

Identificación de locutor a partir de la fonética forense:  
aplicación del *software* SplitsTree4 para una organización  
esquemática de los datos lingüísticos<sup>1</sup>

Speaker Identification from Forensic Phonetics:  
Application of SplitsTree4 Software for Schematic  
Organization of Linguistic Data

Identification du locuteur à partir de la phonétique légale :  
emploi du logiciel SplitsTree4 pour une organisation  
schématique des données linguistiques

Jhon Jimenez Peña  
Oficina de Peritajes del Ministerio Público, Lima, Perú  
[jjimenezp@mpfn.gob.pe](mailto:jjimenezp@mpfn.gob.pe)  
<https://orcid.org/0000-0003-3317-6152>

Fernando Aaron Torres Castillo  
Oficina de Peritajes del Ministerio Público, Lima, Perú  
[fatorres@mpfn.gob.pe](mailto:fatorres@mpfn.gob.pe)  
<https://orcid.org/0000-0002-1432-8811>

---

1 Esta investigación fue impulsada por el Gabinete de Lingüística Forense —del Instituto de Investigación de Lingüística Aplicada de la Universidad Nacional Mayor de San Marcos—, el cual fue creado mediante la Resolución Decanal N° 000623-2021-D-FLCH/UNMSM. Dedicamos el esfuerzo que invertimos en ella a Úrsula Velezmoro Contreras.



**Oscar Esaúl Cueva Sanchez**

Universidad Nacional Mayor de San Marcos, Lima, Perú

[oscar.cueva1@unmsm.edu.pe](mailto:oscar.cueva1@unmsm.edu.pe)

<https://orcid.org/0000-0003-1361-2367>

*Resumen:*

En este artículo se presenta una propuesta de interpretación y organización de las características fonéticas de muestras indubitadas y dubitadas a través del *software* SplitsTree4 con el propósito de esclarecer un presunto delito de cohecho pasivo propio en el ejercicio de la función policial en agravio del Estado. Las muestras dubitadas fueron proporcionadas por el Ministerio Público y las muestras indubitadas se obtuvieron mediante diligencias de toma de muestra de voz; asimismo, se optó por la anonimidad de los datos. Primero, se categorizaron los rasgos fonéticos pertinentes de las muestras; después, se les asignó un valor binario de existencia y no existencia; luego, la información binaria fue procesada por el *software* SplitsTree4 para reagrupar los rasgos de acuerdo con el universo de locutores y mostrar la compatibilidad entre las muestras indicadas. Finalmente, los resultados indican que el *software* SplitsTree4 cumple con el ordenamiento de los datos fonético-articulatorios para la identificación forense de voz.

*Palabras clave:* identificación de locutor, fonética forense, SplitsTree4, procesos fonético-fonológicos

*Abstract:*

This article presents a proposal for the interpretation and organization of the phonetic characteristics of undoubted and doubted samples using SplitsTree4 *software* for the purpose of clarifying an alleged crime of passive bribery in the exercise of the police function to the detriment of the State. The doubted samples were provided by the Public Prosecutor's Office and the undoubted samples were obtained by means of voice sampling; likewise, data anonymity was chosen. First, each sample's relevant phonetic features were categorized; then, they were assigned a binary value of existence and non-existence; then, the binary information was processed by SplitsTree4 *software* to regroup the features according

to the universe of speakers and show the compatibility between the indicated samples. Finally, the results indicate that the SplitsTree4 software complies with the sorting of phonetic-articulatory data for forensic voice identification.

*Key words:* Speaker identification, Forensic Phonetic, SplitsTree4, phonetic-phonological processes

*Résumé:*

Nous proposons dans cet article un procédé d'interprétation et d'organisation des caractéristiques phonétiques d'échantillons contestables et incontestables au moyen du logiciel SplitsTree4, dans le but d'éclaircir un délit présumé de corruption passive dans l'exercice de fonctions de policier, au détriment de l'État. Les échantillons contestables ont été fournis par le Ministère Public et les échantillons incontestables ont été obtenus à travers des prises de voix ; les données ont été traitées dans des conditions d'anonymat. Dans un premier lieu, nous avons catégorisé les traits phonétiques pertinents des échantillons ; ensuite, nous leur avons assigné des valeurs binaires d'existence ou non-existence ; subséquemment, l'information binaire a été traitée par le logiciel SplitsTree4 pour regrouper les traits phonétiques selon l'univers des locuteurs et montrer la compatibilité entre les échantillons indiqués. Finalement, les résultats montrent que le logiciel SplitsTree4 permet l'arrangement de données phonéticos-articulatoires pour l'identification légale de voix.

*Mots clés:* identification du locuteur, phonétique légale, SplitsTree4, procédés phonéticos-phonologiques

---

Recibido: 03/04/2021 Aprobado: 11/02/2022 Publicado: 15/06/2022

## 1. Introducción

El objeto de estudio de la presente investigación forma parte de un caso fiscal. Los hechos ocurridos en este caso estaban relacionados con presuntos actos delictivos en los que los efectivos policiales exigieron el pago de una cantidad de dinero a cambio de dejar en libertad a dos intervenidos que presuntamente comercializaban con drogas. Durante dichos hechos, los familiares de los intervenidos recibieron llamadas telefónicas con el fin de coordinar el depósito bancario del pago; no obstante, ellos grabaron estas llamadas, las cuales sirvieron como evidencia para una posterior identificación de voz. Ahora bien, como el número de efectivos policiales involucrados fue seis, la fiscalía responsable del caso necesitaba saber quién de estos realizó las llamadas para pedir el soborno. Por este motivo, la fiscalía solicitó a la Oficina de Peritajes que se identificara a los actores del delito mediante una pericia de homologación de voz.

Dentro de esta comparación de voz, se consideró dos métodos complementarios para la redacción del informe pericial fonético, los cuales han sido sustentados en los tribunales de justicia en los últimos años. Los métodos complementarios los componen el método clásico o manual y el método automático: el primero está ligado al análisis perceptivo-auditivo de los fragmentos de habla del locutor; el segundo, a la comparación automática a partir de algoritmos que evalúan las características biométricas de las muestras de voz (Garayzábal *et al.*, 2019). Para obtener resultados más fiables, las muestras se compararon con ayuda de un programa de comparación automática, en este caso, el SIS II de la empresa rusa Speech Technology Center (STC), ya que posee *plugins* especializados para la homologación de voz, tal como indican Pop-Dimitrijoska *et al.* (2012). Por otra parte, cabe señalar que también se usó el *software* SplitsTree4, el cual ha sido aplicado por Zariquiey *et al.* (2017) para brindar una clasificación de las variedades lingüísticas pertenecientes a la familia Pano —las lenguas habladas en Purús (Ucayali)—. En el estudio de Zariquiey, se considera un repertorio más amplio de datos, cuyo corpus extraído está compuesto por rasgos léxicos y gramaticales, y el *software* cumple con el ordenamiento de los datos consignados para la elaboración de redes filogenéticas de las variedades dialectales de la familia Pano del Purús. Sin embargo, en

el presente estudio, se emplea el método con fines forenses; es decir, las características que se analizan son de índole fonética y se circunscriben al idiolecto de cada locutor para encontrar grados de compatibilidad entre la muestra de los locutores implicados —voces de las cuales se conoce su autoría— y las muestras dubitadas —voces de las cuales se desconoce su autoría—.

Como se indicó, las muestras fueron proporcionadas por la fiscalía, la cual solicitó que se determinara la pertenencia de la muestra dubitada dentro de un grupo de seis locutores; por ello, se formula la siguiente interrogante: ¿el *software* SplitsTree4 es una herramienta que funciona para organizar datos lingüísticos en el caso de que haya muchos implicados? En ese sentido, el presente estudio tiene como objetivo principal verificar si el *software* SplitsTree4 funciona para la organización de una gran cantidad de datos lingüísticos en casos de identificación en los que participan varios implicados; dicha organización se llevaría a cabo a través del *software* SplitsTree4, el cual se alimenta de datos binarios ya interpretados del análisis fonético-acústico. En suma, brinda los distintos grados de compatibilidad entre muestras; es decir, permite realizar una identificación del hablante de acuerdo con la agrupación de los datos. Además, esta investigación también se enfoca en dar un mayor sustento pericial a la disertación de compatibilidad de muestras de voz desde una perspectiva lingüística. De esta forma, este estudio se justifica científicamente como un aporte a la lingüística forense por la mirada holística y el uso de herramientas innovadoras para la resolución de un caso y la certeza de agrupación de las muestras.

Por otro lado, los locutores implicados en el caso son un total de seis personas, todas del mismo sexo —masculino, en este caso—, hablantes de español peruano que pertenecen a un mismo grupo sociolectal (a la Policía Nacional del Perú). Las muestras de voz de los implicados en el caso se obtuvieron a partir de diligencias fiscales de tomas de muestra de voz, las cuales consistían principalmente en una entrevista libre, en la que el lingüista forense preguntaba al implicado acerca de temas cotidianos, como viajes, pasatiempos, entre otros. Además, se pidió a los implicados leer un protocolo de toma de muestra de voz, el cual está sujeto a la *Guía*

*de elaboración de pericias fonético-acústicas de homologación de voz*, publicada por el Ministerio Público (2020); en dicho protocolo, se consideraron palabras aisladas y un cuento. Por su parte, las grabaciones se llevaron a cabo con una grabadora profesional de marca Zoom H4n Pro y un micrófono de tipo diadema modelo Shure WH20, los cuales permitieron tener audios en formato no comprimido WAV.

Finalmente, Las muestras dubitadas fueron proporcionadas en un CD-R dentro de un sobre lacerado con su respectiva cadena de custodia, remitido por la fiscalía responsable del caso. También, cabe señalar que tanto las muestras dubitadas e indubitadas fueron segmentadas y concatenadas con ayuda del programa Praat a fin de obtener solo las voces de los locutores objetivo en ambas muestras.

Finalmente, se debe indicar que, al tener muchos datos lingüísticos, se necesitaba organizarlos para saber entre qué muestras se mantenía mayor concordancia; por esto, se optó por la aplicación del *software* SplitsTree4, el cual organizó la data lingüística producto del análisis fonético-acústico de todas las muestras.

## 2. Marco conceptual

En este apartado se indican los conceptos y las definiciones necesarias para abordar la investigación.

### 2.1. Fonética forense

Gibbons y Turell (2008) distinguen tres ejes en los que se circunscribe la lingüística forense: el lenguaje jurídico (relacionado con el mundo del derecho y las leyes), el lenguaje judicial (relacionado con los jueces y tribunales) y la lengua como evidencia forense. Es en este último donde se puede abordar

el análisis del discurso legal o jurídico, la interpretación y desambiguación de los textos legales y comerciales (...), la determinación de la autoría y la detección de plagio y de delitos relacionados con la propiedad intelectual,

el control de la traducción e interpretación durante los procesos judiciales (...) y el análisis del habla con el propósito de contribuir al reconocimiento o la identificación del locutor. (Blecua *et al.*, 2014, p. 14)

En estas dos últimas tareas, la fonética forense desarrolla su labor; por ello, se relaciona con el ámbito policial y judicial, y tiene como tarea principal determinar, con el mayor grado de fiabilidad, si existe evidencia suficiente para sostener que dos voces pueden corresponder a una misma persona durante un proceso judicial (Fernández, 2007; Machuca *et al.*, 2014). En otras palabras, se centra en el análisis del habla a partir de la comparación de voces dubitadas (aquellas voces de las cuales no se conoce su autoría) con voces indubitadas (aquellas voces de las cuales se conoce su autoría), en donde se aplican los conocimientos de las tres ramas de la fonética (articulatoria, acústica y perceptiva).

Asimismo, a pesar de que el habla de los individuos varía de acuerdo con diversos factores, las diferencias interhablantes (entre diferentes hablantes) siempre son mayores que las diferencias intrahablantes (entre el mismo hablante), de acuerdo con Fernández (2007). Por esta razón, es que, a partir del análisis del habla, la fonética forense encuentra patrones particulares y recurrentes que responden al idiolecto de un individuo y que contribuyen a su diferenciación respecto a otras voces.

## 2.2. Aplicación del *software* SplitsTree<sup>4</sup>

La aplicación de nuevos métodos de ordenamiento de datos en los estudios que se relacionan con la lingüística ha ido aumentando conforme se ha ido descubriendo nuevos elementos de análisis. Las investigaciones concernientes a la clasificación de familias y variedades lingüísticas se han innovado, tomando una perspectiva de clasificación genética a partir de rasgos lingüísticos.

En ese sentido, uno de los trabajos recientes que emplea el programa SplitsTree4 es el que realizaron McMahon y McMahon (2006). Ambos autores proyectaron la construcción de las familias del índoeuropeo a partir de la lexicoestadística. Sus resultados les permitieron corroborar lo ya conocido respecto al índoeuropeo.

Un estudio posterior de Maguire y McMahon (2011) muestra que el uso de algoritmos representados a partir de esquemas arbóreos también es viable para el inglés. En su investigación, buscan responder cuáles son los orígenes históricos del inglés a partir de similitudes y diferencias entre dialectos.

Ahora bien, un ejemplo peculiar de documentación en el uso de redes y esquemas arbóreos en lenguas peruanas es el que realizó Heggarty (2005). Se trata de un estudio en el que se compara veinte variedades de lenguas andinas de las familias uru-chipaya, quechua y aimara. Específicamente, se analizó una muestra de 150 significados, a partir de los cuales, se produjeron cuantificadores de semejanza. Estos fueron incluidos en los programas de análisis de redes filogenéticas, uno de ellos fue el *NeighborNet*, paquete del *Splits Tree4* que se utiliza en el presente artículo.

Otro estudio es el de Zariquiey *et al.* (2017), en el que se describe el comportamiento de diversificación de dialectos del Purús. En este estudio, los autores sostienen que el uso de redes filogenéticas para la clasificación de lenguas en una familia lingüística es viable. Cabe señalar que no se ha encontrado bibliografía relacionada directamente con el presente artículo, por lo que este estudio supone un primer acercamiento al uso del *software SplitsTree4* para la organización de datos en un caso de identificación de locutor que se circunscribe al ámbito de la fonética forense.

### 3. Metodología

Este estudio es de tipo cualitativo, puesto que se comparan los procesos fonético-acústicos particulares extraídos de las voces de seis imputados —a partir de una entrevista semiestructurada— con los de muestras dubitadas. El fin de la comparación es identificar la compatibilidad de la voz de los imputados con la de las muestras dubitadas. Dicha compatibilidad se obtiene a partir del ordenamiento de los datos fonético-acústicos, el cual se realiza con apoyo del *software SplitsTree4*.

Así, para el análisis de los datos, en primer lugar, se ordenaron las muestras indubitadas; es decir, los audios de cuya autoría no se tiene duda

y que fueron obtenidos mediante la grabación de voz de los imputados. En segundo lugar, se agruparon los archivos de audio de las llamadas telefónicas. Estos audios constituyen la muestra dubitada porque son audios de los que se desconoce su autoría.

Estos dos grupos de muestras se acondicionaron (ver apartado 3.1.) para la segmentación y concatenación de los fragmentos de habla de cada locutor en el programa Praat. Posteriormente, estos audios fueron analizados con el programa SIS II para determinar la calidad de la señal, la cual se basa en la relación señal-ruido (SNR), el ruido tonal, la reverberación, la dispersión, la saturación y la actividad de voz (VAD). Aquellos audios que fueron óptimamente adecuados han sido utilizados para el siguiente análisis de datos con el método fonético-acústico. A partir de este, se pudo extraer características fonéticas preponderantes que sirven para determinar la compatibilidad lingüística de la voz. Sin embargo, por tratarse de un caso de identificación en el que había muchos implicados, el ordenamiento de los resultados del análisis fonético-acústico de todos ellos resultaba engorroso y totalmente desorganizado, por lo que la interpretación de los datos podría caer en error, debido a la gran cantidad de datos que este método había arrojado. Ante la imposibilidad de obtener una conclusión a partir de este método, se tuvo que optar por organizar estos datos con el programa SplitsTree4, el cual permitió realizar un ordenamiento de datos lingüísticos basados en la mayor o menor compatibilidad entre muestras.

### **3.1. Acondicionamiento de las muestras**

#### **3.1.1. Muestras indubitadas**

##### **3.1.1.1. Formato de las muestras indubitadas**

Los audios indubitados se encuentran en formato WAV. Estos archivos de audio presentan un solo canal, una frecuencia de muestreo de 44.1 KHz y 16 bits. Estas características son compatibles con los programas utilizados en el análisis de voz, por lo que no se realizó ninguna adaptación de sus características.

### 3.1.1.2. Extracción de los fragmentos de voz de las muestras indubitadas

En cada archivo de audio, se seleccionaron los intervalos de tiempo que corresponden a los fragmentos de voz de cada persona. Posteriormente, estos segmentos fueron extraídos y concatenados, y se crearon nuevos archivos de muestras de voz indubitada, los mismos que se presentan en la tabla 1, junto a su tiempo de duración. Además, en esta tabla, aparece el código de referencia, el mismo que será utilizado en adelante para señalar el audio indubitado de cada persona. Para el ordenamiento de los datos, se ha considerado colocar de forma anónima los nombres de los implicados en el caso, por lo que se codificó a cada uno de la siguiente manera: la letra «L» de locutor y el número de muestra que le corresponde («00»).

**Tabla 1**

*Valores temporales de los fragmentos de habla de la muestra de los locutores implicados*

Referencia	Duración (s)
L01	321
L02	301
L03	328
L04	284
L05	331
L06	384

*Nota.* L: locutor.

### 3.1.2. Muestras dubitadas

#### 3.1.2.1. Formato de las muestras dubitadas

Los archivos de audio de la muestra dubitada se encontraban en formato M4A. Con la finalidad de utilizar los programas de análisis de voz (Praat

y SIS II), fue necesario transformar el formato original de archivos de audio a WAV y a un solo canal (mono) con el programa FFmpeg en su versión 4.3. Es importante indicar que esta adecuación no afecta a los registros de audio.

### 3.1.2.2. Extracción de los fragmentos de voz de las muestras dubitadas

Se seleccionaron los intervalos de tiempo que corresponden a la voz incriminada en cada uno de los audios. Luego, estos fragmentos fueron extraídos y concatenados para crear un nuevo archivo de sonido por cada audio. Estos conservan el nombre del archivo original más el sufijo «\_C» (concatenado). En la tabla 2, se muestra el listado de los archivos resultantes, junto a sus respectivos tiempos de duración. Es importante señalar que cada archivo está acompañado de un código de referencia «MD» (muestra dubitada) y un número de muestra que le corresponde («00»). Estos códigos son utilizados para referirse a esta clase de audios a lo largo del artículo.

**Tabla 2**

*Valores temporales de los fragmentos de habla de los archivos de audio de la muestra dubitada*

Referencia	Nombre de archivo	Duración (s)
MD_01	AUD-20190103-WA0095_C	72.90
MD_02	AUD-20190103-WA0089_C	32.66
MD_03	AUD-20190103-WA0090_C	16.05
MD_04	AUD-20190103-WA0086_C	11.35
MD_05	AUD-20190103-WA0088_C	09.00
MD_06	AUD-20190103-WA0087_C	07.35
MD_07	AUD-20190103-WA0085_C	02.32

*Nota.* MD: muestra dubitativa.

### 3.1.2.3. Análisis de la calidad de las muestras dubitadas

Según el análisis de calidad para la comparación automática del programa forense SIS II, tal y como se muestra en la tabla 3, la muestra MD\_01 y MD\_03 son apropiadas para el análisis de homologación de voz. De otro lado, según este mismo análisis, las muestras MD\_02, MD\_04, MD\_05, MD\_06 y MD\_07 no son adecuadas para este proceso, por lo que no es posible emitir un resultado concluyente respecto a estas, puesto que no cumplen con los requisitos mínimos para ser evaluados. Por esta razón, se excluyen dichos audios del análisis fonético-acústico.

**Tabla 3**

*Análisis de la calidad de señal de los archivos de audio de las muestras dubitadas*

Referencia	SNR	Ruido Tonal	Reverberación (ms)	Saturación	VAD	Resultado
MD_01	43	0.56	413	0	55.82	Moderadamente apropiada
MD_02	0	0	422	0	20.96	No es apropiada
MD_03	0	2.41	371	0	11.23	Moderadamente apropiada
MD_04	0	0	389	0	7.40	No es apropiada
MD_05	0	0	0	0	4.85	No es apropiada
MD_06	0	0	561	0	2.73	No es apropiada
MD_07	293	0	0	0	1.86	No es apropiada

### 3.2. Métodos de análisis de la fonética forense: fonoarticulatorio y acústico-espectrográfico

El método fonoarticulatorio implica la descripción e identificación de las particularidades sonoras idiolectales (forma de articulación de un sonido, procesos fonológicos, entre otros) del locutor para después compararlas. Por su parte, el método acústico-espectrográfico permite representar la articulación, los procesos y otros fenómenos acústicos en gráficos para

observar las estructuras acústicas de los sonidos y cuantificar sus parámetros, los cuales también son comparados (Garayzabal *et al.*, 2019). Algunos de los procesos que se encontraron en las voces de los locutores correspondientes a las muestras indubitablemente y de la voz de las muestras dubitadas pueden encontrarse en Cueva (2018), tales como elisiones o cambios de la configuración sonora de segmentos consonánticos.

Conforme a los métodos explicados anteriormente, en cada muestra analizada se marcaron las características fonéticas más relevantes de los locutores y fueron reportadas en tablas. Por la extensión del artículo, se presenta un solo ejemplo de tabla resumida con el fin de observar su comportamiento frente a las características fonéticas de la muestra dubitada (MD\_01 y MD\_03). Aquellas características fonéticas reportadas fueron esquematizadas en imágenes a partir del programa Praat con el plugin TgDraw [Praat plug-in] Versión 0.3 (Muñoz, 2020), algunas de las cuales se reportan más adelante.

**Tabla 4**

*Muestra de características lingüísticas coincidentes entre las muestras de L01 y la muestra dubitada*

Procesos fonéticos	L01	MD_01	MD_03
Descenso vocálico	31.78-31.94 Mi [me]	-	-
Aspiración vocálica	9.23-9.80, 11.35-11.86 Espinoza [e <sup>h</sup> pinosa]	8.64-9.33 este [e <sup>h</sup> te:]	20.72-21.09 estás [e <sup>h</sup> ta <sup>h</sup> ]
		11.22-11.43 pues [pwe <sup>h</sup> ]	66.82-67.01 ¿ya? [ja <sup>h</sup> ]
		110.02-110.60 que no haga esto [kena:y <sup>h</sup> to]	

<b>Alargamiento vocálico</b>	-	8.64-9.33 este [e̞te:]  22.83-23.36 ¿por qué? [porke:]  17.28-18.00 indagar [indaya:r]	6.21-6.67 Paredes [pare:ðe]  61.72-62.14 Chorillos [ʃori:yo]
<b>Laringalización</b>	-	29.86-30.23 problema [proleya]  52.17-52.75 teniente [tenjente:]  61.99-62.15 nada [na]	13.19-13.72 saliendo de [saljendoðe]
<b>Disfonía vocálica</b>	-	46.65-46.97 mira [mirɑ]  136.73-137.02 acá [ɑka]	20.19-20.38 O [ø]  15.08-15.60 franco [fran̩kɔ]
<b>Sonorización de oclusivas sordas</b>	40.12-40.25 Por [bor]  64.24-64.72 Supervisor [supeɪβizo]  80.82-81.06 Servir [zeβi]	25.04-25.26 que está [yeða]  80.82-81.06 Servir [zeβi]	12.73-13.03 que iba [giβa]

<b>Debilitamiento de /d/</b>	-	19.82-20.05 nada de [nae]  29.12-29.46 todo [to:]  127.94-128.31 llamado [jamao]	6.69-7.32 ¿Dónde estás? [donehtas]
<b>Elisión de /s/</b>	39.72-39.87 Más [ma]  83.26-83.50 Están [etan]  91.72-92.49 Policiales [polisjale]  287.18- 287.78 Conozco [konoko]	10.67-11.18 policiales [polis- jale]  19.24-19.49 le hemos [lemo]  31.11-31.52 está ese [etaese]  32.45-32.68 estado [etao]	10.62-10.97 Paredes [pareðe]  21.86-22.23 momentos [momento]
<b>Alargamiento de /s/</b>	-	133.48-134.06 ustedes [usteðes:]	6.69-7.32 ¿Dónde estás? [donehtas:]
<b>Realización apical de /s/</b>	59.54-59.71 Su [su]  544.77- 545.11 Subí [suβi]	-	-

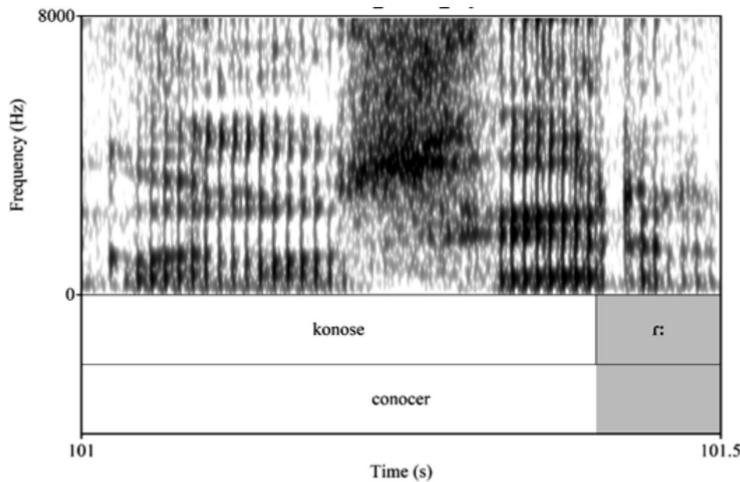
Elisión de la nasal en posición de coda y nasalización de la vocal precedente	-	16.31-16.91 vinieron [binjerō]  41.36-41.68 fueron [fweō]	21.84-22.23 momentos [momēto]
Elisión de vibrante simple	64.24-64.72 Supervisor [supeɪβizo]  80.82-81.06 Servir [zeβi]	41.36-41.68 fueron [fweō]  89.21-89.44 por eso [poeso]  135.57-136.15 sorprenderlos [soprederlo]	22.27-22.35 porque [poe]
Alargamiento de vibrante simple	-	49.89-50.51 podemos decir [poemoesir:]	23.12-23.42 sur [surf:]
Lateralización	59.43-59.54 Era [ela/ela]	-	-
Rotacismo	96.17-96.28 Lo [ɪo]  104.77- 105.14 Solo [soɪo]  179.63- 180.39 Caluroso [kaɾuɾoso]	-	-

Síncopa	59.71-60.01 Trabajo [traxo]  99.58-100.16 Anécdota [aneyota]  107.54- 107.83 Negra [neja]	-	-
Apócope	39.99-40.09 Nada [na]  64.72-64.98 Control [kontio]	-	-
Sinalefa	58.19-58.91 Se dedicaba [sejkaβa:]  74.03-74.46 Me dedique [mejke]	-	-

Una realización característica que se pudo observar en el espectrograma de los dos archivos de audio de la muestra dubitada y aleatoriedad de un locutor (L06) fue el alargamiento de la vibrante simple en final de palabra, tal y como se muestran en las figuras 1, 2 y 3.

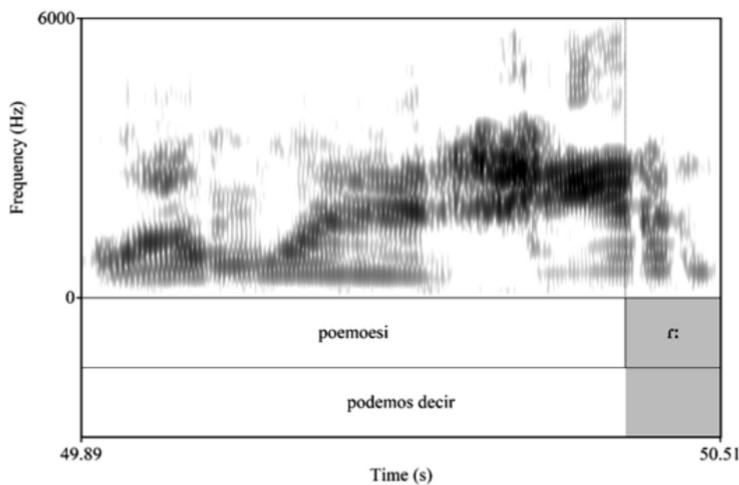
**Figura 1**

*Espectrograma de la palabra conocer en la muestra del L06*

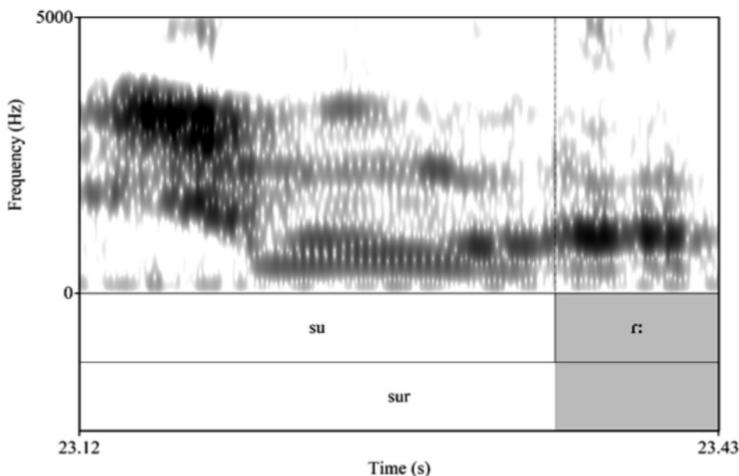


**Figura 2**

*Espectrograma de la frase podemos decir en la MD\_01*



**Figura 3**  
*Espectrograma de la palabra sur en la MD\_03*



Cabe recalcar que, por los límites de espacio, en el artículo, no se colocaron todos los espectrogramas de los procesos fonéticos de la tabla 4; sin embargo, han sido corroborados a partir del análisis acústico-espectrográfico realizado en Praat, documentados en el TextGrid de cada audio.

### 3.3. Método de redes del *software SplitsTree4*

El ordenamiento individual de los datos por investigado, tal como se muestra en la tabla 4, es insuficiente, pues, en primer lugar, el alto número de datos lingüísticos de diversos audios no permite observar la correspondencia entre las muestras; y, en segundo lugar, todos los imputados pertenecen a una misma comunidad sociolectal. Esto último significa que existen rasgos fonéticos coincidentes entre los imputados. Para evitar interpretaciones erróneas por las partes procesales y mostrar una lectura clara de los datos al juez, se tuvo que utilizar un método más refinado de ordenamiento de data.

Las redes del *software SplitsTree4* darían aquella respuesta. En primer lugar, se emplearon valores numéricos binarios de «0» y «1» para la ausencia

o presencia de cada característica fonético-articulatoria, respectivamente. En la tabla 5, se presenta una matriz hipotética de rasgos fonéticos, en la cual cada característica fonético-articulatoria está acompañada de un código de referencia «R» (rasgo) y un número correlativo que le corresponde («0»).

**Tabla 5**

*Matriz hipotética de rasgos fonéticos de las muestras a comparar*

	L01	L02	L03	L04	L05	L06	MD
R1	0	1	0	1	0	1	0
R2	1	0	1	0	0	1	0
R3	0	0	1	1	1	0	1
R4	0	1	1	1	1	1	1
R5	0	0	1	0	1	1	0
Rn...	1	1	1	1	0	0	0

*Nota.* L: locutor; MD: muestra dubitativa; R: rasgo.

En segundo lugar, esta data se tendría que codificar en el programa SplitsTree4 (Hudson y Bryant, 2006). Este *software* ofrece una gama amplia de algoritmos que permite construir diferentes clases de redes a partir de matrices de datos compatibles con lo estructurado en la tabla 2.

#### 4. Análisis

En este apartado, se muestra la tabulación de las características fonéticas de forma binaria y también la interpretación de estas mediante los gráficos que se obtuvieron con el *software* SplitsTree4. Asimismo, se complementa el análisis con los resultados obtenidos a partir de la comparación automática hecha con el *software* SIS II.

#### 4.1. Creación de la matriz de rasgos fonético-articulatorios

Luego de identificar los rasgos fonéticos relevantes para la creación de la matriz de rasgos fonéticos de las muestras de voz de los implicados y de la muestra dubitada, se asignaron los valores correspondientes de «0» y «1». Esto se muestra en la tabla 6. Ahora bien, se indicó como «MD» los rasgos fonéticos encontrados en los dos archivos de audio de la muestra dubitada, debido a que la voz de aquellas muestras pertenece a una misma persona; en ese sentido, también se encontraron los mismos rasgos fonético-articulatorios.

**Tabla 6**

*Matriz de rasgos fonéticos de las muestras de voz de los implicados y la muestra dubitada*

	L01	L02	L03	L04	L05	L06	MD
Elisión vocálica	0	0	0	0	1	0	0
Descenso vocálico	1	0	0	0	0	0	0
Monoptongación	0	0	0	1	0	0	0
Aspiración vocálica	1	0	0	0	1	1	1
Alargamiento vocálico	0	0	0	0	0	1	1
Laringalización	0	0	0	0	0	1	1
Disfonía vocálica	0	0	0	0	0	1	1
Ensordecimiento vocálico	0	0	0	0	1	0	0
Sonorización de oclusivas sordas	1	0	1	1	1	1	1
Fricativización de /k/	0	0	0	0	1	0	0
Debilitamiento de /b/	0	0	0	0	1	0	0
Debilitamiento de /d/	0	0	0	0	1	1	1
Elisión de /s/	1	0	1	1	0	1	1

Realización palatal de /s/	0	0	0	1	0	0	0
Reforzamiento de yod	0	0	1	0	0	0	0
Africación de /s/	0	0	0	0	1	0	0
Alargamiento de /s/	0	0	0	0	1	1	1
Realización apical de /s/	1	0	0	0	0	0	0
Realización de /s/ como fricativa velar	0	0	1	0	0	0	0
Elisión de nasal	0	1	0	0	0	0	0
Elisión de la nasal en posición de coda y nasalización de la vocal precedente	0	0	0	0	0	1	1
Elisión de vibrante simple	1	1	0	0	0	1	1
Alargamiento de vibrante simple	0	0	0	0	0	1	1
Lateralización	1	0	0	0	0	0	0
Rotacismo	1	0	0	0	0	0	0
Realización de /r/ como aproximante	0	1	0	0	0	0	0
Producción de vibrante múltiple /r/ con fricción	0	0	0	0	1	0	0
Supresión consonántica	0	1	1	1	1	0	0
Síncopa	1	0	0	0	0	0	0
Aféresis	0	1	1	1	0	0	0
Sinalefa	1	0	0	0	0	0	0
Apócope	1	1	1	1	1	0	0

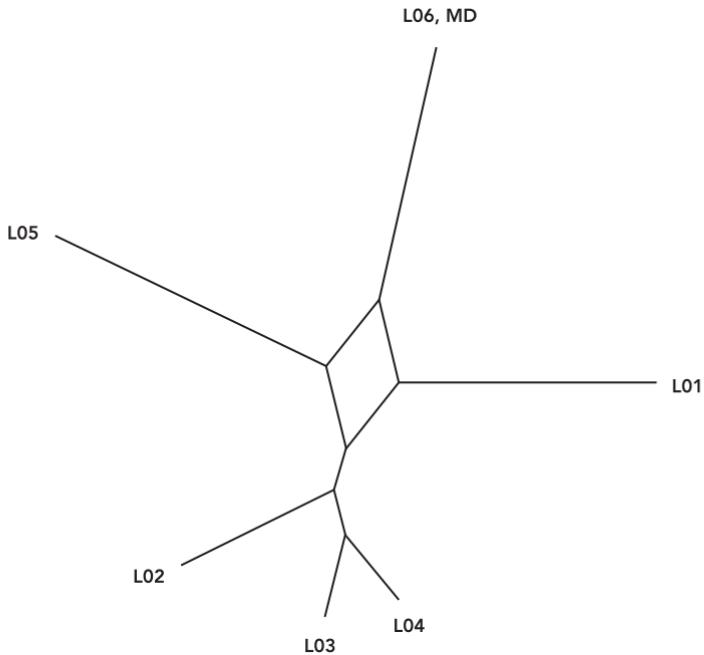
#### 4.2. Codificación de las categorías taxonómicas a partir de los rasgos fonético-articulatorios

La codificación consignada en el programa SplitsTree4 necesita el llenado de datos para su interpretación y esquematización. Al tener especificaciones de una data simple o *simple data*, se utilizó la matriz de rasgos para obtener el «código genético» de cada muestra taxonómica (locutor). En la opción *Enter Data Dialog*, se codificó los comandos correspondientes al formato *nexus*, mientras que el comando *Matrix* se refiere a la matriz de rasgos ya indicados.

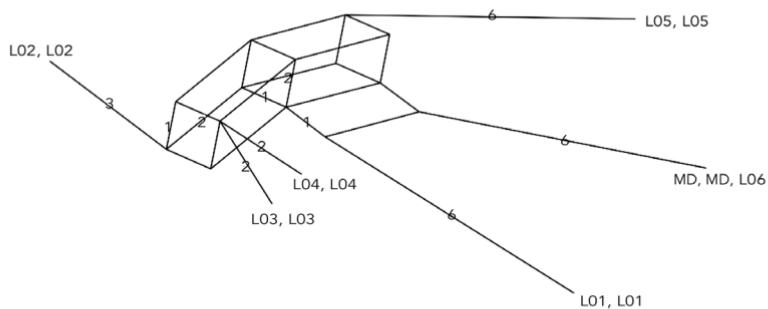
#### 4.3. Análisis de los datos en el *software* SplitsTree4

Para el ordenamiento de las características lingüísticas, se trabajó con la matriz de rasgos de las muestras de voz de los implicados y la muestra dubitada. Se extrajo la codificación binaria de cada locutor y la muestra dubitada. Luego, se generó, a partir del *software* SplitsTree4 (Hudson y Bryant, 2006), un esquema de red o *phylogenetic networks* (Bryant y Moulton, 2004). Como lo solicitado por la fiscalía era responder cuál de aquellas muestras tenía mayor compatibilidad respecto a la muestra dubitada, se optó por esquematizar los datos con el método matemático llamado Neighbor-Net, tal como se muestra en la figura 4, y con *parsimony data*. En general, este principio brinda una explicación simple, con el que se puede aclarar los datos. En el análisis de las redes, la parsimonia significa que una hipótesis de relaciones que requiere el menor número de cambios de carácter tiene más probabilidades de ser correcta (Steven, 2002), tal como se manifiesta en la figura 5.

**Figura 4**  
*Esquema Neighbor-Net*



**Figura 5**  
*Esquema Parsimony data*

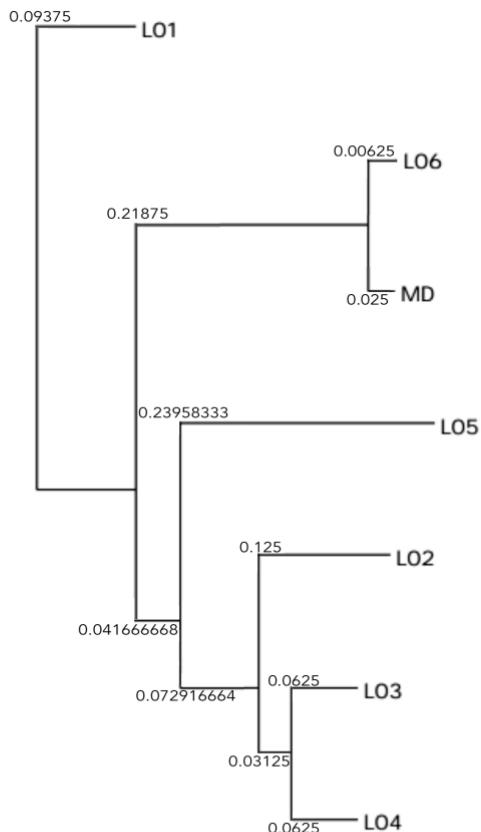


Los resultados obtenidos a partir de la organización de los rasgos fonético-acústicos revelan aspectos esenciales para responder la interrogante planteada por la fiscalía. De acuerdo con el esquema del *Neighbor-Net*, los locutores (L03 y L04) son considerados *outlayers*, es decir, son los que se encuentran mayormente distanciados de las demás muestras taxonómicas (locutores), incluida la MD (taxa n.º 7). Si lo que se busca responder es qué muestra tiene mayor compatibilidad entre todas las muestras indubitadas y la muestra dubitada, se toma a los *outlayers* como punto de distanciamiento máximo.

Considerando lo anterior, se puede inferir los siguientes comportamientos: a) las demás muestras taxonómicas se acercan a la muestra dubitada de forma gradual, b) las muestras taxonómicas (L02, L03 y L04) podrían formar un subgrupo y c) la muestra dubitada (MD) tiene mayor cercanía o compatibilidad con la muestra taxonómica del locutor (L06). Lo mismo se asevera en el esquema arbóreo de consenso basado en la mayoría, como se manifiesta en la figura 6.

**Figura 6**

Árbol de consenso basado en la mayoría de las muestras indubitadas y la muestra dubitada



#### 4.4. Análisis a partir del software SIS II

Con la finalidad de corroborar la fiabilidad de los resultados obtenidos con el software SplitsTree4, se llevó a cabo la comparación automática de las muestras indubitadas y los dos archivos de audio de la muestra dubitada

(MD\_01 y MD\_03) con los tres métodos integrados (estructura formántica o EF, tono o *pitch*, y GMM). Los resultados se muestran en la tabla 7.

Tabla 7

*Resultados de la comparación automática con el SIS II*

Referencia 1	Referencia 2	Método	FR	FA	LR	Resultado	Probabilidad
L01	MD_01	EF (SF), Tono (Pitch) y GMM	0.16	65.2	0.002	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L01	MD_03	EF (SF), Tono (Pitch) y GMM	0.2	79.3	0.003	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L02	MD_01	EF (SF), Tono (Pitch) y GMM	0.16	77.7	0.002	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L02	MD_03	EF (SF), Tono (Pitch) y GMM	0.3	41.5	0.008	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L03	MD_01	EF (SF), Tono (Pitch) y GMM	0.1	67.2	0.001	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L03	MD_03	EF (SF), Tono (Pitch) y GMM	0.1	86.2	0.0016	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L04	MD_01	EF (SF), Tono (Pitch) y GMM	0.2	85.1	0.003	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )

L04	<b>MD_03</b>	EF (SF), Tono (Pitch) y GMM	0.1	74.0	0.002	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L05	<b>MD_01</b>	EF (SF), Tono (Pitch) y GMM	0.1	89.7	0.001	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L05	<b>MD_03</b>	EF (SF), Tono (Pitch) y GMM	0.15	93.6	0.0016	Locutores diferentes ( <i>Different locutor</i> )	Alta probabilidad ( <i>High probability</i> )
L06	<b>MD_01</b>	EF (SF), Tono (Pitch) y GMM	52.4	0.1	494.118	Mismo locutor ( <i>Same speaker</i> )	Alta probabilidad ( <i>High probability</i> )
L06	<b>MD_03</b>	EF (SF), Tono (Pitch) y GMM	53.6	0.5	111.305	Mismo locutor ( <i>Same speaker</i> )	Alta probabilidad ( <i>High probability</i> )

Se puede observar que la muestra indubitable del locutor L06 es convergente con alta probabilidad con los archivos de audio de la muestra duditada.

## 5. Conclusiones

Esta investigación demuestra que el *software* SplitsTree4, además de ser empleado para la clasificación de familias y variedades lingüísticas, puede ser empleado en el ámbito de la fonética forense, específicamente, en la organización de datos provenientes del análisis fonético-acústico. De esta forma, en este estudio se presentan los resultados de la aplicación del *software* SplitsTree4 para el ordenamiento de un corpus compuesto por rasgos fonético-articulatorios de seis locutores presuntamente implicados en un caso de delito de cohecho pasivo que fue remitido a la Oficina de Peritajes del Ministerio Público para que fuera analizado por los lingüistas forenses de dicha entidad.

Los resultados obtenidos muestran que el *software* SplitsTree4 funciona para la clasificación de matrices de rasgos fonéticos en casos de identificación de locutores, en los que se cuenta con una cantidad elevada de posibles autores de llamadas telefónicas en las que solo participa uno o dos de los posibles autores. Además, se ha determinado que se puede postular la correspondencia entre muestras a partir de la organización de matrices de rasgos fonéticos con el *software* SplitsTree4. Así, en este estudio, se pudo determinar que la muestra indubitable del locutor L06 es convergente con la voz contenida en los audios de la muestra dubitada.

Finalmente, se debe recalcar que el análisis de las voces solo se sustentó con datos fonético-acústicos; por ello, abordar otras características —como las características morfológicas, sintácticas, discursivas, entre otras— supondría un análisis complementario para las conclusiones periciales fonético-acústicas forenses. Asimismo, es necesario subrayar que este estudio aplica por primera vez el *software* SplitsTree4 en el ámbito de la fonética forense y demuestra que este programa funciona para la organización de los datos presentados. Sin embargo, para afirmar de forma general que dicho *software* funciona para la clasificación de matrices de rasgos fonéticos, es necesario aplicarlo en otros estudios experimentales que aborden casos más complejos. Solo con estos estudios, se podrá determinar la pertinencia de este *software* para la organización de datos en la fonética forense.

## REFERENCIAS BIBLIOGRÁFICAS

- Blecua, B., Cicres, J., y Gil, J. (2014). Variación en las róticas del español y su implicación en la identificación del locutor. *Revista de Filología Románica*, 31(1), 13-35. [https://doi.org/10.5209/rev\\_RFRM.2014.v31.n1.51021](https://doi.org/10.5209/rev_RFRM.2014.v31.n1.51021)
- Bryant, D., y Moulton, V. (2004). Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular Biology and Evolution*, 21(2), 255-265. <https://doi.org/10.1093/molbev/msh018>
- 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(63), 155-183. <https://doi.org/10.46744/bapl.201801.006>
- Fernández, A. (2007). ¿Para qué sirve la fonética? *Onomázein*, 15(1), 39-51. [http://onomazein.letras.uc.cl/Articulos/15/2\\_Fernandez.pdf](http://onomazein.letras.uc.cl/Articulos/15/2_Fernandez.pdf)
- FFmpeg. (2016). *ffmpeg tool* (version 4.3) [Software]. <http://ffmpeg.org>
- Garayzábal, E., Queralt, S., y Reigosa, M. (2019). *Fundamentos de la lingüística forense*. Síntesis.
- Gibbons, J., y Turell, M. (Eds.). (2008). *Dimensions of Forensic Linguistics*. John Benjamins Publishing Company.
- Heggarty, P. (2005). Enigmas en el origen de las lenguas andinas: aplicando nuevas técnicas a las incógnitas por resolver. *Revista Andina*, 40(1), 9-80. <http://revista.cbc.org.pe/index.php/revista-andina/article/view/304>
- Machuca, M., Ríos, A., y Llisterri, J. (2014). Conocimiento fonético y fonética judicial. *Quaderns de Filología: Estudis Lingüístics*,

19, 95-111. <https://ojs.uv.es/index.php/qfilologia/article/view/5188>

Maguire, W., y McMahon, A. (2011). Quantifying Relations between Dialects. En W. Maguire, y A. McMahon (Eds.), *Analysing Variation in English* (pp. 93-120). Cambridge University Press.

McMahon, A., y McMahon, R. (2006). *Language Classification by Numbers*. Oxford University Press.

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

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

Pop-Dimitrijoska, V., Apostolovska, G., Gerazov, B., Ivanovski, Z., y Jovanovski, J. (2012). Forensic Speaker Identification through Comparative Analysis of the Formant Frequencies of the Vowels in the Macedonian Language. *Physica Makedonica*, 61, 79-84. [https://inis.iaea.org/search/search.aspx?orig\\_q=RN:45082620](https://inis.iaea.org/search/search.aspx?orig_q=RN:45082620)

Speech Technology Center. (2015). SIS II (versión 2.6.357) [Software]. <https://es.speechpro.com/product/analisis/ikarlab#tab3>.

Steven, M. (2002). *Principles of Parsimony Analysis: An Example with Molecular Data*. Memorial University of Newfoundland. [https://www.mun.ca/biology/scarr/2900\\_Parsimony\\_Analysis.htm](https://www.mun.ca/biology/scarr/2900_Parsimony_Analysis.htm)

Zariquiey, R., Vásquez, A., y Tello, G. (2017). Lenguas y dialectos pano del Purús: una aproximación filogenética. *Lexis*, 41(1), 83-120. <https://doi.org/10.18800/lexis.201701.003>