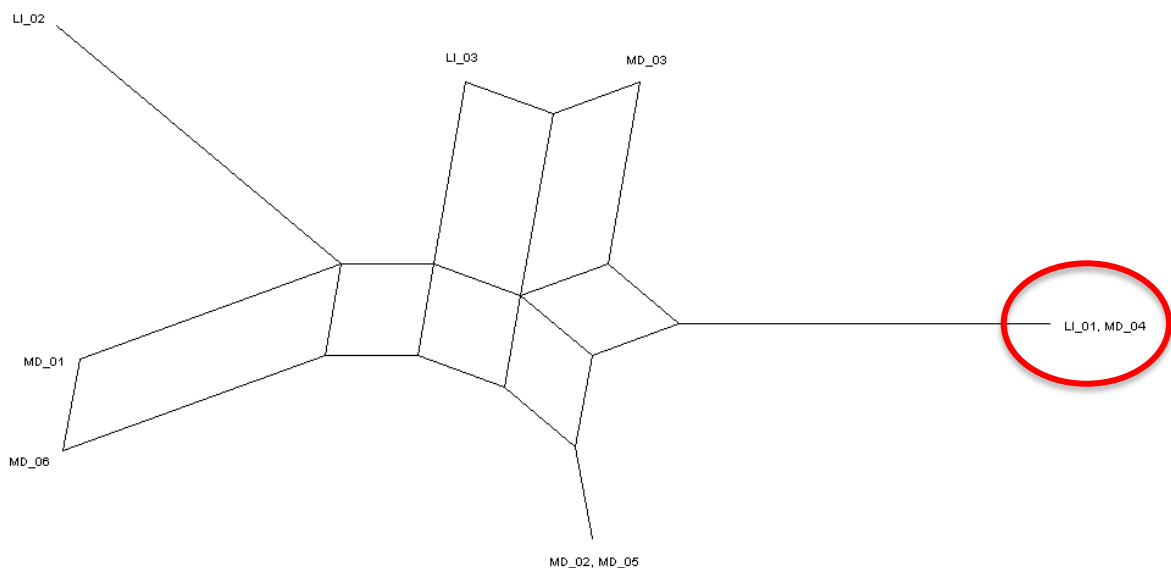
Una vez que se han calculado las distancias entre los elementos o locutores, el algoritmo Neighbor-net utiliza estas distancias para construir una red de fragmentación. Esta red se basa en la idea de que las variedades o elementos que están más cerca (es decir, tienen distancias más pequeñas) se agruparán en clústeres o ramas más cercanas en la red, mientras que aquellos con distancias más grandes se separarán en ramas más separadas. La red de fragmentación generada por el algoritmo Neighbor-net proporciona una representación gráfica de las relaciones entre las variedades o elementos, en este caso, locutores, donde las distancias reflejan la similitud o disimilitud entre ellos. Esta visualización permite analizar y explorar patrones de agrupamiento, estructuras de ramificación y relaciones filogenéticas entre los locutores analizados.

4.1. Resultados del método de ordenamiento con el software SplitsTree 4

Para el análisis de ordenamiento de los rasgos fonético-acústicos se trabajó con la matriz de rasgos de las muestras indubitadas y dubitadas. Como resultado se generó un esquema de red filogenética a partir del algoritmo *Neighbor-net*, tal como se muestra en la Figura 60.

Figura 60

Esquema Neighbor-net



En un esquema Neighbor-net, como el que se observa en la Figura 46, los cuadrados dentro de la representación gráfica tienen un significado especial y se conocen como "cuadrados de cuarteto" o "cuadrados de *split*". Estos cuadrados dentro de la red de fragmentación en un esquema Neighbor-net representan las posibles subdivisiones o agrupaciones de las variedades o elementos que se están analizando. Cada cuadrado indica una posible división o separación entre los elementos que están conectados por líneas en la red.

El tamaño y la posición de los cuadrados de cuarteto en la red indican la fuerza y la importancia de la subdivisión correspondiente. Cuanto más grande sea el cuadrado, más relevante es la división en términos de las diferencias entre los elementos. Además, la posición de los cuadrados dentro de la red puede proporcionar información sobre las relaciones y las ramificaciones entre las subdivisiones.

Los cuadrados de cuarteto son una característica distintiva del esquema Neighbor-net y ayudan a comprender la estructura y la agrupación de las variedades o elementos analizados. Eso quiere decir que las muestras importadas al *software* muestran un comportamiento íntegro.

Mediante la red filogenética se observa una organización más detallada y ordenada de los rasgos fonético-acústicos, se detalla a qué distancias se encuentran cada muestra (indubitada y dubitada); para explicar el resultado, partiremos de la hipótesis fiscal (ver Tabla 13), la cual separaremos por muestra indubitada (LI_01, LI_02 y LI_03).

Para el caso de la muestra LI_01, se observa que tiene menor distancia con la muestra MD_04, mientras que con la MD_03, existe un mayor alejamiento, lo cual indicaría que hay **mayor compatibilidad entre las muestras LI_01 y MD_04**. De la muestra LI_02, se observa que es el más distante de todas las muestras, eso quiere decir que no guarda compatibilidad a nivel fonético-acústico con las muestras analizadas. Por último, la LI_03 no aparece cercana a la muestra MD_03 ni MD_04, lo cual indicaría que no hay compatibilidad entre esas muestras.

En un esquema Neighbor-net, si dos muestras comparten un mismo nodo (LI_01 y MD_04), esto indica que comparten una similitud o una relación más cercana entre ellas en comparación con las otras muestras presentes en la red de fragmentación. Cuando dos muestras comparten un nodo en la representación gráfica, significa que la distancia o disimilitud entre ellas es menor en comparación con las distancias entre otras muestras en la red. Esto sugiere que comparten características lingüísticas o rasgos similares que las agrupan juntas en términos de similitud lingüística.

A partir del análisis con los métodos manual y de comparación automática de las muestras indubitadas y dubitadas, se puede concluir que solo la muestra LI_01 es compatible con una de las voces que se entregó como muestra dubitada (MD_04) por parte de la Fiscalía.

4.2. Balance sobre el método filogenético

El uso del programa SplitsTree 4 presenta ventajas y desventajas que se deben tener en cuenta al momento de utilizarlos en estudios de lingüística, y específicamente, en este estudio que se relaciona con la lingüística forense.

Dentro de sus ventajas, se tiene lo siguiente:

1. Flexibilidad en la representación de datos: Splitstree 4 permite representar y analizar datos que contienen señales contradictorias o que no siguen un modelo de árbol filogenético simple. Esto es especialmente útil en lingüística forense, donde las relaciones entre muestras lingüísticas pueden ser más complejas y no se ajustan a un árbol filogenético tradicional.
2. Visualización de redes filogenéticas: El programa proporciona herramientas para construir y visualizar redes filogenéticas, lo que puede ayudar a comprender y explorar las relaciones entre diferentes muestras de voces. Las redes filogenéticas, como el Neighbor-net, pueden representar de manera más precisa las conexiones reticuladas o híbridas entre diferentes ramas, lo que es especialmente relevante en contextos de compatibilidad o divergencia entre muestras.
3. Interfaz gráfica amigable: Splitstree 4 cuenta con una interfaz gráfica intuitiva y fácil de usar, lo que facilita su uso por parte de los investigadores sin experiencia en programación. Esto hace que el programa sea accesible y práctico para aquellos que deseen realizar análisis filogenéticos.

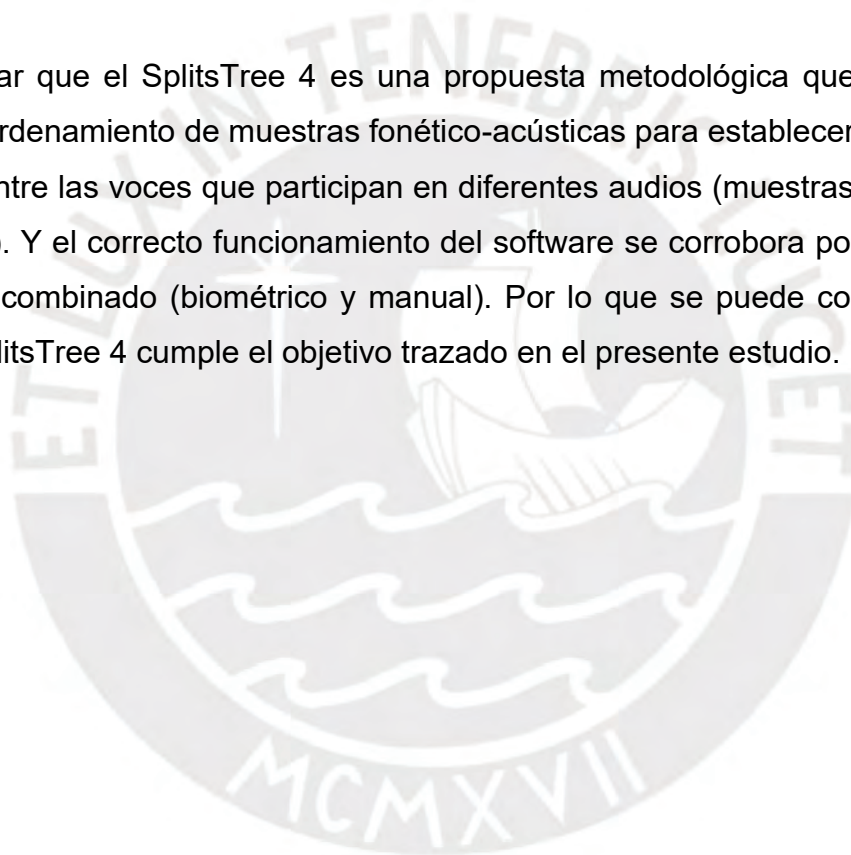
Dentro de las desventajas, se tiene que:

1. Requiere conocimientos técnicos: Aunque Splitstree 4 tiene una interfaz gráfica amigable, el análisis filogenético en sí mismo puede ser complejo y requiere conocimientos técnicos y comprensión de los conceptos subyacentes. Es importante tener un buen entendimiento de los principios filogenéticos y lingüísticos antes de utilizar el programa para obtener resultados precisos y confiables.
2. Sensible a la calidad de los datos: Como cualquier herramienta filogenética, Splitstree 4 depende de la calidad de los datos de entrada. Los resultados

pueden verse afectados por errores en la recolección o codificación de los datos lingüísticos, así como por la falta de datos suficientes para un análisis exhaustivo.

3. Interpretación subjetiva: La interpretación de las redes filogenéticas generadas por SplitsTree 4 puede ser subjetiva y requiere juicio por parte del investigador. La representación visual de las relaciones filogenéticas puede variar según los parámetros utilizados y las decisiones de análisis tomadas, lo que puede llevar a diferentes interpretaciones de los resultados.

Cabe resaltar que el SplitsTree 4 es una propuesta metodológica que tiene como objetivo el ordenamiento de muestras fonético-acústicas para establecer distancias y cercanías entre las voces que participan en diferentes audios (muestras dubitadas e indubitadas). Y el correcto funcionamiento del software se corrobora por el respaldo del método combinado (biométrico y manual). Por lo que se puede concluir que el software SplitsTree 4 cumple el objetivo trazado en el presente estudio.



CONCLUSIONES

La Fiscalía envía la solicitud de pericias de comparación de voz de los investigados implicados en el caso, al haber una tarea que demandaría confusión, como tener un informe pericial por cada investigado, que posiblemente al ser policías, y todos vivan en el mismo departamento, tengan rasgos coincidentes, se optó por la idea de manejar los datos fonético-acústicos desde una perspectiva más unitaria, en el sentido de hacer una sola pericia para identificar a qué voz o voces de la muestra dubitada le pertenece a uno o más investigados. Esta solución conllevaría a incluir un programa que permita realizar este distanciamiento y acercamiento entre muestras, y como propuesta, el programa SplitsTree 4 fue el indicado para cumplir tal función: ordenar los rasgos fonético-acústicos encontrados en las muestras con la finalidad de confirmar la compatibilidad mediante la codificación binaria de rasgos. La propuesta cumple su función a cabalidad, permitiendo al lector (abogados, fiscales o jueces) entender la identificación de voces desde una perspectiva gráfica. Es así que esta propuesta se sustenta y corrobora con los análisis complementarios (análisis biométrico y análisis lingüístico).

Respecto al análisis biométrico, es factible considerar los resultados arrojados por el programa de comparación automática y equiparlos con los del análisis lingüístico, con un porcentaje de 50/50 para la valoración del informe pericial. Debido a que en el Perú no se cuenta con una base de datos de voces recopiladas a nivel nacional, el software SIS II resulta útil para complementar el análisis lingüístico. Por tanto, entidades estatales como la RENIEC, universidades nacionales o entidades privadas que estén relacionadas al estudio de la voz deberían proyectarse a crear una base de datos de voces a nivel nacional.

Respecto al análisis lingüístico, se debe aclarar que, para considerar un rasgo en la tabulación de rasgos fonético-fonológicos, morfológicos, sintácticos, suprasegmentales y extralingüísticos particulares del habla de una persona se debe tener en cuenta la frecuencia en la que aparece en un audio, eso quiere decir que para considerar un rasgo como evidencia de compatibilidad lingüística es necesario contar con más de una aparición del rasgo en el audio, en últimas instancias debe ser representativo.

En cuanto a la aplicación del software SplitsTree 4, se puede aseverar que cumple con su función de ordenamiento de muestras para la identificación forense de voz, tal

y como señalan Jiménez *et al.* (2022). A pesar de ser un estudio diferente, en la cual hubo más de una muestra dubitada (MD), se logró identificar la participación y no participación de los investigados que estaban implicados en el caso, según lo solicitado por la Fiscalía.

Por otro lado, el concepto de “identificación de locutor” descrito por Rose (2002) se ampliaría a este caso, debido a que hay más de una voz dubitada y se comparan con todas las muestras indubitadas. El programa SplitsTree 4 permite organizar las cadenas taxonómicas de rasgos fonético-fonológicos de cada voz, en consecuencia, delimita las distancias entre las muestras, permitiendo observar si existe compatibilidad o no entre las *taxas* incluidas en el programa.

Se evidencia que el software SplitsTree 4 es una herramienta relevante para abordar casos de identificación forense en los cuales sea necesario ordenar rasgos que no se limiten únicamente a aspectos fonéticos. Esta herramienta resulta útil para esclarecer casos delictivos que involucren a más de una persona investigada. Asimismo, se debe aclarar que la interpretación de los códigos binarios de “0” y “1” no pueden describirse en contradicciones de existencia y no existencia, porque puede generar confusiones respecto a la existencia de un solo ejemplo de un rasgo, en este caso, fonético-fonológico, logre incluirse en la tabulación de rasgos para el programa SplitsTree4; por ende se recomienda verificar la frecuencia de rasgos y observar si el comportamiento es mayor o menor, mientras más presencia de ejemplos del rasgo, más porcentaje de consideración tendrá al incluirla en la tabla matriz del software filogenético.

RECOMENDACIONES

Una de las dificultades que se presentaron en el Área de Fonética y Acústica Forense para casos en las que se deba identificar la participación de la voz o voces de un locutor o locutores en un delito en el que hay varias personas implicadas es que no haya existido hasta el momento un método que permita organizar claramente en un solo informe pericial la identificación de locutores. Por ende, el *software* SplitsTree 4 cumplió a cabalidad el objetivo de ordenar los rasgos lingüísticos para identificar la participación y no participación de algunos investigados en un caso de identificación de locutor. Eso no quiere decir que se pueda utilizar otros programas que puedan cumplir el mismo objetivo, sin embargo, se recomienda a quienes estén interesados en el ámbito de la lingüística forense (específicamente en la identificación de locutores), en primer lugar, seguir validando el uso del método de la red filogenética para futuras investigaciones, en segundo lugar, proponer y comparar con otros métodos que cumplan el mismo objetivo de ordenamiento de datos, y por último, que la esquematización sea práctica y concisa al momento de presentar ante el Juez, dejando de lado la ambigüedad de los resultados.

Se recomienda a las entidades nacionales y privadas relacionadas al estudio de la voz o la lingüística forense en crear un banco de datos de voces en diferentes lenguas, sea español u otra lengua indígena, esto con el fin de obtener un registro de comparación que ayude a la creación de programas biométricos de voz y tener una fuente de referencia. Puede que surja la pregunta de ¿por qué tener un banco de voces que no solo sea en español?, pues, el simple hecho de que los delitos también se cometen en otras lenguas.

REFERENCIAS BIBLIOGRÁFICAS

Aguirrezabala, M. (2015) *Estudio de verificación biométrica de voz*. [Tesis de Maestría]. Universidad Politécnica de Madrid, España.

América Noticias (2018, 13-08) “El mil voces”: conoce al extorsionador que se hacía pasar como viceministro del Interior. América Noticias. <https://www.americatv.com.pe/noticias/actualidad/mil-voces-conoce-al-extorsionador-que-se-hacia-pasar-como-viceministro-interior-n334147>

Andrade, L. (2023) El español en el Perú. Publicado en *Dialectología Hispánica*, de Francisco Moreno Fernández y Rocío Caravedo. *Routledge*. 1ra. Ed. p. 333 – 343.

Asociación Fonética Internacional (1999) *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge, Cambridge University Press.

Caravedo, R. (1983) *Estudios sobre el español de Lima. Variación contextual de la sibilante*. Lima: Fondo Editorial PUCP.

Caravedo, R. (1992) Espacio geográfico y modalidades lingüísticas en el español del Perú. En *Historia y presente del español de América*, ed. César Hernández Alonso, p. 719 – 741. Valladolid: Junta de Castilla y León.

Cerrón-Palomino, R. (2020) La presencia puquina en el aimara y en el quechua: aspectos léxicos y gramaticales. *Indiana*, vol 37, p. 129 – 153.

Celis, P. (2009) *Descripción de los métodos utilizados en reconocimiento forense de locutores y su implementación en Chile*. [Tesis para optar el título de ingeniero civil acústico]. Universidad Austral de Chile. Chile. <http://cybertesis.uach.cl/tesis/uach/2009/bmfcic392d/doc/bmfcic392d.pdf>

Chomsky, N. y Halle, M. (1968) *The sound pattern of english*. Harper & Row. https://edisciplinas.usp.br/pluginfile.php/5654485/mod_folder/content/0/Chomsky%20C%20Halle%201968.%20The%20Sound%20Pattern%20of%20English.pdf?forcedownload=1

Ciccía, C., Montezuma, O., Elías, J., Gabillo, G., Echenique, M., Raffo, G., y Seminario, A. (2009) La adquisición de la estructura silábica: El caso del castellano limeño. *Persona*, 12, p. 11 – 28.

Cicres, J. (2007) *Aplicació de l'anàlisi de l'entonació i de l'alineació tonal a la identificació de parlants en fonètica forense*. Tesis doctoral, Universidad Pompeu Fabra. Extraído de <https://www.tdx.cat/bitstream/handle/10803/7504/tjc.pdf.pdf?sequence=1&isAllowed=1>

Cicres, J. (2011) Los sonidos fricativos sordos y sus implicaciones forenses. *Estudios Filológicos*. 48, p. 33 – 48.

Cicres, J. (2011) Transcripció i autenticació de gravacions en contextos judicials. Llengua, Societat i Comunicació. Revista de Sociolingüística de la Universitat de Barcelona 9, p. 23-62. http://www.ub.edu/cusc/revista/lsc/hemeroteca/numero9/articulos/4_JCicres_ling-forense_LSCcorrDEF.pdf

Colls, L. (2016) *Identificación del hablante a partir del análisis acústico de una muestra de voz en condiciones de cansancio*. Facultat de lletres, Universitat de Girona, España.

Cueva, O. (2018) Análisis fonético-fonológico de los procesos que afectan a los segmentos oclusivos en posición de coda en el castellano limeño. *Boletín de la Academia Peruana de la Lengua*. 63, p. 155 - 183. <https://revistas.apl.org.pe/index.php/boletinapl/article/view/165/613>

Cueva, O. (2021) Las vibrantes del español del departamento de Puno: una caracterización acústica. *Letras*. 92 (136), p. 159 - 171.

Dan Levy, Lior Pachter, (2011) The neighbor-net algorithm, *Advances in Applied Mathematics*, Volume 47, Issue 2, p. 240-258, ISSN 0196-8858, <https://doi.org/10.1016/j.aam.2010.09.002>.

Darby, J. (1981) *Speech Evaluation in Medicine and Psychiatry*. Vol II: Medicine. New York.

Delgado, C. (2001) *La identificación de locutores en el ámbito forense*. Tesis. Universidad Complutense de Madrid, España. <https://eprints.ucm.es/id/eprint/4488/1/T25153.pdf>

Elías-Ulloa, J. (2019) Las propiedades acústicas de la fonación no-modal del boruca: evidencia fonética y estatus fonológico. *Revista de Filología y Lingüística de la Universidad de Costa Rica*. 45 (1), p. 185 - 211.

Elías-Ulloa, J. (2011) *Una documentación acústica de la lengua shipibo-conibo (pano): (con un bosquejo fonológico)*. Fondo Editorial PUCP. Lima, Perú. <https://repositorio.pucp.edu.pe/index/handle/123456789/174304>

Escobar, A. (1978) *Variaciones sociolingüísticas del castellano en el Perú*. Lima: Instituto de Estudios Peruanos.

Escobar, A. (1989) Observaciones sobre el interlecto. En *Temas de Lingüística Aplicada*. Concytec, 1ra Ed. p. 147 – 166.

Esquivel, J., Conde, M., Quintanilla, R. y Alonzo, A. (2016) *Aspectos de gramática española*. 1ra Ed. Kartegraf SRL. Lima, Perú.

Garayzábal, E., Queralt, S., y Reigosa, M. (2019). *Fundamentos de la lingüística forense*. Síntesis.

Gil, J. (2014) Más allá del efecto CSI: avances y metas en fonética judicial. En *Fonética experimental, educación superior e investigación : I. Fonética y fonología /*

Congosto, Y. editor. (Yolanda); Montero Curiel, María Luisa, editor; Salvador Plans, Antonio, editor - Madrid : Arco/Libros, p. 63 - 111. Colección Bibliotheca philologica - <http://digital.casalini.it/9788476358832>

Gutierrez, J. (2015) El perito en el nuevo procesal penal, su valoración de la calidad del informe pericial. 2.º Ed. Grijley

Gutierrez, P. (s./f.) Logo-bits de sinfonos para pronunciar bien las consonantes trabadas. Akros. https://www.mumuchu.com/media/ebooks/logo_bits_de_sinfones_para_consonantes_trabadas_akros.pdf

Heggarty, P. (2005) Enigmas en el origen de las lenguas andinas: aplicando nuevas técnicas a las incógnitas por resolver. *Estudios y Debates*. p. 9 – 57.

Heggarty, P., Maguire, W. y McMahon, A. (2010) Split sor waves? Trees or webs? How divergence measures and network analysis can unravel language histories. *Phil. Trans. R. Soc. B*. Vol 365, p. 3829 – 3843. <http://rstb.royalsocietypublishing.org/subscriptions>

Hernández, R., Fernández, C. & Baptista, P. (2014) *Metodología de la Investigación*. 6ta Ed. Mc Graw Hill Education. <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>

Hernández, R., Fernández, C. & Baptista, P. (2014) *Capítulo 4: Estudios de caso*. 6ta Ed. Mc Graw Hill Education. <https://highered.mheducation.com/sites/dl/free/1456223968/1058642/CAPITULO04.pdf>

Holgado, E. (2020) *El informe pericial de fonética y su valoración en juicio oral de los delitos de crimen organizado, sala penal nacional (Lima) del 2015 a 2018*. [Tesis de grado de maestría]. Universidad Norbert Wiener, Lima, Perú.

Hudson, R. (1980). *Sociolinguistics*. London: Cambridge University Press.

Huson, DH. y Bryant, D. (2006) Application of Phylogenetic Networks in Evolutionary Studies, *Mol. Biol. Evol.*, 23(2):254-267. <http://mbe.oxfordjournals.org/cgi/reprint/msj030?ijkey=GvcBFw4QPIORzUE&keytype=ref>

Jimenez, J. (2018). *Los sonidos de la lengua arabela: un bosquejo fonológico*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Letras y Ciencias Humanas, Escuela Profesional de Lingüística]. Repositorio institucional Cybertesis UNMSM.

Jimenez Peña, J., Torres Castillo, F. A., & Cueva Sanchez, O. E. (2022). Identificación de locutor a partir de la fonética forense: aplicación del software SplitsTree4 para una organización esquemática de los datos lingüísticos. *Boletín de la Academia Peruana de la Lengua*, 71(71), 431 - 461. <https://doi.org/10.46744/bapl.202201.014>

Ladefoged, P. (2004) *Phonetic data analysis, an introduction to fieldwork and instrumental techniques*. 1.º Ed. Blackwell Publishing, Australia.

Latina Noticias (2018, 16-09) “*Mil voces*”: *la confesión del rey de las extorsiones*. <https://www.youtube.com/watch?v=dzdFvWrS6OU>

Lazo, V. (2023) La adecuación de la muestra indubitada en la comparación forense de voz. *Escritura y Pensamiento*. 22 (47), p. 179-205. https://www.researchgate.net/publication/373601026_LA_ADECUACION_DE_LA_MUESTRA_INDUBITADA_EN_LA_COMPARACION_FORENSE_DE_VOZ?enrichId=rgreq-082a65b7827ecd481f018573dd85b04a-XXX&enrichSource=Y292ZXJQYWdlOzMzMzYwMTAyNjBUzoxMTQzMTE4MTE4NTM3Nzq2NkAxNjkzNjI2MDAxOTY4&el=1_x_2&esc=publicationCoverPdf

Listerri, J (2022) El análisis acústico del habla mediante Praat: análisis espectrográfico y análisis de formantes. Departamento de Filología, Universidad Autónoma de Barcelona. https://joaquimllisterri.cat/phonetics/fon_anal_acus/Praat_espectrograma.html

López-Escobedo, F. & Solórzano-Soto, J. (2016) Propuesta de clasificación de un banco de voces con fines de identificación forense. *Linguamática*, 8 (1), 33 - 41.
López, H. (1989). *Sociolingüística*. Madrid:Gredos

Macronorte (2019, 10-06) Vergüenza: Roberto Vieira habló como “robot” en pericia fiscal. *Macronorte*. <https://macronorte.pe/2019/06/10/verguenza-roberto-vieira-hablo-como-robot-en-pericia-fiscal/>

Machuca, M., Ríos, A., & Llisterri, J. (2014). Conocimiento fonético y fonética judicial. In A. Hidalgo, C. Hernández, & F. J. Cantero (Eds.), *La fonética como ámbito interdisciplinar. Estudios de fonopragmática, fonética aplicada y otras interfaces. Quaderns de Filologia: Estudis Lingüístics XIX* (p. 95–111). València: Universitat de València. http://liceu.uab.cat/~joaquim/publicacions/Machuca_Rios_Llisterri_15_fonetica_judicial.pdf
Mejía, E. (2005) *Técnicas e instrumentos de investigación*. 1ra Ed. UNMSM. <http://online.aliat.edu.mx/adistancia/InvCuantitativa/LecturasU6/tecnicas.pdf>

McMahon, A., Heggarty, P., McMahon, R. y Maguire, W. (2007) The sound patterns of Englishes: representing phonetic similarity. *English Language and Linguistics*. Vol 11 (1) p. 113 – 142. Cambridge University Press. http://journals.cambridge.org/abstract_S1360674306002139

McMahon, A., & Maguire, W. (2013) “Computing linguistic distances between varieties.” In: Manfred Krug and Julia Schlüter (eds.). *Research Methods in Language Variation and Change*. Cambridge: Cambridge University Press. 421-432.

Maguire, W. & McMahon A. (2011) “Quantifying relations between dialects.” In: Warren Maguire and April McMahon (eds.). *Analysing Variation in English*. Cambridge: Cambridge University Press. 93-120

Ministerio Público (2020) Guía de elaboración de pericias fonético-acústicas de homologación de voz. <https://portal.mpfm.gob.pe/descargas/normas/d66588.pdf>

Miyara, F. (2004) La voz humana. UNR Editora. p. 1 - 10. <https://www.fceia.unr.edu.ar/acustica/biblio/fonatori.pdf>

Monroy-Casas, R. (1980) Cantidad de las vocales acentuadas. *Revista de Filología Española*, Madrid. tomo III, p. 387 - 408.

Morrison, G. (2011) La comparación forense de la voz y el cambio de paradigma. *Estudios Fónicos / Cuadernos de Trabajo*. Vol. 1., p. 1 – 38.

Muñoz, R. (2020) TgDraw {Praat plug-in} (versión 0.3) [Software]. https://rolandomunoz.github.io/praat_tools/tg_draw.html

Napurí, R. (2011) *Categorización de vocales anteriores por hablantes bilingües quechua-español*. Tesis. Pontificia Universidad Católica del Perú. Lima, Perú.

Navarro, T. (1971) Manual de pronunciación española. 16.º Ed., Revista de Filología Española, Madrid.

Ojo Público (2023) “Los Cuellos Blancos: Fiscalía de la Nación desarticula Oficina de Peritajes cuando hay aún audios por analizar”. 7 de marzo de 2023 – 14:00 m. <https://ojo-publico.com/4330/la-desarticulacion-la-oficina-peritajes-y-los-cuellos-blancos>

Olsson, (2008). *Forensic linguistics* (2ª ed.) Continuum.

O'Rourke, E. (2008) Correlating Speech Rhythm in Spanish: Evidence from Two Peruvian Dialects. En *Selected Proceeding of the 10th Hispanic Linguistics Symposium*, eds. Joyce Bruhnde Garavitoy Elena Valenzuela, p. 276 – 287. Somerville MA: Cascadilla Proceedings Project.

Panorama (2019, 09-06) Exclusivo / Vieira y sus palabritas: audio demuestra intento de distorsionar la pericia. <https://panamericana.pe/panorama/politica/267393-exclusivo-vieira-palabritas-audio-demuestra-distorsionar-pericia>

Paz, M., Friedrich, G. y Galasso, C. (2020) Procesamiento de señal visualizado sobre un espectrograma. *Elektron*, 4 (1), p. 35 - 39. <http://elektron.fi.uba.ar/index.php/elektron/article/view/90>

Paul Boersma & David Weenink (2022): *Praat: doing phonetics by computer* [Computer program]. Version 6.3.02, retrieved 14 March 2018 from <http://www.praat.org/>

Peetz, K. (2014) Introduction a SplitsTree4. https://www.uni-bamberg.de/fileadmin/eng-ling/fs/Chapter_22/Index.html?421DataInput.html

Pérez, J. (2004) *Los castellanos del Perú*. 1. ed, PROEDUCA, <https://ropohuaytaespecializacion.files.wordpress.com/2012/08/los-castellanos-del-peru.pdf>

Pérez, J., Acurio, J. y Bendezú, R. (2008) *Contra el prejuicio de la motosidad. Un estudio de las vocales del castellano andino desde la fonética acústica*. Lima, Instituto Riva-Aguero.

Perú21 (2018, 12-08) Descubren cómo extorsionaba “El mil voces” a través de llamadas telefónicas. Perú21. <https://peru21.pe/lima/policiales/descubren-extorsionaba-mil-voces-traves-llamadas-telefonicas-nndc-420660-noticia/>

Ramírez, J. (2015) *El sistema Constelación. Aportes a la lucha contra el crimen desde una perspectiva comparada*. Tesis para optar el grado de magíster, PUCP. Lima, Perú

Redacción RPP (2019, 28-04) Vieira intentó modificar su voz en peritaje de Fiscalía tras difusión de audio que lo acusa. RPP. <https://rpp.pe/politica/congreso/roberto-vieira-intento-modificar-su-voz-en-peritaje-de-fiscalia-tras-difusion-de-audio-que-lo-acusa-noticia-1194129?ref=rpp>

Rosas, C.; Sommerhoff, J.; Sáez, C. & Saavedra, S. (2011) Comparación de voz bajo el cociente de probabilidad en el caso de Luis Tralcal. *Revista de Lingüística Teórica y Aplicada*. Vol. 52 (1), p. 13 – 33. Concepción, Chile.

Rose, P. (2002) *Forensic speaker identification*. London: Taylor & Francis.

Sánchez-Iglesias, J. (2005). El idiolecto y su traducción: tres ejemplos italianos. http://revistas.usal.es/~revistas_trabajo/index.php/1576-7787/article/view/5150

Saavedra-Mendoza, A. y Akaki-Caballero, M. (2014) Guía práctica clínica de disfonía basada en evidencia. *An Orl mex*, 59, p. 195-208

Shuy, R. (1998) *The language of confession, interrogation, and deception*. Thousand Oaks, CA, Sage.

SISII (2001) *Identificación Plugin Suite*. Speech Technology Center Speech Technology Center. IKAR Lab 3. [Versión Trial] <https://es.speechpro.com/product/analisis/ikarlab>

Stensrud, H. (2019) Duración vocálica en el español de la Gran área metropolitana de Costa Rica. *Revista de Filología y Lingüística de la Universidad de Costa Rica*. 45 (1), p. 215 - 224. <https://revistas.ucr.ac.cr/index.php/filyling/article/view/36736>

Trubetzkoy, N. (1973) *Principios de fonología*. Cincel. 1ra ed. España.

Turell, M. (2010) The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *The International Journal of Speech, Language and the Law* 17 (2), p. 211-250.

Uzco, L. (2018) *Estrategias didácticas para la intervención educativa en casos de dislalia funcional en niños menores de 6 años*. Tesis. Pontificia Universidad Católica

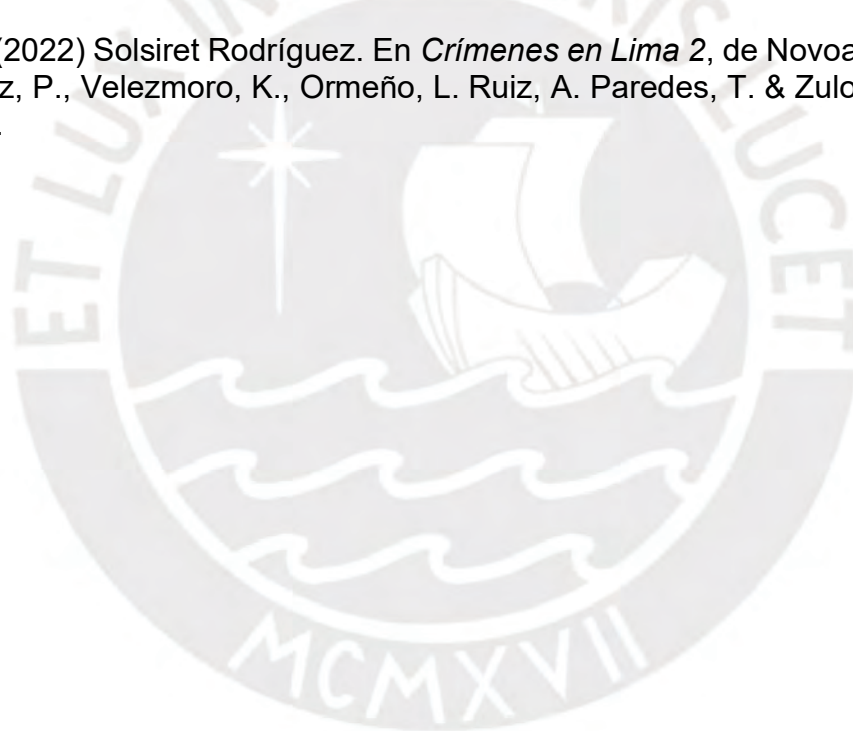
del Perú, Lima.
https://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/16321/Uzco_Rutti_Estrategias_did%C3%A1cticas_intervenci%C3%B3n1.pdf?sequence=4&isAllowed=y

Valenzuela, P. y Jara, M. (2021) “El español peruano amazónico: aportes al conocimiento de su perfil lingüístico”. En *Los castellanos del Perú: historia, variación y contacto lingüístico*, eds. Luis Andrade Ciudad y Sandro Sessarego, p.206 – 242. Londres /Nueva York: Routledge.

Vásquez, A. (2017) Patrones de entonación en construcciones de foco estrecho en el español amazónico de Iquitos. Tesis para optar el título de licenciatura, Pontificia Universidad Católica, Perú.
<https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/8467>

Zariquiey, R., Vásquez, A. y Tello, G. (2017) Lenguas y dialectos pano del Purús: una aproximación filogenética. *Lexis*. Vol. XLI (1) p. 83 – 120.
<http://www.scielo.org.pe/pdf/lexis/v41n1/a03v41n1.pdf>

Zuloeta, C. (2022) Solsiret Rodríguez. En *Crímenes en Lima 2*, de Novoa, K., Muschi, C., González, P., Velezmoro, K., Ormeño, L. Ruiz, A. Paredes, T. & Zuloeta, C. 1 ed. Melquíades.



ANEXOS

Anexo 01: Modelo de Protocolo de toma de muestra de voz

Protocolo de toma de muestra de voz

1. Datos personales

- Lima, de de 2021.
- Mi nombre completo es
- El número de mi DNI es
- Tengo años.
- Nací en el departamento de, provincia de, distrito de
- Me dedico a
- El grado de estudios que poseo es
- Mi lengua materna es el
- Mi papá nació en y mi mamá en
- Mis abuelos se dedicaban a

2. Lista de palabras

ID	Lista de palabras
01	pelota, sapo, taza, pato, casa, vaca
02	vela, bebe, gato, tortuga, dedo, helado
03	mesa, cama, nariz, mano, panda, botón, ñato, piña
04	chaleco, cuchara, chiquito, cachetada, cachos
05	luna, polo, palmera, sol, alrededor
06	foca, elefante, fumar, fino, jabón, ojo, lejía, ceja

07	llave, payaso, zapallo, gallina
08	pera, árbol, collar, ratón perro enrique, roto, gorro
09	blusa, plato, flor, clavo, globo
10	libro, profesor, fresa, cruz, tigre, dragón, tren
11	plano, avión, pie, agua, huevo, pingüino
12	peine, jaula, deuda, aceite, coima, auto.
13	34, 300, 25, 2360, 94, 98, 31, 598
14	aló, chau, porque, por tanto

iiiiiiiiiii, eeeeeee, aaaaaa, oooooo, uuuuuu.

Lunes, martes, miércoles, jueves, viernes, sábado, domingo

1,2,3,4,5,6,7,8,9,10,11,20,24,100,500,999

Alo, alo, alo, ya, ya pe, ya pues, no, no pe. Buenas tardes, buenos días, buenas noches, por favor, si claro, si pe, si que si. Su canto. Su voz.

Comunícate ahora, no, no sé, no quiero. Entiendes.

¿Aló? ¿Me entiendes? ¿Qué has hecho? ¿Cómo es?

¿Qué tal? ¿Cómo estás?

¡Hola! Buenos días. No esta. ¿Quién eres?

No. No es así.

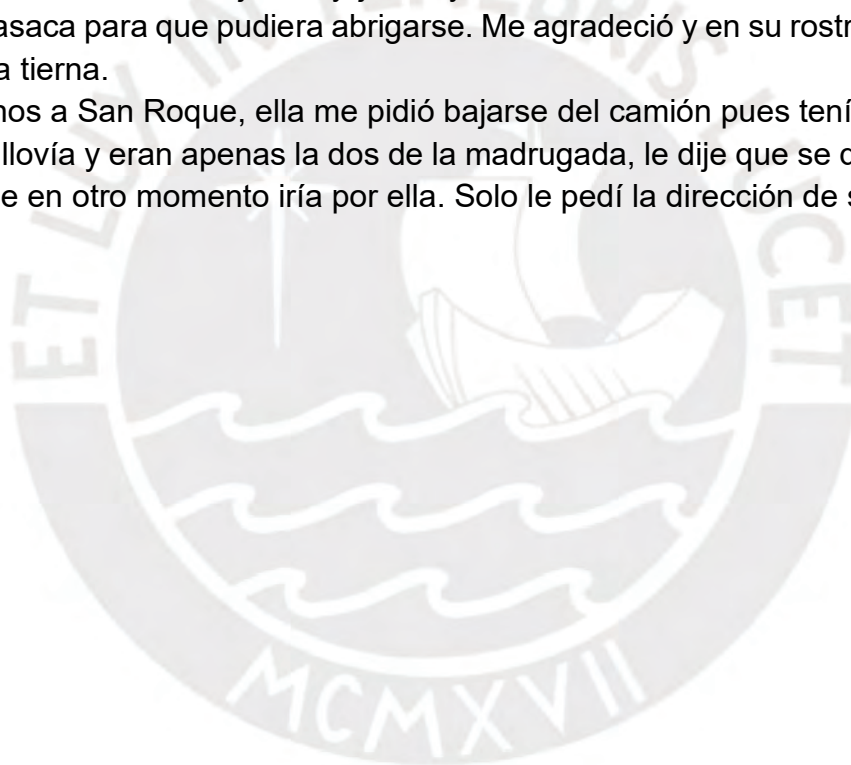
Soy tal o cual ¿y tú? No sé. Si. Así.

3. Lectura

Anécdota en San Roque

Cuando me disponía a venir a Lima, conocí a don Guillermo, quien muy amablemente me invitó a subir a su camión, en donde transportaba gallinas y pescado desde Yurimaguas hasta la capital. Subí en la Oroya. En el trayecto, don Guillermo me contó lo que le pasó en el pueblo de San Roque cuando regresaba de Yurimaguas trayendo carga: “Cuando salía de Yurimaguas, ya muy de noche y bajo una interminable lluvia, pude avistar a una mujer en el camino. Ella iba caminando muy lentamente en la carreta - ¡debiste verla con aquel vestido rojo totalmente embarrado! - frené suavemente y le hice una señal para que suba al camión y así pudiera protegerse de la lluvia. ¡Ella acepto y se sentó en el mismo lugar donde estás tú, sí tú! Era una mujer muy joven y bella. Al verla en esas condiciones, le ofrecí mi casaca para que pudiera abrigarse. Me agradeció y en su rostro vi dibujada una sonrisa tierna.

Al acercarnos a San Roque, ella me pidió bajarse del camión pues tenía familia allí. Como aún llovía y eran apenas las dos de la madrugada, le dije que se quede con mi casaca, que en otro momento iría por ella. Solo le pedí la dirección de su casa”



RASPUTIN EL RATÓN

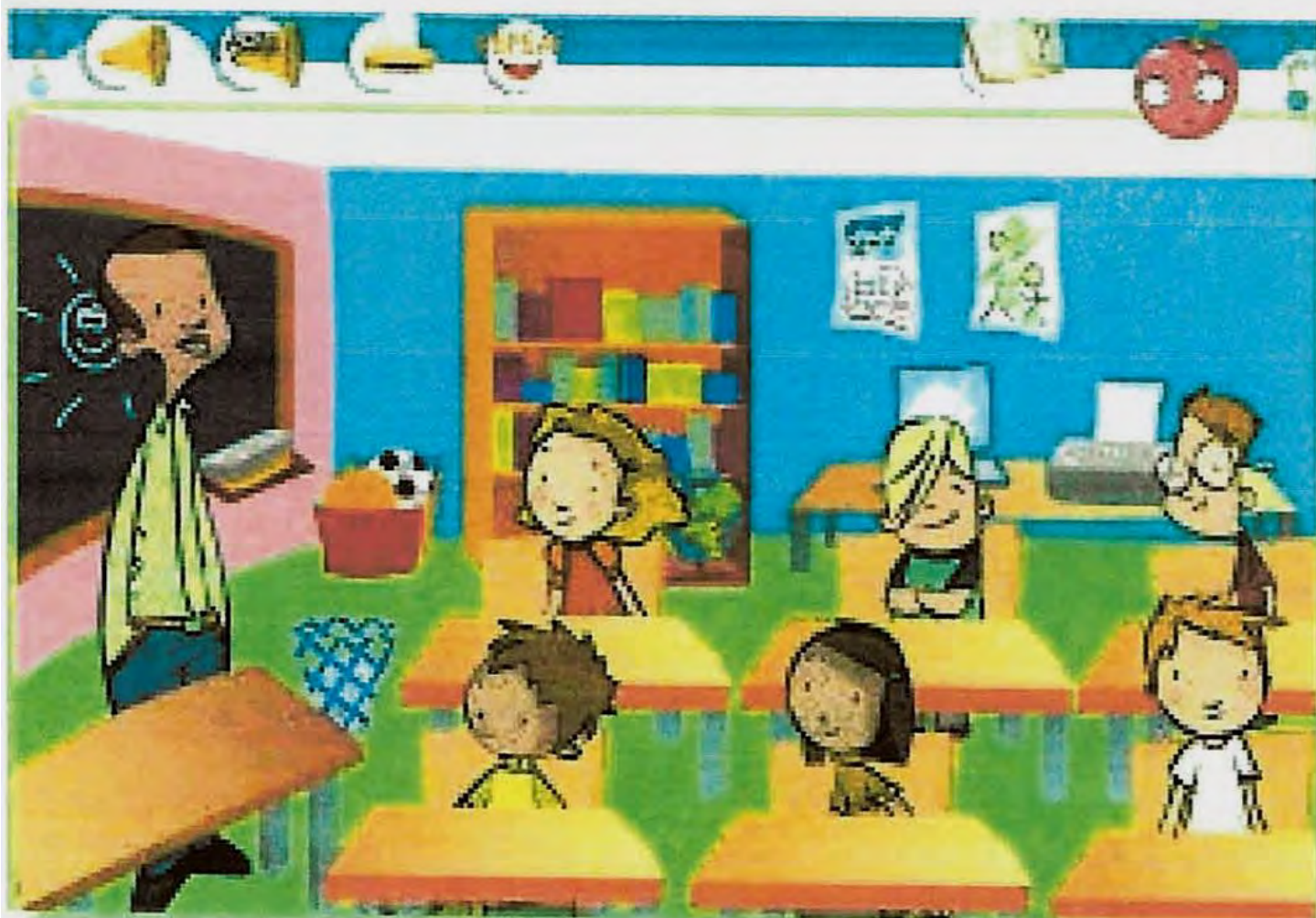
Rasputín era un ratón
que rascaba en una
risca,
con un toscó.

rasca risca rascador,
rasca que rasca,
acabo con el risco,
rasca la risca,
rasca un rincón.

ESCRIBA DETALLADAMENTE LAS IMAGENES



SCRIBA DETALLADAMENTE LAS IMAGENES



Anexo 02: Segmentación de la participación de las voces de las muestras dubitadas

Con el fin de reportar la correspondencia de voces sin superposiciones o ruidos que alteren la muestra dubitada “BVR_2019_07_26_22_30_26_01.wav”, se han generado Intervalos de tiempo de intervención del locutor rotulado como “voz 1”, “voz 2”, voz 3” y “voz 4”.

tmin (s)	Tier 1	tmax (s)
0.41	voz1	1.74
2.29	voz1	3.9
7.42	voz1	8.93
9.18	voz1	9.96
17.58	voz1	20.15
92.32	voz1	93.27
130.94	voz1	132.38
145.71	voz1	147.61
172.59	voz1	174.81
176.37	voz1	178.85
179.17	voz1	180.53
180.93	voz1	185.52
185.98	voz1	190.31
196.27	voz1	197.83
224.81	voz1	226.63

226.83	voz1	228.5
229.02	voz1	230.9
231.44	voz1	233.89
234.35	voz1	235.82
418.19	voz1	419.23

Tmin	Tier 2	tmax
9.94	voz2	11.18
12.22	voz2	13.28
13.51	voz2	13.98
23.07	voz2	24.68
25.63	voz2	27.65
28.13	voz2	29.17
32.09	voz2	33.73
40.71	voz2	44.26
44.69	voz2	46.22
47.98	voz2	48.91
55.71	voz2	58.19
58.74	voz2	60.5

62.4	voz2	64.02
64.88	voz2	67.48
67.92	voz2	72.47
73.25	voz2	75.27
78.5	voz2	84.88
85.31	voz2	88.11
95.58	voz2	99.68
100.79	voz2	111.76
123.61	voz2	125.72
129.18	voz2	130.56
150.64	voz2	155.12
161.29	voz2	162.73
191.14	voz2	193.57
194.18	voz2	195.97
203.92	voz2	206
206.26	voz2	207.3
207.85	voz2	210.3
222.13	voz2	223.54
242.12	voz2	244.71

245.93	voz2	248.09
249.3	voz2	251.55
252.44	voz2	254.03
254.69	voz2	256.59
257.63	voz2	259.13
260.2	voz2	262.31
262.65	voz2	267.9
268.71	voz2	272.09
273.85	voz2	279.62
281.54	voz2	282.47
282.99	voz2	285.12
285.55	voz2	287.75
292.25	voz2	293.46
293.87	voz2	295.63
296.12	voz2	298.95
299.49	voz2	300.36
300.72	voz2	304.53
308.11	voz2	320.18
320.51	voz2	321.67

324.09	voz2	326.92
327.53	voz2	333.7
338.89	voz2	343.77
344.11	voz2	355.39
363.76	voz2	365.17
368.12	voz2	369.41
377.86	voz2	380.78
381.73	voz2	384.15
384.47	voz2	385.31
386.32	voz2	388.65
389.49	voz2	391.43
392.06	voz2	394.02
396.24	voz2	400.05
400.71	voz2	406.45
407	voz2	408.93
410.89	voz2	412.74
423.76	voz2	426.24

Tmin	Tier 3	tmax
-------------	---------------	-------------

36.9	voz3	39.73
49.08	voz3	50.44
51.33	voz3	53.21
137.2	voz3	140.38
166.77	voz3	172.54
306.87	voz3	307.65
369.59	voz3	370.97
371.61	voz3	373.02
414.04	voz3	415.11

Tmin	Tier 4	tmax
359.75	voz4	360.81
568.04	voz4	569.4

Con el fin de reportar la correspondencia de voces sin superposiciones o ruidos que alteren la muestra dubitada "BVR_2019_07_26_21_54_09_01.wav", se han generado Intervalos de tiempo de intervención del locutor rotulado como "voz 1", "voz 2", voz 3" y "voz 4".

Tmin	Tier 1	tmax
227.57	voz 1	231.52
233.75	voz 1	239.31

239.96	voz 1	241.54
242.38	voz 1	245.73
246.17	voz 1	248.64
250.3	voz 1	251.47
268.16	voz 1	269.12
271.19	voz 1	273.63
274.44	voz 1	276.41
276.8	voz 1	279.6
279.81	voz 1	281.03
282.18	voz 1	284.9
285.39	voz 1	288.27
288.27	voz 1	289.72
290.42	voz 1	299.07
299.46	voz 1	305.02
314.04	voz 1	317.13
318.29	voz 1	321.23
321.7	voz 1	323.38
326.08	voz 1	329.2
329.53	voz 1	331.42

339.29	voz 1	341.78
343.99	voz 1	345
360.46	voz 1	362.26
385.14	voz 1	386.78
405.05	voz 1	410.19
410.42	voz 1	416.5
417.41	voz 1	418.19
419.25	voz 1	423.85
424.83	voz 1	425.63
432.57	voz 1	437.37
438.31	voz 1	443.06
445.76	voz 1	447.29
449.63	voz 1	450.62
456.33	voz 1	457.13
465.31	voz 1	466.64
473.52	voz 1	474.74
478.18	voz 1	482.86
501.73	voz 1	503.73
512.05	voz 1	513.38

524.62	voz 1	526.28
537.14	voz 1	540.93
542.43	voz 1	544.3
562.17	voz 1	563.94
622.86	voz 1	625.62
631.31	voz 1	632.64
636.79	voz 1	638.77
647.75	voz 1	648.76
675.94	voz 1	677.88
712.66	voz 1	714.51
715.47	voz 1	718.59
721.36	voz 1	723.93
737.29	voz 1	739.26
750.09	voz 1	750.99
803.19	voz 1	805.99
810.23	voz 1	813.91
815.16	voz 1	818.33
834.32	voz 1	837.56
842.27	voz 1	853.33

856.66	voz 1	858.87
859.11	voz 1	862.51
878.05	voz 1	879.89
894.04	voz 1	896.74
907.35	voz 1	908.94
913.74	voz 1	915.11
922.51	voz 1	923.65
926.15	voz 1	930.01
936.74	voz 1	942.27
943.16	voz 1	947.55
950.59	voz 1	951.88
969.52	voz 1	971.18
972.06	voz 1	973.7
978.09	voz 1	980.53
1011.16	voz 1	1014.33
1015.24	voz 1	1018.09
1018.56	voz 1	1022.64
1023.31	voz 1	1025.62
1026.99	voz 1	1031.38

1039.82	voz 1	1042.16
1044	voz 1	1048.54
1049.89	voz 1	1051.24
1052.41	voz 1	1053.86
1061.81	voz 1	1065.66
1066	voz 1	1066.59
1068.79	voz 1	1072.66
1076.42	voz 1	1079.35
1081.94	voz 1	1084.85
1085.57	voz 1	1086.85
1089.36	voz 1	1090.82
1090.84	voz 1	1092.43
1092.92	voz 1	1095.6
1096.77	voz 1	1098.74
1099.07	voz 1	1100.19
1248.11	voz 1	1249.43
1948.76	voz 1	1950.21
1951.17	voz 1	1952.29
1954.31	voz 1	1955.51

1958.36	voz 1	1961.45
---------	-------	---------

Tmin	tier	tmax
115.75	voz 2	116.12
128.65	voz 2	129.57
161.88	voz 2	163.75
255.12	voz 2	259.39
335.3	voz 2	336.89
344.97	voz 2	346.07
373.85	voz 2	379.66
381.5	voz 2	382.25
383.78	voz 2	385.02
396.76	voz 2	397.97
400.05	voz 2	403.34
426.19	voz 2	428.76
453.43	voz 2	456.4
470.1	voz 2	470.94
482.95	voz 2	483.47

497.83	voz 2	500.46
506.39	voz 2	510.31
514.71	voz 2	516.79
519.07	voz 2	522.94
526.26	voz 2	528.59
529.75	voz 2	532.63
533.79	voz 2	535.26
540.88	voz 2	542.44
544.98	voz 2	545.32
551.67	voz 2	552.91
555.1	voz 2	558.08
570.73	voz 2	572.49
574.42	voz 2	576.38
577.11	voz 2	579.01
580.38	voz 2	583.5
585.1	voz 2	586.19
586.77	voz 2	593.21
593.5	voz 2	594.91
597.23	voz 2	600.08

601.56	voz 2	604.79
605.22	voz 2	609.81
610.44	voz 2	612.43
612.69	voz 2	619.27
664.21	voz 2	673.88
677.99	voz 2	681.34
692.83	voz 2	698.02
728.5	voz 2	730.4
746.35	voz 2	748.2
753.97	voz 2	756.74
757.17	voz 2	758.16
760	voz 2	765.86
766.48	voz 2	767.46
773.1	voz 2	774.83
777.4	voz 2	779.91
780.2	voz 2	781.9
782.57	voz 2	785.89
786.38	voz 2	787.36
788.92	voz 2	794.4

795.85	voz 2	799.31
870.21	voz 2	872.2
873.88	voz 2	875.52
879.95	voz 2	889.21
889.85	voz 2	891.12
896.66	voz 2	897.15
903.79	voz 2	905.89
986.1	voz 2	992.91
997.95	voz 2	1001.87
1032.2	voz 2	1033.44
1102.64	voz 2	1104.2
1243.61	voz 2	1246.38
1253.28	voz 2	1254.69
1270.98	voz 2	1275.19
1276.48	voz 2	1285.89
1324.35	voz 2	1325.65
1480.11	voz 2	1482.94
1486.26	voz 2	1490.61
1492.11	voz 2	1493.44

1495.95	voz 2	1500.92
1504.33	voz 2	1505.57
1506.46	voz 2	1512.43
1513.53	voz 2	1516.73
1519.4	voz 2	1521.11
1567.42	voz 2	1569.13
1591.59	voz 2	1593.9
1604.28	voz 2	1605.9
1606.42	voz 2	1610.57
1757.4	voz 2	1761.13
1761.84	voz 2	1763.2
1764.18	voz 2	1765.88
1766.34	voz 2	1768.6
1769.14	voz 2	1770.93
1794.91	voz 2	1796.69
1813.25	voz 2	1817.66
1827.27	voz 2	1828.82
1829.81	voz 2	1832.26
1838.2	voz 2	1842.3

1889.95	voz 2	1892.92
1893.73	voz 2	1895.37
1936.27	voz 2	1946.17
1952.86	voz 2	1953.81
1956.5	voz 2	1958
1962.06	voz 2	1966.56
1967.54	voz 2	1971.01
1975.1	voz 2	1982.35
2019.99	voz 2	2025.13
2028.27	voz 2	2031.5
2033.72	voz 2	2038.28
2047.52	voz 2	2055.83
2057.47	voz 2	2062.81

Tmin	tier	tmax
643.81	voz 3	646.75
656.71	voz 3	659.65
686.21	voz 3	688.51

724.53	voz 3	725.28
770.44	voz 3	772.14
867.38	voz 3	868.91
1132.88	voz 3	1134.32
1134.9	voz 3	1136.29
1137.21	voz 3	1138.77
1147.87	voz 3	1149.77
1149.95	voz 3	1152.98
1153.58	voz 3	1154.45
1160.48	voz 3	1163.85
1174.78	voz 3	1176.46
1176.51	voz 3	1184.29
1184.98	voz 3	1186.3
1210.01	voz 3	1211.45
1212.72	voz 3	1214.22
1287.46	voz 3	1296.06
1341.69	voz 3	1344.09
1350.66	voz 3	1351.4
1353.06	voz 3	1354.79

1362.63	voz 3	1364.3
1374.57	voz 3	1378.53
1422.98	voz 3	1424.83
1436.96	voz 3	1438.63
1446.85	voz 3	1448.7
1478.12	voz 3	1478.67
1539.51	voz 3	1541.24
1563.46	voz 3	1564.27
1575.71	voz 3	1583.96
1633.68	voz 3	1635.96
1647.4	voz 3	1648.74
1661.38	voz 3	1665.37
1669.52	voz 3	1672.81
1673.54	voz 3	1676.48
1727.08	voz 3	1729.44
1787.11	voz 3	1788.49
1843.85	voz 3	1847.94
1857.15	voz 3	1858.85
1866.99	voz 3	1868.72

Tmin	tier	tmax
1158.32	voz 4	1160.25
1205.87	voz 4	1208.22
1222.35	voz 4	1223.77
1241.11	voz 4	1242.45
1302.66	voz 4	1304.23
1357.59	voz 4	1359.09
1365.74	voz 4	1370.61
1414.6	voz 4	1416.1
1418.01	voz 4	1420.07
1425.97	voz 4	1428.72
1432.14	voz 4	1435.31
1460.28	voz 4	1461.25
1528.42	voz 4	1532.69
1559.52	voz 4	1561.45
1697.81	voz 4	1699.36
1699.82	voz 4	1701.44
1711.36	voz 4	1712.25
1723.69	voz 4	1726.2

1737.41	voz 4	1738.73
1745.22	voz 4	1746.26
1782.14	voz 4	1785.16
1871.4	voz 4	1876.23
1884.3	voz 4	1888.4
1899.45	voz 4	1902.65
1904.05	voz 4	1907.56
1912.47	voz 4	1916.99
1918.5	voz 4	1925.41



Anexo 03: Capturas de pantalla de los resultados de análisis de la calidad de la señal de las muestras indubitadas y dubitadas

LI_01	LI_02
<p>Emotional State: Regular</p> <p>Spectrum Characteristics: Frequency response 0 --- 3521 Hz, SNR 100 --- 3300 Hz 28 dB</p> <p>Tonal Noise: Stationary tonal noise detected Total 9.66 s</p> <p>Reverberation Time: Average Time 308 ms, Dispersion 9</p> <p>Clipping: Clipping detected Total 0.43 s, Total 0.14 %</p> <p>Voice: Total 176.22 s, Total 56.74 %</p> <p>The recording can be accepted for voice ID analysis</p> <p>Process: <input checked="" type="radio"/> Entire Signal 310.55 seconds</p> <p>Source: <input checked="" type="radio"/> Microphone</p> <p>Buttons: Save to project, Copy, Calculate, Close</p>	<p>Emotional State: Regular</p> <p>Spectrum Characteristics: Frequency response 0 --- 4221 Hz, SNR 100 --- 3300 Hz 29 dB</p> <p>Tonal Noise: Stationary tonal noise detected Total 23.22 s</p> <p>Reverberation Time: Average Time 298 ms, Dispersion 10</p> <p>Clipping: Clipping detected Total 0.09 s, Total 0.03 %</p> <p>Voice: Total 152.61 s, Total 52.65 %</p> <p>The recording can be accepted for voice ID analysis</p> <p>Process: <input checked="" type="radio"/> Entire Signal 289.85 seconds</p> <p>Source: <input checked="" type="radio"/> Microphone</p> <p>Buttons: Save to project, Copy, Calculate, Close</p>
LI_03	LI_04
<p>Emotional State: Regular</p> <p>Spectrum Characteristics: Frequency response 105 --- 4875 Hz, SNR 100 --- 3300 Hz 46 dB</p> <p>Tonal Noise: No stationary tonal noise detected</p> <p>Reverberation Time: Average Time 229 ms, Dispersion 15</p> <p>Clipping: No clipping detected</p> <p>Voice: Total 46.59 s, Total 59.55 %</p> <p>The recording can be accepted for voice ID analysis</p> <p>Process: <input checked="" type="radio"/> Entire Signal 78.24 seconds</p> <p>Source: <input checked="" type="radio"/> Microphone</p> <p>Buttons: Save to project, Copy, Calculate, Close</p>	<p>Emotional State: Regular</p> <p>Spectrum Characteristics: Frequency response 117 --- 3773 Hz, SNR 100 --- 3300 Hz 46 dB</p> <p>Tonal Noise: Stationary tonal noise detected Total 0.19 s</p> <p>Reverberation Time: Average Time 281 ms, Dispersion 68</p> <p>Clipping: No clipping detected</p> <p>Voice: Total 10.27 s, Total 62.17 %</p> <p>The recording cannot be accepted for voice ID analysis</p> <p>Process: <input checked="" type="radio"/> Entire Signal 16.52 seconds</p> <p>Source: <input checked="" type="radio"/> Microphone</p> <p>Buttons: Save to project, Copy, Calculate, Close</p>
MD_01	MD_02

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be partially accepted for voice ID analysis

Process

Entire Signal **312.00 seconds**

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_03

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

Stationary tonal noise detected

Total s

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be partially accepted for voice ID analysis

Process

Entire Signal **256.18 seconds**

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_04

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

Stationary tonal noise detected

Total s

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be partially accepted for voice ID analysis

Process

Entire Signal **100.36 seconds**

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_05

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be partially accepted for voice ID analysis

Process

Entire Signal **70.67 seconds**

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_06

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be partially accepted for voice ID analysis

Process

Entire Signal seconds

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_07

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording can be accepted for voice ID analysis

Process

Entire Signal seconds

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

MD_08

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording cannot be accepted for voice ID analysis

Process

Entire Signal seconds

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

Emotional State Regular ?

Spectrum Characteristics ?

Frequency response --- Hz Plot a chart

SNR --- Hz dB Plot a chart

Tonal Noise ?

No stationary tonal noise detected

Reverberation Time ?

Average Time ms Dispersion

Clipping ?

No clipping detected

Voice ?

Total s Total %

The recording cannot be accepted for voice ID analysis

Process

Entire Signal seconds

Selected Area seconds

Source

Microphone Telephone

Save to project Copy Calculate Close

Anexo 04: Capturas de pantallas de los resultados de la comparación automática con el programa SIS II

LI_01 vs MD_03

Automatic Comparison: restored

Files

Different emotional states

File 1: [Redacted] Mark Group: [Redacted]

Use as pitch model

Source: Microphone Telephone

File 2: [Redacted] Mark Group: [Redacted]

Use as pitch model

Source: Microphone Telephone

Confidence level: 95%

Method	FR [min,max], %	FA [min,max], %	LR [min,max]	P [min,max], %	P# [min,max], %	DET
<input checked="" type="checkbox"/> SF	43.2 [37.0, 49.5]	34.9 [33.3, 36.5]	1.2 [1.1, 1.4]	54.2 [50.9, 57.4]	45.8 [42.6, 49.1]	DET
<input checked="" type="checkbox"/> Pitch	90.1 [86.3, 93.9]	9.1 [8.1, 10.1]	9.9 [8.9, 11.0]	90.5 [88.6, 92.5]	9.5 [7.5, 11.4]	DET
<input checked="" type="checkbox"/> GMM	74.7 [69.2, 80.2]	0.6 [0.3, 0.9]	125.3 [85.9, 164.8]	87.1 [84.3, 89.8]	12.9 [10.2, 15.7]	DET

Summary:

False rejection percentage FR: **88.6%**
 False acceptance percentage FA: **0.1%**
 Likelihood ratio LR: **1 023.313**
 Same speaker (High probability)

DET

Compare Copy results Save to project Close

LI_01 vs MD_04

Automatic Comparison: restored

Files

Different emotional states

File 1: [Redacted] Mark Group: [Redacted]

Use as pitch model

Source: Microphone Telephone

File 2: [Redacted] Mark Group: [Redacted]

Use as pitch model

Source: Microphone Telephone

Confidence level: 95%

Method	FR [min,max], %	FA [min,max], %	LR [min,max]	P [min,max], %	P# [min,max], %	DET
<input checked="" type="checkbox"/> SF	2.3 [0.6, 3.9]	79.6 [77.8, 81.4]	0.03 [0.01, 0.05]	11.3 [10.1, 12.5]	88.7 [87.5, 89.9]	DET
<input checked="" type="checkbox"/> Pitch	71.0 [65.9, 76.2]	20.0 [18.3, 21.8]	3.5 [3.2, 3.9]	75.5 [72.8, 78.2]	24.5 [21.8, 27.2]	DET
<input checked="" type="checkbox"/> GMM	0.5 [0.0, 1.3]	58.9 [56.7, 61.1]	0.01 [0.0, 0.02]	20.8 [19.7, 21.9]	79.2 [78.1, 80.3]	DET

Summary:

False rejection percentage FR: **0.17%**
 False acceptance percentage FA: **50.8%**
 Likelihood ratio LR: **0.003 (1 / 299)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_01

Automatic Comparison: restored

Files

Different emotional states

File 1: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

File 2: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

Confidence level: 95%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	0.16 [0.0, 0.6]	80.5 [79.1, 82.0]	0.0 [0.0, 0.01]	9.8 [9.1, 10.6]	90.2 [89.4, 90.9]	DET
<input checked="" type="checkbox"/> Pitch	35.2 [30.5, 40.0]	23.4 [21.8, 24.9]	1.5 [1.3, 1.7]	55.9 [53.4, 58.4]	44.1 [41.6, 46.6]	DET
<input checked="" type="checkbox"/> GMM	0.16 [0.0, 0.6]	60.8 [59.0, 62.5]	0.0 [0.0, 0.01]	19.7 [18.8, 20.6]	80.3 [79.4, 81.2]	DET

Summary:

False rejection percentage FR: **0.16%**
 False acceptance percentage FA: **65.4%**
 Likelihood ratio LR: **0.0025 (1 / 397)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_02

Automatic Comparison

Files

Different emotional states

File 1: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 44100 Hz; 289.85 s;
Clear voice signal: more than 124 s;

File 2: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 48000 Hz; 256.18 s;
Clear voice signal: more than 124 s;

Confidence level: 99%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	0.6 [0.0, 2.1]	87.2 [84.4, 90.0]	0.01 [0.0, 0.02]	6.7 [5.2, 8.2]	93.3 [91.8, 94.8]	DET
<input checked="" type="checkbox"/> Pitch	26.2 [17.5, 34.8]	20.2 [16.8, 23.6]	1.3 [0.8, 1.8]	53.0 [48.3, 57.6]	47.0 [42.4, 51.7]	DET
<input checked="" type="checkbox"/> GMM	0.6 [0.0, 2.1]	61.1 [57.0, 65.2]	0.01 [0.0, 0.03]	19.7 [17.6, 21.9]	80.3 [78.1, 82.4]	DET

Summary:

False rejection percentage FR: **0.6%**
 False acceptance percentage FA: **66.8%**
 Likelihood ratio LR: **0.009 (1 / 115)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_03

Automatic Comparison: restored

Files

Different emotional states

File 1: A: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

File 2: B: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

Confidence level: 95%

Method	FR [min,max], %	FA [min,max], %	LR [min,max]	P [min,max], %	P# [min,max], %	DET
<input checked="" type="checkbox"/> SF	29.1 [23.0, 35.1]	46.2 [43.7 , 48.7]	0.6 [0.5, 0.8]	41.4 [38.2, 44.7]	58.6 [55.3 , 61.8]	DET
<input checked="" type="checkbox"/> Pitch	65.0 [58.7 , 71.4]	17.2 [15.3, 19.0]	3.8 [3.3 , 4.3]	73.9 [70.6 , 77.3]	26.1 [22.7, 29.4]	DET
<input checked="" type="checkbox"/> GMM	0.6 [0.0, 1.7]	77.0 [74.9 , 79.1]	0.01 [0.0, 0.02]	11.8 [10.7, 13.0]	88.2 [87.0 , 89.3]	DET

Summary:

False rejection percentage FR: **0.4%**
 False acceptance percentage FA: **57.4%**
 Likelihood ratio LR: **0.008 (1 / 131)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_04

Automatic Comparison: restored

Files

Different emotional states

File 1: A: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

File 2: B: [redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

Confidence level: 95%

Method	FR [min,max], %	FA [min,max], %	LR [min,max]	P [min,max], %	P# [min,max], %	DET
<input checked="" type="checkbox"/> SF	1.6 [0.2, 3.1]	82.9 [81.2 , 84.5]	0.02 [0.0, 0.04]	9.4 [8.3, 10.5]	90.6 [89.5 , 91.7]	DET
<input checked="" type="checkbox"/> Pitch	78.7 [74.0 , 83.3]	15.4 [13.8, 17.0]	5.1 [4.5 , 5.7]	81.7 [79.2 , 84.1]	18.3 [15.9, 20.8]	DET
<input checked="" type="checkbox"/> GMM	0.4 [0.0, 1.1]	58.4 [56.2 , 60.5]	0.01 [0.0, 0.02]	21.0 [19.9, 22.1]	79.0 [77.9 , 80.1]	DET

Summary:

False rejection percentage FR: **0.2%**
 False acceptance percentage FA: **47.7%**
 Likelihood ratio LR: **0.004 (1 / 262)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_05

Automatic Comparison

Files

Different emotional states

File 1: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 44100 Hz; 289.85 s;
Clear voice signal: more than 124 s;

File 2: B: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 48000 Hz; 188.53 s;
Clear voice signal: 112.62 s;

Confidence level: 99%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	0.4 [0.0, 1.0]	66.4 [64.6, 68.1]	0.01 [0.0, 0.01]	17.0 [16.1, 17.9]	83.0 [82.1, 83.9]	DET
<input checked="" type="checkbox"/> Pitch	53.2 [48.2, 58.2]	15.9 [14.5, 17.2]	3.3 [2.9, 3.8]	68.6 [66.1, 71.2]	31.4 [28.8, 33.9]	DET
<input checked="" type="checkbox"/> GMM	0.16 [0.0, 0.6]	57.2 [55.4, 59.0]	0.0 [0.0, 0.01]	21.5 [20.6, 22.4]	78.5 [77.6, 79.4]	DET

Summary:

False rejection percentage FR: **0.16%**
 False acceptance percentage FA: **56.4%**
 Likelihood ratio LR: **0.003 (1 / 343)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_02 vs MD_06

Automatic Comparison

Files

Different emotional states

File 1: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 44100 Hz; 289.85 s;
Clear voice signal: more than 124 s;

File 2: B: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 48000 Hz; 39.37 s;
Clear voice signal: 23.25 s;

Confidence level: 99%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	2.3 [0.0, 5.4]	40.4 [37.6, 43.1]	0.06 [0.0, 0.1]	31.0 [29.1, 32.9]	69.0 [67.1, 70.9]	DET
<input checked="" type="checkbox"/> Pitch	98.8 [96.7, 99.99]	1.1 [0.5, 1.7]	86.1 [56.4, 115.8]	98.8 [98.0, 99.7]	1.2 [0.3, 2.0]	DET
<input checked="" type="checkbox"/> GMM	0.6 [0.0, 2.1]	76.3 [73.9, 78.7]	0.01 [0.0, 0.03]	12.1 [10.8, 13.4]	87.9 [86.6, 89.2]	DET

Summary:

False rejection percentage FR: **0.6%**
 False acceptance percentage FA: **63.6%**
 Likelihood ratio LR: **0.009 (1 / 113)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_03 vs MD_03

Automatic Comparison: restored

Files

Different emotional states

File 1: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

File 2: B: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

Confidence level: 95%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	1.2 [0.0, 3.1]	97.3 [96.6, 98.0]	0.01 [0.0, 0.03]	2.0 [1.1, 2.8]	98.0 [97.2, 98.9]	DET
<input checked="" type="checkbox"/> Pitch	1.5 [0.0, 3.6]	92.9 [91.7, 94.0]	0.02 [0.0, 0.04]	4.3 [3.3, 5.4]	95.7 [94.6, 96.7]	DET
<input checked="" type="checkbox"/> GMM	12.1 [6.5, 17.8]	38.2 [36.1, 40.4]	0.3 [0.17, 0.5]	36.9 [33.9, 40.0]	63.1 [60.0, 66.1]	DET

Summary:

False rejection percentage FR: **0.7%**
 False acceptance percentage FA: **79.6%**
 Likelihood ratio LR: **0.008 (1 / 121)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

LI_03 vs MD_04

Automatic Comparison

Files

Different emotional states

File 1: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 48000 Hz; 78.24 s;
Clear voice signal: -46.54 s;

File 2: B: [Redacted] Mark Group: []

Use as pitch model

Source

Microphone Telephone

16 bit; mono; 48000 Hz; 70.67 s;
Clear voice signal: -42.30 s;

Confidence level: 99%

Method	FR [min,max] %	FA [min,max] %	LR [min,max]	P [min,max] %	P# [min,max] %	DET
<input checked="" type="checkbox"/> SF	1.0 [0.0, 2.5]	76.5 [74.7, 78.3]	0.01 [0.0, 0.03]	12.3 [11.2, 13.3]	87.7 [86.7, 88.8]	DET
<input checked="" type="checkbox"/> Pitch	0.3 [0.0, 1.1]	81.9 [80.3, 83.6]	0.0 [0.0, 0.01]	9.2 [8.3, 10.1]	90.8 [89.9, 91.7]	DET
<input checked="" type="checkbox"/> GMM	0.3 [0.0, 1.1]	72.9 [71.0, 74.8]	0.0 [0.0, 0.02]	13.7 [12.7, 14.7]	86.3 [85.3, 87.3]	DET

Summary:

False rejection percentage FR: **0.3%**
 False acceptance percentage FA: **84.3%**
 Likelihood ratio LR: **0.004 (1 / 268)**
 Different speakers (High probability)

DET

Compare Copy results Save to project Close

Anexo 06: Script para generar redes filogenéticas con *Simple Data*

```
#nexus
```

```
BEGIN Taxa;  
DIMENSIONS ntax=9;  
TAXLABELS
```

```
[1] 'LI_01'
```

```
[2] 'LI_02'
```

```
[3] 'LI_03'
```

```
[4] 'MD_01'
```

```
[5] 'MD_02'
```

```
[6] 'MD_03'
```

```
[7] 'MD_04'
```

```
[8] 'MD_05'
```

```
[9] 'MD_06'
```

```
;
```

```
END; [Taxa]
```

```
BEGIN Unaligned;  
DIMENSIONS ntax=9;  
FORMAT
```

```
    datatype=STANDARD  
    missing=?  
    symbols="0 1 2 3 4 5 6 7 8 9"  
    labels=left;
```

```
MATRIX
```

```
'LI_01' 0000111111111100000000000,  
'LI_02' 1111000000000000000000000,  
'LI_03' 0000100000000111110000000,  
'MD_01' 000000000000001001011100,  
'MD_02' 00000100010110001000011,  
'MD_03' 000011000100111100000000,  
'MD_04' 0000111111111100000000000,  
'MD_05' 00000100010110001000011,  
'MD_06' 000000000000000001011101;
```

```
END; [Unaligned]
```

Anexo 07: Transcripción de las muestras dubitadas

BVR_2019_07_26_21_54_09_01 (voz1) MD_01				
Línea	Inicio (s)	Tier	Contenido	Final (s)
1	115.75	Voz_2	No.	116.12
2	128.65	Voz_2	Se fueron a (...)	129.57
3	161.88	Voz_2	Y yo le dije y me dijo que estaba malogrado.	163.75
4	255.12	Voz_2	El como se llama, el superior creo que había dicho que si habían visto unas bolsas.	259.39
5	335.3	Voz_2	Vine a cargar el móvil.	336.89
6	344.97	Voz_2	Si, estás con (...)	346.07
7	373.85	Voz_2	Disculpe jefe, eh, mira, yéndonos al grano, perdone, no quiero faltarle el respeto.	379.66
8	381.5	Voz_2	Podemos...	382.25
9	383.78	Voz_2	Hacer un arreglo.	385.02
10	396.76	Voz_2	Negociamos.	397.97
11	399.82	Voz_2	Yo no, no estoy negociando, yo ofrezco mi, yo ofrezco mi cariño.	403.34
12	426.19	Voz_2	Ellos son de la vieja escuela, son personas.	428.76
13	453.43	Voz_2	Ya, ah, me dices eso y desaparecemos acá.	456.4
14	470.1	Voz_2	Cinco, ¿no?	470.94
15	482.95	Voz_2	Si.	483.47
16	497.83	Voz_2	Los trámites han iniciado, pero no es así, yo...	500.46
17	506.39	Voz_2	Esa cantidad ahorita que usted me dice no lo tengo disponible a la, en efectivo.	510.31
18	514.71	Voz_2	Otra cosa sería digamos, que se queden ellos.	516.79
19	519.07	Voz_2	O tendría que llamar a Tacna para que digamos, me puedan traer el efectivo.	522.94
20	526.26	Voz_2	En Tacna tengo yo un negocio, tengo cosas que hacer en Tacna.	528.59
21	529.75	Voz_2	Y la verdad yo ahorita, digamos, hace poco tiempo que me separé de mi esposa.	532.63
22	533.79	Voz_2	Por eso me metí, pues.	535.26
23	540.88	Voz_2	No, tengo un negocio en Tacna, yo.	542.44
24	544.98	Voz_2	Si.	545.32
25	551.67	Voz_2	Ahorita tengo...	552.91
26	555.1	Voz_2	Cuánto ININT, debo tener...	558.08
27	570.73	Voz_2	Ahorita acá, acá, acá, exactamente.	572.49
28	574.42	Voz_2	Usted también está acá, no tengo, digamos.	576.38
29	577.11	Voz_2	En otra palabras, prácticamente, no tengo nada.	579.01
30	580.38	Voz_2	Pues, para qué te voy a mentir, digamos, jefe, decir que si tengo.	583.5
31	585.1	Voz_2	Y yo se que...	586.19

32	586.77	Voz_2	Digamos, de todas las personas que he conocido, la verdad me pareció unas personas correctas, porque...	593.21
33	593.5	Voz_2	Lo que es, lo que es, ya...	594.91
34	597.23	Voz_2	Pero, yo en Tacna te puedo conseguir porque tengo un negocio.	600.08
35	601.56	Voz_2	Justamente hoy día me he ido a entregar la cuenta la chica, como he estado acá estos días.	604.79
36	605.22	Voz_2	Comenzó comentando que yo me quité de Tacna porque tenía un negocio primero en común.	609.81
37	610.44	Voz_2	Nos separamos, empecé mi propio negocio.	612.43
38	612.69	Voz_2	Estaba yendo allá y como había mucho, está que jode la mamá de mi hija, decidí venirme para olvidarme.	619.27
39	664.21	Voz_2	En Tacna, sino iríamos, iríamos hasta a más tardar, hasta las diez, doce, exagerando, yo ten consigo esa cantidad, dos lucas te consigo al toque.	673.88
40	677.99	Voz_2	El resto te puedo dar, digamos, el diez mas o menos.	681.34
41	692.83	Voz_2	Y al toque voy a disponer del efectivo, voy, digamos, mañana, como te digo para las doce, te puedo tener esa cantidad exactamente.	698.02
42	728.5	Voz_2	Mañana yo puedo ir en el carro de las cuatro.	730.4
43	746.35	Voz_2	La verdad para ser sinceros no pasa de cien soles.	748.2
44	753.97	Voz_2	Que te puedo decir, la persona que en Tacna que me lo consigue.	756.74
45	757.17	Voz_2	No pasa de cien soles.	758.16
46	760	Voz_2	Como justamente le estaba diciendo hace rato, se de unas personas ININT que están con RQ.	765.86
47	766.48	Voz_2	Y ya, pues.	767.46
48	773.1	Voz_2	Espere, la verdad no estamos alquilados.	774.83
49	777.4	Voz_2	No estamos, pues, lo que pasa es que, digamos.	779.91
50	780.2	Voz_2	Ahora yo vine para cargar el celular.	781.9
51	782.57	Voz_2	Y cuando estaba cargando mi celular y prendí el celular, no entraba Entel.	785.89
52	786.38	Voz_2	No había señal de Entel.	787.36
53	788.92	Voz_2	Jalo un poco y, justamente en eso llamo y mi papá llama, mi papá me contesta.	794.4
54	795.85	Voz_2	Pero he estado todo el día acá abajo, estoy acá esperando, ¿puedes venir? ya.	799.31
55	870.21	Voz_2	No, lo veo yo solo, digamos.	872.2
56	873.88	Voz_2	Esperando el carro para irme.	875.52
57	879.95	Voz_2	Donde este carro de acá del, del minibus, baja a las cuatro de la mañana, del que va subido, creo que le llaman, un bus grande.	889.21
58	889.85	Voz_2	Y baja a las cuatro de la mañana.	891.12
59	896.66	Voz_2	Si.	897.15
60	903.79	Voz_2	Es, mira, así medio verde creo que es, ¿no?	905.89

61	986.1	Voz_2	Yo no estaba durmiendo, pensé que íbamos a regresar arriba, a esperar, digamos, como viendo un colchoncito que duermen ahí hasta la noche.	992.91
62	997.95	Voz_2	No, no, nos hemos hospedado, estamos ahí, estamos sentaditos, ni si quiera estamos hospedados.	1001.87
63	1032.2	Voz_2	Puede desaparecer, nada más.	1033.44
64	1102.64	Voz_2	¿Puedo conversar con el un ratito?	1104.2
65	1243.61	Voz_2	Yo se que no va a querer esperar hasta mañana, pues, ¿no?	1246.38
66	1253.28	Voz_2	Le voy a decir, que venga (...)	1254.69
67	1270.98	Voz_2	Perdone que dude de la palabra, yo se que ustedes son buenas personas.	1275.19
68	1276.48	Voz_2	Ya me sorprendió cuando, digamos, vi que mi hermano se estaba corriendo y no le han dado un golpe, eso es un policía, porque cualquier otra persona le hubiese metido un par de tacasos.	1286.59
69	1324.35	Voz_2	¿Quién es el que me ha llamado de la comisaría?	1325.65
70	1480.11	Voz_2	A veces tengo, pero sabes como les estoy comentando.	1482.94
71	1486.26	Voz_2	Me pasó un caso una vez, que estaba trayendo carne de, mi papá también era antes carnicero de burro, de repente puede saber.	1490.61
72	1492.11	Voz_2	En Tarata me bajaron.	1493.44
73	1495.95	Voz_2	Ironías del destino, el ININT, el policía ya sabía.	1500.92
74	1504.33	Voz_2	Te la apuesto que ahorita me para.	1505.57
75	1506.46	Voz_2	Ni si quiera arriba, yo ya sabía que me iba a ir por el otro estique, iba a dar mi vuelta y no iba a pasar por el control.	1512.43
76	1513.53	Voz_2	Es para abajo ya, o sea, ¿cómo es que sabía que estaba viniendo así?	1516.73
77	1519.4	Voz_2	Pero al otro yo ya le había dado su plata, pues.	1521.11
78	1567.42	Voz_2	Ustedes hasta cuándo van a estar acá.	1569.13
79	1591.59	Voz_2	Hasta cuándo tú crees que pueda sacar la plata.	1593.9
80	1604.28	Voz_2	Con el otro paradero ININT	1605.9
81	1606.42	Voz_2	El otro, digamos, de las dos lucas si eran de garantía, yo podía trabajar tranquilo.	1610.57
82	1757.4	Voz_2	Las cosas las puedo un ININT por ahí.	1761.13
83	1761.84	Voz_2	O lo subo arriba.	1763.2
84	1764.18	Voz_2	Si me bajo ahorita en, a Tacna.	1765.88
85	1766.34	Voz_2	Y mañana yo subiría.	1768.6
86	1769.14	Voz_2	Le timbraría.	1770.93
87	1794.91	Voz_2	Ah, si, jefe, si.	1796.69
88	1813.25	Voz_2	Para que lleguen a dar, ININT, encuentren mi huella ahí, pero com oyo no toqué, no hay problema.	1817.66
89	1827.27	Voz_2	Es que ahorita, digamos, yo me bajaría para Tacna.	1828.82
90	1829.81	Voz_2	Mañana yo les traería en la mañana.	1832.26

91	1838.2	Voz_2	Pues, igual, tiene ahorita la garantía de cualquier momento puede subir arriba y bajarse todo.	1842.3
92	1889.95	Voz_2	Es una buena, una buena persona porque...	1892.92
93	1893.73	Voz_2	Me dicen, buen acta.	1895.37
94	1936.27	Voz_2	Yo quiero seguir chambeando arriba. Pero, yo, me bajo mañana temprano a las cuatro de la mañana a Tacna, y yo regreso antes del medio día para acá.	1946.17
95	1952.86	Voz_2	Dos lucas.	1953.81
96	1956.5	Voz_2	Chambear arriba, pues, jefe.	1958
97	1962.06	Voz_2	Me deja chambear... unos cuantos días, y yo te consigo, digamos los dos lucas ININT	1966.56
98	1967.54	Voz_2	Déjam chambear unos... diez días, diez días.	1971.01
99	1975.1	Voz_2	Pues, ya, aparte yo también tengo la garantía, que tengo su confianza, ¿no? de que usted me está ayudando.	1982.35
100	2019.99	Voz_2	Esto es bien irónico, el subprefecto, hace creo que quince días o menos fue creo, le pidió plata a mi papá.	2025.13
101	2028.27	Voz_2	El subprefecto, el un señor que vive en la quebrada.	2031.5
102	2033.72	Voz_2	Ya, el le pidió plata, pues, el subprefecto creo que fue, o no se quién, pero le pidió.	2038.28
103	2047.52	Voz_2	Parece que tuvieron un intercambio de palabras, algo así, y ya pues, y por eso él cree que por lo que hace una cosa está sanamente trabajando.	2055.83
104	2057.47	Voz_2	Pero no es así, pues, yo se que no es así, que las cosas están mal, una persona, digamos, yo he estudiado, yo.	2062.81

BVR_2019_07_26_21_54_09_01 (voz1) MD_02				
Línea	Inicio (s)	Tier	Contenido	Final (s)
1	227.57	voz_1	Mira, el superior ya se ha arrancado, nos ha dejado toda la chamba de mierda.	232.01
2	233.75	voz_1	El fiscal ha dicho que se haga lo que se vea, la webada esa cómo es, que se vaya al lugar otra vez,	239.31
3	239.96	voz_1	Hemos ido a recoger la bolsa.	241.54
4	242.38	voz_1	Putas, y hay una docena de cartuchos.	245.73
5	246.17	voz_1	Hay mecha y hay fulminantes.	248.64
6	250.3	voz_1	Tú sabes esa webada o ¿no?	251.47
7	268.16	voz_1	Ni mierda de esa huevada.	269.12
8	271.19	voz_1	Yo ahorita lo voy a ubicar, lo voy a traer a tu hermano.	273.63
9	274.44	voz_1	Se quedan ustedes los tres, con tu viejo más.	276.41
10	276.8	voz_1	La gente que ININT que hemos subido o no hemos subido.	279.6
11	279.81	voz_1	Has visto que hemos subido o ¿no?	281.03
12	282.18	voz_1	¿Dónde está la huevada? ya lo hemos sacado, ya lo hemos revisado.	284.9
13	285.39	voz_1	Así que se va a proceder, ya.	288.27
14	288.27	voz_1	Yo no creo en tus huevadas.	289.72

15	290.42	voz_1	Ya, esas son huevadas, compadre, no, yo a tu viejo le he dicho ya, puta le he sacado Reniec, toda la huevada, ya váyase, que se vaya le he dicho.	299.07
16	299.46	voz_1	Y el superior ya se ha quitado, ahorita lo llamamos, mira compadre, de seguro siguen ahí, ¿han ido a recoger la huevada? si.	305.02
17	314.04	voz_1	ININT, bien claro que se vayan hasta antes que yo saque esa huevada, ¿si o no?	317.13
18	318.29	voz_1	Si o no, han pasado como cinco, como cuatro carros han pasado arriba.	321.23
19	321.7	voz_1	El superior dice que te ha dado una vez.	323.38
20	326.08	voz_1	Y ust+D21:F76edes siguen en esa huevada, hay tantas cosas botadas.	329.2
21	329.53	voz_1	Yo pensé que estabas por ahí, estaba todo botado.	331.42
22	339.29	voz_1	Tu viejo no está, si yo lo he visto a tu viejo ahorita en la, subiendo gradas.	341.78
23	343.99	voz_1	Yo he subido con otro huevón.	345
24	360.46	voz_1	Yo lo saco de ahí, huevón.	362.26
25	385.14	voz_1	¿Qué cosas quieres? Habla, pues, huevón.	386.78
26	405.05	voz_1	Esa huevada, yo te digo ha sido, esa huevada es tu cana para los tres cojudos que están involucrados en eso, les saco la mierda.	410.19
27	410.42	voz_1	Yo te digo, tu que haces, qué, negociar qué, cinco lucas, así como probremente, ah, como amigos.	416.5
28	417.41	voz_1	¿Qué dices?	418.19
29	419.25	voz_1	Y no hay que pensar mucho, oye, yo le he dicho antes a tu hermano, tu viejo es una cagada.	423.85
30	424.83	voz_1	¿Sí o no?	425.63
31	432.57	voz_1	El no entiende que ustedes, los tres se van a ir cana, te van a cagar tu defensa profesional, concha de su madre, se va a la mierda.	437.37
32	438.31	voz_1	Así que tú dirás. Yo en la mañana, yo, a las cinco de la mañana no los quiero encontrar acá.	443.06
33	445.76	voz_1	Esa huevada, puta, no, pe.	447.29
34	449.63	voz_1	Habla.	450.62
35	456.33	voz_1	Ahorita.	457.13
36	465.31	voz_1	Yo no hablo más, ah.	466.64
37	473.52	voz_1	No estás grabando, ¿no?	474.74
38	478.18	voz_1	Mira, en ti confío un poquito, lo más tranquilo, ININT tu hermano es medio cruzado, ¿no?	482.86
39	501.73	voz_1	Tú te quieres coger una cana cojuda, no, ¿si o no?	503.73
40	512.05	voz_1	Esa es la huevada.	513.38
41	524.62	voz_1	Pero quién va a traer, ¿de dónde vas a traer?	526.28
42	537.14	voz_1	Ya, ¿y? qué y ¿qué y que de dónde vas a traer el billete, te vsa a prestar, qué vas a hacer?	540.93

43	542.43	voz_1	Ya, vas a ir a traer el billete ahí.	544.3
44	562.17	voz_1	Al toque, nada más, tú sabes cuánto tienes.	563.94
45	622.86	voz_1	Ya ves, estamos en nada tu viejo, ¿qué dice?	625.62
46	631.31	voz_1	¿Cuál es la solución?	632.64
47	636.79	voz_1	Te vas a Tacna y te apoyo, por las huevas.	638.77
48	647.75	voz_1	Lo llevas en tu carro.	648.76
49	675.94	voz_1	Tu virueta, estamos hablando de cinco palos, pe, huevón.	677.88
50	712.66	voz_1	No, o sea, dice ahorita él no tiene, dice.	714.51
51	715.47	voz_1	Y si para mañana al medio día te queda conseguir los dos lucas, ¿no?	718.59
52	721.36	voz_1	Ya, dos lucas, nada más. Para mañana al medio día, dice.	723.93
53	737.29	voz_1	Ni cagando, huevón, y ¿qué hacemos con esa huevada?	739.26
54	750.09	voz_1	No creo, ah.	750.99
55	803.19	voz_1	Que el mismo que vaya y que lo saque a su viejo, a su hermano.	805.99
56	810.23	voz_1	Ya empecé a chambear, pues, hoy día, ya está perdido el día, mañana, pasado.	813.91
57	815.16	voz_1	Hasta pasado mañana, son dos, cuarenta y ocho horas, ¿no?	818.33
58	834.32	voz_1	Ustedes no piensan en la magnitud de los hechos, huevón.	837.56
59	842.27	voz_1	Ya tu dirás de dónde has sacado sus manifestaciones, eso, la procedencia de esa huevada, ya en Tacna, canearán a esa gente también, no se, porque yo acá por lo que tengo, sobre esa base.	853.33
60	856.66	voz_1	De ese concha de su madre también le van a reventar el culo, ve.	858.87
61	859.11	voz_1	Si el canta, pues, ¿no? va a decir de dónde lo ha comprado, tiene que decir de dónde lo ha comprado.	862.51
62	878.05	voz_1	¿Qué carro te vas a ir, pues, si de la policía han bajado carros?	879.89
63	894.04	voz_1	Ah, ya, ¿cómo se llama esa huevada? señor (...)	896.74
64	907.35	voz_1	Con eso abajo nos puede.	908.94
65	913.74	voz_1	Pueden sacar a más de uno.	915.11
66	922.51	voz_1	No hay, pues, un.	923.65
67	926.15	voz_1	No lo creo yo huevón, esa huevada, aparte eso ya es, trae espinas, ya.	930.01
68	936.74	voz_1	Hay que hacer otra acta para incluir esa huevada, esa máquina de mierda que tiene. Haz otra acta, huevón.	942.27
69	943.16	voz_1	Huevón de (...) se mete en huevadas, oye, por eso quiero que se largue, huevón.	947.55
70	950.59	voz_1	Acá estoy con toda la chamba.	951.88
71	969.52	voz_1	El (...) de mierda que se deje de huevadas también.	971.18
72	972.06	voz_1	Toca la puerta defrente, compadre, y que salga.	973.7
73	978.09	voz_1	Tú también estás ahí con ellos, ¿no? ahí está (...) también.	980.53
74	1011.16	voz_1	Porque es por las huevas, huevón, vas a esperar hasta mañana por gusto, y ¿después?	1014.33
75	1015.24	voz_1	Las cosas son calientes, pe, este, mi querido (...)	1018.09

76	1018.56	voz_1	Uno te quiere apoyar, pero no se puede, si no se puede, caballero nada más, lo que venga.	1022.64
77	1023.31	voz_1	Esa huevada de mañana es una huevada.	1025.62
78	1026.99	voz_1	Putita, mañana nos quedamos con la huevada esa allí, qué hacemos con esa cochizada.	1031.38
79	1039.82	voz_1	Huevón vamos a entrar hasta abajo, la gente ha salido huevón.	1042.16
80	1044	voz_1	Nadie, de acá nadie es tonto huevón. Que chucha, el huevón fue tu hermano, para que mierda fue a botar.	1048.54
81	1049.89	voz_1	Para qué fue a botar.	1051.24
82	1052.41	voz_1	Por cojudo ha ido a botar, pues.	1053.86
83	1061.81	voz_1	Están los fulminantes, están los cachorros, está la huevada esa, la...	1065.66
84	1066	voz_1	la mecha.	1066.59
85	1068.79	voz_1	Envuelta en dos pantalones ININT, ah.	1072.66
86	1076.42	voz_1	Vamos nada más te dije, ya no hay nada, ¿qué vas a hacer?	1079.35
87	1081.94	voz_1	Vamos a continuar, nada más, pues, este, (...)	1084.85
88	1085.57	voz_1	Caballero, pues, huevón.	1086.85
89	1089.36	voz_1	Allá no me hables nada.	1090.82
90	1090.84	voz_1	Esa huevada, graba.	1092.43
91	1092.92	voz_1	Graba, tiene audio y firma.	1095.6
92	1096.77	voz_1	Eso también, ¿ya?	1098.74
93	1099.07	voz_1	Calladito nada más.	1100.19
94	1148.76	voz_1	Antes del medio día.	1150.21
95	1151.17	voz_1	Y ¿cuánto vas a traer?	1152.29
96	1154.31	voz_1	Y ¿qué quieres con eso?	1155.51
97	1158.36	voz_1	Ah, tu cana, es tu hermano.	1161.45

BVR_2019_07_26_21_54_09_01 (voz4) MD_03

Línea	Inicio (s)	Tier	Contenido	Final (s)
1	643.81	Voz_04	Pero, igual, pues, estamos hoy día, mañana, pasado, pues, tiene tres días, normal.	646.75
2	656.71	Voz_04	No se, pues, ¿qué dirá el hombre? pues, si es que tiene, sino va a ir por gusto.	659.65
3	686.21	Voz_04	No voy a, me voy a ir y voy a estar, este.	688.51
4	724.53	Voz_04	Si bajamos ahorita.	725.28
5	770.44	Voz_04	Ese fulano donde están alquilando ahí.	772.14
6	867.38	Voz_04	O se han metido en un hueco, al menos.	868.91
7	1133.11	Voz_04	Dígale que te regrese, pues.	1134.32
8	1134.9	Voz_04	¿Espera a que subas arriba?	1136.29
9	1137.21	Voz_04	Si, esa es, esa es.	1138.77
10	1147.87	Voz_04	Ahorita el... el día.	1149.77
11	1149.95	Voz_04	Y el también es buena punta, sino, te he dicho, como te hemos dicho.	1152.98

12	1153.58	Voz_04	Dale su vaina.	1154.45
13	1160.48	Voz_04	Probar, nada más le ha dicho, quiero que le de su voluntad, ¿cuánto ya?	1163.85
14	1174.78	Voz_04	No, dale su vaina y...	1176.46
15	1176.51	Voz_04	Normal, pues, te apoyamos, y te vas arriba, pues, al toque te ININT, esta noche.	1184.29
16	1184.98	Voz_04	Que subes arriba y ya no bajas.	1186.3
17	1210.01	Voz_04	Te internas esta semana, y...	1211.45
18	1212.72	Voz_04	Ahí arriba, nada más te quedas, no bajas.	1214.22
19	1287.46	Voz_04	Dale un tiempo, no te vas a molestar, te damos la palabra, anda, chambea, compadre, métete ahorita nada más, vete y chapa tus cosas y vete arriba.	1296.06
20	1341.69	Voz_04	Por eso te estoy diciendo, dale, y vete ahora mismo.	1344.09
21	1350.66	Voz_04	Jefe, tengo esto.	1351.4
22	1353.06	Voz_04	Ya, compadre, nos vamos.	1354.79
23	1362.63	Voz_04	Tengo esto, gracias.	1364.3
24	1374.57	Voz_04	Sales y te vas, como a irte, fun, vete arriba nada más, ya.	1378.53
25	1422.98	Voz_04	Están comprando y están subastando, normal.	1424.83
26	1436.96	Voz_04	Anda al tío, dale tu cariño, ya...	1438.63
27	1446.85	Voz_04	Dale al tío, dale al tío, normal.	1448.7
28	1478.12	Voz_04	¿Tienes, no?	1478.67
29	1539.51	Voz_04	Devolvemos tu vaina, vete a trabajar ya...	1541.24
30	1563.46	Voz_04	Cuatro de la mañana vas.	1564.27
31	1575.54	Voz_04	El domingo para comprarte agua, ya, pues, te quedas si quiera hasta el jueves, quédate porque nosotros subimos jueves y miércoles, subo, miércoles subo y me quedo todavía hasta el domingo	1583.96
32	1633.68	Voz_04	Quiero agua, hasta agua de pasada te puedo botar, yo.	1635.96
33	1647.4	Voz_04	O puedo botarte acá.	1648.74
34	1661.38	Voz_04	Dale al tío, porque ese tío te va a paoyar bastante, cualquier cosa que haya, te va a apoyar, huevón.	1665.37
35	1669.52	Voz_04	Así como patas te hablo, yo te hablo y...	1672.81
36	1673.54	Voz_04	Yo soy una persona diferente, no, no me gusta ese tipo de cosas.	1676.48
37	1727.08	Voz_04	Yo vine recién, porque todavía tengo para rato.	1729.44
38	1787.11	Voz_04	Tenía las mochilas.	1788.49
39	1843.85	Voz_04	No ya, lleva tus cosas, ahorita vamos a decirle que se, no tiene, por las huevas es.	1847.94
40	1857.15	Voz_04	¿Cómo pensabas hablarle al comisario?	1858.85
41	1866.99	Voz_04	Habla (...) al toque, sin mentiras.	1868.72

BVR_2019_07_26_21_54_09_01 (voz3) MD_04				
Línea	Inicio (s)	Tier	Comentario	Final (s)

1	1205.87	Voz_03	Tienes estos días para chambear, ven conmigo en la noche trabajas (ININT)	1208.22
2	1222.35	Voz_03	Darle su vaina ya.	1223.77
3	1241.11	Voz_03	Dale su vaina a ese.	1242.45
4	1302.63	Voz_03	La condición de ustedes ahorita.	1304.23
5	1357.59	Voz_03	Mi tío no es el primero que trabaje.	1359.09
6	1365.74	Voz_03	Mire tío, ahorita, te va a apoyar, usted se va a ir a ese (...) le va a decir que no hable huevadas.	1370.61
7	1414.6	Voz_03	Ya no importa quédate toda la semana.	1416.1
8	1418.01	Voz_03	Hasta el día sábado, nosotros estamos el viernes acá.	1420.07
9	1425.97	Voz_03	Lo trabajen con (...) pero ya no revienten si es reventón.	1428.72
10	1432.14	Voz_03	Trabajen con cuidado, no vaya ser que se les caiga encima, puta, y te van a bancar.	1435.31
11	1460.28	Voz_03	Eso fue en tu tiempo.	1461.35
12	1528.42	Voz_03	Pasa cualquier cosa nos van a cagar a nosotros. Ya (...) el trato es entre nosotros, nada más.	1532.69
13	1559.52	Voz_03	Si quieres bajar a Tacna, baja en la madrugada.	1561.45
14	1697.81	Voz_03	Vamos a chambear como patas.	1699.36
15	1699.82	Voz_03	Y que te va a decir, ya te va a decir.	1701.44
16	1711.36	Voz_03	Tres palos.	1712.25
17	1723.69	Voz_03	Esperemos que cuatro o cinco meses más, quién sabe, huevón.	1726.2
18	1737.41	Voz_03	Tu tío también, recién.	1738.73
19	1745.22	Voz_03	Tu tío tiene para rato, pues.	1746.26
20	1782.14	Voz_03	Oye, entonces ese floro que metiste arriba que tenías ININT.	1785.16
21	1871.4	Voz_03	Solo te arregla por, esos hijos de puta, cobran de diez para arriba, por eso el comisario está allá, pues, la huevada no había.	1876.23
22	1884.3	Voz_03	Tu tío es fuerte, todos los que trabajan en esas áreas, puta, que no.	1888.4
23	1899.45	Voz_03	Y al final es el tío, simplemente el que yo esté a carago de esa huevada.	1902.65
24	1904.05	Voz_03	Porque estoy, por ellos nada más estamos viniendo, sino, puta, no, no vendría.	1907.56
25	1912.47	Voz_03	O como sea estoy pensandole, déjalo ahí nomás, puta, desaparecer o nada.	1916.99

26	1918.5	Voz_03	Esos huevones de, tengo que dar trámite todo, si tu vas, puta, llévalos como solo, pues, huevón, si lo vas a comer solo, si pasa cualquier cosa.	1925.41
----	--------	--------	--	---------

BVR_2019_07_26_22_30_26_01 (voz 2) MD_05				
Línea	Inicio (s)	Tier	Contenido	Final (s)
1	10.1	voz2	¿Qué tal el superior?	11.18
2	12.22	voz2	A ver, vamos a ver, compadre.	13.28
3	13.51	voz2	Hoy día.	13.98
4	23.07	voz2	Él ha subido hasta donde dicen que han llegado antes de ayer.	24.68
5	25.63	voz2	Antes de ayer han subido me han dicho, pues.	27.65
6	28.13	voz2	Y están cerca.	29.17
7	32.09	voz2	El ha subido por eso es que él conoce.	33.73
8	40.71	voz2	Cómo si acá dice (...) que esa huevada es, para mi ya no...	44.26
9	44.69	voz2	Esa forma no me bacila.	46.22
10	47.98	voz2	Es peligroso.	48.91
11	55.71	voz2	Huele a mancada, son huevadas.	58.19
12	58.74	voz2	Eso es, hasta las huevas es esa huevada.	60.5
13	62.4	voz2	¿Qué tal cojudo como te digo?	64.02
14	64.88	voz2	Así como tú dices, estamos hablando en serio.	67.48
15	67.92	voz2	Ya, mañana te vas a dejar tu huevada, que acá dices, de acá a una semana, diez días.	72.47
16	73.25	voz2	Y qué pasaría si el superior sube (ININT).	75.27
17	78.5	voz2	Acá la cosa como te digo tiene que ser seria, tú te la comes, tú, nada que tu hermano, tu papá, no tiene que saber ni mierda.	84.88
18	85.31	voz2	Peor la gente acá donde estás, no sabe que es (ININT)	88.11
19	95.58	voz2	Tu vienes un día en la madrugada, en la noche, no se, compadre, un carro particular.	99.68
20	100.79	voz2	Baja tu huevada, baja, baja la noche, ININT en la noche, no se, que no te vea, o busca otra ruta, no se, no, pua, tu ve, pero que no te vean.	111.76
21	123.61	voz2	Esa huevada, no se, ¿qué es tan temprano?	125.72
22	129.18	voz2	¿Mañana a qué hora es que vas a venir dices?	130.56
23	150.64	voz2	Que no venga tu hermano ni tu viejo, porque ese huevón que está en la comisaría lo conoce muy bien, a ti no te conoce.	155.12
24	161.29	voz2	A tu hermano lo conoce muy bien.	162.73
25	191.14	voz2	Este papeles tengo, como te he dicho, el objeto es bien claro.	193.57
26	194.18	voz2	Ellos por	195.97

27	203.92	voz2	Autorizo al señor (...)	206
28	206.26	voz2	Para que pueda... ¿no?	207.3
29	207.85	voz2	Cavar, sacar (...) de tal sitio, del lugar.	210.3
30	222.13	voz2	Ahí está, huevón, tiene documentos.	223.54
31	242.12	voz2	Ahora si yo lo, tú quieres quitar un billete como dices ¿no?	244.71
32	245.93	voz2	No lo hagas a la maldad ni me vengas con huevadas.	248.09
33	249.3	voz2	Mañana firmo, hago una huevada, estoy firmando.	251.55
34	252.44	voz2	Que uno también se cuide de esa huevada, huevón.	254.03
35	254.69	voz2	Son jóvenes, yo todavía por irme a la calle	256.59
36	257.63	voz2	Por salvarte una huevada que te guste.	259.13
37	260.2	voz2	Profeo como dice, comote digo lo escaneo también.	262.31
38	262.65	voz2	Eso habrá responsabilidad, de repente uno de ustedes se come, ¿no? ¿qué va a decir el fiscal? los tres, ¿los tres han estado? los tres.	267.9
39	268.71	voz2	A la mierda ha sido, como tu viejo puede ser porque ha estado arriba, han estado los dos.	272.09
40	273.85	voz2	Y ahora el (...) los testigos van a decir ¿quién ha botado? este ha botado, acta de reconocimiento, ni lo va a ver él, él lo ha botado.	279.62
41	281.54	voz2	Cagado.	282.47
42	282.99	voz2	Ustedes los dos, entonces los dos saben.	285.12
43	285.55	voz2	Y a ti ya te preguntarán que tú lo apoyaste a tu superior.	287.75
44	292.25	voz2	Y eso está penado.	293.46
45	293.87	voz2	Creo que es de seis a doce años.	295.63
46	296.12	voz2	Ya, pues, que te comas, pues, unos, la mitad nada más que te comas, ya, pues, seis.	298.95
47	299.49	voz2	Te das cuenta o ¿no?	300.36
48	300.72	voz2	Al dos por uno y la buena conducta, mira te comes seis años.	304.53
49	308.11	voz2	Peligro común es esa huevada. Es este, eh, hacer uso de explosivos de uso civil sin tener la debida autorización y, ¿no?, por manipulación, es más, ¿por dónde lo has comprado? legalmente, huevón.	320.18
50	320.51	voz2	Esa es otra pena.	321.67
51	324.09	voz2	Hay, este..., entidades.	326.92
52	327.53	voz2	Exclusivas y especiales para comprar, ¿no?, en la tienda no te van a vender, pues, una docena así el de aquí, ¿cómo se llama esa huevada?	333.7
53	338.89	voz2	Es una cola, te va a hasta el Ministerio contra el Terrorismo, otra huevada.	343.77

54	344.11	voz2	¿De dónde has sacado? ¿quién te ha dado? No, eso he comprado en actividad, estás, pom, rompen esa huevada, pom, en cana, ah, no, concha de su madre, a ti, ati te he vendido, puta, tuú me has soplado, jwa!	355.39
55	363.76	voz2	Tú vienes y me timbras.	365.17
56	368.12	voz2	No te acerques a la comisaría.	369.41
57	377.86	voz2	Mañana vamos a dejarlo como está a las nueve, diez, por lo ININT	380.78
58	381.73	voz2	Ya, y me dejas tu mensaje, a la mierda.	384.15
59	384.47	voz2	Esa huevada no es.	385.31
60	386.32	voz2	Llegas y ya sabes que estás en tal sitio y listo.	388.65
61	389.49	voz2	Como dice el hombre, ¿ya? a la mierda.	391.43
62	392.06	voz2	Y, saca tu merca como dices.	394.02
63	396.24	voz2	De acá a los diez días que dices, cumples, si estás en tu huevada, normal.	400.05
64	400.71	voz2	Pero si, lo que yo quiero que tu estés, una señal donde chucha, o compadre, ¿sabes qué? arráncate.	406.45
65	407	voz2	Están subiendo, el día que yo no estoy de repente.	408.93
66	410.89	voz2	Te van a encontrar y vas a sonar.	412.74
67	423.76	voz2	Ya hasta hoy día, pero por mi, ya te has ido, ya.	426.24

BVR_2019_07_26_22_30_26_01 (voz 1) MD_06				
Línea	Inicio (s)	Tier	Contenido	Final (s)
2	0.41	voz1	No, generalmente, digamos.	1.74
3	2.29	voz1	No, a esa zona no entra el señor.	3.9
4	7.42	voz1	Se me ha acercado una persona.	8.93
5	9.18	voz1	Que estoy viendo.	9.96
6	17.58	voz1	A ciencia cierta, no creo que pueda llegar el superior tan lejos.	20.15
7	92.32	voz1	Para irme.	93.27
8	130.94	voz1	A más tardar, a las dos estaré acá.	132.38
9	145.71	voz1	Por eso, ¿cómo me voy a jugar con algo que?	147.61
10	172.59	voz1	El jefe dijo que lo que le quería decir es diferente.	174.81
11	176.37	voz1	Si... están digamos allá.	178.85
12	179.17	voz1	Ya estaríamos saliendo.	180.53
13	180.93	voz1	Saliendo, yo te presento que íbamos allá, que allá, ustedes están con ese conocimiento.	185.52
14	185.98	voz1	Con ese conocimiento ya iríamos, estaríamos, todavía un poco más tranquilo para yo poder trabajar, ¿no?	190.31
15	196.27	voz1	El trámite ¿cómo es?, digamos.	197.83
16	224.81	voz1	Falta la gente de verdad.	226.63

17	226.83	voz1	Se siente... menos.	228.5
18	229.02	voz1	Siempre yo se, manejo carro.	230.9
19	231.44	voz1	Y la única persona que siempre te ayuda.	233.89
20	234.35	voz1	Malos o buenos, son policías.	235.82
21	418.19	voz1	Usted me dice cuando ya los	419.23



Anexo 08: Transcripción de las muestras indubitadas

LI_01_O.C.M.				
Línea	Inicio (s)	tier	Contenido	Final (s)
1	0.6	Entrevistador	Empecemos, Tacna...	4.92
40	5.09	LI_01_O.C.M.	Tacna, hoy día estamos 27, ¿no?	7.71
2	7.72	Entrevistador	Si.	7.94
41	8.33	LI_01_O.C.M.	Veintisiete de abril de 2022, mi nombre completo es (...). El número de mi DNI es (...), tengo 27 años, nací en el departamento de Puno, provincia de Chucuito, distrito de Huacuyani, me dedico a la Policía Nacional del Perú, el grado de estudios que poseo es técnico superior, mi lenguaje materno es el español, mi papá nació el once de abril de 1968 y mi mamá nació el 23 de abril de 1968 también, mis abuelos se dedican, yo tengo una abuela nada más.	62.81
3	65.05	Entrevistador	Ya, vamos, ahora si vamos a una entrevista te voy a hacer antes de que prosigas, ¿ya?	68.67
42	68.75	LI_01_O.C.M.	Ya.	68.98
4	70.4	Entrevistador	Usted me dice que nació en ¿dónde?	72.13
43	72.48	LI_01_O.C.M.	En Puno.	73.25
5	73.33	Entrevistador	Puno.	73.65
44	73.76	LI_01_O.C.M.	Si.	73.98
6	74.1	Entrevistador	¿Qué ha sido, cómo así que usted? ¿Radica en Puno o en Tacna?	79.25
45	79.26	LI_01_O.C.M.	No, de Puno, nos vinimos a Tacna.	81.4
7	82.13	Entrevistador	¿A qué edad usted tenía? ¿a los cuántos años?	83.73
46	83.74	LI_01_O.C.M.	Ah, no, yo me vine recién a cuando culminé la secundaria, ahora estoy en Puno. Ya toda mi vida he estado en Puno, primaria, secundaria todo en Puno. Y acá llegué cuando terminé la secundaria en el año dos mil doce, trece, en el trece he llegado a Tacna.	98.66
8	98.94	Entrevistador	Dos mil trece.	99.75
47	99.84	LI_01_O.C.M.	Si.	100.06
9	100.18	Entrevistador	Eh, según usted, ¿cuál es la diferencia, por ejemplo, la vida en Puno y la vida en Tacna?	108.05

48	108.2	LI_01_O.C.M.	Ah, no, lo que pasa es que toda mi familia se vino acá, o sea, todos están viniendo acá, mis padres también se vinieron, por eso yo también me vine de Puno. Yo estaba viviendo en Puno con mis abuelos, nadie más, con mis abuelos, este, por parte, ¿no? de mi mamá. Como todos, mi papá, mi mamá, todos se vinieron a Tacna, entonces yo me vine de Puno, ya. Terminé la secundaria, un año más me quedé y me vine.	124.72
10	125.12	Entrevistador	¿Dónde hace más frío, en Puno o en Tacna?	126.87
49	127.3	LI_01_O.C.M.	En Puno, pues.	127.89
11	128.6	Entrevistador	Puno.	128.85
50	129	LI_01_O.C.M.	Claro.	129.33
12	129.43	Entrevistador	Y, usted vivía en el mismo Puno o, o sea, en la misma ciudad.	132.06
51	132.22	LI_01_O.C.M.	Ah, no, no, no, no, yo vivía en el campo. Vivía en el distrito de Huacuyani.	135.98
13	136.21	Entrevistador	Huacuyani, ¿Huacuyani está hacia lo que llaman desahuadero?	139.85
52	140.24	LI_01_O.C.M.	Claro. Si.	141.18
14	141.35	Entrevistador	Hacia ese lado.	142.1
53	143.11	LI_01_O.C.M.	Es este, colindante a ver con Mazocruz, no se si conoce Mazocruz, distrito de Mazocruz.	148.74
15	149.36	Entrevistador	¿Ya?	149.66
54	149.9	LI_01_O.C.M.	Si.	150.18
16	151.28	Entrevistador	Toda esa parte lo sigue abarcando el lago Titicaca o ya, hasta cierta parte.	154.97
55	155.32	LI_01_O.C.M.	No, lo que es Huacuyani, no. Solamente la parte de desahuadero, eso esta alejado ya, esta... a una hora, una hora de dedesaguadero está, ya no...	164.28
17	164.45	Entrevistador	O sea, está más lejos.	165.07
56	165.54	LI_01_O.C.M.	Claro	166.03
18	166.74	Entrevistador	Ah, mira, y eso no pertenece a dónde, ¿a Bolivia?	169.97
57	170.4	LI_01_O.C.M.	No, no, no.	171.04
19	171.15	Entrevistador	¿Perú?	171.53
58	171.75	LI_01_O.C.M.	Perú, Huacuyani pertenece a Perú, es un distrito.	174.07
20	175.26	Entrevistador	Bien	175.54
59	176.06	LI_01_O.C.M.	Claro.	176.31
21	176.46	Entrevistador	Y... ¿cómo era ahí el día a día con tus amigos?	180.81

60	182.06	LI_01_O.C.M.	Ah, no, como le dije, vivíamos en el campo, mayormente ¿no? la pasaba con mis animalitos jugando, más que todo ayudaba a mis abuelos, como estaba en el colegio, está igual, ¿no? iba al colegio, después del colegio salía, me iban donde mis abuelos, mi abuelo también era profesor. Y, entonces, salía del colegio me iba al toque a la casa a ayudar a mi abuela, pues. Con los animalitos.	200.99
22	200.62	Entrevistador	Y de qué nivel era profesor tu abuelo.	203.67
61	203.88	LI_01_O.C.M.	Ah, era de primaria,	204.75
23	204.97	Entrevistador	Ah, de primaria.	205.59
62	205.74	LI_01_O.C.M.	Ah, si profesor de primaria, ya falleció ya él, en el dos mil veinte, pues, apenas empezó el covid, en dos mil veinte falleció, igual vino acá a Tacna y acá falleció.	217.37
24	218.7	Entrevistador	Y, dígame, este, qué le parece la diferencia entre las comidas de Tacna y Puno. O tiene mucho de central...	228.64
63	229.24	LI_01_O.C.M.	No, para mi hay un poco la diferencia, allá en Puno se come más natural, pues. En Puno es más natural las cosas. En ca...	236.66
25	236.64	Entrevistador	Comen más condimentado, algo así.	237.91
64	237.96	LI_01_O.C.M.	Claor, casi un poco más condimentado. Allá lo que encuentras, pues, no, lo del campo, la quinua, la cañihua, la papa, el chuño, ¿no? eso es, en cambio acá, no, acá un poco está la vida, está caro la ciudad,	253.18
26	253.94	Entrevistador	Uhm, y... le gusta los picantes de, de Tacna.	258.75
65	259.63	LI_01_O.C.M.	De Tacna, sí, si he probado el picante.	262.41
27	263.5	Entrevistador	Y díganos, este... cuando usted quiere recrearse en algún lugar aquí de Tacna ¿a dónde va? o ¿cómo, hace deporte?	272.72
66	273.68	LI_01_O.C.M.	Si, a veces hago deporte con mis colegas, ¿no? a veces también con los familiares. Con ellos hago deporte.	281.67
28	282.23	Entrevistador	Fútbol.	282.57
67	282.81	LI_01_O.C.M.	Sí, futbol.	283.61
29	283.93	Entrevistador	¿Y en qué posición usted juega?	285.63
68	285.71	LI_01_O.C.M.	Futbol fluído ¿no? eh, delantero.	287.6
30	287.99	Entrevistador	¿Delantero?	288.6
69	288.68	LI_01_O.C.M.	Si.	288.96
31	289.27	Entrevistador	Y recuerda alguna vez donde usted había hecho, no se, muchos goles, quizá a un equipo.	293.9

70	294.51	LI_01_O.C.M.	Bueno, ya, ahora no mucho, sino que hace cuando estaba en la sierra con mis abuelos viviendo, un tiempo me enfermé y ya no volvía a jugar como antes.	303.48
32	304.9	Entrevistador	¿Fácil eres goleador?	305.71
71	305.78	LI_01_O.C.M.	Si, cuando estaba en el colegio, la primaria, si. Me gustaba jugar de delantero bastante.	310.42
33	310.99	Entrevistador	Y has llegado a jugar así, intercolegios.	314.86
72	315.03	LI_01_O.C.M.	Claro, allá en la sierra si, la sierra es muy distinto que acá en la ciudad. Por ejemplo allá en Puno se valora bastante el deporte, hay bastante intercambio, ¿no? de lo que es entre escolares. Se juega bastante el fútbol allá.	327.94
34	328.06	Entrevistador	Y son más competitivos.	328.76
73	328.78	LI_01_O.C.M.	Claro, más competición, si, allá se juega bastante.	331.33
35	332.99	Entrevistador	Bien, hemos terminado esa parte de la entrevista, ahora vamos a continuar aquí y tiene que leer desde la siete a la diez, por ejemplo, llave, payaso, todo eso, hasta ahí.	342.18
74	344.72	LI_01_O.C.M.	Llave, payaso, zpallo, gallina, pera, árbol, collar, ratón, perro, Enrique, roto, gorro, blusa, plato, flor, clavo, globo, libro, profesor, fresa, cruz, tigre, dragón, tren.	369.55
75	373.28	LI_01_O.C.M.	Trentaicuatro, trescientos veinticinco, dos mil trescientos sesenta, noventa y cuatro, noventa y ocho, treinta y uno, quinientos noventa y ocho.	385.28
76	396.09	LI_01_O.C.M.	i	396.88
77	397.63	LI_01_O.C.M.	e	398.43
78	399.09	LI_01_O.C.M.	a	399.97
79	400.62	LI_01_O.C.M.	o	401.54
80	402.1	LI_01_O.C.M.	u	403.03
81	406.51	LI_01_O.C.M.	Lunes, martes, miercoles, jueves, viernes, sábado, domingo, uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, once.	424.03
82	425.35	LI_01_O.C.M.	Veinte, veinticuatro, cien, quinientos, novecientos noventa y nueve.	429.96

83	431.21	LI_01_O.C.M.	Aló, aló, aló, ya, ya, pe, ya, pues, no, no, pe, buenas tardes, buenos días, buenas noches, por favor, si, claro, si, pe, si que si, su canto, su voz. Comunícate ahora, no, no se, no quiero, entiendes, aló, me entiendes, ¿qué has hecho? ¿qué tal? ¿cómo estás? Hola, buenos días, no está, ¿quién eres? no, no es así, soy tal o cual, y tú. No se, si, así, cuando tú me digas la razón del sí, como también el porqué, tú podrás comprender si tengo la razón y no tú.	495.68
36	496.35	Entrevistador	Ahora leemos el párrafo de aquí hasta aquí.	501.43
84	501.57	LI_01_O.C.M.	Cuando me disponía a venir a Lima, conocí a Don Guillermo, quien muy amablemente me invitó a subir a su camión, en donde transportaba gallinas y pescado desde Yurimaguas hasta la capital, subí a la Oroya, en el trayecto Don Guillermo me contó lo que le pasó en el pueblo de San Roque cuando regresaba de Yurimaguas trayendo carga.	527.54
37	528.97	Entrevistador	Muy bien, vamos ahora con el trabalenguas, repite de aquí a acá.	537.05
85	537.73	LI_01_O.C.M.	Rasputín el ratón. Rasputín era un ratón que rascaba en una risca, en un toscó, rasca, risca, rascador, rasca que rasca, acabo con el risco, rasca la risca, rasca un rincón.	552.54
38	553.72	Entrevistador	Más rápido, como ya la sabe.	558.77
86	559.48	LI_01_O.C.M.	Rasputín era un ratón que rascaba en una risca, en un toscó, rasca, risca, rascador, rasca que rasca, acabo con el risco, rasca la risca, rasca un rincón.	570.7
39	570.85	Entrevistador	Para finalizar le voy a pedir que describa la primera imagen con eso acabamos. Usted observa, usa la imaginación, la observación y va describiendo qué es lo que hay.	581.7
87	582.79	LI_01_O.C.M.	Ah, ya, están en una ciudad, ya, creo que hay un colegio si no me equivoco, hay bastantes niños que están jugando al fútbol, dos niños, hay una niña que está en el columpio, acá dos niñas jugando con el globo, hay un resbaladizo, hay un señor con su hijo, dos ancianos conversando, hay una pileta, hay áreas verdes y árboles.	613.41

LI_02_M.E.M.V.				
Línea	Inicio (s)	Tier	Contenido	Final (s)
39	0.4	Entrevistador	Cuenta del 1 al 5 primero.	1.51
1	1.88	LI_02_M.E.M.V.	Uno, dos, tres, cuatro, cinco.	5.81
40	6.27	Entrevistador	Mantiene ese nivel, ese volumen de voz.	8.03

2	12.55	LI_02_M.E.M.V.	Ah, ya.	12.9
41	13.84	Entrevistador	Ahora vamos con los datos. Tacna, la fecha de hoy y sus datos personales.	18.02
3	19.15	LI_02_M.E.M.V.	Tacna, 27 de abril del 2022. Mi nombre completo es (...). El número de DNI es (...). Tengo 55 años. Nací en el departamento de Madre de Dios. Provincia Tambopata, distrito Puerto Maldonado. Me dedico soy suboficial superior de la Policía Nacional en actividad. El grado de estudios que poseo es Técnico Superior. Mi lengua materna es español. Mi papá nació en Cusco. Y mi mamá en Madre de Dios. Mis abuelos se dedican, están fallecidos.	65.32
42	66.73	Entrevistador	Bien, ahora comenzamos con la parte de la entrevista. Usted me dice que es de Madre de Dios, ¿de qué parte de Madre de Dios es usted?	76.03
4	76.47	LI_02_M.E.M.V.	No conozco Madre de Dios, nací en Madre de Dios y a los meses, antes de cumplir el año, eh, me trajeron mis padres a Arequipa es donde he vivido y me he criado.	85.31
43	86.17	Entrevistador	¿Arequipa?	86.66
5	86.8	LI_02_M.E.M.V.	Si.	87.03
44	87.8	Entrevistador	En Arequipa, ¿en qué barrio creció usted?	90.22
6	90.72	LI_02_M.E.M.V.	Inicialmente crecí en el distrito de Alto Selva Alegre y posteriormente, en Cerro Colorado.	97.88
45	98.93	Entrevistador	Selva Alegre, ¿eso está por?	100.17
7	104.68	LI_02_M.E.M.V.	Bueno, este, esa es la parte alta de Arequipa.	106.65
46	106.84	Entrevistador	Es la parte alta de Arequipa, ¿no?	107.94
8	107.94	LI_02_M.E.M.V.	Si.	108.21
47	108.5	Entrevistador	Y, ¿qué tal es vivir por ahí, ¿cómo lo describiría usted?	113.59
9	114.07	LI_02_M.E.M.V.	Ah, bueno, este, cuando yo he vivido por ahí ha sido hasta cuando era adolescente, el clima es el mismo que tiene Arequipa, de 2400 metros de altura sobre el nivel del mar. En tiempo de invierno es, este... frígido. Casi helado. En tiempo de verano, es, este..., que son entre los meses de diciembre a marzo, son de lluvia.	136.39
48	137.24	Entrevistador	¿Y las lluvias son intensas, o cómo es la lluvia?	139.64
10	139.7	LI_02_M.E.M.V.	En algunos, algunos años, las lluvias son intensas, algunas son escasas y en algunos años inclusive no llueve, y si llueve, llueve poquísimo.	147.26
49	147.51	Entrevistador	Ah.	147.66
11	148.01	LI_02_M.E.M.V.	Es variado.	148.76

50	149.02	Entrevistador	Es variado, y uno, por ejemplo, este..., si va a Arequipa, tiene un día, ¿qué debería conocer en un día? ¿qué le recomendaría?	160.75
12	161.74	LI_02_M.E.M.V.	Si solamente tienen un día, en todo caso tendría que ser solamente el centro de la ciudad, recomendaría el monasterio de Santa Catalina.	168.61
51	171.16	Entrevistador	Quizá la catedral también.	172.78
13	172.83	LI_02_M.E.M.V.	Podría, si, la catedral, porque están relativamente cerca a unas cuadras nada más.	177.85
52	178.31	Entrevistador	Claro, y el volcán Misti, creo que hay un mirador.	183
14	183.9	LI_02_M.E.M.V.	Mire, no conozco el volcán, es que el volcán Misti lo conozco de lejos. Pero no se si existe un mirador en el volcán. Pero el volcán, si hablamos a la inversa, de todo Arequipa se lo ve.	195.47
53	196.38	Entrevistador	Claro, en el caso, me parece que es Yanahuara.	199.11
15	199.34	LI_02_M.E.M.V.	Ah ya, existe un mirador, que es el mirador de Yanahuara, lugar turístico que queda en la plaza de Yanahuara. Y si, de ahí se ve el volcán. Y parte de Arequipa, ¿no?, de la ciudad de Arequipa.	210.85
54	211.64	Entrevistador	Y a usted le gusta, por ejemplo, hay, como, una especie de sarsa con cebolla, con pansa, de patita, de cerdo, ¿algo así?	221.68
16	222.66	LI_02_M.E.M.V.	Sarsa de patitas. Claro, preparado en base a patitas de cerdo y...	228.34
55	228.96	Entrevistador	¿Es una comida tradicional de Arequipa?	230.04
17	230.09	LI_02_M.E.M.V.	Es una comida tradicional.	231.52
56	231.73	Entrevistador	¿Cuándo se come, sabe eso, o sea, alguna festividad, algún día particular?	236.87
18	237.08	LI_02_M.E.M.V.	No, en cualquier. Este, cualquier momento del año se puede comer, es un plato tradicional que lo sirven en las picanterías Arequipeñas, y lo pueden servir todos los días.	250.35
57	251.62	Entrevistador	Arequipa también es conocida por su picante, entonces.	254.03
19	254.37	LI_02_M.E.M.V.	Claro, tiene un plato típico que es el picante, pero ese plato es una mezcla de distintos, de otros platos, ¿no?, lo que tiene ahí es, por ejemplo, sus habas, perdón, sancochadas, rocoto relleno, un arroz amarillo, algo así más o menos, una mezcla de otros platitos.	278.19
58	279.42	Entrevistador	Y entre el picante tacneño o el arequipeño, ¿por cuál de los dos usted le gusta?	284.66
20	286.79	LI_02_M.E.M.V.	Obviamente que el tacneño es un poquito más dañino porque contiene más ají, ¿no?. Pero, si se trata del gusto, yo creo que es ambos.	294.52

59	295.28	Entrevistador	Hemos terminado esta parte, ahora vamos a continuar, le voy a pedir que lea de la número siete, como llave, payaso, hasta la número diez.	302.19
21	305.41	LI_02_M.E.M.V.	Llave, payaso, zapallo, gallina, pera, árbol, collar, ratón. Perro, Enrique, roto, gorro. Bluza, plato, flor, clavo, globo. Libro, profesor, fresa, cruz, tigre, dragón, tren. Trenta cuatro, trescientos, veinticinco, veintitrés, sesenta, noventa y cuatro, noventa y ocho, treinta y uno, cinco, noventa y ocho.	334.76
22	346.32	LI_02_M.E.M.V.	i	347.48
23	347.88	LI_02_M.E.M.V.	e	348.75
24	349.18	LI_02_M.E.M.V.	a	350.1
25	350.74	LI_02_M.E.M.V.	o	351.46
26	352.01	LI_02_M.E.M.V.	u	352.61
27	355.84	LI_02_M.E.M.V.	Lunes, martes, miércoles, jueves, viernes, sábado, domingo. Uno, dos, tres, cuatro, cinco, seis. Siete, ocho, nueve, diez, once, veinte, veinticuatro, cien, quinientos, novecientos noventa y nueve. Aló, aló, aló.	375.51
28	376.29	LI_02_M.E.M.V.	Ya, ya, ya, pe.	377.68
29	378.95	LI_02_M.E.M.V.	Ya, pues.	379.93
30	380.51	LI_02_M.E.M.V.	No	381.23
31	381.46	LI_02_M.E.M.V.	No, pe.	382.27
32	382.53	LI_02_M.E.M.V.	Buenas tardes.	383.74
33	383.85	LI_02_M.E.M.V.	Buenos días.	384.95
34	404.68	LI_02_M.E.M.V.	Hola, buenos días.	406.31
35	416.11	LI_02_M.E.M.V.	Cuando tú me digas la razón del sí, como también el porqué, tú podrás comprender si tengo la razón o no.	423.62
60	424.11	Entrevistador	Ahora seguimos aquí, de aquí hasta aquí.	426.34
36	431.8	LI_02_M.E.M.V.	Cuando me disponía venir a Lima, conocí a Don Guillermo, que muy amablemente me invitó a subir a su camión, en donde transportaba gallinas y pescado desde Yurimaguas hasta la capital. Subí en la Oroya, en el trayecto Don Guillermo me contó lo que le pasó en el pueblo de San Roque, cuando regresaba de Yurimaguas trayendo carga. Rasputín el ratón, Rasputín era un ratón que rascaba en una risca, rasca, risca, rascador, rasca que rasca, acabo con el risco, rasca la risca, rasca un rincón.	475.3
61	475.85	Entrevistador	Una vez más lo repetimos, con más velocidad.	477.89

37	478.94	LI_02_M.E.M.V.	Rasputín el ratón, Rasputín era un ratón que rascaba en una risca, rasca, risca, rascador, rasca que rasca, acabo con el risco, rasca la risca, rasca un rincón.	489.05
62	489.77	Entrevistador	Ahora, lo último, ah, este de aquí descríballo, use la imaginación un poco, la observación y con eso finalizamos.	497.73
38	498.84	LI_02_M.E.M.V.	Bueno lo que yo observo aquí, en la primera gráfica, es este, me parece un parque. Al fondo hay, algunas, este, ehm, unos edificios, al parecer, antiguos, y en el parque están, este, existen personas adultas y también niños, y los niños están jugando, el parque tiene un columpio, tiene un, un resbalón, y me parece que también tiene una pila para el agua.	530.95

LI_03_E.A.C.M.				
Línea	Inicio	Tier	Contenido	Final
12	194.39	Entrevistador	Prenda su audio, por favor, señor, que no le estamos escuchando.	197.92
1	198.98	LI_03_E.A.C.M.	Muy buenas tardes, doctora, presente.	202.03
13	203.71	Entrevistador	Su nombre completo, su número de DNI, su domicilio.	206.92
2	208.45	LI_03_E.A.C.M.	(...) número (...)	214.53
14	219.4	Entrevistador	¿Cuál es su domicilio?	220.87
3	222.76	LI_03_E.A.C.M.	(...)	226.85
15	229.35	Entrevistador	¿A dónde pertenece?	230.46
4	231.16	LI_03_E.A.C.M.	Tacna.	231.58
5	339.66	LI_03_E.A.C.M.	Buenas tardes, doctora. Me permite unas palabras, doctora, acerca de mi abogado. Hace media hora me llamó, doctora, diciendo de que, que no, que se había suspendido y me dio a conocer sobre esta audiencia también doctora. Yo creo que, doctora, no se si fuera factible, por favor, pedirle, este, un tiempito o unos días, doctora, para poder contratar un abogado. Y, y ya usted doctora, puede fijar la fecha y la hora, doctora.	379.13
16	472.03	Entrevistador	Señor (...) ¿cuál es en si su domicilio real, donde usted vive, por favor, puede repetirme?	478.26
6	479.2	LI_03_E.A.C.M.	(...) Tacna.	487.31
17	498.9	Entrevistador	(...) ¿así es la ciudad o el distrito?	503.11

7	504.41	LI_03_E.A.C.M.	Es cercado, doctora, o sea, es (...) pero pertenece a (...). Es cercado.	509.39
18	510.54	Entrevistador	¿Me repite cuál es el nombre?	512.54
8	513.74	LI_03_E.A.C.M.	(...)	514.23
19	514.64	Entrevistador	(...)	515.79
9	516.15	LI_03_E.A.C.M.	Correcto.	516.82
20	528.74	Entrevistador	Dígame, ¿usted no ha recibido o reflexionado la notificación remitida a su domicilio?	533.02
10	534.76	LI_03_E.A.C.M.	Eh, doctora, aún no todavía, pero me llamó mi ex abogado comunicándome de que tenía una audiencia, doctora, a las tres de la tarde y media. Por la cual, me hago presente, este, no, bueno yo quisiera, doctora, por favor. Poder, no, darme un tiempo para poder contratar a un nuevo abogado, doctora.	555.6
21	3928.84	Entrevistador	Está notificado.	3929.49
11	3931.04	LI_03_E.A.C.M.	Correcto, doctora magistrada.	3933.07
22	3933.46	Entrevistador	Ya.	3933.62