ORIGINAL

RESERVADO

DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

PERICIA NRO:	CARACTERÍSTICA:
149.586	**FACTIBILIDAD DEL MATERIAL SONORO**

REALIZADA EN: SUBDIRECCIÓN DE FORENSIA DIGITAL – DIVISIÓN ACÚSTICA FORENSE

EXPEDIENTE ELECTRONICO NRO: EX-2026-68535568-APN-DICRIEFOR#GNA

SOLICITADA POR: JUZGADO NACIONAL EN LO CRIMINAL Y CORRECCIONAL FEDERAL N°11, SECRETARÍA N° 21

CAUSA NRO: CFP 3402/2025

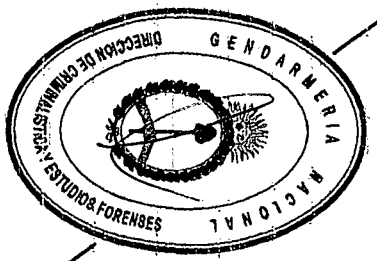
CARATULA: "MILEI JAVIER Y OTROS S/ DEFRAUDACIÓN POR ADMINISTRACIÓN FRAUDULENTA"

SÍNTESIS DE ESTUDIOS REALIZADOS: FACTIBILIDAD DEL MATERIAL SONORO

PERITOS ACTUANTES:
ALFÉREZ JOHANA MARÍA BELÉN LUGO
ALFÉREZ STEFANÍA MARIANELLA SULIGOY
CABO GABRIEL EDUARDO JOAQUÍN GAETE

PERITO DE PARTE:
LIC. DIEGO BERMÚDEZ GOLINELLI
ING. ADRIÁN NOREMBERG
ING. MARCELO ANTONIO TOROK
LIC. CECILIA BEATRIZ VERA
GALA IGLESIAS BRICKLES
DR. ROBERTO LEONARDO LEONARDO
BRUNO MOGLIA
ING. NICOLÁS EDUARDO SANGUINETI
ING. JOSÉ GABRIEL POLLOLA BARÓN

AÑO: 2.026


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
1
DIR CRIMINALÍSTICA Y EST FORENSES

PERITACIÓN: FACTIBILIDAD DEL MATERIAL SONORO

Nro.: 149.586

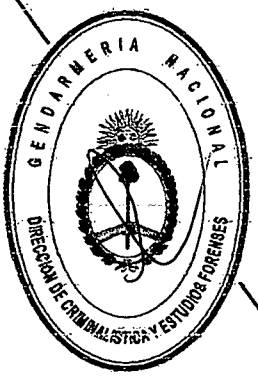
SEÑOR JUEZ:

Quienes suscriben, Alférez Johana María Belén LUGO, Licenciada en Criminalística, Alférez Stefanía Marianella SULIGOY, Licenciada en Fonoaudiología, y Cabo Gabriel Eduardo Joaquín GAETE, Técnico en Grabación y Sonido, pertenecientes a la División Acústica Forense, dependiente del Departamento Informática y Multimedia Forense de la Subdirección Forensia Digital, designados como Peritos oficiales, con la participación de los Peritos de parte, LIC. DIEGO BERMÚDEZ GOLINELLI, ING. ADRIÁN NOREMBERG, ING. MARCELO ANTONIO TOROK, LIC. CECILIA BEATRIZ VERA, GALA IGLESIAS BRICKLES, DR. ROBERTO LEONARDO LEONARDO, BRUNO MOGLIA, ING. NICOLÁS EDUARDO SANGUINETI, ING. JOSÉ GABRIEL POLLOLA BARÓN, en el marco de Juzgado Nacional en lo Criminal y Correccional Federal Nro. 11 interinamente a cargo del Dr. Ariel Lijo, Secretaría Nro. 21 a cargo del Dr. Alan VENCE, a su digno cargo, elevan a su consideración el presente informe pericial, a los efectos de exponer los estudios realizados y conclusiones correspondientes.

1. OBJETO:

El presente estudio pericial tiene por objeto, conforme con lo solicitado en el Oficio Judicial de fecha 13 de julio de 2026, lo siguiente:

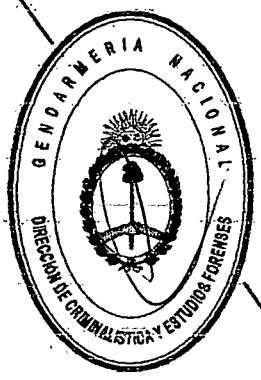
1. ***Determinar el formato del contenedor, codec utilizado, sample rate, bitrate,número de canales, duración y todos los metadatos técnicos extraíbles de los archivos de audio objeto de análisis.***
2. ***Efectuar un análisis comparativo con material indubitado correspondiente a Diego Orlando Spagnulo, debiendo convocarse al nombrado a fin de que aporte el material que resulte necesario para la realización del estudio.***
3. ***Realizar una comparación de parámetros acústico-fonéticos entre los segmentos dubitados y el material indubitado, informando el grado de compatibilidad existente mediante metodologías reconocidas.***
4. ***Determinar si los archivos presentan indicios técnicos de operaciones de edición posteriores a su captura original, identificando en su caso:***
   a) ***la naturaleza de las ediciones detectadas;***
   b) ***los puntos temporales que presenten indicios de manipulación;***
   c) ***los segmentos que no presenten tales indicios;***
   d) ***los archivos que no exhiban ningún indicio de manipulación; y***
   e) ***el tipo y metodología técnica de edición eventualmente empleada.***
5. ***Analizar si los segmentos atribuidos a Diego Orlando Spagnulo presentan características compatibles con voz humana de captura natural o con audio generado mediante sistemas de síntesis automática, incluyendo sistemas de "text-to-speech", "voice cloning" o tecnologías similares, pudiendo considerarse a tal fin:***
   a) ***análisis acústico-fonético;***
   b) ***análisis espectral;***


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 2
DIR. CRIMINALÍSTICA Y EST. FORENSES

c) *análisis de coherencia ambiental;*
d) *análisis de respiración natural y coarticulación; y*
e) *aplicación de detectores algorítmicos de audio sintético reconocidos en la literatura forense.*

6. *Emitir una conclusión acerca de:*
   a) *si los segmentos atribuidos a Diego Orlando Spagnulo resultan compatibles con voz humana real de origen no sintético;*
   b) *si dichos segmentos resultan atribuibles a la persona de Diego Orlando Spagnulo conforme al material indubitado analizado;*
   c) *y las limitaciones metodológicas del estudio y los rangos de confianza de las conclusiones alcanzadas*
7. *Determinar si el material presenta indicios de:*
   a) *clonación de voz;*
   b) *modificación mediante inteligencia artificial;*
   c) *procesamiento mediante sistemas de síntesis de voz;*
   d) *"deep voice manipulation"; o*
   e) *utilización de herramientas de postproducción automatizada.*
8. *Establecer:*
   a) *la cantidad de recompresiones detectables;*
   b) *el tipo de códec utilizado;*
   c) *la existencia de alteraciones derivadas de exportaciones sucesivas;*
   d) *eventuales incompatibilidades entre metadatos y estructura digital del archivo.*
9. *Intentar reconstruir la secuencia cronológica de eventuales modificaciones, identificando:*
   a) *puntos de corte;*
   b) *superposición de capas;*
   c) *inserciones no naturales; y*
   d) *posibles pérdidas de continuidad lógica y acústica.*
10. *Comparar técnicamente los archivos obrantes en la causa con los difundidos en medios de comunicación y cualquier otra versión circulante, a fin de detectar divergencias estructurales o modificaciones entre copias.*
11. *Determinar, en la medida en que ello resulte técnicamente posible:*
    a) *si existen elementos que permitan evaluar la continuidad e integridad de la cadena de custodia del material analizado;*
    b) *si de los archivos examinados surgen datos que permitan establecer fecha, lugar o circunstancias de obtención de la grabación;*
    c) *si resulta posible recuperar o identificar diálogos, voces o manifestaciones presentes en sonidos de fondo;*
    d) *si en algún tramo de las grabaciones se registra la mención del nombre y/o el apellido de Diego Orlando Spagnulo;*
    e) *las razones técnicas que permitan explicar la eventual ausencia de registro audible del interlocutor del hablante principal;*
    f) *si se observan variaciones significativas en las características técnicas del audio durante su desarrollo, indicando en su caso, si se advierte la*


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

existencia de distintos formatos o modalidades de captura del audio durante su desarrollo, sus características técnicas diferenciales y su eventual compatibilidad con una grabación continua; y

g) toda otra observación técnica vinculada con las cuestiones precedentemente señaladas que resulte de interés para la adecuada valoración del material analizado.

2. ELEMENTOS DE JUICIO:

A tal efecto, el elemento de juicio recibido para la realización del estudio encomendado es el que se describe a continuación:

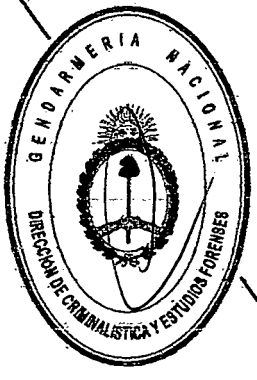
A. UN (1) sobre de papel color blanco, cerrado con cinta adhesiva transparente, el cual presenta inscripción manuscrita que reza "CFP 3402/2025 PEDRIVE MAURO FEDERICO", además inscripción impresa que reza "MINISTERIO PÚBLICO DE LA NACIÓN", con su respectiva planilla de cadena de custodia.
De su apertura, realizada mediante corte de cinta adhesiva transparente, resulta que contiene UN (1) dispositivo de almacenamiento portátil tipo pendrive de color blanco y amarillo, que reza "Kingston" y el modelo "Datatraveler G4" de 8 gigabytes de capacidad.

El control y apertura del elemento de juicio, correspondiente a evidencia digital, fueron realizados por la División Informática Forense, la cual asimismo efectuó la correspondiente imagen forense del dispositivo previamente mencionado. Como resultado de dicho procedimiento, la División Informática Forense puso a disposición de esta División un disco externo marca Western Digital número de serie "WX51AB493L9C" que contenía la imagen forense obtenida, a los fines de llevar a cabo los análisis inherentes a la División Acústica Forense.

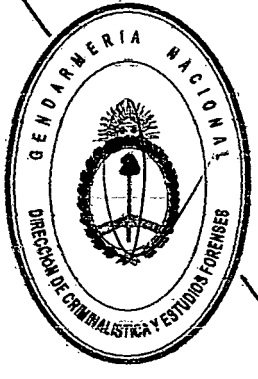
A los fines de dar integridad al material a estudiar, se calcularon las funciones HASH (MD5 y SHA256), de la totalidad del material sonoro que almacena el disco externo, cuyos valores se muestran en la siguiente tabla:

Tabla 1. Valores HASH (MD5 y SHA256) del material aportado.

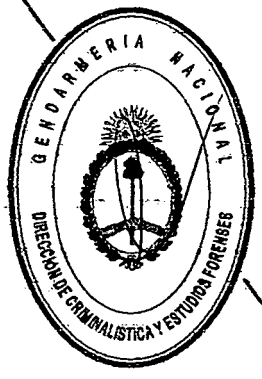
Filename	MD5	SHA-256
800 lucas por mes.mp4	a421397c01906bf30 28c1279775d79f5	2c68cff6719d6ff73e6da09c17e8d52723 4826e9e5846dd2024d268b65088b54
a la primera que se van a llevar puesta.mp4	d9044648ac17b39d 1b397497eccb0902	aae5fdedc1c1e3b5e30ea0b99a4b4818 2b19bf9c3627c396f08ab72812046e5c
a mi me pusieron un tipo que maneja toda la caja mia.mp4	eb908d93c186f87c3 d419053ecff005e	a24d85b5a3b25d3b9261ae6dc54ba832 686a7ed5c27b2b77a54a9c2f63b8ef17
Chupa pija.m4a	173ea6f16ecf568c5 b3434282363c45d	1b7b5e185e0491cafa0ba9ce6c073529 879f9d1f3aaa11dcdd4732b38817bd98
daniel garbellini.mp4	9a4cc260e1bdb082 bc38cff2475845b5	021065062cd6fead84f534333b62bf904f b3e88fdb95602a2183329c32f60260


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

Denuncia.m4a	f001e28aa09006078803aecab3e611ae	9449cdb1f425d1b47de1f2dfb89af46777b95d752c029c7abc7df86a73e45d85
después viene Karina con Santiago y los Menem.mp4	5fd63a83664f8071d1282bd58fb1f4cc	909b4900b80be895b6f92c4230f2a24b2233e965edf3de3eb4e28d08434b8bae
diego sos bueno.mp4	7e677c750cecde60bacd3f71b4795190	9c13a784e6ed12294e13e7a7f872a722a30e78d3d9ba4bbf507c12f8b9e4da54
Ejemplo torta .m4a	f200e9a7b956f1611ce396f763395b5d	156cb6ddb80ebbbf37741953ff30f6e2bc615a20fcffb00b31b621da86d807ae
el 3 es para Karina.mp4	7ad1ad528686e3596452abdcb37c3b23	9b83e27c842b96f5622e4ee9cab6d14fffaabeb1bbaeefb58bc4b688f3bffa79
el año pasado pagué 80 lucas de luz.mp4	7ea5aa463ea222aaf73e18da72a31bd5	72684f9fc26b51f137057f52f4ea52600868fa642a89f3b1d48810b84f89a7b5
el daño que puede hacer victoria.mp4	dcc9af9ffe95c7c32de787de39709bbd	66fe8b3503ba498d444d1e6d8eaacfe7206d4edf57d9e1534ee063ab845002c0
el dia que karina tenga que salir a hablar en la tele.mp4	edc5e0d48855532aca2c9d4c24c30d84	9ee98793918e8d5929adc0fdacf1461d58aa04edf3054f48c8f00cda3160d5f0
el no está metido pero.mp4	dae7bb17dd9326912b0a5ca7e1290b49	4f87a7516989e061e1807443d1d49af0a9d49c4563b89d35586087e6a03c11a6
ella iba a manejar seguridad y defensa.mp4	6beecf066921ce9fe8131fc32b089b16	6003e01e9b3677c7e03fa5406a92833d5b7688653f543b99f14f508b5674000e
ella se lleva medio palo.mp4	3e45738418591740a605c16a53a71d72	422b82bdb198c3d94b4451537b376630ed976f2f9f9b1548b4fc6662301e9ccd
es insostenible este nivel de interna.mp4	e3135515fd09d39937aff478347f671c	41a842df00968a075cd31dc6be6d5853df311a7c16902b8b632663e3e3f1b04b
es un cachito asi.mp4	cd8ccf4c7d294434cbcdf9448a355142	44a1943e15492f0163ce1da18918b083e8813e4bd8160e9dc534e00536eab6cc
es un error tremendo el que cometieron con villarruel.mp4	b93bac3d73744912e9d21f830e474980	c5a34fce9eb595e098f3ab0658f0b609117fdc1ffab5ed87ec9d1066c5c86960
ese es un delincuente.mp4	da2c60a923f08066baed712dda9dc52f	e332fc6ad91e15099958ba1d395257d1f7fb118d9efc0a7f83847dd3be0ccee5
espert pedazo de hijo de puta.mp4	12acbf1fb16dcfbbcc303b5584096633	f041bbec2ff6cdfdb22eb95285399deb38ce9a9708ba49640eab9158f824b78f
ESTO ESTÁ IMPLOSIONANDO DE ADENTRO MEDIO PELOTUDA SANDRA.mp4	1e1bb7b8bf523df20441dedd560937ce	f92fe8f13a8e74ea7f81bc973fe6c4aef2103e883e06185a0f984869e494c207
esto lo hacen de ratas.mp4	bffb1862fd83ea291eea5d2e4e09ef4e	aa37d02c4b05b38206b17f5b6a8f4f069450f34eec95e7ee58e27272d328a0be
estos dos ya tienen mas ruido que una lata llena de monedas.mp4	f92dd9f0b442946dc45a9222a31a737e	bc3a6b2d172e59135d6896fe07317566d957c90a159ed8ee2926257daa107313
estos tipos aprovechan y quieren armar su kioskito.mp4	c32c86cb4be48d20622b0a19dd7b1b2b	46d534668b8340784a174e19da092526a51a54e28d6d5fbc337d4e286b7ae747
Fake News adormi.m4a	4dab99d62ffea0556ab872f23a49d791	794d9298621c88fb6024e00abbb3ac224d39164fef470357318e5b61ec425ee2
Javi que hago.mp4	3f59724491fbd133ade2ffe43b0e518e	f47ec40f050cf03641a0ba79a1eb635345730b357c711c92f4e4b0e3ac47e325
javier me dice yo voy a hablar con Lule.mp4	fab4f7388d47252a7ea78eaed5bc22ba	926b05be05aa1f90a0393040d4787de5cd77fac1cf4b382238f890e1a815ea67
javier piensa que estamos haciendo el mejor gobierno de la historia.mp4	6967feecdb300a52853c87e3bb6518b6	db2a2b456898735d3bf0065fddd27cc18dc1bb60ca664c838d70a11c68bc739f
karina lo mima a guillermo para pegarle a victoria.mp4	baf6180c0ce28cbba28b316aef92b4f5	a96bb0933631ecb7c9b916d503acc8f24202c5068d95f762e1b75f781dc55444
KILOMBOS GROSOS CON DIANA MONDINO.mp4	88715656d8e86d86f0ab754955a5d81e	fc4d88d00ec4c34435d99af16e4d771ef-bfcf5096da40ce55d0ff7168f851d71


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

La caja..m4a	2378751e4fd6c88fb42e252e0a81f87c	6dacd04a79e571a6c268d148b372f453b1233b17f108255d62531b9f412fe881
la gente de la suizo.mp4	1f25115cd1a314c359c6552f8008168c	aa3d29700524561606450653be0fda09a8264f68aa6e3944308c628d051f1394
le afanaron la caja también a Mondino.mp4	6828439d1082cd017ce71a1023967942	cac35a76bee9570c4b55811b203462bbb25e1a74c457a29c75ba449f02f7efb6
Lule vino de una manera a mi me estan desfalcando a la gente.mp4	985985d0917e7ebdc98e3751e7bd2f3a	9847370a79bb3730599ae71110306118 58cb0e3db1bf16405521dd0985a3dce6
lule y karina.mp4	d33000c48190f61c6300932549aa08c6	9f42e69c4a0aeaf03cba8e93dad098f0614a30eb5c9f2795ebcb15d72c53b73f
me metio el mas importante de todos.mp4	3c5f33a8c6cb2cd26cdefbcf5f62d606	566afc0f436db27ed18196469ef532dc46e7a9368f52e6590a807b542f2fc553
no lo quiere nadie, pero nadie.mp4	511f58c9cfb8bece63f8758ef82b6cf6	8f84476f29c9d5b6e4cd0704977d89fc069abf13a0cad3892c108218bfa76086
nosotros la votamos a victoria.mp4	35253f136ffd9b87ec8befe05f7364bd	5df6ead8adb68ca3ff45824aabf9374290fc1753328605a928456ca7dbbcea88
nosotros lo subimos a presidencia.mp4	1cf3a981f7cd007e12a788910b10a343	63da4857845f7f1c4dc588eebbd5acdf1ac0567cb2f16cfd2ee2c4bd7cb70c68
ponela a Karina de candidata y que vaya a hablar.mp4	818894a310aed98c00372ca06e356d25	61d7135780c5b39bdd1e5f1a642e69912f9ec361e4712149e0b5ce7240c9a5d8
que conocimiento tiene karina.mp4	2a30a3aa-daee7f9a19945514187d809f	973e23ef64c103e0a3a6ead4513b68cbb35bf8de774f3eafead519d8d8e923f4
que tiene lule menem con Karina.mp4	f257340b4671eff3b0a7503ff6454975	978914390cf252cabfd5ff78ae14774c352fcab596de8a7ebc778a276007339b
SANDRA SE HACE LA PELOTUDA.mp4	9fb837f6ca600720e4536ab42ae0c98a	03730f9c94d65acf820de922577549b298688862574374594bb6abeadd0a2ee7
SE HACE LA RE PELOTUDA JAVIER DESENTENDIDO DE TODO.mp4	580d5cd9dfab-faca02575a3514cd7ca3	2456e82271a4029f363f9ab03c8897db139306ba499c205d9d0a416d04083b03
se lo dijo sandra, se lo dije yo.mp4	1715e7c12d3754bbf53879da5dcc609f	e32dabeb1e0cae73d439e6978f8a67131965fbd256f594e3dde2f35d4273573a
sturzzeneger.mp4	5f93ac9fc05f03214c754af133b12a18	3fdb07d1d65702b61f859b259173797c1547276d71bad76eb68be237bed9605a
sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4	42496665464367f9906f003a38899948	d06bf6d3d54d2dee249a20bccfe995a9de98f72dfe14a80a82c820670d879ff0
sueldos.mp4	0eaff8bed43ecaeae9edda73964cdb3d	16ac564b539d67dd39a9d0941a13a3c5f4c0471bdf0669c9d315d360f23a1ec5
tenes chorros en lule tucumán.mp4	b265d96cd4406d9e5d681ee6b44d6e69	55ada4783224a1f7c4301a1fe14d16fd46dec79c63286389054d0d353339d457
victoria claramente se cansó.mp4	77a48d459d9b3202ca3395236b619868	daca1c5283c2bc6562d0f12bb89b1428f6e7b05cf760b1be6ea25792ee0d76a9
victoria se canso.mp4	8ac1b99584cab18dbc7e4b160e5afb07	af59efc93af632c9605fa86687cf37ec88937323a91d90449f348602e7c552ef
villarruel ampliado y sandra pettovello.mp4	f257340b4671eff3b0a7503ff6454975	978914390cf252cabfd5ff78ae14774c352fcab596de8a7ebc778a276007339b
voracidad genética.mp4	ceb49403a4c8d254b1779455d167c943	1297037f2bf20168dd09110d2548fcf086c928cce51289bcfe004e8ff6d7fb32
vos viste la entrevista que me hizo fantino.mp4	20e691a2b9f95bd38323b74212105b7f	1dc94165598432b0dfc379ed5c237faba11318f41a16a4a3b4ee69936e635ab0
yo hablé con el presidente.mp4	90eaea583a6c78ef09654953d2397bd9	31829567c7edfcb0ed19e025fa66aea3df83ccd23a21f4ca4afd6ead35b7537e
yo no se como va a terminar la economia.mp4	82da1d7f0788530fe-aacaaf9e1fb7ee2	b4e9a68013f4650353f45f0d1995e33eec810b1eebde53a34d44b192f362f4f1


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

Zafarrancho.m4a	aeecf16cfb1916cfc3021c61351969d4	bde108ece2aa4a0a5643edd2c552e4a4482470f369da8e539f0a0f1364c61b6e

3. DESARROLLO FUNDAMENTADO

Se describen los incisos relevantes para el estudio de factibilidad del material sonoro; tanto el pre-procesamiento como la verificación de las condiciones mínimas, los cuales se encuentran debidamente detallados y formalmente establecidos en el "Protocolo Único para la Comparación Forense de Voces"[1], aplicable a este tipo de análisis.

3.1. PRE-PROCESAMIENTO DEL MATERIAL SONORO

1) **Conversión del formato de audio:** en caso que la grabación esté en formato analógico se verificará si es factible de digitalizar previamente. Debe convertirse al formato requerido por el método de comparación de voces (por ejemplo, formato RIFF conocido como WAV). Se ha de informar si el formato original es con pérdida de información (por ej. MP3 (grabación comprimida), OPUS (WhatsApp) o sin pérdida (por ej. WAV, FLAC).

2) **Separación del habla de interés:** se debe separar el habla de la persona de interés para la causa (dubitado) de la de su interlocutor, incluido el/los hablantes/s irrelevantes si existiesen. Las conversaciones superpuestas no deben ser incluidas en el análisis. Se eliminan silencios, de manera de procesar exclusivamente los segmentos de habla. Se deben marcar y borrar los ruidos, distorsiones locales de la señal de audio que puedan interferir el análisis, como así también los segmentos de audio que incluyan aspectos extra-lingüísticos. Por ej.: risas, toses, carraspeos, respiración, estilos de habla no usuales, tales como susurros, falsetes, gritos, etc.

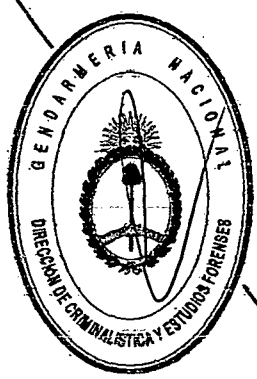
3) **Filtrado**: La señal no será modificada bajo ninguna técnica digital con el fin de preservar la información original

3.2. VERIFICACIÓN DE FACTIBILIDAD DEL MATERIAL SONORO

A efectos de determinar la aptitud de un audio para ser sometido a un posterior cotejo, se analiza si éste reúne un conjunto mínimo de condiciones técnicas indispensables. En tal sentido, se analizan los siguientes aspectos:

1) **Duración efectiva del segmento de habla:** se debe determinar si su duración (posterior al pre-procesamiento) cumple con el mínimo requerido por el método de comparación a emplear.

[1] Ministerio de Seguridad de la Nación, secretaría de lucha contra el narcotráfico y la criminalidad organiza (2024) – "Protocolo único para la comparación forense de voces" RM 526/2024 https://www.boletinoficial.gob.ar/detalleaviso/primera/309533/20240624


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

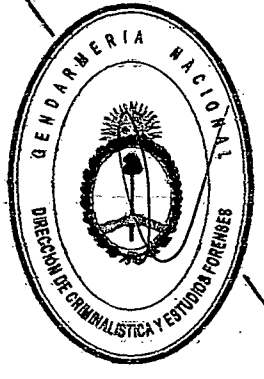
2) **Frecuencia de muestreo**: como mínimo se aceptará un valor equivalente a 8.000 Hz, que permite garantizar el procesamiento y la visualización de la información presente hasta 4.000 Hz.
3) **Relación señal ruido**: debe ser mayor a 10 dB. La SNR indica la relación de la energía de la señal con respecto a la energía del ruido de fondo, de tal manera que una SNR mayor, indica un ruido de fondo menos perceptible que enmascara en menor medida la información.
4) **Reverberación y eco**: se analiza perceptualmente su presencia y magnitud en el material, a fin de determinar si este es aceptable o no para trabajar.
5) **Frecuencia fundamental (F0)**: debe estar presente y variar lentamente. Una F0 apropiada a cada sonido garantizará una mejor inteligibilidad del habla.
6) **Audios no-contemporáneos**: considerar si es factible llevar adelante la comparación de acuerdo a la diferencia temporal entre las grabaciones.
7) **Diferencia de idioma o dialecto:** diferencias de sociolecto, idiolecto, regiolecto o acento extranjero y el tipo de lenguaje en casos de regiones de diferentes lenguas o multilingüismo, quedando a consideración del perito, si es factible trabajar con dichas muestras.
8) **Distorsión de la voz:** enmascaramiento y/o alteración evidente de una emisión vocal.
9) **Presencia de señales espurias:** Cuando éstas son continuas y están presentes en el rango de las frecuencias de habla.
10) **Estado emocional y de salud:** Se observará que no demuestre estados emocionales o de salud alterados.

3.3. MODELO DE PRODUCCIÓN DE VOZ

El sonido que da origen a la voz se produce cuando el aire procedente de los pulmones pasa con una determinada presión a través de una pequeña apertura (glotis) que dejan entre sí los pliegues vocales al aproximarse. La frecuencia que se genera durante el proceso de cierre y apertura de los pliegues vocales, es llamada frecuencia glótica o frecuencia fundamental (F0), por consiguiente, los pliegues vocales son la fuente de producción de la voz.

Los armónicos generados en la laringe van perdiendo energía acústica en la medida en que se van alejando de la frecuencia fundamental, es decir, en la medida en que los armónicos son más agudos se vuelven más débiles (pendiente espectral). En otras palabras, el primer armónico que acompaña a la F0 siempre es el más intenso o "fuerte", mientras que cada armónico adicional es progresivamente más débil.

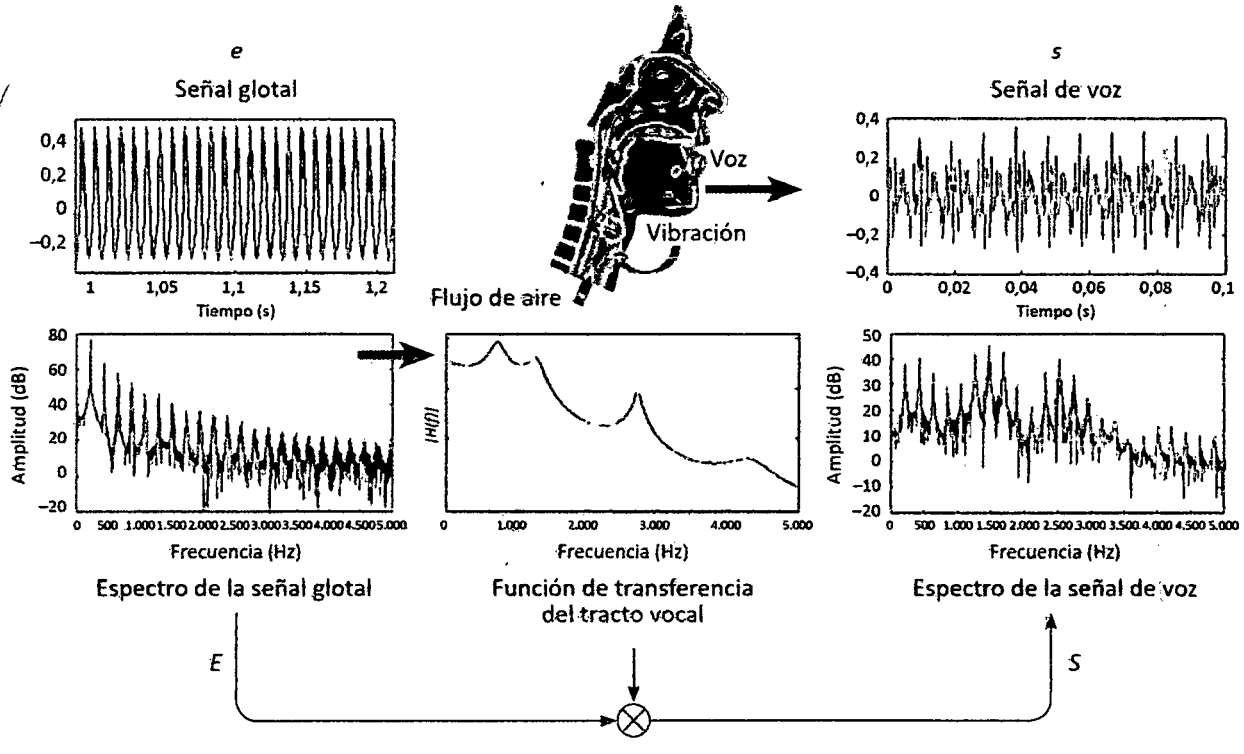
Lo anterior cambia considerablemente en el momento en el que el sonido ingresa al tracto vocal, para finalmente irradiarse al exterior. En este sentido, los armónicos que componen el sonido de la voz serán amplificados o atenuados en el tracto vocal, de acuerdo a las características anatómicas que éste adopte y por lo tanto a las características resonanciales que presente en dicha configuración anatómica. Po lo cual, los armónicos que coincidan con la resonancia de una cavidad determinada serán amplificados, mientras que los demás serán atenuados. De lo anteriormente descripto nace el concepto de "formante": los formantes corresponden a las resonancias del tracto vocal,


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR CRIMINALÍSTICA Y EST FORENSES
Foja
8

es decir, se asocian a la amplificación y mejora del sonido vocal. Por lo tanto, los formantes corresponden a una acumulación de energía acústica en zonas determinadas del espectro de la señal. El tracto vocal genera varios formantes, denominados generalmente F1, F2, F3, etc., desde aquellos que se encuentran en frecuencias más bajas, hacia los que se encuentran en frecuencias altas. Los más bajos a nivel frecuencial (F1, F2), determinarán las características o cualidades de las vocales, mientras que los demás (F3 , F4) se asocian a las características individuales de la voz de cada sujeto (timbre) y pueden aumentar el poder vocal o proyección del sonido.[2] La siguiente figura[3] ilustra el proceso de producción de la voz, mostrando la relación entre la fuente glótica y la señal acústica resultante.

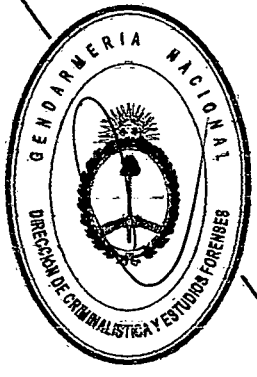
Figura 1. En la parte superior, forma de onda de la fuente glótica y señal de voz. En la parte inferior, espectro de la señal glotal, contribución del tracto y espectro de la señal de voz.. Fuente: Cobeta, Núñez y Fernández (2013, Fig. 15).



3.4. DESARROLLO DE LAS TAREAS PERICIALES

Las tareas periciales se iniciaron en presencia de los peritos de parte y las actividades desarrolladas en esas instancias quedaron documentadas en las Constancias de Trabajo Nro. 1 a Nro. 6. En dichas reuniones, determinadas operaciones se realizaron en presencia de los peritos de parte, a fin de posibilitar el control de las tareas efectuadas.

[2] Jackson-Menaldi, M. C. (2019). *La voz normal y patológica: Diagnóstico y tratamiento de la patología vocal*. Editorial Médica Panamericana.
[3] Cobeta, I., Núñez, F., & Fernández, S. (Coords.). (2013). *Patología de la voz*. Marge Médica Books.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

Atendiendo a la extensión del material por examinar y de las tareas pendientes, se acordó por unanimidad que el equipo de peritos oficiales continuara el examen de los archivos sin la presencia de los peritos de parte. Una vez finalizadas las tareas, se realizará una reunión final destinada a la puesta en común y al debate de los procedimientos y resultados obtenidos, previo a la emisión del informe pericial.

A efectos expositivos, los resultados se sistematizan conforme a las etapas metodológicas de los estudios realizados y no necesariamente según el orden cronológico en que se desarrollaron las tareas.

Participación y distribución de funciones del equipo de peritos oficiales

El examen se desarrolló de manera interdisciplinaria. Los tres integrantes del equipo oficial participaron en la escucha y el análisis de los archivos, su segmentación, la detección de posibles discontinuidades y el etiquetado temporal de los segmentos de interés. Asimismo, plantearon interrogantes sobre las características del material y realizaron consultas bibliográficas para profundizar los aspectos que requirieron.

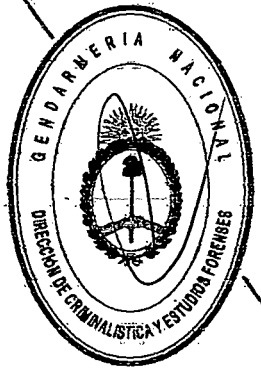
En ese marco, las principales funciones específicas se distribuyeron de la siguiente manera:

Coordinación general e integración del análisis. Alférez Johana María Belén LUGO tuvo a su cargo la organización y coordinación de las tareas, la conducción de las reuniones de trabajo y la articulación de las distintas etapas del examen. Participó en la toma de decisiones relativas al abordaje de los puntos periciales y en la valoración e integración de los resultados obtenidos para la elaboración del informe.

Análisis técnico de las señales de audio. Cabo Gabriel Eduardo Joaquín GAETE tuvo a su cargo la ejecución de los procedimientos de cálculo y de los algoritmos empleados para estimar la relación señal-ruido y analizar la medición H1*-A1*. Asimismo, realizó el examen e interpretación técnica de los metadatos de los archivos, la búsqueda y evaluación de alternativas metodológicas ante la falta de disponibilidad de determinadas herramientas, recurriendo para ello a bibliografía del estado del arte.

Análisis perceptual. Alférez Stefanía Marianella SULIGOY tuvo a su cargo el análisis perceptual especializado del material sonoro y participó en la búsqueda y consulta de bibliografía del estado de arte relacionada con los aspectos comprendidos en su intervención.

Las funciones específicas señaladas no implicaron una división excluyente de las tareas. Los procedimientos, observaciones y resultados fueron compartidos y considerados en el marco del trabajo interdisciplinario, de acuerdo con la naturaleza de cada cuestión y las áreas de especialización de los integrantes.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
10

Cada integrante interviene y se expide conforme a los conocimientos propios de su ciencia, arte o profesión, dentro de su respectivo ámbito de especialización. Ello se integra en el marco del trabajo interdisciplinario realizado por el equipo, sobre cuya base se elaboró el presente informe pericial, dejando identificadas las intervenciones de cada profesional.

3.4.1. METADATOS DE LOS ARCHIVOS

A fin de obtener metadatos de cada uno de los archivos objeto de estudio, se utilizó inicialmente el software MediaInfo, versión 26.05, registrando la totalidad de la información obtenida al archivo de texto "METADATOS ARCHIVOS EXPORTADOS" que posteriormente se organizó en las Tablas A1 y A2 (ver Anexo I).

En reuniones de peritos posteriores, se complementó la información obtenida mediante las herramientas mp4dump, perteneciente a Bento4 SDK 1.6.0-641, y ExifTool 13.59, ambas ejecutadas en un entorno Windows x86_64 sobre la copia de trabajo del material analizado. Sobre cada archivo se ejecutaron comandos como los que se exponen en los siguientes ejemplos, generándose en cada caso un archivo de texto cuyo nombre conserva el del archivo analizado e incorpora sufijos identificativos del tipo de información obtenida y de la herramienta empleada.

Bloque de código 1. Comando de ejemplo para extraer información del archivo mediante mp4dump.

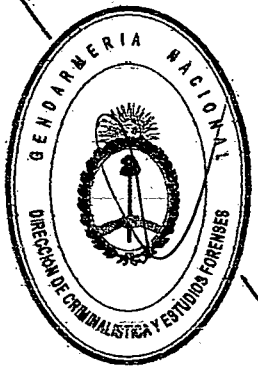
```
</>
mp4dump "Nombre de archivo.m4a" > "Nombre de archivo_metadata_mp4dump.txt"
```

Bloque de código 2. Comando de ejemplo para extraer información del archivo mediante ExifTool.

```
</>
exiftool -time:all -a -G1 -s "Nombre de archivo.m4a" > "Nombre de archivo_meta-
data_exiftool.txt"
```

En la reunión de fecha 09 de septiembre del 2026 de trabajo se procedió a ejecutar estos comandos a la totalidad de archivos en una sola ejecución, para lograr organizar y sistematizar el material sonoro, completando la extracción de metadata de la totalidad del mismo. Para el caso de archivos en cuyo nombre se encuentran caracteres con tilde y/o con la letra "ñ", se procedió a la ejecución del comando de mp4dump sobre una copia del archivo reemplazando estos caracteres no admitidos en la herramienta por el caracter "_". Se acompañan al presente Informe todos los archivos generados.

Finalizada la extracción de metadata a partir de las herramientas mencionadas, se procedió a registrar la información extraída en las Tablas A1 y Tabla A2 (ver Anexo I).


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
11
DIR. CRIMINALÍSTICA Y EST. FORENSES

Vale destacar que los archivos de audio "que tiene lule menem con Karina.mp4" y "villarruel ampliado y sandra pettovello.mp4" presentan idéntico contenido y el mismo valor hash (ver Tabla 1). Por tal motivo, se realizó el análisis sobre el archivo de audio "que tiene lule menem con Karina.mp4", por tratarse de registros sonoros idénticos.

3.4.2. FACTIBILIDAD

Considerando que el punto pericial Nro. 2 requería la obtención de un registro indubitado de voz de la persona de referencia para su posterior cotejo, se procedió al examen inicial del archivo de audio de mayor duración. Esta selección tuvo por objeto determinar si dicho material cumplía con las condiciones mínimas necesarias de aptitud para fines de comparación y, consecuentemente, evaluar la necesidad de requerir la correspondiente toma de muestra indubitada de voz.

Dicha metodología de trabajo respondió a una decisión adoptada en el marco de las reuniones con peritos de parte y tuvo carácter instrumental respecto de la planificación de las etapas posteriores.

Como primera etapa del examen se procedió a evaluar la factibilidad del material sonoro, a los fines de determinar cuáles de los registros presentaban condiciones suficientes para ser considerados susceptibles de un posterior análisis comparativo de voz.

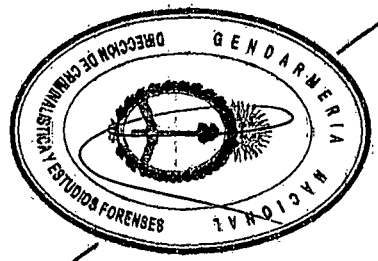
A tal efecto, se estableció como instancia inicial de evaluación la determinación de la duración efectiva del habla contenida en cada registro. El Protocolo Único para la Comparación Forense de Voces[4], en adelante, "el Protocolo", contempla la duración efectiva del habla como una de las condiciones de factibilidad para la comparación forense de voces y permite la determinación de su suficiencia a los requerimientos del método o herramienta de comparación empleados.

En ese marco, y considerando que el informe técnico correspondiente a la herramienta de comparación disponible en la División Acústica Forense [5] establece un umbral mínimo de 8 segundos de habla efectiva para su procesamiento, se adoptó dicho valor como criterio operativo de evaluación. En consecuencia, aquellos registros cuya duración total era inferior al mencionado umbral fueron considerados NO APTOS. Ello se debe a que, si el archivo de audio tenía inicialmente una duración inferior a 8 segundos, resultaba imposible que, mediante su análisis, pudiera alcanzar la duración mínima de habla efectiva requerida. Por tal motivo, dichos registros fueron descartados automáticamente, sin que resultara necesario evaluar las restantes condiciones de factibilidad.

---

[4] Ministerio de Seguridad de la Nación, Secretaría de Lucha contra el Narcotráfico y la Criminalidad Organizada (2024). *Protocolo Único para la Comparación Forense de Voces*. Resolución Ministerial N.º 526/2024.

[5] Martínez Soler et al. (2020), *FORENSIA 2.0: Sistema de Identificación Forense de Hablantes*, apartado "Características técnicas", BlackVOX-CONICET.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
12
DIR. CRIMINALÍSTICA Y EST. FORENSES

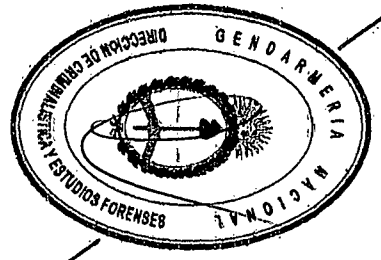
Tabla 2. Archivos de audio de duración inferior al umbral mínimo de 8 segundos, considerados NO APTOS.

Nombre del archivo	Duración de la pista de audio (hh:mm:ss)
es insostenible este nivel de interna.mp4	00:00:07,988
estos dos ya tienen mas ruido que una lata llena de monedas.mp4	00:00:07,523
SE HACE LA RE PELOTUDA JAVIER DESENTENDIDO DE TODO.mp4	00:00:07,407
el no está metido pero.mp4	00:00:07,275
800 lucas por mes.mp4	00:00:06,421
Lule vino de una manera a mi me estan desfalcando a la gente.mp4	00:00:06,080
victoria se canso.mp4	00:00:05,632
esto lo hacen de ratas.mp4	00:00:05,611
ESTO ESTÁ IMPLOSIONANDO DE ADENTRO MEDIO PELOTUDA SANDRA.mp4	00:00:05,550
karina lo mima a guillermo para pegarle a victoria.mp4	00:00:05,333
KILOMBOS GROSOS CON DIANA MONDINO.mp4	00:00:05,271
ella se lleva medio palo.mp4	00:00:04,971
a mi me pusieron un tipo que maneja toda la caja mia.mp4	00:00:04,693
La caja .m4a	00:00:04,365
victoria claramente se cansó.mp4	00:00:04,309
que conocimiento tiene karina.mp4	00:00:04,011
voracidad genética.mp4	00:00:03,989
es un cachito asi.mp4	00:00:03,968
diego sos bueno.mp4	00:00:03,755
se lo dijo sandra, se lo dije yo.mp4	00:00:03,158
vos viste la entrevista que me hizo fantino.mp4	00:00:02,176
a la primera que se van a llevar puesta.mp4	00:00:02,005
daniel garbellini.mp4	00:00:01,173
lule y karina.mp4	00:00:01,045

3.4.2.1. Duración efectiva de habla

Respecto de los registros que presentaban una duración igual o superior a 8 segundos, se estableció inicialmente como tarea de pre-procesamiento la segmentación de la voz de interés, contemplándose asimismo la identificación y eventual eliminación de aquellos segmentos que presentaran ruidos interferentes, superposición de voces u otras señales ajenas a la voz objeto de análisis. Ello con el fin de disponer de una señal adecuada para la posterior verificación de condiciones de factibilidad del Protocolo.

Cabe mencionar que, en esta instancia, no se aplicaron técnicas de procesamiento digital destinadas a modificar las características acústicas de la voz de interés, tales como filtros digitales o algoritmos de reducción de ruido. Obedeciendo a que, si bien estas técnicas pueden mejorar percepción en la inteligibilidad de una grabación, también pueden modificar las características espectrales y temporales de la


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

señal e introducir artefactos adicionales, afectando los parámetros acústicos que posteriormente se extraen y utilizan en los sistemas automáticos de comparación forense de voces, entre ellos, los coeficientes cepstrales de frecuencia de Mel (MFCC) junto con sus derivadas temporales de primer y segundo orden. En este sentido, una mejora en la calidad perceptiva del audio no implica necesariamente una mejora en su aptitud para el análisis, ni garantiza que se conserven inalteradas las características relevantes para la comparación. Por tal motivo, se procuró preservar la información acústica original del material recibido, conforme a lo establecido en el apartado "4.3 Pre-procesamiento" del Protocolo.

La evaluación de factibilidad se inició durante las reuniones periciales, en presencia de los peritos de parte intervinientes. En el transcurso de dichas jornadas se examinó progresivamente el material sonoro y se evaluaron las condiciones acústicas de los archivos abordados en cada instancia.

A partir de la quinta reunión de peritos de parte, celebrada el 23 de septiembre de 2026, y conforme a lo acordado por unanimidad, se estableció que los peritos de la División Acústica Forense continuarán las tareas de preprocesamiento y evaluación de factibilidad de los archivos de audio sin la presencia de los peritos de parte, a fin de agilizar el desarrollo del examen.

En este contexto, las tareas de preprocesamiento, en particular la segmentación de la voz de interés, requerían la escucha detallada de cada registro. Esta instancia permitía advertir de manera preliminar características del material sonoro que debían evaluarse durante la etapa de verificación de factibilidad. Cuando, durante dicha escucha, se identificaba una condición acústica cuya evaluación podía determinar por sí misma la aptitud del registro (por ejemplo, una relación señal-ruido insuficiente), se procedía a realizar la evaluación correspondiente. Si el resultado no cumplía con el criterio establecido, el registro se consideraba como NO APTO y se daba por concluido su estudio de factibilidad, dado que el incumplimiento de ese criterio era suficiente para determinar su no aptitud, sin necesidad de completar las restantes verificaciones previstas en el Protocolo.

Se detallan en la Tabla 3, los archivos de audios cuya duración es mayor a 8 segundos, sobre los cuales se llevan a cabo los procedimientos de evaluación de aptitud para una posterior comparación forense de voz.

Tabla 3. Archivos de audio de duración superior al umbral mínimo de 8 segundos, sometidos a evaluación de APTITUD.

Nombre del archivo	Duración de la pista de audio (hh:mm:ss)
que tiene lule menem con Karina.mp4	00:03:11,000
villarruel ampliado y sandra pettovello.mp4	00:03:11,000
sturzzeneger.mp4	00:02:42,000
sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4	00:01:42,000
Zafarrancho.m4a	00:01:36,000
sueldos.mp4	00:01:33,000
nosotros lo subimos a presidencia.mp4	00:00:36,800
Chupa pija.m4a	00:00:33,530
el dia que karina tenga que salir a hablar en la tele.mp4	00:00:32,322

GENDARMERIA NACIONAL

DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
14
DIR. CRIMINALÍSTICA Y EST. FORENSES

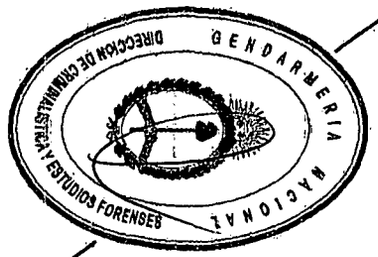
yo hablé con el presidente.mp4	00:00:30,272
javier piensa que estamos haciendo el mejor gobierno de la historia.mp4	00:00:29,884
le afanaron la caja también a Mondino.mp4	00:00:28,672
estos tipos aprovechan y quieren armar su kioskito.mp4	00:00:27,349
Ejemplo torta .m4a	00:00:24,000
es un error tremendo el que cometieron con villarruel.mp4	00:00:23,684
el año pasado pagué 80 lucas de luz.mp4	00:00:22,430
Fake News adorni.m4a	00:00:22,407
el 3 es para Karina.mp4	00:00:19,989
Denuncia.m4a	00:00:18,603
Javi que hago.mp4	00:00:17,835
SANDRA SE HACE LA PELOTUDA.mp4	00:00:17,786
me metio el mas importante de todos.mp4	00:00:16,981
javier me dice yo voy a hablar con Lule.mp4	00:00:14,763
el daño que puede hacer victoria.mp4	00:00:14,357
después viene Karina con Santiago y los Menem.mp4	00:00:13,978
ella iba a manejar seguridad y defensa.mp4	00:00:13,781
no lo quiere nadie, pero nadie.mp4	00:00:13,421
tenes chorros en lule tucumán.mp4	00:00:13,355
nosotros la votamos a victoria.mp4	00:00:11,925
la gente de la suizo.mp4	00:00:10,731
ponela a Karina de candidata y que vaya a hablar.mp4	00:00:10,704
ese es un delincuente.mp4	00:00:10,368
espert pedazo de hijo de puta.mp4	00:00:09,636
yo no se como va a terminar la economia.mp4	00:00:08,171

Como resultado del procedimiento de segmentación, se generaron nuevos archivos de audio correspondientes a los fragmentos de interés, identificados mediante el agregado del sufijo "_limpia" al nombre original de cada archivo y utilizando la extensión ".wav".

A los fines de dar integridad los archivos de audio generados, se calcularon sus funciones HASH (SHA-256), valores que se exhiben en la siguiente tabla junto con la duración de cada archivo de audio limpio.

Tabla 4. Valores HASH (MD5 y SHA256) y duración de los archivos de audio

Archivo de audio	MD5	SHA256	Duración
espert pedazo de hijo de puta_limpia.wav	823b620c0a3019cd e4f217492f7386d0	37379e97652e4ee81cbacc085afeea3a db4253bc6f7ec3b8e4023d56789eb8f2	00:00:06,257
javier piensa que estamos haciendo el mejor gobierno de la historia_limpia.wav	ac68eb9b120f69dc 9070574cd6c59fea	bd9442d87e13769d31951db78a4b4c4 588fb6f3ad79214f16038338ea9a5f37f	00:00:13,569
nosotros la votamos a victoria_limpia.wav	870eb74bbc782d49 eba8a8cdebf86545	1fd08248ca94a763c5439c77c7573ad8 a51fe4b5ff389bdbd4821f0d4e71cefe	00:00:09,357
que tiene lule menem con Karina_limpia.wav	dccf652c040865dcb 09b2e9a9d9ba2aa	4d2c1058cc7ba023d5b84d8ce30d11c5 a597d4fb106725b420c48ea57d744ca4	00:02:01,110
SANDRA SE HACE LA PELOTUDA_limpia.wav	99b1f5ef53d5da71e 54c595740db5351	a3eeaac752629f6bb046fdd74af0a4441 897e0d9db0d26a443d02179ab6f5df0	00:00:11,072


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y CIENCIAS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
15

sturzzeneger_limpia.wav	c30cbcf2deb95ef24d8176aaedcaa54a	57f4a588b8611ca3ea3655a236e76b77199eb8477d28821782a8b355319f926a	00:02:00,962
sucalesca y proyecto discapacidad lilia y romina puto de mierda_LIMPIA.wav	e6f60be52fe7f494a61a2013cbf34f3f	8a1a27d6883031fe0b9e01d22b3e9ee2ec740ac53ea226bc7082d39c2c871fdc	00:00:58,630
yo no se como va a terminar la economia_limpia.wav	11405998f10ee9457db347d5651c6153	9da87279245a12218217e74b9192accaf5e0efe1e01f6a2c13e16e3ab717a702	00:00:04,296
después viene Karina con Santiago y los Menem_limpia.wav	9f5fb58f2997252ca9437172e556896e	b38c51570e3d8d03cd9b7e13e9ac9981919e54cdbe45281b888be8bc87307193	00:00:08,765
el año pasado pagué 80 lucas de luz_limpia.wav	b5e024a8c045cb13f0cf04d6132a13e7	640cf8b60e607a72815f1520f90192c070a0f2b1a9d1ad3fcdcb702e9a398e0f	00:00:15,778
el daño que puede hacer victoria_limpia.wav	6ba4f7b662e86bc28edb5062d606306e	764213ce7d460ab05a2112dd77815be5c0f8296f3fcdea6216d9f0c7f5e99c98	00:00:11,622
el dia que karina tenga que salir a hablar en la tele_LIMPIA.wav	fb2ad240684411073eff1c1b4b1d43c5	cedeb527d41e222dacfbc6d577626c3ec532a2253fb514b3a48efe095ccfc9f0	00:00:16,322
ella iba a manejar seguridad y defensa_limpia.wav	2a2b5576233519b7ddd19b8f71b2b9eb	b2fa78a34c8d5fb99186aee76d42d4c128fcd0c0c41d9c8baa1efa18dcc546d4	00:00:09.780

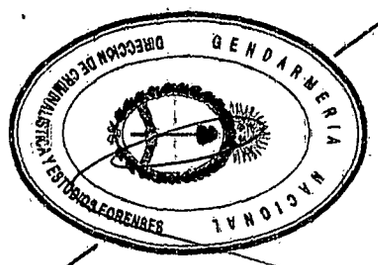
3.4.2.2. Estudio de la relación señal-ruido y dinámica de la señal

Se efectuó la estimación de la relación señal-ruido (SNR, por sus siglas en inglés *Signal to Noise Ratio*) mediante la ejecución de los algoritmos disponibles en SNREval v0.5, del proyecto "Medidas Objetivas de Calidad del Habla/SNR" de LabROSA[6]. Este conjunto de herramientas comprende diversas de funciones desarrolladas en MATLAB (entorno de programación y cálculo numérico), orientadas principalmente a obtener medidas objetivas relacionadas con la proporción entre la energía correspondiente al habla y la atribuida a componentes no pertenecientes a ella.

La ejecución de SNREval bajo la configuración utilizada, proporciona tres estimaciones: NIST-STNR, WADA-SNR y SNRvad, expresadas en decibelios (dB), que constituyen aproximaciones al grado de separación entre la señal de habla y el ruido presente en el registro, obtenidas mediante procedimientos y supuestos diferentes. El primer estimador se basa en el análisis estadístico de la distribución de la energía de la señal en intervalos temporales breves, a partir de la cual se estiman los niveles correspondientes al habla y al ruido. Por su parte, WADA-SNR[7] (Waveform Amplitude Distribution Analysis) , utiliza la distribución estadística de las amplitudes de la forma de onda, considerando un modelo en el cual las amplitudes del habla limpia siguen aproximadamente una distribución gamma, mientras que el ruido aditivo presenta una distribución gaussiana. En condiciones experimentales evaluadas por

[6] Laboratory for the Recognition and Organization of Speech and Audio. (s. f.). LabROSA. Columbia University. https://www.ee.columbia.edu/~dpwe/LabROSA/

[7] Kim, C., & Stern, R. M. (2008). *Robust signal-to-noise ratio estimation based on waveform amplitude distribution analysis*. En Proceedings of Interspeech 2008 (pp. 2598–2601). International Speech Communication Association (ISCA). https://doi.org/10.21437/Interspeech.2008-644


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
16
DIR. CRIMINALÍSTICA Y ESTUDIOS FORENSES

sus autores, este último presentó menor sesgo y variabilidad frente a distintos tipos de ruido en comparación con NIST-STNR.

La utilización de estas medidas como indicadores de las condiciones acústicas de las grabaciones encuentra antecedentes en investigaciones sobre reconocimiento automático de hablantes (FASR). En particular, se ha empleado WADA-SNR para caracterizar y seleccionar conjuntos de grabaciones según sus condiciones acústicas, así como para examinar la relación entre dichas condiciones y el rendimiento de sistemas[8,9,10].

De forma complementaria, SNREval proporciona la estimación SNRvad, basada en la detección de actividad vocal (VAD, por sus siglas en inglés Voice Activity Detection). Este procedimiento identifica automáticamente intervalos con y sin actividad vocal a partir del análisis espectral de la señal, considerando la banda comprendida entre 100 y 1000 Hz y aplicando un umbral adaptativo para su clasificación. Una vez identificados dichos intervalos, se calcula la energía media de los segmentos considerados activos, que comprenden tanto habla como ruido, y la correspondiente a los intervalos inactivos comprendidos entre ellos, utilizada para estimar la energía del ruido de fondo. Para esta última estimación, no se consideran los intervalos anteriores al primer segmento activo ni posteriores al último[11]. A partir de estos valores, la relación señal-ruido se estima conforme a la ecuación 1.

$$SNRVad = 10log_{10}\left(\frac{E_{S+N} - E_N}{E_N}\right) \quad (1)$$

Donde E_{S+N} representa la energía media de los segmentos activos y E_N la energía media de los segmentos inactivos, a partir de lo cual se calcula la relación señal-ruido, expresada en decibeles (dB). Cabe aclarar que el cálculo de estas energías se realiza a partir del espectrograma de la señal, previamente normalizado respecto de su valor máximo y limitado a un rango dinámico de 80 dB, considerando todo el espectro analizado y no exclusivamente la banda de 100 a 1000 Hz utilizada para la detección de actividad vocal.

Es necesario señalar que los procedimientos descriptos presentan limitaciones inherentes a sus respectivos modelos y supuestos de estimación, por lo que pueden producir resultados superiores o inferiores a la relación señal-ruido efectiva del material examinado. En consecuencia, los mismos deben interpretarse como indicadores orientativos de las condiciones acústicas del material, cuya valoración debe realizarse conjuntamente con el criterio del examinador.

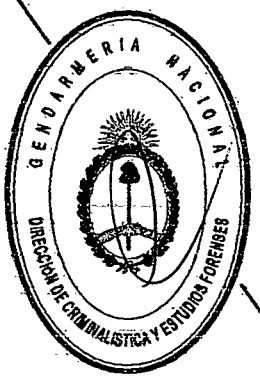
---

[8] Atreya, A., Forth, O., Kent, S., Kelly, F., & Alexander, A. (2018). *Estimating the good, the bad and the ugly in speech recordings*. International Association for Forensic Phonetics and Acoustics (IAFPA) Conference 2018, Huddersfield, United Kingdom.

[9] Bakst, S., Cobo-Kroenke, C., Lawson, A., McLaren, M., & Stauffer, A. (2022). *Time-varying score reliability prediction in speaker identification*. The Speaker and Language Recognition Workshop (Odyssey 2022). https://doi.org/10.21437/Odyssey.2022-29

[10] Hughes, V., Xu, C., Foulkes, P., Harrison, P., Welch, P., Kelly, F., & van der Vloed, D. (2024). *Exploring individual speaker behaviour within a forensic automatic speaker recognition system*. The Speaker and Language Recognition Workshop (Odyssey 2024), 1–8. https://doi.org/10.21437/Odyssey.2024-1

[11] Véase Ellis (2014), snreval.m, líneas 227–228.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. COMUNICACIÓN, CRIMINALÍSTICA Y ESTUDIOS FORENSES
Foja
17

Para su implementación, en esta etapa del análisis, se utilizó una máquina virtual y la versión ejecutable de SNREval disponible en la página oficial del proyecto[12]. Las características del sistema y del software utilizado se consultaron mediante la ejecución, en la terminal del sistema operativo Ubuntu, del Bloque de código 3 el cual brindó los detalles expuestos en el Cuadro 1.

Bloque de código 3. Comando ejecutado para consulta de características.

```
</>
snreval@ACUSTICA:~$ {
> echo "===== SISTEMA ====="
> date
> cat /etc/os-release
> uname -a
> echo
> echo "===== SNREVAL ====="
> file ~/SNREval/snreval_GLNXA64/snreval_prj
> echo
> echo "===== SHA256 PAQUETE ORIGINAL ====="
> sha256sum ~/Downloads/snreval_GLNXA64.zip
> echo
> echo "===== SHA256 ARCHIVOS SNREVAL ====="
> sha256sum ~/SNREval/snreval_GLNXA64/*
> echo
> echo "===== MATLAB COMPILER RUNTIME ====="
> ls -ld /opt/MATLAB/MATLAB_Compiler_Runtime/v714
> } | tee ~/SNREval/informacion_instalacion.txt
```

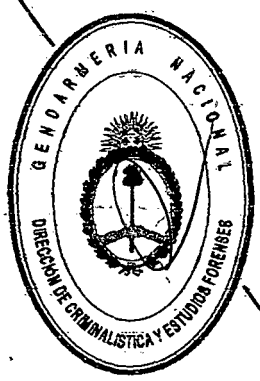
Cuadro 1. Características consultadas del sistema y del código SNREval para evaluación de relación señal-ruido.

```
===== SISTEMA =====
mié 02 sep 2026 10:50:52 -03
NAME="Ubuntu"
VERSION="20.04.6 LTS (Focal Fossa)"
ID=ubuntu
ID_LIKE=debian
PRETTY_NAME="Ubuntu 20.04.6 LTS"
VERSION_ID="20.04"
HOME_URL="https://www.ubuntu.com/"
SUPPORT_URL="https://help.ubuntu.com/"
BUG_REPORT_URL="https://bugs.launchpad.net/ubuntu/"
PRIVACY_POLICY_URL="https://www.ubuntu.com/legal/terms-and-policies/privacy-policy"
VERSION_CODENAME=focal
UBUNTU_CODENAME=focal
Linux ACUSTICA 5.15.0-139-generic #149~20.04.1-Ubuntu SMP Wed Apr 16 08:29:56 UTC 2025 x86_64 x86_64
x86_64 GNU/Linux

===== SNREVAL =====
/home/snreval/SNREval/snreval_GLNXA64/snreval_prj: ELF 64-bit LSB executable, x86-64, version 1 (SYSV), dy-
namically linked, interpreter /lib64/ld-linux-x86-64.so.2, for GNU/Linux 2.6.8, stripped

===== SHA256 PAQUETE ORIGINAL =====
d7c2023dfe1b8ee18c61251af8eea22cc738c07866c8bf33e72121346bca4555                /home/snreval/Down-
loads/snreval_GLNXA64.zip
```

[12] Ellis, D. P. W. (2011). *Objective measures of speech quality/SNR*. LabROSA, Columbia University. https://www.ee.columbia.edu/~dpwe/LabROSA/projects/snreval/


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 18
DIR. CRIMINALÍSTICA Y ESTUDIOS FORENSES

```
===== SHA256 ARCHIVOS SNREVAL =====
8872f354dbf5eae88d5e5bb6f4bfa097872a690b297341b54feb87c720b45656
/home/snreval/SNREval/snreval_GLNXA64/libXp.so.6
14dd60f12376f3b2f5cd0c0833c77d59db0d146b42d54dee81f5aaf632cad35a
/home/snreval/SNREval/snreval_GLNXA64/readme.txt
572be7800f88775196153fed1d2889378812e29c05ba6d1f311edcfdc130b2c6
/home/snreval/SNREval/snreval_GLNXA64/run_snreval_prj.sh
8165a2e759a270f058e051301b4efee2ddee333682523b33dbf12119acd1da24
/home/snreval/SNREval/snreval_GLNXA64/snreval_prj

===== MATLAB COMPILER RUNTIME =====
drwxr-xr-x 12 root root 4096 ago 31 19:16 /opt/MATLAB/MATLAB_Compiler_Runtime/v714
```

El procesamiento de los archivos mediante SNREval requería disponer de una versión de cada uno en formato WAV, con una frecuencia de muestreo de 16.000 Hz y una profundidad de 16 bits. Estas versiones se generaron mediante la herramienta FFmpeg, utilizando el siguiente comando.

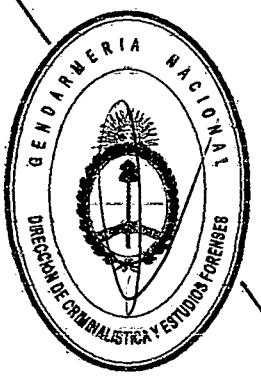
Bloque de código 4. Comando de ejemplo ejecutado obtención de archivos ".wav" en 16.000 Hz.

```
</>
ffmpeg -i "Nombre de archivo.m4a" -map 0:a:0 -ar 16000 -c:a pcm_s16le "Nombre de archivo_wav.wav"
```

Como resultado, se obtuvo una versión de cada uno de los archivos que se encuentran consignados en la Tabla Nro. 5, con las características indicadas, identificados mediante el sufijo "_wav" añadido al nombre original.

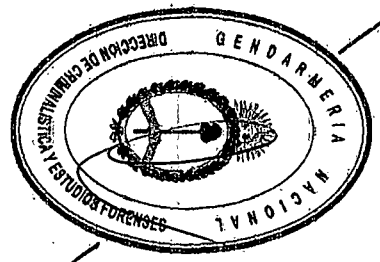
Tabla 5. Archivos obtenidos con frecuencia de muestreo de 16.000 Hz y 16 bits de profundad.

Nombre del archivo	Frecuencia de muestreo	Profundidad de bit
que tiene lule menem con Karina_wav	16000	16
villarruel ampliado y sandra pettovello_wav	16000	16
sturzzeneger_wav	16000	16
sucalesca y proyecto discapacidad lilia y romina puto de mierda_wav	16000	16
Zafarrancho_wav	16000	16
sueldos_wav	16000	16
nosotros lo subimos a presidencia_wav	16000	16
Chupa pija_wav	16000	16
el dia que karina tenga que salir a hablar en la tele_wav	16000	16
yo hablé con el presidente_wav	16000	16
javier piensa que estamos haciendo el mejor gobierno de la historia_wav	16000	16
le afanaron la caja también a Mondino_wav	16000	16
estos tipos aprovechan y quieren armar su kioskito_wav	16000	16
Ejemplo torta_wav	16000	16
es un error tremendo el que cometieron con villarruel_wav	16000	16


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
19
DIR. CRIMINALÍSTICA Y ESTUPEFACIENTES

el año pasado pagué 80 lucas de luz_wav	16000	16
Fake News adorni_wav	16000	16
el 3 es para Karina_wav	16000	16
Denuncia_wav	16000	16
Javi que hago_wav	16000	16
SANDRA SE HACE LA PELOTUDA_wav	16000	16
me metio el mas importante de todos_wav	16000	16
javier me dice yo voy a hablar con Lule_wav	16000	16
el daño que puede hacer victoria_wav	16000	16
después viene Karina con Santiago y los Menem_wav	16000	16
ella iba a manejar seguridad y defensa_wav	16000	16
no lo quiere nadie, pero nadie_wav	16000	16
tenes chorros en lule tucumán_wav	16000	16
nosotros la votamos a victoria_wav	16000	16
la gente de la suizo_wav	16000	16
ponela a Karina de candidata y que vaya a hablar_wav	16000	16
ese es un delincuente_wav	16000	16
espert pedazo de hijo de puta_wav	16000	16
yo no se como va a terminar la economia_wav	16000	16
es insostenible este nivel de interna_wav	16000	16
estos dos ya tienen mas ruido que una lata llena de monedas_wav	16000	16
SE HACE LA RE PELOTUDA JAVIER DESENTENDIDO DE TODO_wav	16000	16
el no está metido pero_wav	16000	16
800 lucas por mes_wav	16000	16
Lule vino de una manera a mi me estan desfalcando a la gente_wav	16000	16
victoria se canso_wav	16000	16
esto lo hacen de ratas_wav	16000	16
ESTO ESTÁ IMPLOSIONANDO DE ADENTRO MEDIO PELOTUDA SANDRA_wav	16000	16
karina lo mima a guillermo para pegarle a victoria_wav	16000	16
KILOMBOS GROSOS CON DIANA MONDINO_wav	16000	16
ella se lleva medio palo_wav	16000	16
a mi me pusieron un tipo que maneja toda la caja mia_wav	16000	16
La caja_wav	16000	16
victoria claramente se cansó_wav	16000	16
que conocimiento tiene karina_wav	16000	16
voracidad genética_wav	16000	16
es un cachito asi_wav	16000	16
diego sos bueno_wav	16000	16
se lo dijo sandra, se lo dije yo_wav	16000	16
vos viste la entrevista que me hizo fantino_wav	16000	16
a la primera que se van a llevar puesta_wav	16000	16
daniel garbellini_wav	16000	16
lule y karina_wav	16000	16


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
20

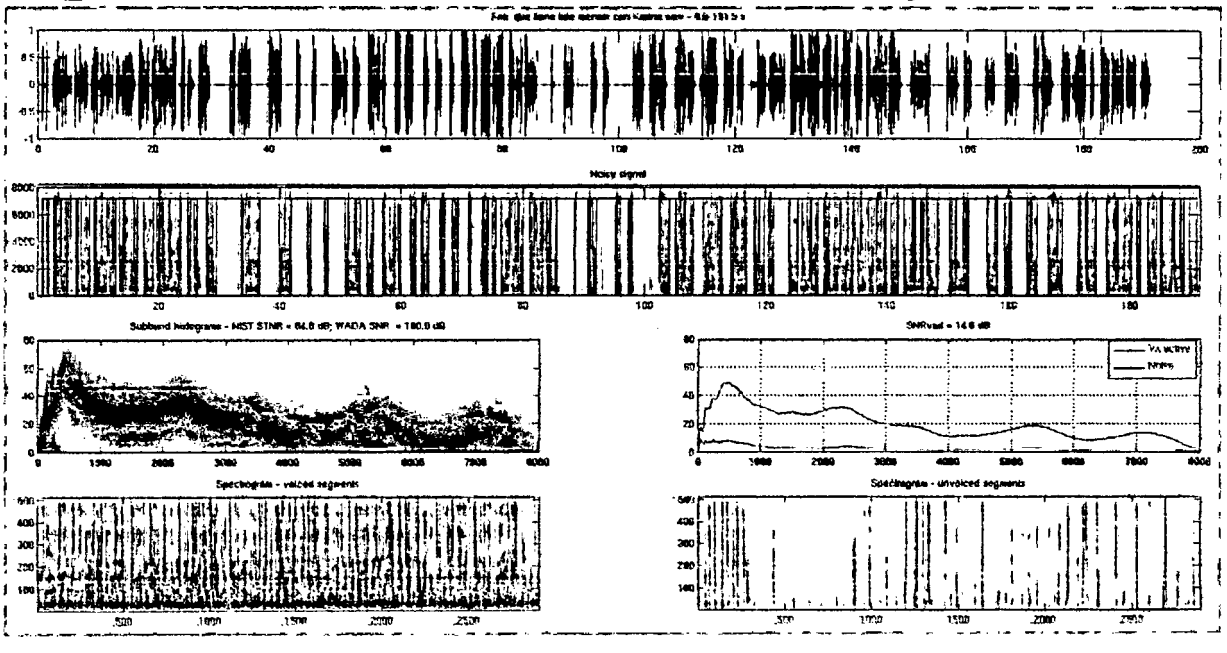
Esta configuración se adoptó con el propósito de disponer de archivos de audio con características uniformes y compatibles con la versión ejecutable de SNREval utilizada. A continuación, se presenta un ejemplo del comando ejecutado.

Bloque de código 5. Comando de ejemplo ejecutado para medición de SNR.

```
</>
cd ~/SNREval/snreval_GLNXA64
./run_snreval_prj.sh "/media/sf_VM_SNREval_input_file/Nombre de archivo_wav.wav" -
guessvad 1 -disp 1
```

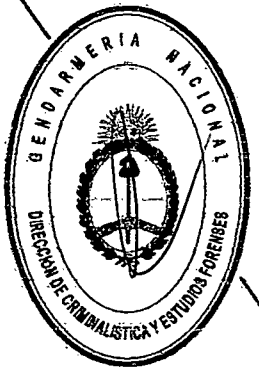
Para el caso del archivo de mayor duración, denominado "que tiene lule menem con Karina", se efectuó el procesamiento conforme al procedimiento previamente descripto y se obtuvieron los resultados que se exponen en la siguiente figura.

Figura 2. Resultados obtenidos de SNREval del audio "que tiene lule con Karina_wav.wav", NIST-STNR=64.8 dB, WADA-SNR=100.0 dB y SNRvad=14.6 dB.



Los valores obtenidos mediante los dos primeros métodos de SNREval indican una diferencia extremadamente elevada entre la potencia atribuida a la señal y la atribuida al ruido. Para dimensionar estos valores, una relación señal-ruido de 60 dB equivale a una potencia de señal un millón de veces mayor que la del ruido ($1:10^6$), mientras que una relación de 100 dB equivale a una diferencia de diez mil millones a uno ($1:10^{10}$).

Con el objetivo de obtener una valoración más representativa de la relación señal-ruido efectiva del audio, se continuó el análisis mediante la inspección de la forma de onda y espectrograma de la señal. Para tal fin, se utilizó el software Audacity versión 3.7.9, para la visualización de la forma de onda y del espectrograma de


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
21
DIR. CRIMINALÍSTICA Y EST. FORENSES

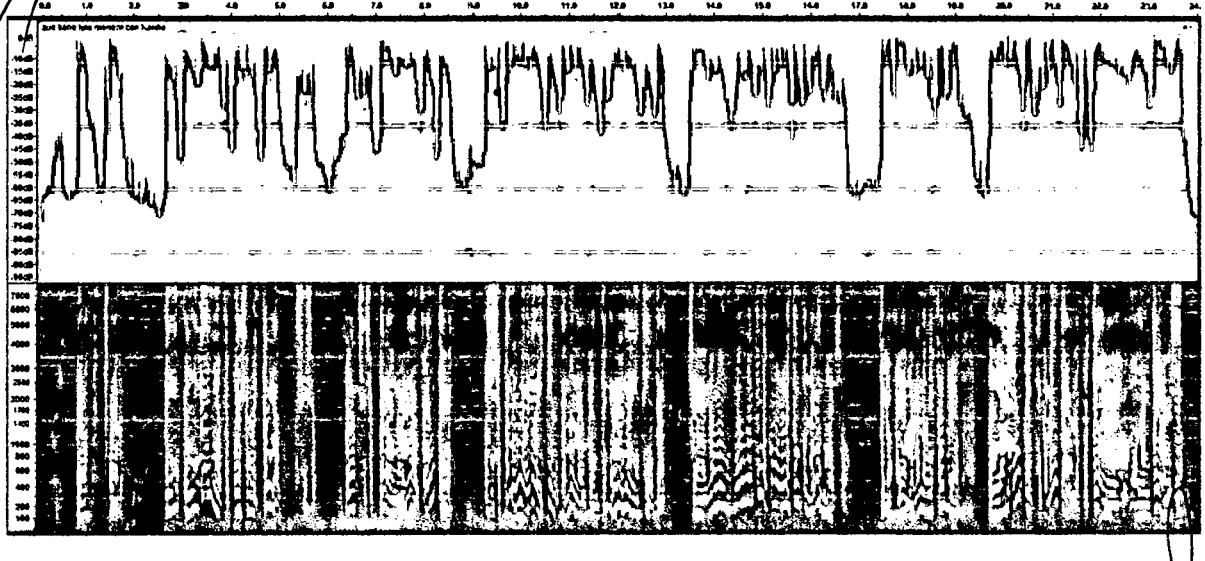
la pista de audio, empleando escala de Mel, ventana de Hamming de 1024 muestras y factor de zero-padding[13] de 2, como muestra la siguiente figura.

Figura 3. Forma de onda en escala logarítmica (arriba) y espectrograma de la pista de audio (abajo) del archivo "que tiene lule menem con Karina.mp4".

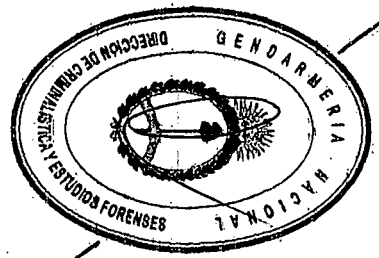


A modo de ejemplo, para una observación de mayor detalle, la siguiente figura expone los instantes iniciales con habla entre los 00:00:00,000 y los 00:00:24,000 segundos de la pista de audio, utilizando visualización de la forma de onda en media onda.

Figura 4. Instantes iniciales de la forma de onda en escala logarítmica y media onda (arriba) y espectrograma de la pista de audio (abajo) del archivo "que tiene lule con Karina.mp4".



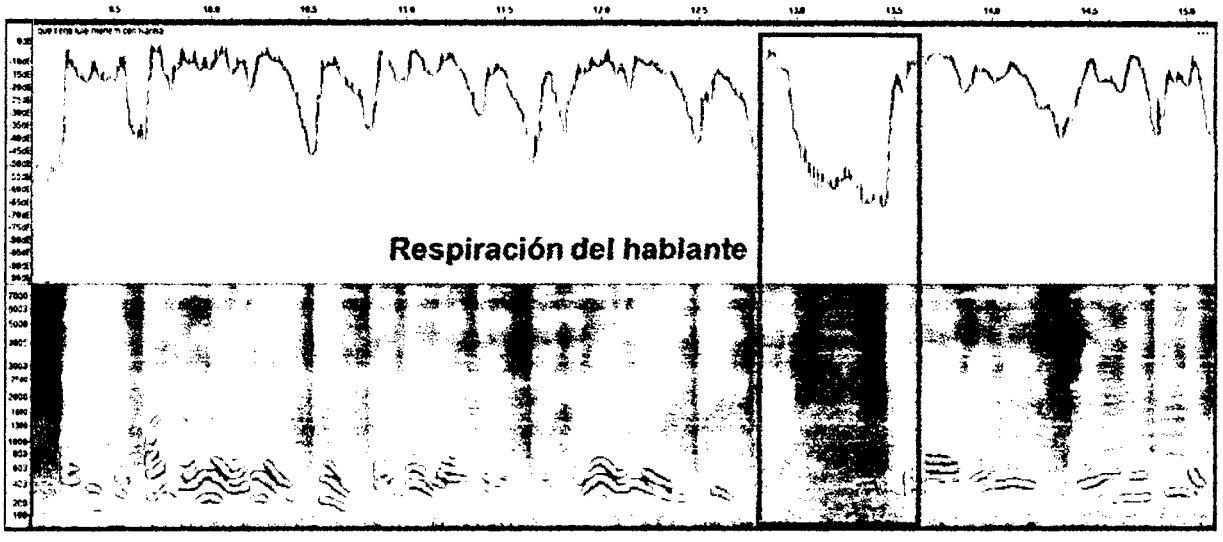
[13] Factor que determina la cantidad de ceros añadidos a la señal antes del cálculo de la FFT, aumentando la cantidad de puntos del espectro y su densidad de representación, sin modificar la resolución espectral real.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

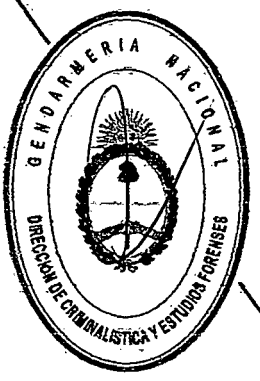
GENDARMERÍA NACIONAL
Foja
22
DIR. CRIMINALÍSTICA Y EST. FORENSES

La escucha perceptual y la inspección del espectrograma permiten identificar componentes acústicos en periodos sin fonación, entre ellos un evento compatible con la respiración del hablante, aproximadamente en el instante 00:00:13,232. Para examinar estas propiedades con mayor profundidad, se seleccionó el segmento comprendido entre los 00:00:09,088 a 00:00:15,141 segundos de la grabación. La siguiente figura presenta la ampliación mencionada, donde se observa la distribución espectral de la energía durante la fonación y el evento respiratorio identificado sin actividad vocal. Para facilitar la visualización de los componentes espectrales de menor nivel, se configuró el espectrograma con un rango de representación de 120 dB.

Figura 5. Forma de onda en media en escala logarítmica y espectrograma de la pista de audio entre los instantes 00:00:09,088 a 00:00:15,141 del archivo "que tiene lule con Karina.mp4" con el recuadro del sonido correspondiente a la respiración.

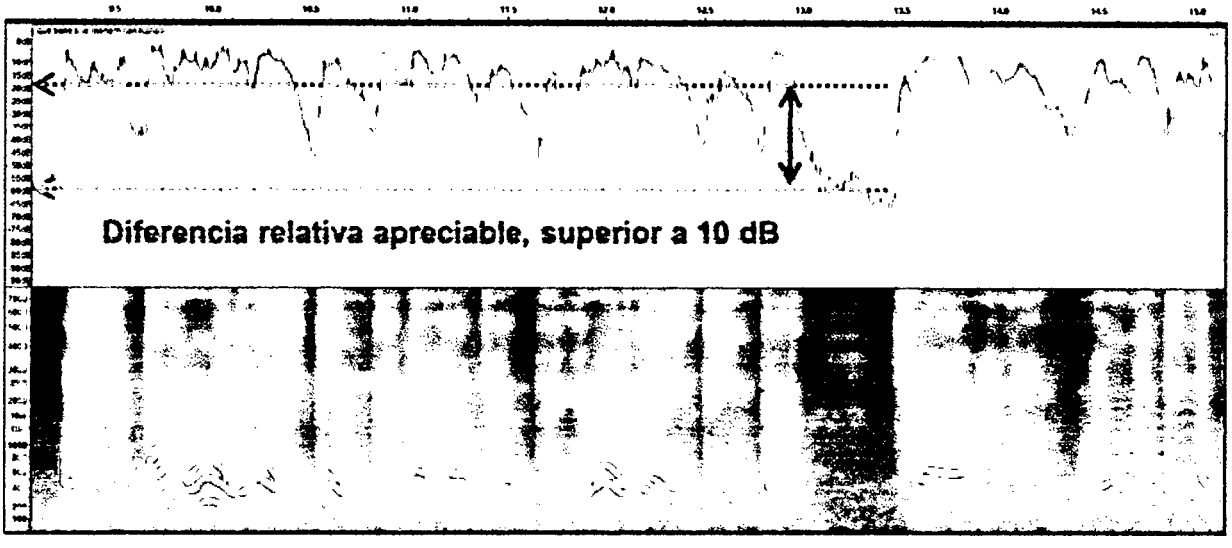


Independientemente de las medidas de SNREval obtenidas, la observación de la forma de onda y espectrograma permite apreciar cualitativamente una diferencia considerable de nivel entre los segmentos con habla y los intervalos sin actividad vocal, como se expone en la Figura 6.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

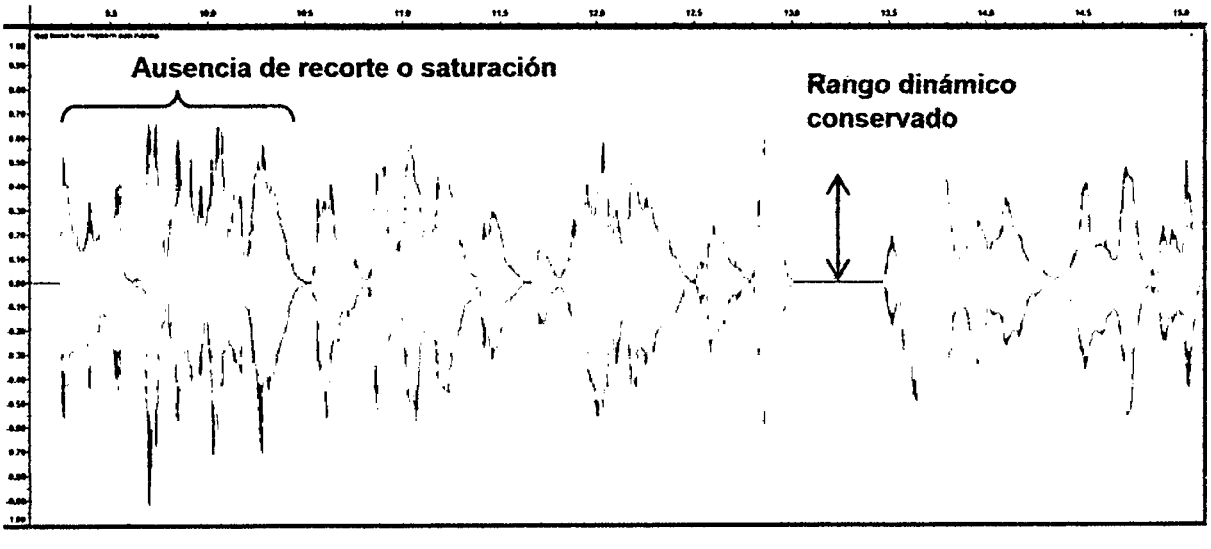
GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
23

Figura 6. Forma de onda en media onda en escala logarítmica y espectrograma de la pista de audio entre los instantes 00:00:09,083 a 00:00:15,115 del archivo "que tiene lule *menem con Karina.mp4*" con la diferencia relativa observada.



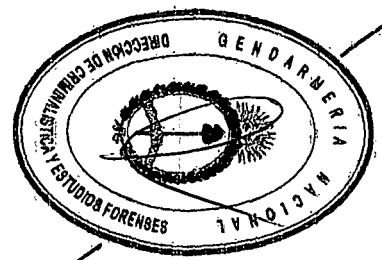
Por otra parte, respecto de posibles distorsiones de la forma de onda atribuibles a saturación de nivel, no se observan crestas con recorte ni pérdida apreciable de rango dinámico en la representación examinada. A modo de ejemplo, la Figura 7 presenta la forma de onda en escala lineal correspondiente al segmento temporal previamente analizado.

Figura 7. Forma de onda en escala lineal de la pista de audio entre los instantes 00:00:09,083 a 00:00:15,115 del archivo "que tiene lule menem con Karina.mp4".



Se continuó con la evaluación de la relación señal ruido (SNR) en los restantes archivos de audio, siguiendo el procedimiento expuesto anteriormente. Las observaciones efectuadas para cada archivo se consignan en la Tabla Nro. 6.

La totalidad de las imágenes correspondientes a los resultados obtenidos mediante SNREval se presentan en el Anexo I.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

3.4.2.3. Reverberación acústica

En el marco de la evaluación de factibilidad del material sonoro, se examinó además la presencia de reverberación acústica y su posible efecto sobre la señal de habla del audio dentro del archivo "que tiene lule menem con Karina.mp4". El espectrograma (ver Figura 4) permite distinguir la frecuencia fundamental (F0) asi como sucesivos armónicos del habla durante los segmentos de voz, comportamiento que se mantiene a lo largo del resto de contenido de audio del archivo.

Por otra parte, la escucha de la grabación permite establecer que el contenido del habla de los registros resulta, en términos generales, inteligible. Desde el campo de la acústica, tanto la reverberación como el ruido de fondo constituyen factores que pueden afectar negativamente la inteligibilidad. En este sentido, la ausencia de dichos fenómenos en el archivo de audio mencionado es concordante con lo observado durante el examen acústico realizado.

Tanto la escucha crítica como la visualización del espectrograma no permitieron establecer la existencia de un grado elevado o suficiente de reverberación, ni fenómenos físicos asociados, que resulten en la afectación de la distribución de la energía espectral de los armónicos.

Cabe mencionar que, del mismo modo, no se observan pérdidas de información producto de artefactos de procesamiento en el rango espectral del habla (ver Figuras Nros. 3, 4 y 5).

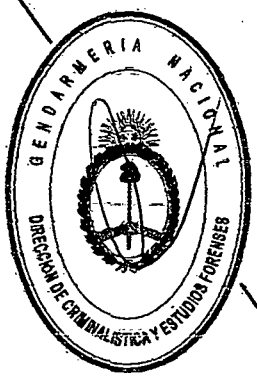
La metodología de análisis descripta se aplicó al resto de los archivos de audio mediante escucha crítica y examen de las señales, a fin de evaluar la presencia de reverberación acústica, ruido de fondo y su eventual incidencia en la inteligibilidad del habla. Las observaciones efectuadas para cada archivo se consignan en la Tabla Nro. 6.

3.4.2.4. Resultados de Factibilidad

La evaluación de factibilidad efectuada sobre el archivo denominado "que tiene lule menem con Karina.mp4" permitió determinar que reunía las condiciones previstas en el Protocolo para ser considerado APTO. Este resultado fue comunicado al finalizar la reunión de trabajo del 2 de septiembre de 2026, documentada en la Constancia de Trabajo Nro. 2.

En esa oportunidad, a solicitud de los peritos de parte y con la conformidad de los peritos oficiales, se acordó comunicar al Juzgado interviniente, el resultado a los fines de gestionar la convocatoria de la persona de referencia para la obtención de muestras indubitadas de su voz.

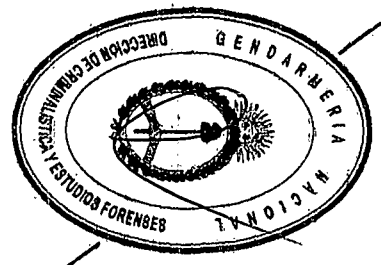
La siguiente tabla muestra el resumen de los resultados obtenidos, detallando el cumplimiento o no de la condición de factibilidad, para cada audio analizado.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
25
DIR CRIMINALÍSTICA Y EST FORENSES

Tabla 6. Resumen de resultados obtenidos del material objeto de estudio conforme la etapa "4.4. Verificación de factibilidad de la señal de audio" del Protocolo.

Nombre del archivo	Observación	Resultado del análisis
que tiene lule menem con Karina.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación.	**cumple con las condiciones de aptitud**
villarruel ampliado y sandra pettovello.mp4	Archivo con contenido de audio exacto de "que tiene lule menem con Karina.mp4"	Idéntico contenido y mismo valor hash que "que tiene lule menem con Karina.mp4"
sturzzeneger.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación.	**cumple con las condiciones de aptitud**
sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación.	**cumple con las condiciones de aptitud**
Zafarrancho.m4a	SNR insuficiente con presencia de bable noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
sueldos.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
nosotros lo subimos a presidencia.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
Chupa pija.m4a	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
el dia que karina tenga que salir a hablar en la tele.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
yo hablé con el presidente.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
javier piensa que estamos haciendo el mejor gobierno de la historia.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
le afanaron la caja también a Mondino.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
estos tipos aprovechan y quieren armar su kioskito.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y ESTUDIOS FORENSES
Foja
26

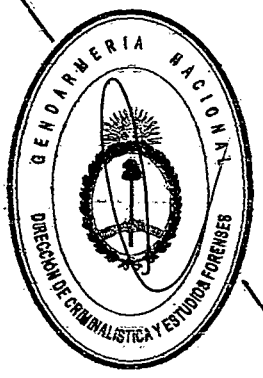
Ejemplo torta .m4a	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
es un error tremendo el que cometieron con villarruel.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
el año pasado pagué 80 lucas de luz.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
Fake News adorni.m4a	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
el 3 es para Karina.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
Denuncia.m4a	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
Javi que hago.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
SANDRA SE HACE LA PELOTUDA.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
me metio el mas importante de todos.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
javier me dice yo voy a hablar con Lule.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
el daño que puede hacer victoria.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
después viene Karina con Santiago y los Menem.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
ella iba a manejar seguridad y defensa.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
no lo quiere nadie, pero nadie.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud

GENDARMERÍA NACIONAL

DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
27
DIR CRIMINALÍSTICA Y EST FORENSES

tenes chorros en lule tucumán.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
nosotros la votamos a victoria.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
la gente de la suizo.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
ponela a Karina de candidata y que vaya a hablar.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
ese es un delincuente.mp4	SNR insuficiente con presencia de babble noise, sin grado elevado de reverberación, tiempo de habla efectivo suficiente	no cumple con las condiciones de aptitud
espert pedazo de hijo de puta.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
yo no se como va a terminar la economia.mp4	SNR y tiempo de habla efectivo suficiente, sin grado elevado de reverberación	**cumple con las condiciones de aptitud**
es insostenible este nivel de interna.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
estos dos ya tienen mas ruido que una lata llena de monedas.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
SE HACE LA RE PELOTUDA JAVIER DESENTENDIDO DE TODO.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
el no está metido pero.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
800 lucas por mes.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
Lule vino de una manera a mi me estan desfalcando a la gente.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
victoria se canso.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
esto lo hacen de ratas.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
ESTO ESTÁ IMPLOSIONANDO DE ADENTRO MEDIO PELOTUDA SANDRA.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
karina lo mima a guillermo para pegarle a victoria.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
28
DIR. CRIMINALÍSTICA Y EST. FORENSES

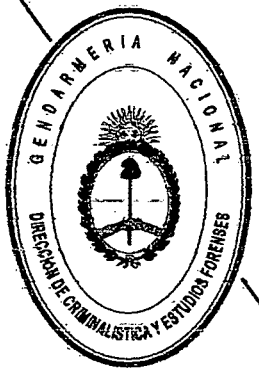
KILOMBOS GROSOS CON DIANA MONDINO.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
ella se lleva medio palo.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
a mi me pusieron un tipo que maneja toda la caja mia.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
La caja .m4a	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
victoria claramente se cansó.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
que conocimiento tiene karina.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
voracidad genética.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
es un cachito asi.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
diego sos bueno.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
se lo dijo sandra, se lo dije yo.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
vos viste la entrevista que me hizo fantino.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
a la primera que se van a llevar puesta.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
daniel garbellini.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud
lule y karina.mp4	tiempo de habla efectivo insuficiente	no cumple con las condiciones de aptitud

3.4.3. INTEGRIDAD DEL MATERIAL SONORO

Respecto del análisis de integridad, esta División continuó el examen del material sonoro con posterioridad a las reuniones de trabajo, lo que permitió realizar una revisión exhaustiva de los registros. Dicho examen estuvo orientado a la identificación de posibles cortes, discontinuidades y otras anomalías observables, así como al estudio de características y fenómenos acústicos relevantes para la evaluación de su integridad.

3.4.3.1. Alcance y limitaciones del análisis

La finalidad de esta etapa es identificar y localizar eventuales indicios técnicos de modificaciones posteriores a la captación original o inconsistentes con la grabación y, cuando los hallazgos lo permitan, describir su naturaleza y evaluar su compatibilidad con determinadas operaciones de edición. El abordaje adoptado toma


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
29
DIR CRIMINALÍSTICA Y EST FORENSES

como referencia las buenas prácticas establecidas para el examen forense de autenticidad de audio[14,15], particularmente aquellas relacionadas con la identificación de inconsistencias en las trazas del proceso de grabación y de posibles indicios de procesamiento posterior. En este sentido, no se dispone de información contextual independiente aportada relativa al dispositivo o sistema empleado para la captura, las condiciones acústicas ni circunstancias de grabación, ni de elementos que permitan verificar la correspondencia de los archivos examinados con aquellos generados originalmente durante su adquisición. Tampoco se cuenta con antecedentes acerca de eventuales procesos de tratamiento a los que pudieran haber sido sometidos con anterioridad a su entrega para examen. En consecuencia, y de acuerdo con las limitaciones contempladas en el estado del arte, las conclusiones del presente análisis se condicionan a los hallazgos técnicos obtenidos mediante los procedimientos aplicados a los datos disponibles y a las conclusiones que estos permitan sustentar.

3.4.3.2. Generación de archivos complementarios

A fin de disponer de material complementario necesario para la realización de análisis posteriores, inicialmente se generaron archivos adicionales a partir de los registros sonoros. En primera instancia, mediante el programa Audacity, se duplicó la pista de audio de cada archivo y se aplicaron las expresiones en lenguaje Nyquist que se presentan en los Bloques de código 4 y 5.

Bloque de código 4. Expresión Nyquist aplicada para el cálculo de la primera diferencia discreta.

```
</>
(diff *track* (at-abs (/ 1.0 *sound-srate*) (cue *track*))))
```

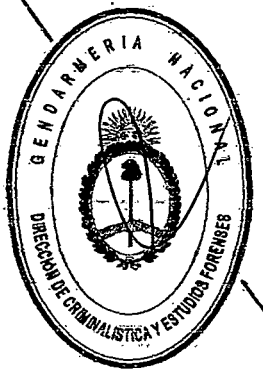
Bloque de código 5. Expresión Nyquist aplicada para obtener el valor absoluto de la señal resultante.

```
</>
(s-abs *track*)
```

La implementación sucesiva de ambas expresiones permite obtener el valor absoluto de la primera diferencia discreta hacia atrás de la señal, como expresa en la ecuación 1, con el propósito de destacar variaciones abruptas de amplitud entre muestras consecutivas.

$$|\nabla x[n]| = |x[n] - x[n-1]| \quad (2)$$

[14] European Network of Forensic Science Institutes. (2022). *Best practice manual for digital audio authenticity analysis* (Version 01, ENFSI-FSA-BPM-002).
[15] Scientific Working Group on Digital Evidence. (2018). *SWGDE best practices for digital audio authentication* (Version 1.3).


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
30
DIR. CRIMINALÍSTICA Y EST. FORENSES

Donde $x[n]$ representa la amplitud de la señal original en la muestra n, $x[n-1]$la amplitud de la muestra inmediatamente anterior, ∇ el operador de primera diferencia discreta hacia atrás y $|.|$ la operación de valor absoluto o módulo.

Este procedimiento permite en determinadas circunstancias destacar variaciones abruptas de amplitud entre muestras consecutivas, que pueden constituir puntos de interés a examinar. Las señales resultantes fueron exportadas y almacenadas en archivos independientes, identificados mediante el nombre original seguido del sufijo "_derivada".

Asimismo, para la totalidad de los archivos (los cuales contienen audio codificado mediante AAC), se utilizó la herramienta FFmpeg mediante el comando presentado en el Bloque de código 6, para extraer el flujo de audio codificado del contenedor y almacenarlo en un archivo con extensión ".aac", utilizando el formato ADTS, sin efectuar una nueva codificación de la señal. Los archivos resultantes fueron identificados mediante el nombre original seguido del sufijo "_raw".

Bloque de código 6. Ejemplo de código para desempaquetado.

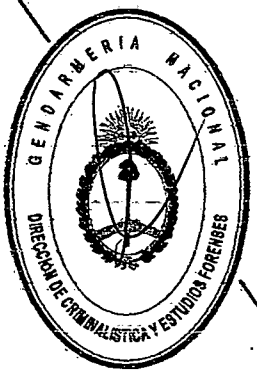
```
</>
ffmpeg -i "Nombre de archivo.m4a" -map 0:a:0 -c:a copy -f adts "Nombre de archivo_raw.aac"
```

Se deja constancia que la totalidad de los archivos generados se aportan junto al presente informe pericial.

3.4.3.3. Identificación y marcación temporal

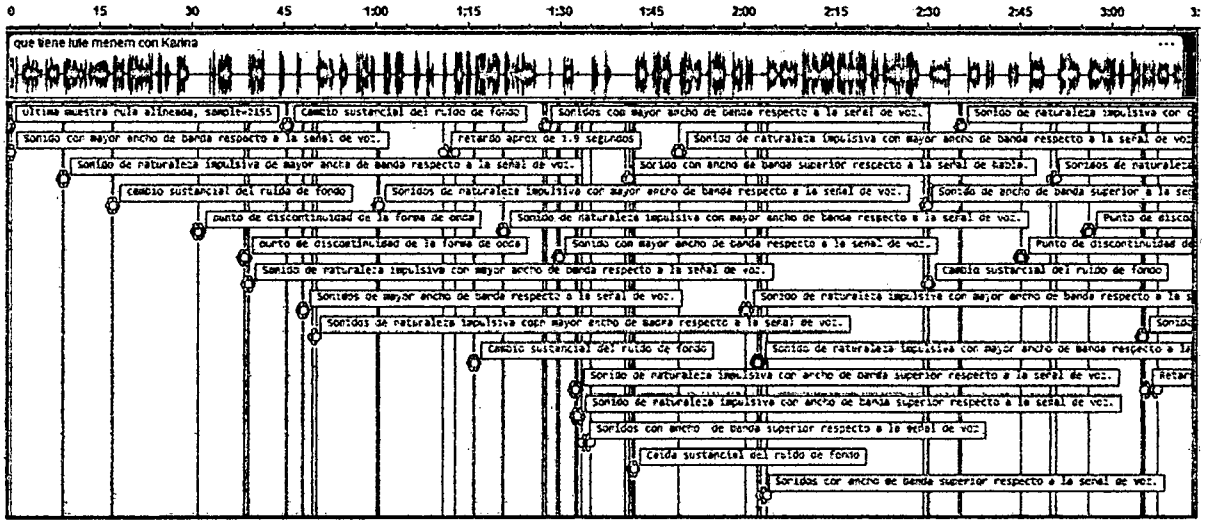
De cada registro se procedió a identificar y consignar temporalmente aquellos segmentos que presentaban características acústicas diferenciadas respecto del desarrollo general de la señal, con atención a posibles interrupciones, discontinuidades y variaciones sustanciales del ruido de fondo.

Para la mencionada tarea se utilizó la pista de etiquetado dentro del software Audacity, generando para cada audio un archivo de texto con extensión ".txt", cuyo nombre se compone del nombre original seguido del sufijo "_etiquetas". En estos archivos, los eventos identificados se registraron mediante su ubicación temporal en la pista (expresada en formato hh:mm:ss) y una descripción del evento, a fin de facilitar su posterior localización. A modo de ejemplo la siguiente figura muestra la pista de audio y la correspondiente pista de etiquetas del archivo de mayor duración, "que tiene lule menem con Karina.mp4".


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
31

Figura 8. Pista de audio del archivo "que tiene lule menem con Karina.mp4" y pista de etiquetas generada.

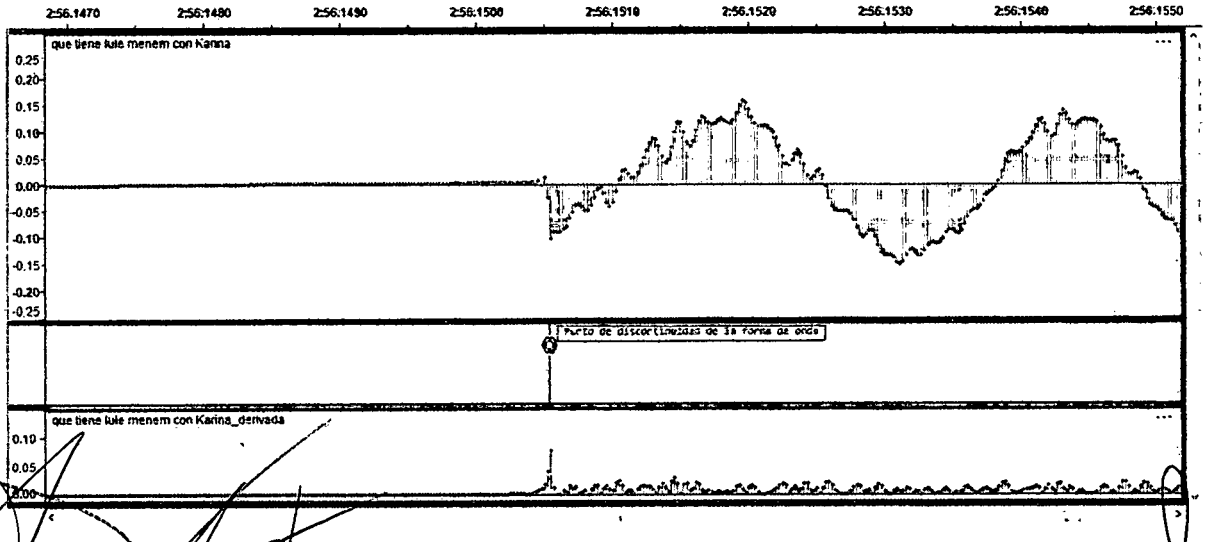


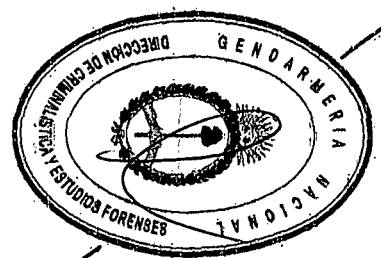
Ante la ausencia de información contextual independiente sobre las condiciones de obtención del registro sonoro, se adoptó como hipótesis de trabajo que este corresponde a una adquisición continúa realizada mediante un mismo sistema de grabación, tomando como referencia la frecuencia de muestreo de 44.100 Hz declarada en sus metadatos, expuestos en las tablas A1 a A4 (ver Anexo I).

Bajo este supuesto, se identificaron múltiples instantes en los que se observaron eventos acústicos cuyas características espectrales y/o temporales presentan posibles inconsistencias con la hipótesis adoptada, conforme se expone en la Figura 8.

Entre los eventos hallados se encuentran discontinuidades de la forma de onda, a modo de ejemplo se expone la Figura 9 entre los instantes 00:02:56,147 a 00:02:56,155.

Figura 9. Forma de onda (escala lineal), pista de etiquetas y primera diferencia del audio "que tiene lule menem con Karina".

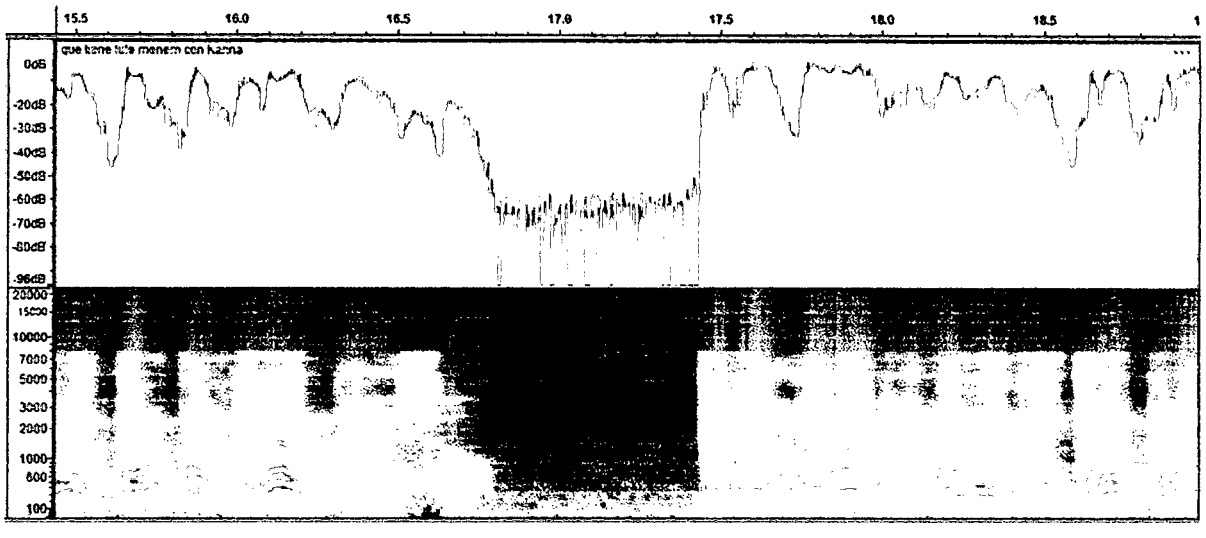



GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
32
DIR CRIMINALÍSTICA Y EST FORENSES

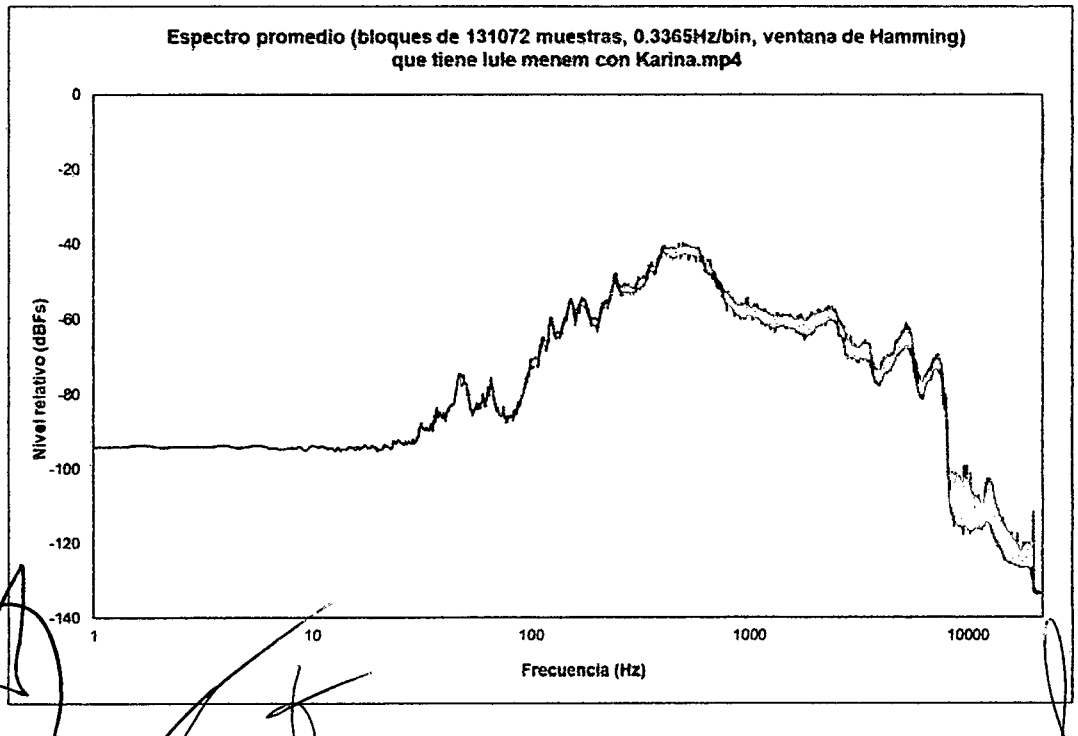
Asimismo, la escucha crítica y la inspección del espectrograma permite establecer intervalos de audio donde el ruido de fondo presenta cambios sustanciales, de forma instantánea, de la energía distribuida en el espectro. Esta condición es inconsistente con el supuesto de que el ruido de fondo dentro en un entorno acústico varía de forma gradual. A modo de ejemplo la siguiente figura expone un instante de cambio sustancial en el ruido de fondo, representada a los 00:00:17,111 segundos.

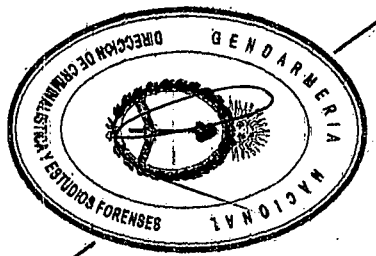
Figura 10. Espectrograma y pista de etiquetas del audio "que tiene lule menem con Karina" en el instante 00:00:17,111 segundos.



En este sentido y continuando con la inspección del espectrograma de la señal completa, se observa que los segmentos con contenido de habla se encuentran limitados hasta los 8.000 Hz. Para ilustrarlo las siguientes figuras contienen el espectro promedio de la señal y el espectrograma en escala lineal de la señal.

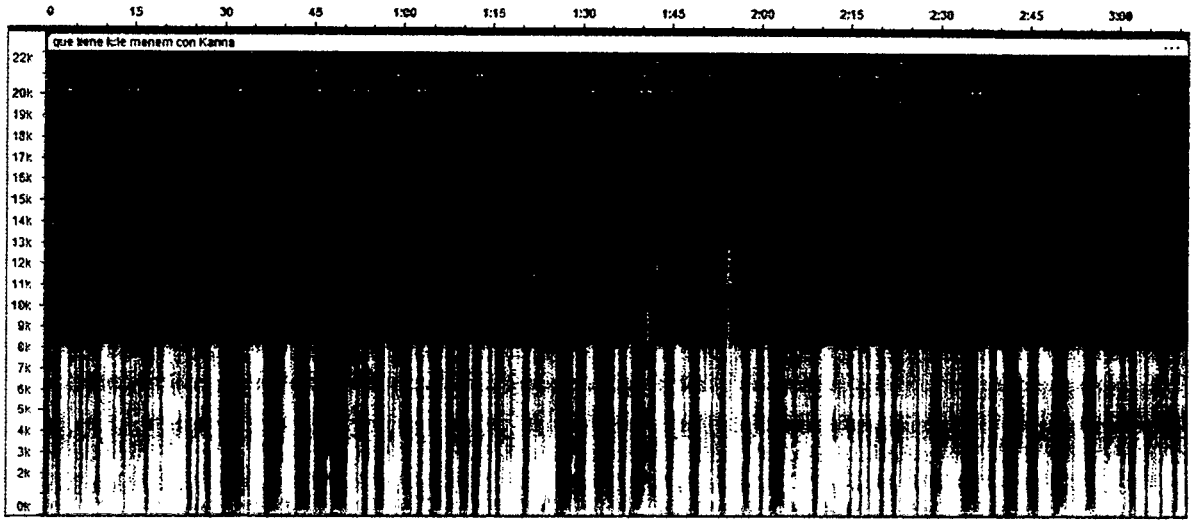
Figura 11. Espectro promedio del audio del archivo "que tiene lule menem con Karina".




GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

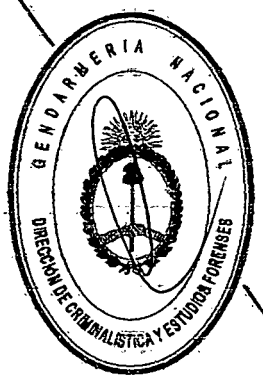
GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
33

Figura 12. Espectrograma de la señal en escala lineal, rango de frecuencias en la visualización hasta 22.050 Hz.



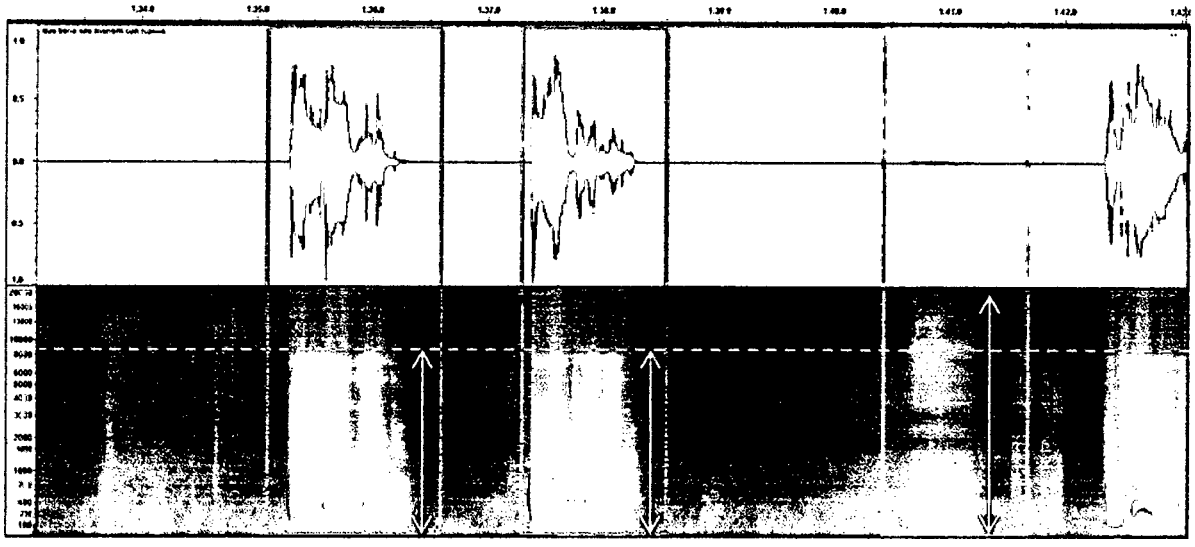
En contraste con la señal de habla, se identificaron múltiples sonidos cuyo contenido espectral se extiende por encima de los 8.000 Hz. Esta circunstancia permite establecer que el registro contiene información espectral por encima de dicha frecuencia y que, por lo tanto, la atenuación observada en la señal de habla no resulta atribuible, por sí sola, a una limitación general de rango frecuencial del archivo. Asimismo, la frecuencia de muestreo de 44.100 Hz declarada en sus metadatos permite, en términos teóricos, representar componentes espectrales de hasta 22.050 Hz, valor correspondiente a la frecuencia de Nyquist[16]. Mediante la escucha perceptual se establece que las características acústicas de mencionados eventos resultan compatibles con respiraciones cerca del micrófono y un plano diferente al hablante, con componentes espectrales observables mayores a los 8.000 Hz previamente señalados. Para ejemplificar lo mencionado, se expone la siguiente entre los instantes 00:01:33,000 a 00:01:43,000.

[16] La frecuencia de Nyquist corresponde a la mitad de la frecuencia de muestreo y representa la frecuencia máxima que, en términos teóricos, puede ser representada sin ambigüedad en una señal digital.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
34
DIR. CRIMINALÍSTICA Y EST. FORENSES

Figura 13. Espectrograma de la señal en escala de Mel, con rango de frecuencia de la visualización hasta 22.050 Hz, y forma de onda en escala lineal.

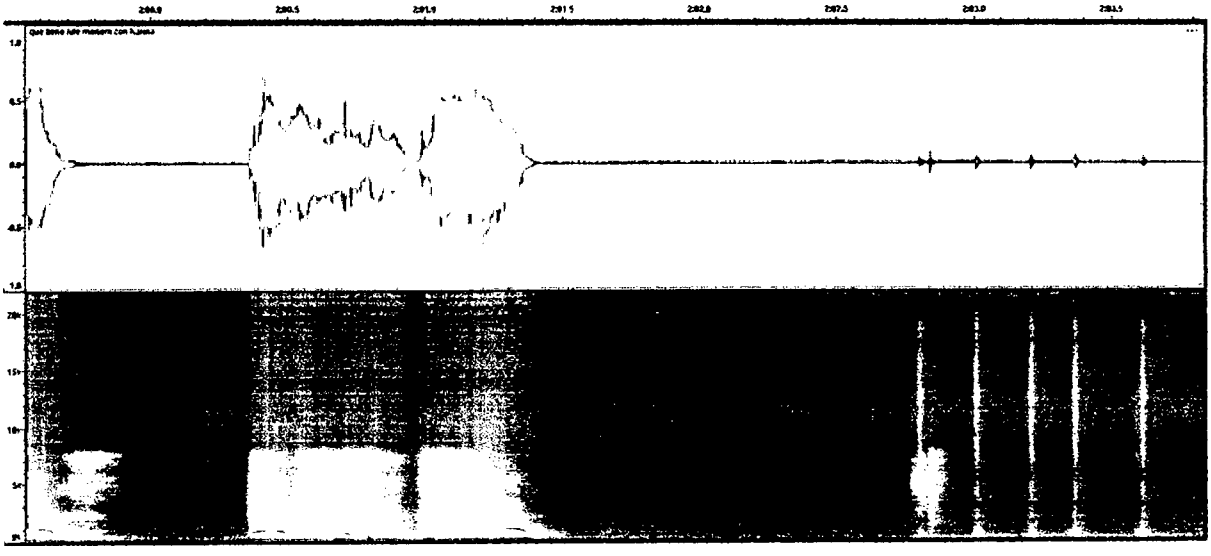


Señales de habla (azul) con contenido espectral predominante por debajo de 8.000 Hz

Sonido (magenta) con contenido espectral que se extiende por encima de 8.000 Hz

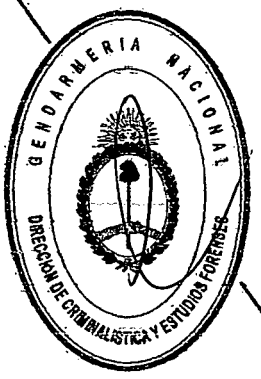
La pista de etiquetado para el audio "que tiene lule menem con Karina" indica, en el mismo sentido, reiterados sonidos cuyo ancho de banda resulta superior a la limitación en 8.000 Hz observada sobre la señal de habla. Estos sonidos son de naturaleza impulsiva y perceptualmente se encuentran en el mismo plano de adquisición de las respiraciones, que también exhiben un ancho de banda superior a la voz. La siguiente figura expone ejemplos de sonidos con esta naturaleza.

Figura 14. Espectrograma de la señal en escala lineal, rango de frecuencias en la visualización hasta 22.050 Hz.



3.4.3.4. Observaciones adicionales

En el archivo denominado "que tiene lule menem con Karina" se identificaron dos intervalos que presentan repeticiones de contenido sonoro, acompañadas de una degradación de la calidad acústica y una limitación del ancho de banda. Estas


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y ESTUDIOS FORENSES
Foja
35

características resultan compatibles con artefactos asociados a la comunicación. Las siguientes figuras ilustran ambos eventos.

Figura 15. Repetición degradada observada en el audio a los 00:01:11,000 segundos. Espectrograma de la señal (arriba), pista de etiquetas (medio) y espectros desplazados para mejor comparación de segmentos (abajo).

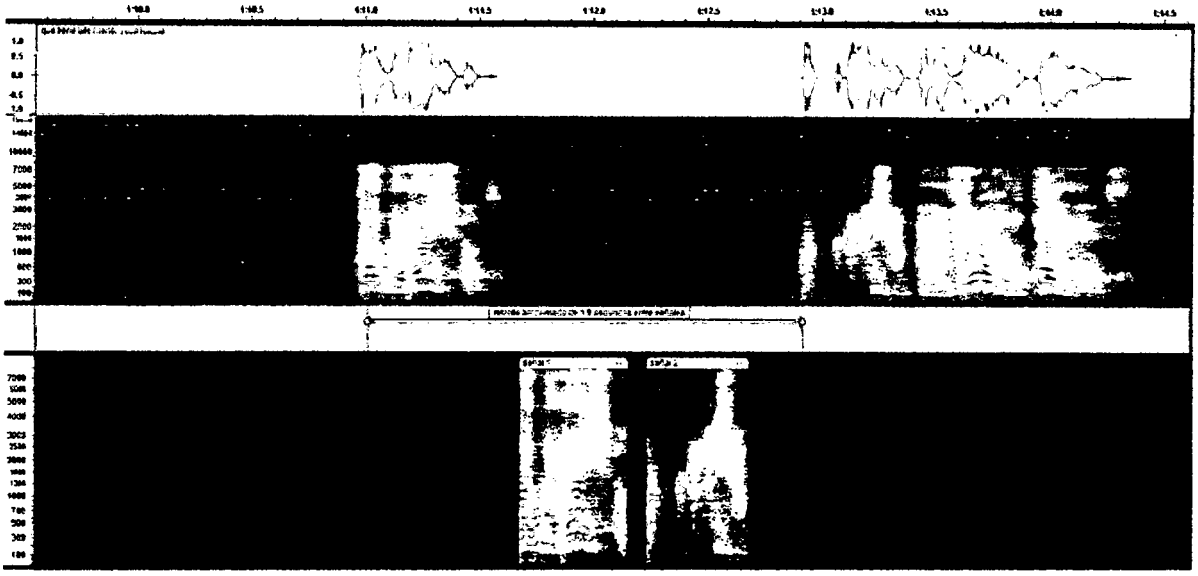
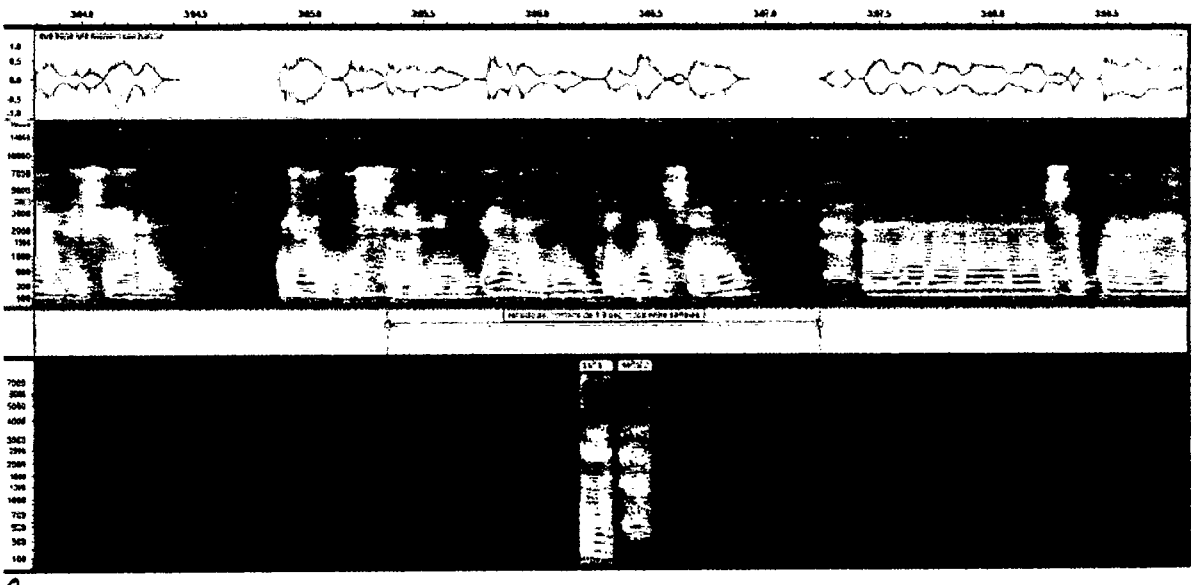
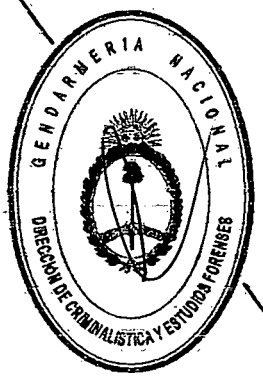


Figura 16. Repetición degradada observada en el audio a los 00:03:05,000 segundos. Espectrograma de la señal (arriba), pista de etiquetas (medio) y espectros desplazados para mejor comparación de segmentos (abajo).



Del análisis sobre los restantes audios, se identificó un conjunto de ellos con presencia de ruido de fondo de características compatibles con babble noise[17], música de fondo, observándose su superposición con la señal de habla principal en determinados segmentos de los registros sonoros examinados. Asimismo, se observaron eventos acústicos transitorios asociados a sonidos incidentales, lo que sería compatible a impactos de objetos y/o vajilla, que presentan solapamiento temporal

[17] Ruido compuesto por la superposición simultánea de múltiples voces.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

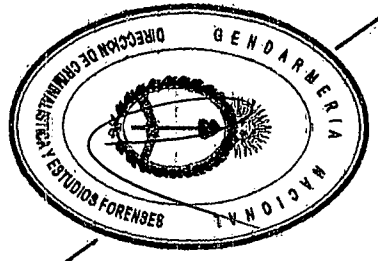
con segmentos de habla. Los audios en los cuales se detectaron las mencionadas características se encuentran consignados en la siguiente tabla:

Tabla 7. Audios con presencia de ruido interferente constante.

Nombre del archivo
Chupa pija
Denuncia
Ejemplo torta
el 3 es para Karina
es un error tremendo el que cometieron con villarruel
ese es un delincuente
Fake News adorni
Javi que hago
javier me dice yo voy a hablar con Lule
la gente de la suizo
le afanaron la caja también a Mondino
me metio el mas importante de todos
no lo quiere nadie, pero nadie
nosotros lo subimos a presidencia
ponela a Karina de candidata y que vaya a hablar
sueldos
tenes chorros en lule tucumán
yo hablé con el presidente
Zafarrancho
800 lucas por mes
a la primera que se van a llevar puesta
a mi me pusieron un tipo que maneja toda la caja mia
daniel garbellini
diego sos bueno
el no está metido pero
ella se lleva medio palo
es un cachito asi
esto lo hacen de ratas
La caja
Lule vino de una manera a mi me estan desfalcando a la gente
lule y karina
se lo dijo sandra, se lo dije yo
voracidad genética
vos viste la entrevista que me hizo fantino
estos tipos aprovechan y quieren armar su kioskito

3.4.3.5. Muestras nulas iniciales asociadas al comportamiento de priming de la codificación AAC.

Del examen de los metadatos técnicos se observó, en algunos de los archivos de audio, la presencia del valor "com.apple.iTunes". Asimismo, se constató


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 37
DIR CRIMINALÍSTICA Y EST FORENSES

que la totalidad de los registros presentaba codificación AAC-LC, utilizada, entre otros, por dispositivos y aplicaciones de Apple. Estas observaciones motivaron la consulta de documentación técnica y bibliografía científica relacionada con el comportamiento inicial de la codificación AAC.

En virtud de ello, se examinó el comportamiento inicial de los archivos codificados mediante AAC-LC, contenidos en los formatos MP4 y M4A, cuyo detalle se presenta en la Tabla Nro. A5. (ver Anexo I). Para ello, mediante FFmpeg y utilizando el comando previamente expuesto en el Bloque de código 6, se desempaquetó y copió (sin recodificación) el *stream*[18] de audio AAC contenido en cada archivo, generándose un nuevo archivo ".aac" en formato ADTS, desprovisto de la estructura y de la información temporal propia del contenedor MP4/M4A. El estudio se orientó particularmente a observar las muestras iniciales asociadas al *priming*[19] del codificador, es decir, al retardo introducido por el proceso de codificación AAC antes del comienzo efectivo de la señal de audio. A los fines comparativos, las señales obtenidas fueron alineadas y se determinó, para cada registro, la posición de la primera muestra con valor distinto de cero. El detalle del valor establecido para cada archivo de audio se encuentra dentro la tabla de etiquetas para cada archivo disponible en la Tabla A5 del Anexo.

Los resultados evidenciaron un comportamiento variable entre los archivos analizados. En algunos casos, la primera muestra distinta de cero se encontró desde el inicio de la señal decodificada, mientras que en otros se observó un conjunto inicial de muestras nulas cuya extensión resultó diferente entre registros. La distribución observada no resultó uniforme entre los archivos analizados. Si bien la cantidad de muestras de amplitud nula asociadas al priming de la codificación AAC puede presentar variaciones, Apple ha documentado históricamente valores próximos a 2.112 muestras[20], mientras que estudios experimentales recientes sobre grabaciones AAC generadas en dispositivos iOS informaron valores centrales igualmente próximos a dicho orden de magnitud[21]. En este contexto, la reducción marcada o ausencia de este conjunto de muestras iniciales en algunos de los registros constituye una característica de interés. También se ha señalado que, en grabaciones AAC generadas mediante Voice Memos, la ausencia de la latencia inicial puede constituir un indicio de supresión de la porción inicial de la grabación[22].

Los comportamientos observados, pueden resultar compatibles con una eventual supresión de muestras o segmentos iniciales efectuada con posterioridad a

---

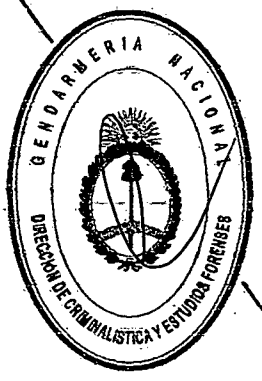
[18] Flujo de datos que contiene la información de audio codificada dentro de un archivo o contenedor multimedia

[19] Retardo inicial propio de la codificación AAC, asociado al procesamiento necesario al comienzo de la señal.

[20] Apple Inc. (2012). QuickTime file format specification. https://developer.apple.com/documentation/quicktime-file-format

[21] Wales, G. S. (2025). *Quantitative study of zero-amplitude sample padding in iOS AAC encoding: Consistency across devices and the impact of background noise. Journal of Forensic Sciences, 70*(6), 2309–2323. https://doi.org/10.1111/1556-4029.70157

[22] Park, N. I., Lee, J. W., Shim, K.-S., Byun, J. S., & Jeon, O.-Y. (2021). *A method of forensic authentication of audio recordings generated using the Voice Memos application in the iPhone. Forensic Science International, 320*, 110702. https://doi.org/10.1016/j.forsciint.2021.110702


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
38
DIRECCIÓN CRIMINALÍSTICA Y ESTUDIOS FORENSES

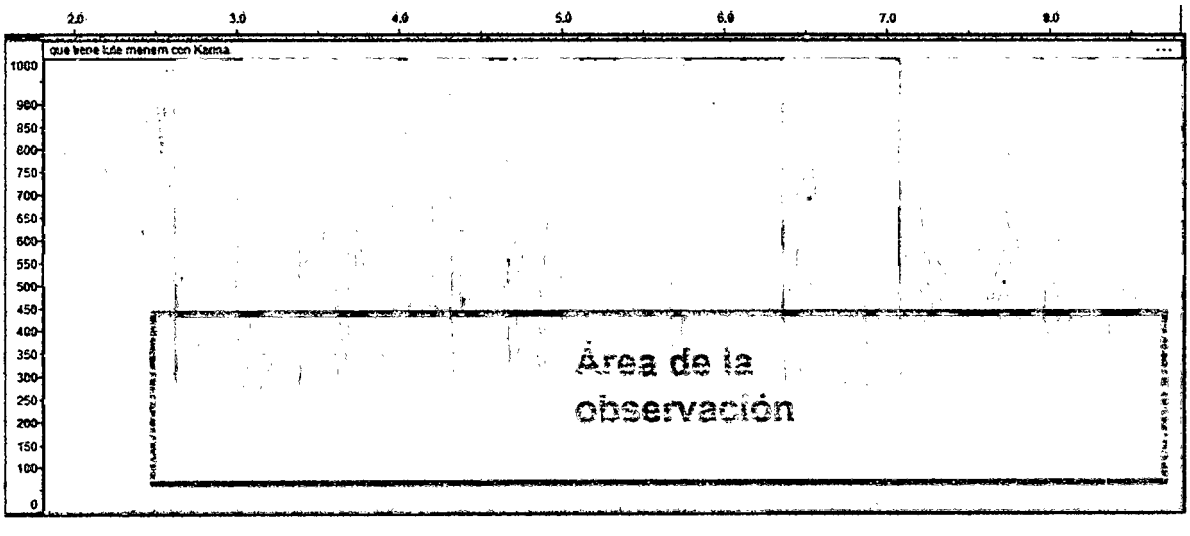
la codificación original. No obstante, la interpretación de esta observación se encuentra limitada por la ausencia de información contextual independiente respecto del origen, dispositivo de adquisición, condiciones de generación y el eventual procesamiento previo de los archivos analizados. En consecuencia, no resulta posible determinar inequívocamente el origen de las características observadas ni atribuirlas a una operación específica de corte o supresión.

3.4.3.6. Estudio de atenuación en el rango de baja frecuencia y transferencia acústica.

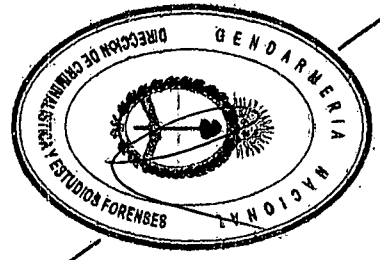
El conjunto de observaciones relativas a la presencia de sonidos con diferentes anchos de banda dentro del mismo audio, la limitación superior (entorno a 8.000 Hz) del espectro de la señal de habla y la percepción de distintos planos de adquisición de la señal, se complementó con la identificación de un nivel relativo menor de la componente espectral correspondiente a la frecuencia fundamental (F0) de la voz, en relación a los armónicos superiores. Estas observaciones se corresponden al grupo de audios que no presentaron ruido de fondo interferente (es decir, aquellos no incluido de la Tabla Nro. 7), entre los cuales se encuentran el archivo "que tiene lule menem con Karina.mp4".

A modo de ilustración la siguiente figura expone el espectrograma de la totalidad del audio limitado hasta los 1.000 Hz.

Figura 17. Espectrograma de la señal en escala lineal con la región donde se observa la atenuación relativa.



A fin de profundizar en esta observación, se optó por estudiar cuantitativamente la distribución relativa de energía entre componentes armónicos de la señal, particularmente mediante la medida H1*-A1*. Esta expresa, en decibeles, la diferencia entre la amplitud del primer armónico (H1) y la amplitud del armónico situado próximo al primer formante (A1), luego de aplicar sobre ambas magnitudes una corrección destinada a compensar la influencia de las resonancias del tracto vocal. Se


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
39
DIR. CRIMINALÍSTICA Y EST. FORENSES

trata de una medida acústica empleada en el estudio de la calidad vocal y de las características espectrales asociadas a la fuente de voz.[23]

Para el desarrollo de la tarea, y considerando el volumen de material a analizar, se implementó el procesamiento mediante scripts ejecutados dentro del software Praat, versión 7.0.02, destinado al análisis acústico del habla. La Tabla 8 detalla los scripts utilizados, su función y su origen.

Tabla 8. Scripts utilizados para la medición y análisis de H1*–A1*.

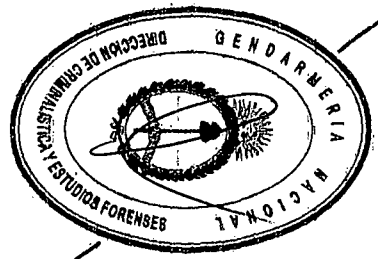
Nombre del script	Función	Origen
praat-script-syllable-nuclei-v2file.praat	Estimar automáticamente la cantidad de núcleos silábicos de una grabación, sin necesidad de transcripción fonética	Repositorio público de github[24], método basado la publicación[25] de Jong y Wempe.
praat-script-extract-syllables.praat	Extraer como objetos Sound los segmentos correspondientes a cada núcleo silábico previamente identificado en el TextGrid, con ventanas de 70 ms centradas en cada núcleo y conservando sus coordenadas temporales respecto del audio fuente.	Desarrollo propio.
praat-script-stats-harmonics.praat	Analizar individualmente un núcleo silábico mediante la estimación de F0, formantes, armónicos H1 y H2 y amplitudes A1, A2 y A3; calcular sus diferencias de amplitud observadas y corregidas (H1*–H2*, H1*–A1*, H1*–A2* y H1*–A3*) y generar una representación gráfica del espectro de la señal y su envolvente LPC.	Desarrollo propio, implementando corrección por formantes.[26]
praat-script-stats-harmonics-multiple.praat	Aplicar por lote a todos los núcleos silábicos el mismo procedimiento de análisis implementado en praat-script-stats-harmonics.praat, consolidando los resultados obtenidos y generando un diagrama de dispersión de H1*–A1* en función de la frecuencia de H1.	Desarrollo propio.

[23] Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). *VoiceSauce: A program for voice analysis*. En Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), 1846–1849.

[24] FieldDB. *Praat Scripts – praat-script-syllable-nuclei-v2file.praat*. GitHub. Consultado el 23 de octubre de 2026. https://github.com/FieldDB/Praat-Scripts/blob/main/praat-script-syllable-nuclei-v2file.praat

[25] Jong, N. H., & Wempe, T. (2009). *Praat script to detect syllable nuclei and measure speech rate automatically*. Behavior Research Methods, 41(2), 385–390. https://doi.org/10.3758/BRM.41.2.385

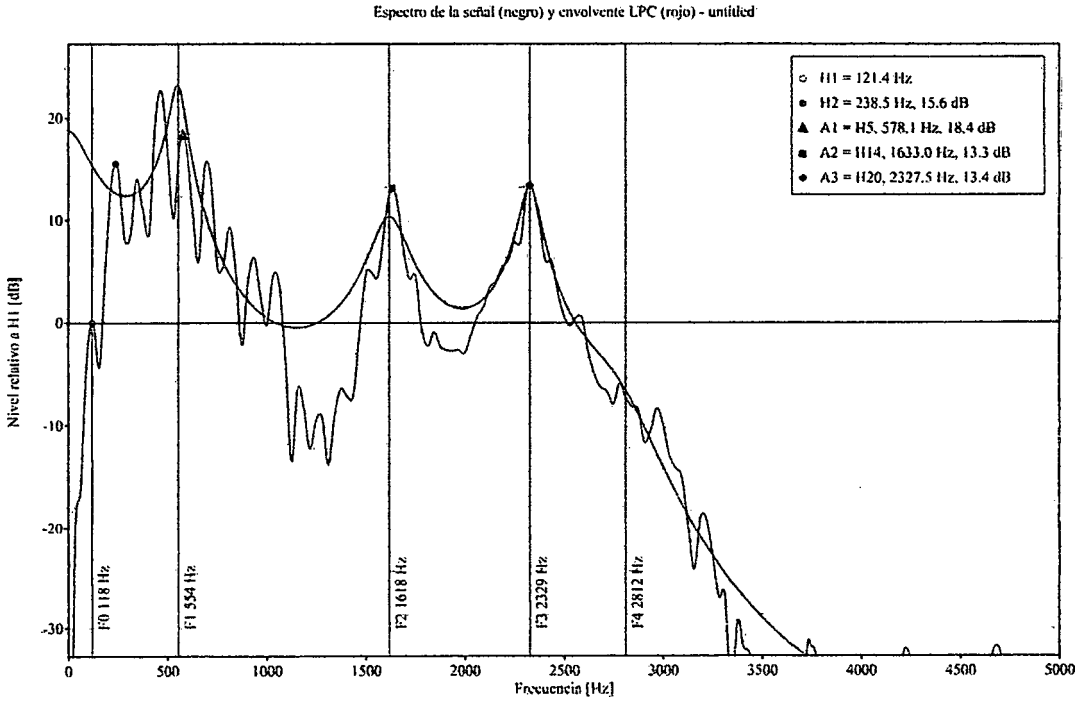
[26] Iseli, M., Shue, Y.-L., & Alwan, A. (2007). *Age, sex, and vowel dependencies of acoustic measures related to the voice source*. *The Journal of the Acoustical Society of America, 121*(4), 2283–2295. https://doi.org/10.1121/1.2697522


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

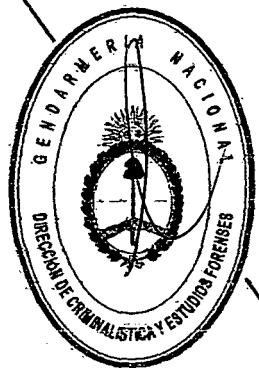
GENDARMERÍA NACIONAL
DIR CRIMINALÍSTICA Y EST FORENSES
Foja
40

Durante la reunión de trabajo de fecha 23 de septiembre del 2026, se expusieron los resultados obtenidos para el audio del archivo denominado "sturzzeneger" mediante el análisis de H1*–A1*. Inicialmente, luego de presentar la metodología de medición adoptada, se ilustró gráficamente el procedimiento sobre un único núcleo silábico que correspondiente al primero detectado dentro de este audio. Para ello se utilizó la salida gráfica del script "praat-script-stats-harmonics.praat", en la cual se representaron el espectro de la señal, la envolvente espectral obtenida mediante LPC y las posiciones estimadas de H1, H2, A1, A2 y A3, conjuntamente con las frecuencias de los primeros formantes utilizadas en el análisis.

Figura 18. Espectro de la señal, envolvente LPC y ubicación de H1, H2, A1, A2 y A3 correspondiente al primer núcleo silábico del archivo "sturzzeneger.mp4".



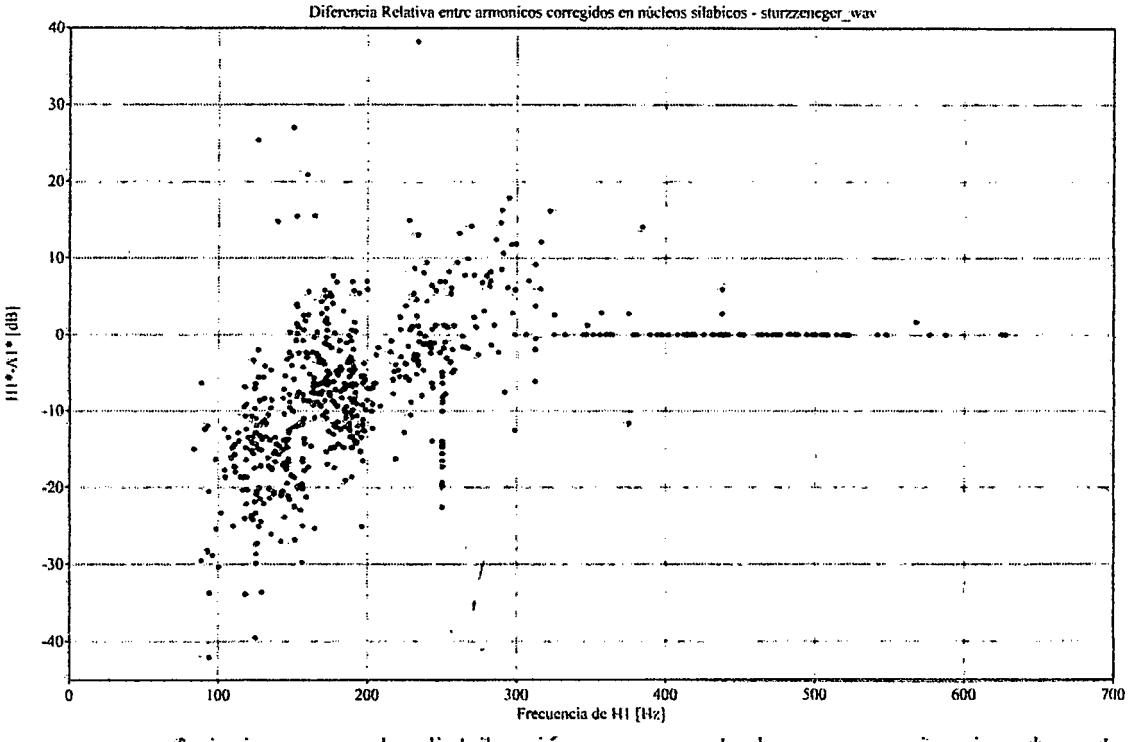
Posteriormente, se efectuó el análisis de H1*–A1* sobre la totalidad de los núcleos silábicos extraídos del archivo mediante el script "praat-script-stats-


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
41
DIR CRIMINALÍSTICA Y EST FORENSES

harmonics-multiple.praat". Los resultados individuales se representaron mediante un diagrama de dispersión, en el que cada punto corresponde a un núcleo silábico analizado.

Figura 19. Diagrama de dispersión de H1*–A1* en función de la frecuencia de H1 para los núcleos silábicos analizados del archivo "sturzzeneger.mp4".

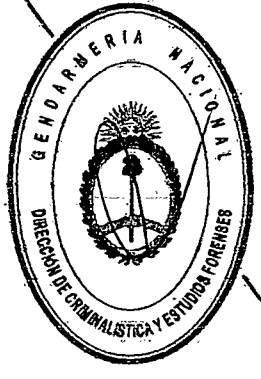


Asimismo, en la distribución representada se aprecia visualmente una tendencia de correlación entre ambas variables, observándose que, a medida que disminuye la frecuencia de H1 (correspondiente a la frecuencia fundamental (F0) del núcleo silábico), también tienden a disminuir los valores de la diferencia relativa H1*–A1*. La distribución de los puntos muestra, por lo tanto, un grado de correlación visual no despreciable entre ambas magnitudes. Este comportamiento resulta compatible con las limitaciones en frecuencia propias de un sistema de reproducción electroacústica mediante altavoz.

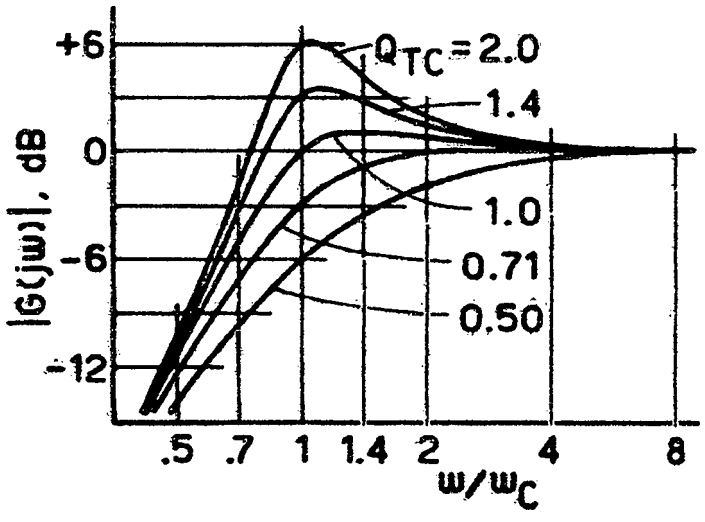
En particular, para sistemas de altavoz en caja cerrada, Small[27] describe la respuesta de baja frecuencia como la correspondiente a un filtro pasa-altos de segundo orden, con una pendiente asintótica de 12 dB por octava por debajo de la región de transición. La siguiente figura reproduce las curvas teóricas normalizadas presentadas por el autor para distintos valores del factor de calidad total del sistema Q_{TC}, donde se aprecia la disminución progresiva de la respuesta hacia las bajas frecuencias.

Figura 20. Respuesta de amplitud normalizada en función de la frecuencia normalizada para un sistema de altavoz en caja cerrada, considerando diferentes valores del factor de calidad total Q_{TC}. Fuente: Small (1972, Fig. 4).

[27] Small, R. H. (1972). *Closed-box loudspeaker systems-Part 1: Analysis*. Journal of the Audio Engineering Society, 20(10), 798–808.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR CRIMINALÍSTICA Y EST FORENSES
Foja 42



Donde ω es la frecuencia angular, ω_c es la frecuencia angular de resonancia (f_c) del sistema altavoz-caja cerrada, $|G(j\omega)|$ representa la magnitud de la respuesta en frecuencia del sistema, expresada en decibelios (dB) y Q_{TC} es el factor de calidad total del sistema a la frecuencia de resonancia, considerando todas las pérdidas del sistema.

Se aplicó el análisis de cuantitativo de la diferencia relativa H1*−A1* para el resto de audios con características semejantes, es decir, aquellos que denotaban atenuación en el rango de baja frecuencia. La totalidad de los resultados obtenidos se encuentran en el respectivo Anexo I.

3.4.3.7. Identificación de segmentos sonoros coincidentes entre archivos de audio

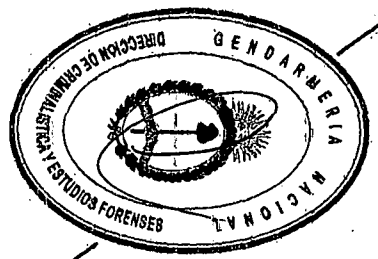
Durante el examen de la totalidad del material sonoro aportado, se identificaron segmentos que presentan coincidencias de contenido en distintos archivos de audio. A continuación, se detallan las correspondencias observadas y sus respectivas ubicaciones temporales aproximadas.

En el archivo "el daño que puede hacer victoria.mp4" se observó un segmento sonoro coincidente con el intervalo comprendido entre 00:00:02,5 y 00:00:16,8 del archivo "que tiene lule menem con Karina.mp4".

Respecto del archivo "no lo quiere nadie, pero nadie.mp4", se observó que el segmento final, comprendido aproximadamente entre 00:00:11,1 y 00:00:13,4, coincide con el segmento inicial del archivo "Lule vino de una manera a mi me están desfalcando a la gente.mp4", comprendido entre 00:00:00 y 00:00:02,2.

Por otra parte, se observó que la totalidad del contenido sonoro del archivo "victoria claramente se cansó.mp4" coincide con el segmento comprendido aproximadamente entre 00:00:00,6 y 00:00:04,9 del archivo "victoria se canso.mp4". Asimismo, ambos registros presentan coincidencia en su contenido verbal.

3.4.3.8. Resultados del análisis de integridad


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
43

A partir del análisis de integridad efectuado sobre la totalidad de los archivos de audio aportados, se identificaron dos grupos diferenciados según las características acústicas observadas.

El primero comprende aquellos registros que presentan ruido de fondo interferente a lo largo de su extensión, cuyo efecto de enmascaramiento dificulta la identificación de posibles discontinuidades y otros eventos acústicos de interés. El segundo comprende archivos en los que se identificaron características compatibles con un posible proceso de transferencia acústica, entre ellas, eventos respiratorios y ruidos impulsivos con contenido espectral superior al ancho de banda observado en la señal de habla, así como discontinuidades abruptas en la forma de onda.

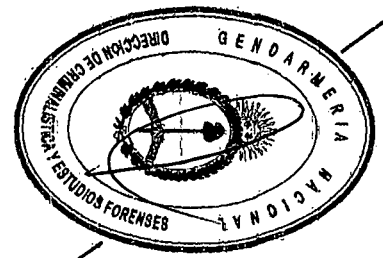
Asimismo, se identificaron segmentos sonoros coincidentes entre distintos archivos examinados, conforme a lo desarrollado en el apartado 3.4.3.7. Estas correspondencias evidencian la presencia de contenido sonoro repetido entre diferentes registros, sin que sea posible establecer la secuencia de generación de los archivos ni los eventuales procedimientos de edición empleados.

La metodología de análisis previamente descripta se aplicó a la totalidad de los registros, individualizando los fenómenos acústicos detectados y consignando su ubicación temporal y características relevantes. Las observaciones correspondientes a cada archivo se detallan en la Tabla A5 del Anexo I.

3.4.4. CONSIDERACIÓN CONJUNTA DE LOS RESULTADOS Y LIMITACIÓN PARA EL PROCESAMIENTO MEDIANTE FORENSIA V2.0

La evaluación de factibilidad y el análisis de integridad constituyen instancias de examen diferenciadas, orientadas a finalidades específicas. Mientras que la evaluación de factibilidad tiene por objeto determinar si el material reúne condiciones acústicas que permitan su utilización en un eventual análisis comparativo de voces con fines de identificación, el análisis de integridad se orienta a la identificación y caracterización de posibles cortes, discontinuidades, ediciones, manipulaciones u otros fenómenos presentes en los registros sonoros. En consecuencia, la identificación de características compatibles con modificaciones o alteraciones en un registro no determina, por sí misma, la imposibilidad de efectuar un análisis comparativo de voces, del mismo modo que la determinación de factibilidad para comparación no implica, por sí misma, la ausencia de modificaciones o discontinuidades en el material.

Del análisis realizado se identificaron, en distintos registros, eventos y características acústicas que fueron individualizados y consignados temporalmente. En particular, los hallazgos expuestos en el apartado "3.4.3.6. Estudio de atenuación en el rango de baja frecuencia y transferencia acústica", considerados en conjunto, constituyen indicios acústicos que proporcionan fundamentos razonables para plan-


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
44

tear la hipótesis de que la señal de habla registrada podría ser producto de un proceso de regrabación mediante transferencia acústica; esto es, de la reproducción de una grabación previa a través de un sistema electro-acústico y su posterior captación mediante otro dispositivo de grabación. Si bien las observaciones efectuadas no permiten establecer de manera concluyente dicho mecanismo de producción, justifican considerar esta posible condición de adquisición al momento de evaluar la procedencia de una comparación forense de voces mediante el software Forensia v2.0.

Los sistemas automáticos de comparación forense de voces (FASR) requieren que su desempeño y la calibración de sus resultados hayan sido evaluados bajo condiciones representativas de aquellas presentes en las grabaciones objeto de comparación, particularmente respecto de las diferencias entre las condiciones y canales de adquisición. Sobre este aspecto, las directrices metodológicas de ENFSI para el reconocimiento automático y semiautomático forense de hablantes[28] establecen que la validación debe efectuarse sobre muestras representativas del material encontrado en la práctica forense y, en lo posible, bajo condiciones de grabación similares o estrechamente comparables a las del caso analizado. De acuerdo con la documentación técnica[29,30] correspondiente a Forensia, el procedimiento comparativo contempla una etapa de calibración asociada a las condiciones de canal, conforme se representa en su diagrama de funcionamiento y descripción de etapas.

Figura 21. Diagrama de funcionamiento de Forensia v2.0. Fuente: Martínez Soler et al. (2020, Figura 2).

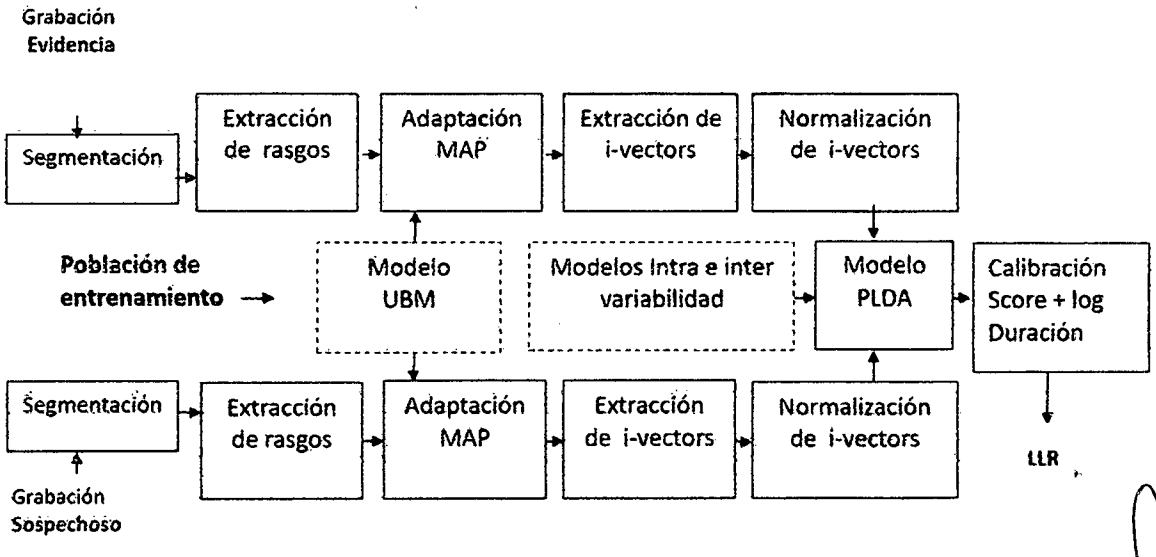
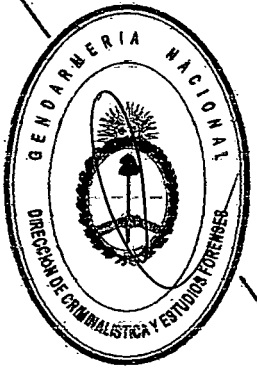


Tabla 9. Características técnicas de FORENSIA. Fuente: Univaso et al. (2020, Tabla 1).

[28] Drygajlo, A., Jessen, M., Gfroerer, S., Wagner, I., Vermeulen, J., & Niemi, T. (2015). *Methodological guidelines for best practice in forensic semiautomatic and automatic speaker recognition: Including guidance on the conduct of proficiency testing and collaborative exercises*. European Network of Forensic Science Institutes (ENFSI), Forensic Speech and Audio Analysis Working Group.

[29] Martínez Soler, M., Univaso, P., & Gurlekian, J. (2020). *FORENSIA 2.0: Sistema de identificación forense de hablantes* [Informe técnico]. BlackVOX-CONICET.

[30] Univaso, P., Gurlekian, J., Martínez Soler, M., & Stalker, G. (2020). FORENSIA: un sistema de identificación forense por voz. En *Anales del XX Simposio Argentino de Informática y Derecho (SID 2020), 49 JAIIO* (pp. 116–130). Sociedad Argentina de Informática e Investigación Operativa.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
45
DIR. CRIMINALÍSTICA Y EST. FORENSES

Etapa	Descripción
Pre-procesamiento	Tramado TPDF (dithering) + filtros Wiener y RASTA-LDA
Parámetros acústicos	39 parámetros = 12 MFCC + Energía + Δ + Δ2
Detección de actividad de habla (VAD)	Basado en energía + algoritmo heurístico de restricción de duraciones
Post-procesamiento de parámetros	Normalización cepstral media (CMN)
Modelo de referencia universal (UBM)	Modelo de mezclas gaussianas (GMM): 1024 mezclas independientes del género
Modelado de i-vectors	400 dim
Post-procesamiento de i-vectors	Normalización esférica de atributos perjudiciales (sph-Norm) + análisis de discriminación lineal (LDA)
Compensación y puntuación	Análisis discriminativo lineal probabilístico (PLDA): 300 dim p/hablante y 100 dim p/canal
Calibración	Regresión logística: duración y Cllr
Post-calibración	Hombres/mujeres: micrófono dinámico, micrófono electroestático, celular, teléfono, Whatsapp y Youtube.

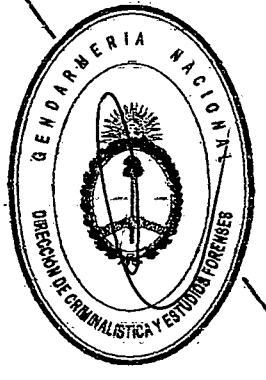
La eventual existencia de una transferencia acústica incorpora una etapa adicional de reproducción y captación de la señal, susceptible de modificar sus características acústicas y, por lo tanto, afectar los resultados de la comparación. Ante la presencia de indicios compatibles con esta condición y no encontrándose contemplada en la validación disponible del sistema una condición de canal equivalente (situación que puede encuadrarse dentro de las denominadas *mismatched recording conditions* descriptas por ENFSI[31]), no resulta metodológicamente adecuado proceder con la comparación automática. Esta circunstancia es producto de que no puede establecerse que los resultados obtenidos conserven las condiciones de fiabilidad requeridas para su interpretación en un contexto forense.

Las condiciones de canal efectivamente contempladas, conforme la documentación, por el sistema disponible se exponen en la Tabla 10.

Tabla 10. Condiciones de canal contempladas para la calibración en Forensia v2.0. Fuente: adaptado de Univaso et al. (2020, Fig. 1).

Canales de calibración disponibles
Celular - Celular
Teléfono - Teléfono
Micrófono - Micrófono
Hi micrófono - Hi micrófono
Celular - Micrófono
Teléfono - Hi micrófono
WhatsApp - WhatsApp
YouTube - YouTube
WhatsApp - YouTube
WhatsApp - Micrófono

[31] Drygajlo et al. (2015), p. 15, apartado 3.7, "Mismatched recording conditions".


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
46
DIR. CRIMINALÍSTICA Y EST. FORENSES

Cabe señalar que esta limitación deriva de las características particulares detectadas durante el análisis de integridad, compatibles con una grabación realizada mediante transferencia acústica, y no de una modificación de la valoración de factibilidad acústica (conforme el Protocolo) previamente realizada.

Por lo expuesto, respecto de los registros que presentan indicios compatibles con esta modalidad de adquisición, no resulta técnicamente procedente efectuar la comparación forense de voces mediante Forensia v2.0 bajo las condiciones disponibles.

Para el caso específico del audio denominado "que tiene lule menem con Karina", esta conclusión fue comunicada durante la reunión de trabajo celebrada el 16 de septiembre de 2026 y documentada en la correspondiente acta Constancia de Trabajo Nro. 04.

En virtud de ello, se dejó sin efecto la toma de muestra indubitada de voz del Sr. Diego Orlando Spagnuolo, destinada a la realización de la comparación forense de voces solicitada.

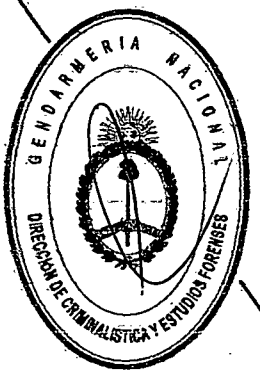
3.4.5. ESTUDIOS RELATIVOS AL ANÁLISIS DE AUDIO SINTÉTICO

Se efectuó un examen desde la perspectiva fonoaudiológica, a través de procedimientos auditivo-perceptuales y de exploración acústica de la señal vocal. Para ello, se tuvo en consideración, como información complementaria, las observaciones previamente documentadas respecto de la continuidad de los registros, a fin de integrar dichos antecedentes con los hallazgos obtenidos durante el presente análisis.

Se deja constancia que, se solicitó al magistrado interviniente la individualización de los segmentos específicamente atribuidos a Diego Orlando Spagnuolo, conforme lo expresado en el acta Constancia de Trabajo Nro. 03 de fecha 09 de septiembre del 2026, sin que se aportara información al respecto. No obstante, durante el examen de los archivos objeto de estudio, se constató perceptualmente la presencia de una única voz masculina discernible, sin advertirse la intervención de otro hablante. En consecuencia, ante la ausencia de una individualización temporal de los segmentos referidos, el análisis perceptual se efectuó sobre la única voz presente en los registros.

Los procedimientos aplicados se orientaron a identificar y describir propiedades observables de las señales vocales examinadas, sin constituir un procedimiento específico de detección de habla sintética. Al respecto, las investigaciones sobre percepción humana de voces sintéticas han documentado limitaciones en la capacidad de los oyentes para discriminar entre emisiones naturales y artificiales, aún cuando sus valoraciones consideran diferentes características acústico-fonéticas y fenómenos no lingüísticos.[32] El alcance de los procedimientos realizados y sus limitaciones metodológicas se desarrollan en el apartado 3.4.5.6.

[32] San Segundo, E., López-Jareño, A., Wang, X., & Yamagishi, J. (2026). *Human perception of audio deepfakes: The role of language and speaking style*. Speech Communication, 185, Article 103487


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 47
DIR CRIMINALÍSTICA Y EST FORENSES

De la totalidad de los archivos de audio que integran el material de estudio, se seleccionaron únicamente CINCO (5), debido a que la duración de habla continua en dichos registros supera los TREINTA (30) segundos, estableciéndose dicho criterio temporal como parámetro operativo para delimitar el material objeto de análisis. Esta decisión se fundamentó en la conveniencia de disponer de una cantidad de material de voz que favoreciera la observación de diferentes fenómenos acústico-fonéticos y de sus variaciones a lo largo de la emisión vocal.

Vale destacar que, de los CINCO (5) archivos seleccionados, TRES (3) de ellos ya se encontraban segmentados producto de los análisis realizados anteriormente, detallado en el apartado 3.4.2.1, verificándose la duración de habla.

Asimismo, respecto a los DOS (2) archivos de audio restantes, se analizó la duración del habla mediante escucha y análisis auditivo-perceptual.

En el ámbito de la fonética forense, se ha señalado la importancia de la extensión de las muestras para el análisis fonético y se ha propuesto, cuando resulta necesario establecer un umbral mínimo orientativo, una duración de treinta (30) segundos de habla o cien (100) palabras.[33] La selección respondió, por lo tanto, a un criterio operativo de suficiencia del material disponible, considerando que una mayor extensión del registro puede favorecer la identificación de características acústicas, fonéticas y temporales, así como el examen de sus variaciones durante la emisión vocal.

Cabe precisar que la duración del archivo constituye un criterio de selección del material y no un umbral específicamente validado para la detección de audio sintético.

A continuación, se mencionan los archivos de audio seleccionados para el presente análisis:

- **que tiene lule menem con Karina.mp4**
- **sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4**
- **sturzzeneger.mp4**
- **Zafarrancho.m4a**
- **sueldos.mp4**

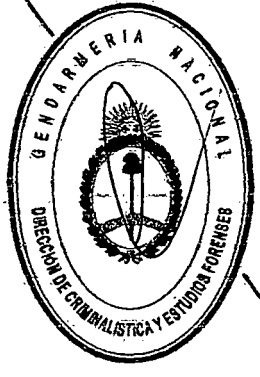
3.4.5.1. Análisis auditivo-perceptual de características acústico-fonéticas

El análisis auditivo-perceptual constituye un procedimiento basado en una escucha activa y sistemática de la voz presente en los archivos de audio, a fin de examinar características acústicas, elementos segmentales y suprasegmentales, entonación, pausas, velocidad de articulación y determinados fenómenos paralingüísticos.[34]

---

[33] Gil Fernández, J., Fernández Trinidad, M., Infante Ríos, P., & Lahoz-Bengoechea, J. M. (2017). *Obtaining speech samples for research and expertise in forensic phonetics*. En F. Orletti & L. Mariottini (Eds.), *Forensic communication in theory and practice: A study of discourse analysis and transcription* (pp. 27–46). Cambridge Scholars Publishing, p. 31.

[34] San Segundo Fernández, E. (2024). *Profundizando en los deepfakes: ¿Qué hace humana a una voz?* En Anuario AC/E de Cultura Digital 2024 (pp. 28–41). Acción Cultural Española.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
48
DIR. CRIMINALÍSTICA Y EST. FORENSES

En la interpretación de los atributos identificados deben considerarse la variabilidad natural de la producción de la voz, tanto entre distintos hablantes como dentro de las emisiones de un mismo individuo, así como los factores articulatorios, fonatorios, emocionales y expresivos que pueden modificar las características del habla. Asimismo, deben contemplarse las condiciones acústicas y técnicas de los registros, incluyendo posibles alteraciones derivadas del ruido ambiental, la compresión y la reverberación, dado que estos factores pueden condicionar la calidad del material disponible y su análisis.[35]

A continuación, se detalla el análisis perceptual de los CINCO (5) archivos de audio seleccionados:

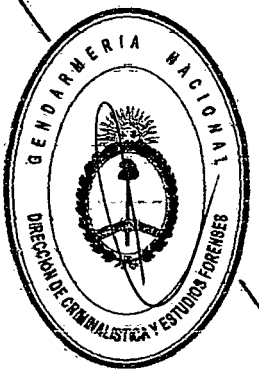
- **que tiene lule menem con Karina.mp4**

1. Características fonatorias:
   1.1. Tipo: se percibe una voz de características masculina
   1.2. Altura tonal: predominantemente media
   1.3. Calidad vocal: modal
2. Características prosódicas (suprasegmentales):
   2.1. Contorno entonativo: movimientos tonales ascendentes y descendentes durante la emisión
   2.2. Modalidades entonativas: entonación interrogativa y exclamativa
   2.3. Acentuación enfática: presente en determinados segmentos del discurso
   2.4. Ritmo y cadencia: se observa un ritmo general relativamente estable, con variaciones en la cadencia en determinados segmentos de la emisión
   2.5. Velocidad del habla: variaciones durante determinados segmentos de la emisión
3. Fluidez y organización temporal del habla:
   3.1. Pausas: se perciben pausas breves de respiración durante la emisión
   3.2. Repetición silábica: presente durante la emisión
   3.3. Repetición léxica: presente durante la emisión
4. Eventos orales no lingüísticos:
   4.1. Chasquidos linguales: presentes durante determinados segmentos de la emisión.

- **sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4**

1. Características fonatorias:
   1.1. Tipo: se percibe una voz de características masculina
   1.2. Altura tonal: predominantemente media
   1.3. Calidad vocal: modal
2. Características prosódicas (suprasegmentales):
   2.1. Contorno entonativo: movimientos tonales ascendentes y descendentes durante la emisión
   2.2. Modalidades entonativas: entonación interrogativa y exclamativa

[35] García Antuña, M. (2026). Deepvoice en el ámbito criminal: Retos metodológicos para la fonética forense. *Lingüística en la Red, 22*, 1–19. https://doi.org/10.37536/linred.2025.XXII.3294


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 49
DIR CRIMINALÍSTICA Y EST FORENSES

2.3. Acentuación enfática: presente en determinados segmentos del discurso
2.4. Ritmo y cadencia: se observa un ritmo general relativamente estable, con variaciones en la cadencia en determinados segmentos de la emisión
2.5. Velocidad del habla: variaciones durante determinados segmentos de la emisión.

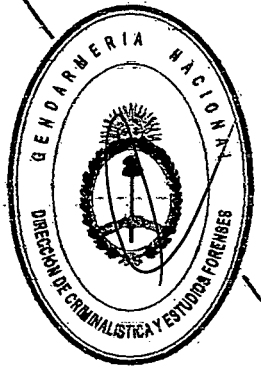
3. Fluidez y organización temporal del habla:
   3.1. Pausas: se perciben pausas breves de respiración durante la emisión
   3.2. Repetición silábica: presente durante la emisión
   3.3. Repetición léxica: no presente durante la emisión
4. Eventos orales no lingüísticos:
   4.1. Chasquidos linguales: presentes durante determinados segmentos de la emisión.

- **sturzzeneger.mp4**

1. Características fonatorias:
   1.1. Tipo: se percibe una voz de características masculina
   1.2. Altura tonal: predominantemente media
   1.3. Calidad vocal: modal
2. Características prosódicas (suprasegmentales):
   2.1. Contorno entonativo: movimientos tonales ascendentes y descendentes durante la emisión
   2.2. Modalidades entonativas: entonación interrogativa y exclamativa
   2.3. Acentuación enfática: presente en determinados segmentos del discurso
   2.4. Ritmo y cadencia: se observa un ritmo general relativamente estable, con variaciones en la cadencia en determinados segmentos de la emisión
   2.5. Velocidad del habla: variaciones durante determinados segmentos de la emisión
3. Fluidez y organización temporal del habla:
   3.1. Pausas: se perciben pausas breves de respiración durante la emisión
   3.2. Repetición silábica: presente durante la emisión
   3.3. Repetición léxica: presente durante la emisión
4. Eventos orales no lingüísticos:
   4.1. Chasquidos linguales: presentes durante determinados segmentos de la emisión.

- **Zafarrancho.m4a**

1. Características fonatorias
   1.1. Tipo: se percibe una voz de características masculina
   1.2. Altura tonal: predominantemente media
   1.3. Calidad vocal: modal
2. Características prosódicas (suprasegmentales):
   2.1. Contorno entonativo: movimientos tonales ascendentes y descendentes durante la emisión
   2.2. Modalidades entonativas: entonación interrogativa y exclamativa


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
50
DIR. CRIMINALÍSTICA Y EST. FORENSES

2.3. Acentuación enfática: presente en determinados segmentos del discurso
2.4. Ritmo y cadencia: se observa un ritmo general relativamente estable, con variaciones en la cadencia en determinados segmentos de la emisión
2.5. Velocidad del habla: variaciones durante determinados segmentos de la emisión
3. Fluidez y organización temporal del habla:
3.1. Pausas: se perciben pausas breves de respiración durante la emisión
3.2. Repetición silábica: presente durante la emisión
3.3. Repetición léxica: presente durante la emisión
4. Eventos orales no lingüísticos:
4.1. Chasquidos linguales: presentes durante determinados segmentos de la emisión.

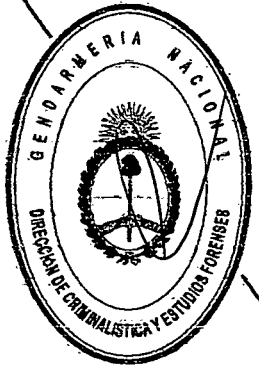
- **sueldos.mp4**

1. Características fonatorias
1.1. Tipo: se percibe una voz de características masculina
1.2. Altura tonal: predominantemente media
1.3. Calidad vocal: modal
2. Características prosódicas (suprasegmentales):
2.1. Contorno entonativo: movimientos tonales ascendentes y descendentes durante la emisión.
2.2. Modalidades entonativas: entonación interrogativa y exclamativa.
2.3. Acentuación enfática: presente en determinados segmentos del discurso.
2.4. Ritmo y cadencia: se observa un ritmo general relativamente estable, con variaciones en la cadencia en determinados segmentos de la emisión.
2.5. Velocidad del habla: variaciones durante determinados segmentos de la emisión.
3. Fluidez y organización temporal del habla
3.1. Pausas: se perciben pausas breves de respiración durante la emisión
3.2. Repetición silábica: presente durante la emisión
3.3. Repetición léxica: presente durante la emisión
4. Eventos orales no lingüísticos
4.1. Chasquidos linguales: presentes durante determinados segmentos de la emisión.

La valoración de estas observaciones frente a las hipótesis de producción vocal natural o sintética, se presenta conjuntamente en el apartado 3.4.5.6.

3.4.5.2. Análisis sobre el espectro de la señal de habla

El análisis espectrográfico constituye un procedimiento de examen acústico de la señal vocal que permite representar y estudiar la distribución de sus componentes frecuenciales a lo largo del tiempo. Mediante la utilización de representaciones espectrales es posible identificar y caracterizar determinadas propiedades acústicas del habla, incluyendo la distribución de los armónicos, el comportamiento


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja 51
DIR CRIMINALÍSTICA Y EST FORENSES

de la frecuencia fundamental y la distribución de energía asociada a las resonancias de los segmentos vocálicos.[36,37,38]

Para el análisis espectrográfico se llevará a cabo el examen de los siguientes parámetros:

- **Estructura armónica**: presencia, disposición y continuidad de los componentes armónicos en los segmentos sonoros de la emisión, con el propósito de caracterizar la organización periódica de la señal vocal.
- **Frecuencia fundamental (F0)**: estimación y variaciones a lo largo del tiempo, con la finalidad de caracterizar el comportamiento tonal de la emisión vocal.
- **Formantes (F1, F2, F3 y F4)**: distribución en determinados segmentos vocálicos, con el propósito de caracterizar sus propiedades de resonancia y su relación con la configuración articulatoria del tracto vocal.

Las consideraciones relativas a las condiciones acústicas y técnicas de los registros, mencionadas en el apartado 3.4.5.1, resultan igualmente aplicables al presente análisis. En particular, estos factores pueden modificar la representación espectral de la señal, dificultar la estimación de determinados parámetros acústicos y condicionar la interpretación de las características observadas.

Se efectuó un análisis espectral de los CINCO (5) archivos de audio mediante el software Praat, para ello se utilizaron los archivos en formato ".wav" previamente obtenido conforme lo descripto en el apartado 3.4.2.2, para poder analizarlo.

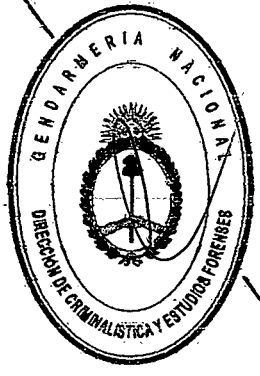
A continuación, se presentan, con carácter ilustrativo, capturas de pantalla correspondientes a un segmento de cada uno de los archivos de audio analizados.

---

[36] San Segundo Fernández (2024), p. 36

[37] Yang, T., Sun, C., Lyu, S., & Rose, P. (2026). *Forensic deepfake audio detection using segmental speech features. Forensic Science International, 379*, Article 112768. https://doi.org/10.1016/j.forsciint.2025.112768

[38] Li, M., Ahmadiadli, Y., & Zhang, X.-P. (2022). *A comparative study on physical and perceptual features for deepfake audio detection*. En *Proceedings of the 1st International Workshop on Deepfake Detection for Audio Multimedia* (pp. 35–41). Association for Computing Machinery. https://doi.org/10.1145/3552466.3556523, secciones 3.1–3.2, pp. 36–37.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
52

Figura 22. Presencia de armónicos, formantes y frecuencia fundamental. Archivo de audio: "que tiene lule menem con Karina.mp4"

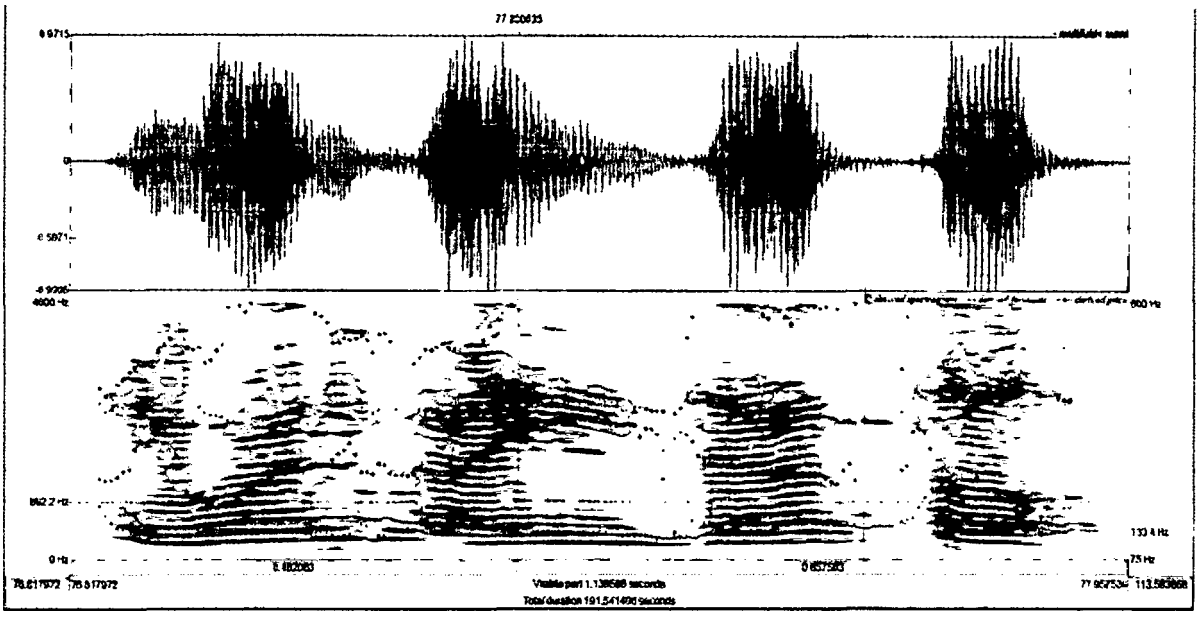
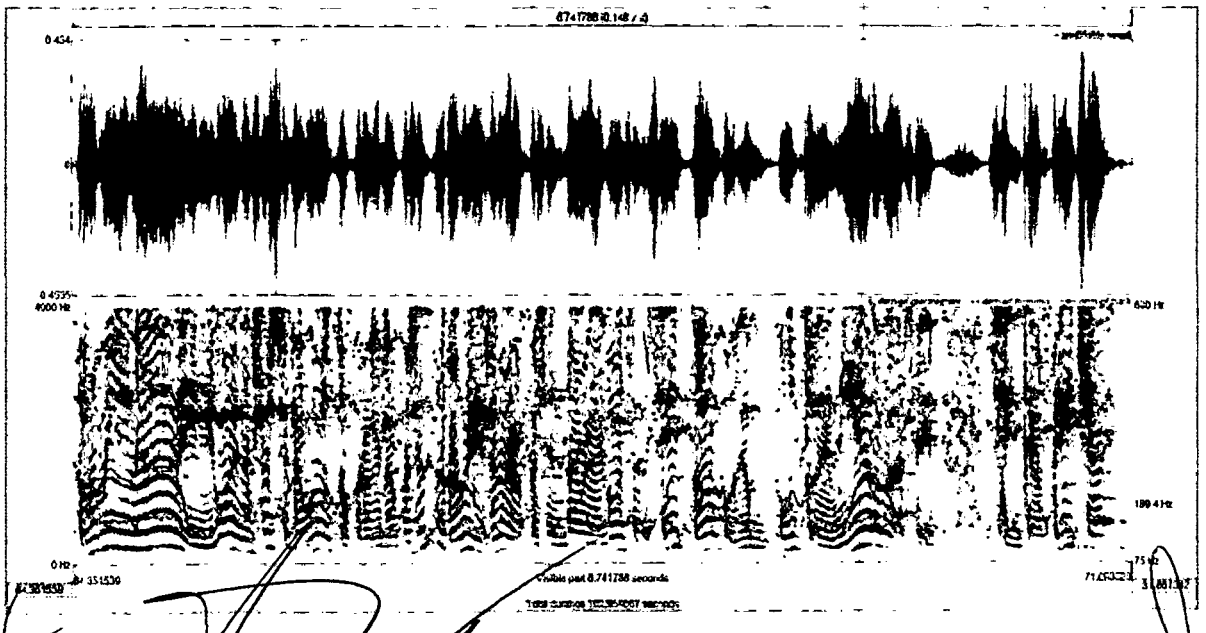
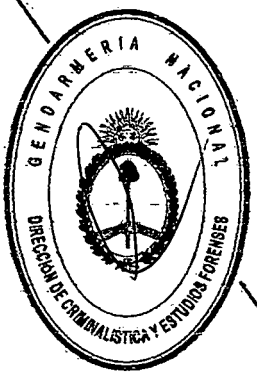


Figura 23. Presencia de armónicos, formantes y frecuencia fundamental. Archivo de audio: "sucalesca y proyecto discapacidad lilia y romina puto de mierda.mp4"




GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
53

Figura 24. Presencia de armónicos, formantes y frecuencia fundamental. Archivo de audio: "Zafarrancho.m4a".

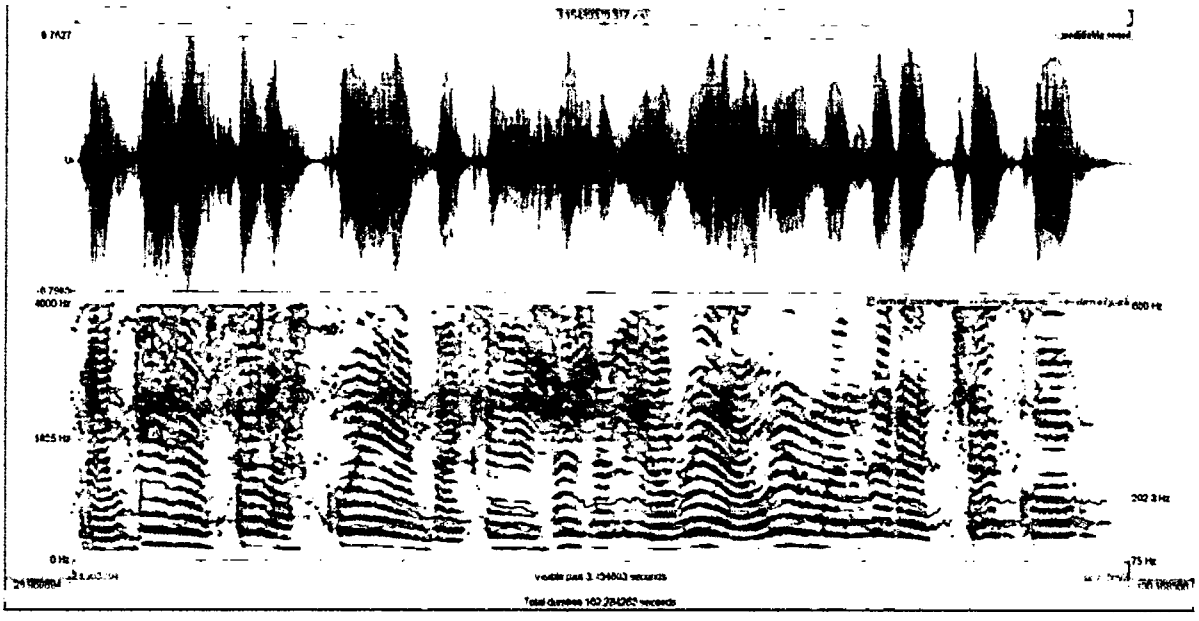
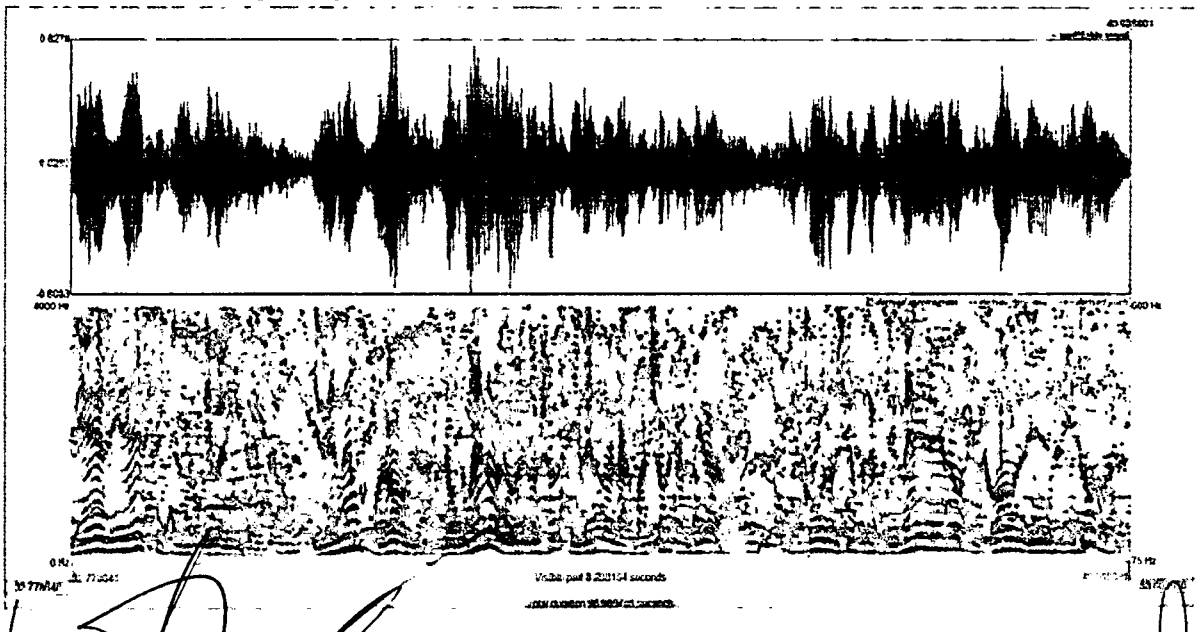
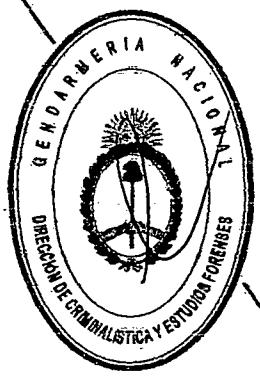


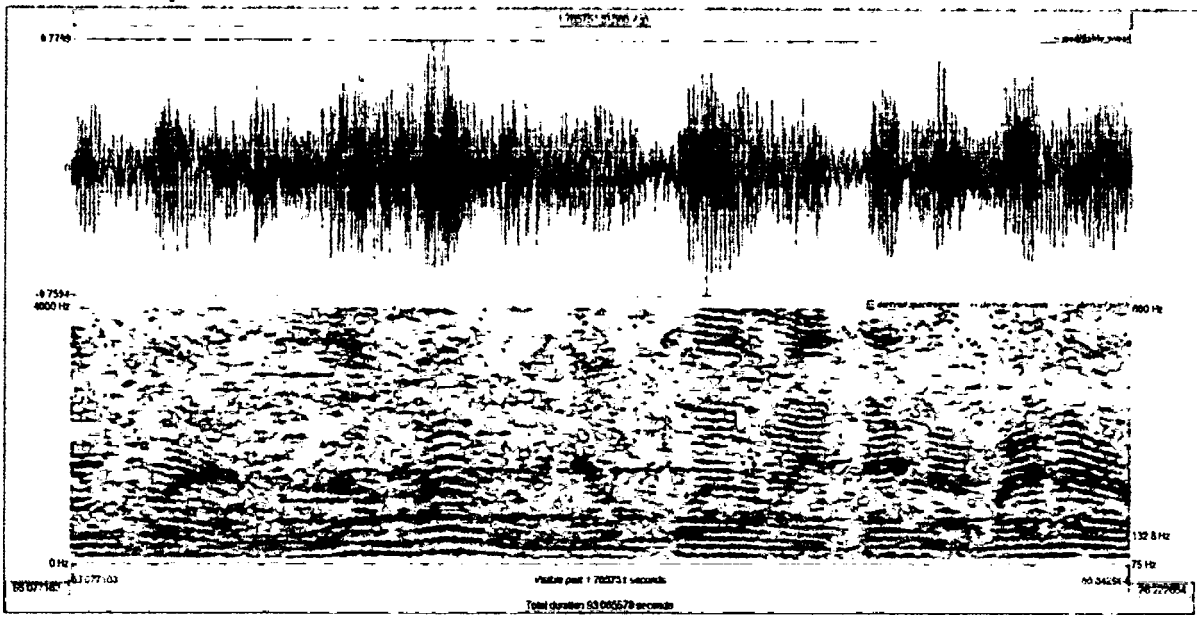
Figura 25. Presencia de armónicos, formantes y frecuencia fundamental. Archivo de audio: "sturzzeneger.mp4"




GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
DIR. CRIMINALÍSTICA Y ESTUDIOS FORENSES
Foja
54

Figura 26. Presencia de armónicos, formantes y frecuencia fundamental. Archivo: "sueldos.mp4"



Del examen visual, se observaron componentes armónicos y trazados correspondientes a las estimaciones de frecuencia fundamental y frecuencias de formantes obtenidas mediante Praat. Por el contrario, no se identificaron anomalías y/o irregularidades manifiestas en las características consideradas.

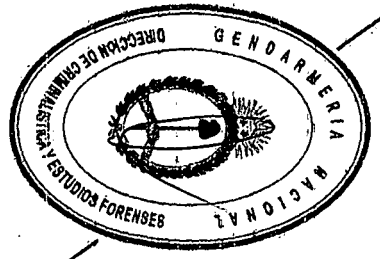
Los trazados automáticos pueden mostrar fluctuaciones o interrupciones asociadas a cambios fonatorios y prosódicos, a intervalos de baja intensidad o a limitaciones del procedimiento de estimación. Por consiguiente, una discontinuidad en el trazado estimado no debe confundirse directamente con una discontinuidad física del registro.

3.4.5.3. Coherencia ambiental

Como se describió en el apartado 3.4.3.3, durante el examen de los registros sonoros se identificaron diversos eventos acústicos cuyas características espectrales y/o temporales presentan posibles inconsistencias con la hipótesis de trabajo adoptada, entre ellos variaciones sustanciales del ruido de fondo. Estas características aportaron información relevante para el estudio de la continuidad de una grabación, dado que las propiedades acústicas del entorno constituyen trazas que pueden ser examinadas para identificar posibles inconsistencias en el desarrollo temporal del registro.[39,40]

[39] Zhao, H., Chen, Y., Wang, R., & Malik, H. (2014). *Audio source authentication and splicing detection using acoustic environmental signature*. En Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security (pp. 159–164). Association for Computing Machinery. https://doi.org/10.1145/2600918.2600933

[40] ENFSI, 2022, apartado 5.4.1, p. 13


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

Estas variaciones ambientales también pueden producirse por cambios en el entorno acústico, en las condiciones de captación o por procesos implementados por el sistema de grabación, incluso durante una adquisición continua[41] . El alcance interpretativo de los eventos observados, dadas las limitaciones contextuales ya establecidas en el apartado 3.4.3.1, se considera en la valoración conjunta del apartado 3.4.5.6.

3.4.5.4. Análisis de eventos respiratorios y consideraciones sobre la coarticulación

En este apartado se describe el análisis efectuado para identificar y localizar temporalmente eventos acústicos compatibles con respiraciones audibles durante la emisión vocal, mediante la escucha crítica de cada archivo de audio examinado. Las respiraciones audibles constituyen fenómenos no lingüísticos que pueden presentarse durante el habla espontánea y ser objeto de descripción dentro del análisis fonético forense[42,43].

Asimismo, se presentan consideraciones sobre la coarticulación y su posible utilidad en el estudio de las señales vocales, exponiendo los fundamentos metodológicos por los cuales no se efectuó su evaluación específica en el presente estudio.

3.4.5.4.1. Respiraciones

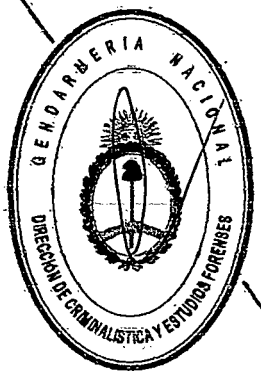
Se procedió a efectuar el análisis de la respiración en los CINCO (5) archivos de audio seleccionados previamente. Mediante la escucha minuciosa y el análisis perceptual, se consignaron los segmentos atribuibles a respiraciones, encontrados para cada archivo, en forma de etiquetas mediante el uso de la herramienta informática Audacity. Por lo cual, se ha generado para cada archivo, uno nuevo en formato ".txt" agregando al nombre original, el sufijo "_etiquetas_respiracion", detallados en la Tabla A6 (ver Anexo I). Se deja constancia, los archivos ".txt" generados para cada audio, será aportado junto con el presente informe pericial.

Las etiquetas generadas documentan la ubicación temporal de eventos acústicos compatibles con respiraciones audibles. Su valoración respecto del origen fisiológico de esos sonidos y de la eventual generación vocal mediante inteligencia artificial se desarrolla en el apartado 3.4.5.6.

3.4.5.4.2. Coarticulación

La coarticulación es un fenómeno de la producción del habla caracterizado por la influencia articulatoria y acústica entre segmentos fonéticos próximos. Sus manifestaciones temporales y espectrales, incluidas determinadas transiciones

[41] ENFSI, 2022, apartados 5.4.1–5.4.2, pp. 13–14
[42] San Segundo Fernández (2024), apartado 2.2, p. 36
[43] San Segundo et al. (2026), apartados 3.3.1 y 4.2.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
56
DIR CRIMINALÍSTICA Y ESTUDIOS FORENSES

en formantes, han sido estudiadas en el estado del arte como posibles características para distinguir habla natural y sintética. En este contexto, se evaluaron experimentalmente indicadores de coarticulación y del tiempo de inicio de la sonoridad, sin que sus resultados constituyan una validación generalizable a cualquier material forense.[44]

En el presente estudio no se efectuó un análisis específico de coarticulación, dado que no se dispone de un procedimiento de aplicación suficientemente establecido ni de la experiencia técnica específica necesaria para garantizar su adecuada ejecución e interpretación. Por tal motivo, no se consideró metodológicamente procedente incorporar dicho examen para evaluar el posible origen natural o sintético de las señales vocales.

3.4.5.5. Detectores algorítmicos de audio sintético

Al momento de la realización del presente Informe Pericial, la División Acústica Forense no dispone de un sistema de contramedidas destinado a la detección algorítmica de habla sintética que haya sido validado para su aplicación en las condiciones de los presentes registros. Esta clase de herramientas debe distinguirse de los sistemas de reconocimiento automático de hablantes (FASR), cuya finalidad es la comparación de muestras de voz para evaluar hipótesis relacionadas con la identidad del hablante, y no la determinación del posible origen sintético de las muestras. En este sentido, continúan en el ámbito académico, el desarrollo de numerosos métodos orientados a la discriminación entre habla natural y habla generada o modificada mediante técnicas de síntesis y/o conversión de voz. Esto incluye la creación de conjuntos de datos y protocolos experimentales destinados al entrenamiento, desarrollo y evaluación de estos sistemas, contemplando diferentes escenarios de ataque, entre ellos los denominados ataques de acceso lógico[45] (*Logical Access*, LA), acceso físico[46] (*Physical Access*, PA) y detección de audio sintético (Deepfake, DF).[47,48]

No obstante, las investigaciones han evidenciado dificultades relacionadas con la capacidad de generalización de los detectores frente a algoritmos de generación no presentes en sus conjuntos de entrenamiento, así como variaciones en su rendimiento ante diferentes condiciones acústicas y técnicas de registro. El ruido

---

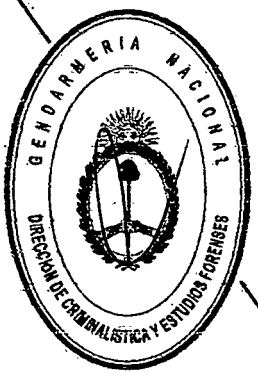
[44] Dhamyal, H., Ali, A., Qazi, I. A., & Raza, A. A. (2021). *Fake audio detection in resource-constrained settings using microfeatures*. En Proceedings of Interspeech 2021 (pp. 4149–4153). ISCA. https://doi.org/10.21437/Interspeech.2021-524

[45] Escenario de suplantación donde señales de habla generadas mediante síntesis, o conversión de voz, se introducen directamente al sistema de comunicación o reconocimiento, sin propagación acústica a través de un entorno físico.

[46] Escenario de suplantación basado en la reproducción de señales de habla, previamente registradas, mediante un dispositivo electro-acústico (como un altavoz), y su posterior captación mediante un micrófono, incorporando los efectos de la transferencia acústica.

[47] Liu et al. (2023). *ASVspoof 2021: Towards spoofed and deepfake speech detection in the wild. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31*, 2507–2522. https://doi.org/10.1109/TASLP.2023.3285283.

[48] Wang et al. (2026). *ASVspoof 5: Evaluation of spoofing, deepfake, and adversarial attack detection using crowdsourced speech. IEEE Transactions on Audio, Speech and Language Processing, 34*, 2354–2367. https://doi.org/10.1109/TASLPRO.2026.3682962.


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
57
DIR. CRIMINALÍSTICA Y EST. FORENSES

ambiental, la compresión del audio, las características del canal de transmisión y determinados procesos de degradación de la señal pueden afectar la capacidad discriminante de estos sistemas.[49] En el contexto forense, la aplicación de un procedimiento de detección algorítmica requiere disponer de información suficiente sobre su funcionamiento, rendimiento y limitaciones, así como de estudios de validación que permitan establecer su aptitud para las condiciones de análisis consideradas. Entre otros aspectos, resulta necesario conocer las tasas de error, el comportamiento frente a diferentes técnicas de generación o modificación de voz y la influencia de las condiciones de adquisición y procesamiento de las grabaciones.[50] En consecuencia, si bien existen herramientas de detección algorítmica de habla sintética de acceso libre o disponibles para uso público, su disponibilidad no implica que resulten adecuadas para su aplicación en el ámbito forense. Por tal motivo, al no contar con herramienta específica adecuada, no se encuentra esta División en capacidad de realizar detección mediante algoritmos.

3.4.5.6. Valoración conjunta de resultados y limitaciones

A partir de los procedimientos auditivo-perceptuales y de exploración acústica aplicados a los CINCO (5) archivos seleccionados, se identificaron y documentaron diversas propiedades de las señales vocales examinadas, entre ellas características fonatorias, segmentales y suprasegmentales, componentes armónicos, estimaciones de frecuencia fundamental y formantes, así como eventos acústicos compatibles con respiraciones audibles.

Durante el examen auditivo-perceptual se reconocieron variaciones entonativas, pausas, cambios en la velocidad de articulación, repeticiones y fenómenos vocales no lingüísticos. Asimismo, mediante la observación espectrográfica de los segmentos seleccionados no se identificaron irregularidades manifiestas en las características acústico-fonéticas examinadas. Los resultados obtenidos permitieron reconocer características habituales de la producción vocal humana, sin que se identificaran anomalías evidentes que fundamentaran una atribución de generación o modificación sintética de las emisiones. No obstante, la literatura científica ha documentado limitaciones en la capacidad humana para distinguir voces naturales y sintéticas, por lo que estas observaciones no permiten establecer el origen natural de las señales examinadas[51].

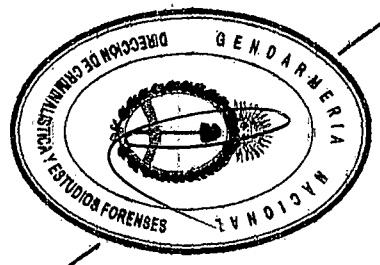
En relación con los eventos acústicos compatibles con respiraciones audibles, se identificaron y documentaron sus características y ubicaciones temporales en los audios examinados. Sin embargo, mediante los procedimientos aplicados no

---

[49] Univaso, P., & San Segundo, E. (2025). *Detection of cloned voices in realistic forensic voice comparison scenarios. Frontiers in Artificial Intelligence, 8*, Article 1678043. https://doi.org/10.3389/frai.2025.1678043.

[50] Swofford et al. (2025). *Validation in forensic science: Guiding principles for the collection and use of validation data* (NISTIR 8589). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8589.

[51] San Segundo et al. (2026), apartados 4.2 y 5.


GENDARMERIA
NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
58
DIR CRIMINALÍSTICA Y EST FORENSES

fue posible establecer la procedencia de dichos eventos ni su relación con la producción de las emisiones vocales registradas. En consecuencia, su presencia no permite fundamentar una atribución de origen natural o sintético de las emisiones vocales.

Por otra parte, las variaciones del ruido de fondo, previamente identificadas y documentadas durante el análisis de integridad, constituyen características de interés para evaluar la continuidad de los registros sonoros. Sin embargo, dichas variaciones pueden responder a diferentes causas y no permiten establecer, de manera inequívoca, la existencia de operaciones de edición. Asimismo, debe distinguirse la posible alteración de la continuidad de una grabación del origen natural o sintético de las emisiones vocales, dado que se trata de aspectos técnicamente diferentes.

Respecto de las limitaciones metodológicas, conforme a lo desarrollado en los apartados 3.4.5.4.2. y 3.4.5.5., no se efectuaron análisis específicos de coarticulación ni se aplicaron detectores algorítmicos de habla sintética[52], por los fundamentos técnicos y metodológicos allí expuestos.

En virtud de los resultados obtenidos, en las señales vocales examinadas no se identificaron irregularidades manifiestas en los parámetros acústico-fonéticos evaluados que permitan sustentar una atribución de generación o modificación mediante sistemas de inteligencia artificial. Sin embargo, dado que los procedimientos empleados no cuentan con capacidad discriminativa validada para establecer el origen natural o sintético de las emisiones, los resultados tampoco permiten descartar la intervención de dichos sistemas.

La valoración precedente se circunscribe a los CINCO (5) archivos seleccionados, a las características acústico-fonéticas examinadas y a las limitaciones metodológicas previamente expuestas.[53,54]

3.4.6. RELEVAMIENTO DE LAS CONDICIONES DE PRESERVACÓN Y CADENA DE CUSTODIA DEL ELEMENTO DE JUICIO

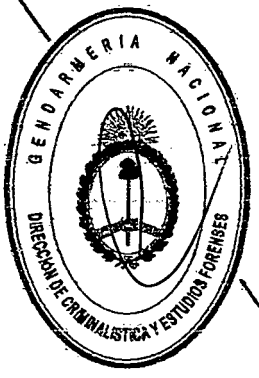
Conforme a la documentación obrante en las actuaciones, el día 28 de agosto de 2025 el Juzgado interviniente recepcionó el elemento de juicio y confeccionó la correspondiente planilla de cadena de custodia.

Al momento de su recepción en esta Dirección de Criminalística y Estudios Forenses, el elemento de juicio se encontraba contenido en UN (1) sobre de papel color blanco, cerrado mediante cinta adhesiva transparente, con una inscripción manuscrita que reza "CFP 3402/2025 PEDRIVE MAURO FEDERICO" y una inscripción impresa que reza "MINISTERIO PÚBLICO DE LA NACIÓN", acompañado de su respectiva planilla de cadena de custodia.

El mismo almacenaba en su interior, UN (1) dispositivo de almacenamiento portátil tipo pendrive, de colores blanco y amarillo, identificado con la marca "Kingston", modelo "DataTraveler G4", de 8 GB de capacidad.

---

[52] ENFSI (2022), apartados 5.4.1–5.4.2 y 5.4.5, pp. 13–16.
[53] ENFSI (2022), apartado 12, p. 21
[54] SWGDE (2018), versión 1.3, apartado 5.2, pp. 12–13.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

Respecto de las condiciones de preservación observadas, si bien el embalaje descripto permite contener e identificar físicamente el elemento de juicio, el cierre mediante cinta adhesiva transparente no ofrece garantías suficientes de inviolabilidad ni permite verificar de manera fehaciente la ausencia de accesos previos al dispositivo de almacenamiento.

Asimismo, la documentación de cadena de custodia permite examinar la trazabilidad del material a través de las intervenciones registradas, aunque no acredita necesariamente la integridad de su contenido digital.

3.4.7. IDENTIFICACIÓN DE MENCIONES DE NOMBRES PROPIOS

Se efectuó la escucha de la totalidad de los archivos de audio aportados, a fin de identificar posibles menciones del nombre y/o apellido de la persona de referencia, conforme a lo solicitado en el punto pericial correspondiente.

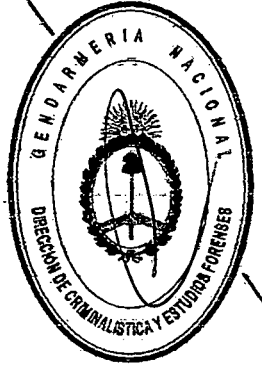
Durante el examen del archivo denominado "diego sos bueno.mp4", se identificó una mención audible del nombre propio "Diego", sin que se detectaran menciones de su apellido en los registros examinados.

SE DEJA CONSTANCIA QUE LA TOTALIDAD DEL MATERIAL GENERADO PARA EL DESARROLLO DEL PRESENTE INFORME PERICIAL, INCLUIDO EL ANEXO I, ES AGRUPADO EN UN ÚNICO ARCHIVO DE EXTENSIÓN ".RAR" IDENTIFICADO COMO "MATERIAL DEFINITIVO DEL INFORME PERICIAL.RAR", AL CUAL SE LE CALCULARON LOS RESPECTIVOS VALORES HASH:
MD5: 238fc54ee1684fa97d3222665a23120a
SHA256: 8d67d75a8cd02d9a8a912e818bf38530c87fc46de973f0d641661a68849e44be

ASIMISMO, SE INFORMA QUE TODA LA INFORMACIÓN, INCLUYENDO LA IMAGEN FORENSE, EXTRACCIONES DE ARCHIVOS, LOS VIDEOS DE APERTURA Y MATERIAL GENERADO ANTES MENCIONADO, FUERON ALMACENADOS EN UN DISCO RIGIDO MARCA SEAGATE NÚMERO DE SERIE NACSGC74 DE 1 TERABYTE DE CAPACIDAD APORTADO POR LOS PERITOS DE PARTE DEL MINISTERIO PÚBLICO FISCAL, EL CUAL ES PRESERVADO MEDIANTE BOLSA DE EVIDENCIA NRO "B0006459" (NIVEL DE SEGURIDAD 4), IDENTIFICADA ADEMÁS CON LOS DATOS DE LA PRESENTE CAUSA Y CERRADA CON PEGAMENTO. LA MISMA ES FIRMADA POR LA TOTALIDAD DE LOS PERITOS INTERVINIENTES. ADEMÁS SE GENERA UNA PLANILLA DE CADENA DE CUSTODIA

4. CONCLUSIONES

Luego de los estudios realizados sobre los archivos de audio analizados, quienes suscriben informan lo siguiente, Que:


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL - DIR. CRIMINALÍSTICA Y EST. FORENSES
Foja
60

1.A. EL FORMATO DEL CONTENEDOR, CÓDEC UTILIZADO, FRECUENCIA DE MUESTREO, TASA DE BITS, NÚMERO DE CANALES, DURACIÓN Y DEMÁS METADATOS TÉCNICOS EXTRAÍDOS DE LA TOTALIDAD DE LOS ARCHIVOS DE AUDIO APORTADOS PARA ESTUDIO, EXTRAÍDOS MEDIANTE LAS HERRAMIENTAS MEDIAINFO, MP4DUMP Y EXIFTOOL, SE DETALLAN EN LAS TABLAS A1, A2, A3 Y A4 DEL ANEXO I. - - - - - - - - - - - -

1.B. LOS ARCHIVOS DE AUDIO "que tiene lule menem con Karina.mp4" Y "villa-rruel ampliado y sandra pettovello.mp4" PRESENTAN IDÉNTICO CONTENIDO Y EL MISMO VALOR HASH, EXPUESTO EN LA TABLA 1. POR TAL MOTIVO, EL ANÁLISIS SE REALIZÓ SOBRE EL ARCHIVO DE AUDIO “que tiene lule menem con Karina.mp4”, POR TRATARSE DE REGISTROS CON IDÉNTICO CONTENIDO. -

2. DEL ESTUDIO DE FACTIBILIDAD REALIZADO, RESULTARON CATORCE (14) ARCHIVOS DE AUDIO APTOS Y CUARENTA Y TRES (43) NO APTOS PARA SER SOMETIDOS A UNA COMPARACIÓN FORENSE DE VOCES, CONFORME A LOS CRITERIOS ESTABLECIDOS EN EL PROTOCOLO DE COMPARACIÓN FORENSE DE VOCES. -

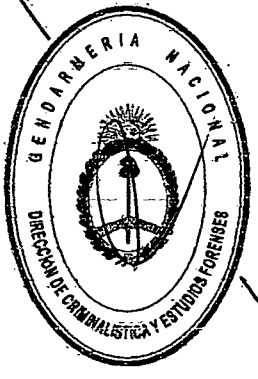
NO OBSTANTE, EN LOS ARCHIVOS QUE RESULTARON APTOS SE IDENTIFICARON CARACTERÍSTICAS ACÚSTICAS COMPATIBLES CON UN PROCESO DE REGRABACIÓN MEDIANTE TRANSFERENCIA ACÚSTICA, CONFORME LO DESARROLLADO EN EL APARTADO 3.4.3.6. - - - - - - - - - - -

DADO QUE EL SISTEMA AUTOMÁTICO FORENSIA V2.0 NO CONTEMPLA LA TRANSFERENCIA ACÚSTICA ENTRE LAS CONDICIONES DE CALIBRACIÓN, SE DESCONOCE SU DESEMPEÑO Y LAS TASAS DE ERROR ASOCIADAS A DICHA CONDICIÓN, POR LO QUE NO RESULTA METODOLÓGICAMENTE PROCEDENTE SOMETER A SU PROCESAMIENTO LOS CATORCE (14) ARCHIVOS QUE RESULTARON APTOS, SIN QUE ELLO MODIFIQUE LA CONDICIÓN DE APTITUD ESTABLECIDA CONFORME AL PROTOCOLO VIGENTE. -

EN CONSECUENCIA, ANTE LA IMPOSIBILIDAD DE EFECTUAR EL ANÁLISIS AUTOMÁTICO Y CONSIDERANDO QUE UN ANÁLISIS EXCLUSIVAMENTE AUDITIVO-PERCEPTUAL NO SATISFACE LOS REQUISITOS METODOLÓGICOS ESTABLECIDOS PARA LA COMPARACIÓN FORENSE DE VOCES, NO FUE POSIBLE REALIZAR EL COTEJO SOLICITADO. - - - - - -

3. POR LAS LIMITACIONES TÉCNICAS Y METODOLÓGICAS EXPUESTAS EN EL PUNTO ANTERIOR (2.), NO FUE POSIBLE EFECTUAR LA COMPARACIÓN DE PARÁMETROS ACÚSTICOS – FONÉTICOS SOLICITADA NI, POR CONSIGUIENTE, ESTABLECER UN GRADO DE COMPATIBILIDAD. - - - - - - -

4. SE EFECTUÓ UN EXAMEN EXHAUSTIVO DE LA TOTALIDAD DE LOS ARCHIVOS DE AUDIO APORTADOS, CONSTATÁNDOSE LA PRESENCIA DE SEGMENTOS SONOROS COINCIDENTES ENTRE DISTINTOS REGISTROS, CONFORME A LO DETALLADO EN EL APARTADO 3.4.3.7. DEL DESARROLLO DEL INFORME PERICIAL. ASIMISMO, SE IDENTIFICARON DISCONTINUIDADES ABRUPTAS EN LA FORMA DE ONDA Y EVENTOS SONOROS CON CONTENIDO ESPECTRAL SUPERIOR AL ANCHO DE BANDA OBSERVADO EN LA SEÑAL DE HABLA, ENTRE OTRAS


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

Foja 61

CARACTERÍSTICAS ACÚSTICAS QUE RESULTAN INCONSISTENTES CON LA HIPÓTESIS DE "SER UNA ADQUISICIÓN SONORA CONTINUA SIN PROCESAMIENTO POSTERIOR A SU CAPTURA ORIGINAL". - - - - - - - - - - - -

4.A. NO OBSTANTE, A PARTIR DEL MATERIAL APORTADO Y LAS HERRAMIENTAS DE ANÁLISIS DISPONIBLES, NO FUE POSIBLE ESTABLECER DE MANERA CONCLUYENTE EL ORIGEN DE DICHAS CARACTERÍSTICAS, NI DETERMINAR SU NATURALEZA O LA METODOLOGÍA TÉCNICA EVENTUALMENTE EMPLEADA, CONFORME A LO DETALLADO EN EL APARTADO 3.4.3.1. -

4.B. LA TOTALIDAD DE LOS HALLAZGOS FUERON INDIVIDUALIZADOS TEMPORALMENTE Y CONSIGNADOS EN LA TABLA A5 DEL ANEXO I. COMO ASI TAMBIÉN, ALMACENADOS EN ARCHIVOS DE TEXTO (.TXT) CORRESPONDIENTES A CADA AUDIO, LOS CUALES SE REMITEN JUNTO AL PRESENTE INFORME PERICIAL. -

4.C. EN LOS SEGMENTOS EVALUADOS EN LOS QUE NO SE IDENTIFICARON LAS CARACTERÍSTICAS ACÚSTICAS DESCRIPTAS, NO SE REGISTRARON HALLAZGOS DE INTERÉS PARA EL ANÁLISIS DE INTEGRIDAD, CONFORME A LA METODOLOGÍA APLICADA. -

4.D. DEL ESTUDIO EXHAUSTIVO DEL ARCHIVO DE AUDIO "LA CAJA .M4A" NO SE DETECTÓ INDICIO ALGUNO DE MANIPULACIÓN EVIDENTE, CONFORME LA METODOLOGÍA DE ANÁLISIS APLICADA. -

4.E. CONFORME YA FUE EXPUESTO EN EL PUNTO 4.A DE LA PRESENTE CONCLUSIÓN NO ES POSIBLE DETERMINAR EL TIPO NI METODOLOGÍA TÉCNICA DE EDICIÓN EVENTUALMENTE EMPLEADA. - - - - - - - - - - - - - - -

5. ANTE LA FALTA DE RESPUESTA AL REQUERIMIENTO DE INDIVIDUALIZAR LOS SEGMENTOS ATRIBUIDOS A DIEGO ORLANDO SPAGNUOLO, Y CONSIDERANDO QUE EN CADA UNO DE LOS ARCHIVOS SE CONSTATÓ PERCEPTUALMENTE LA PRESENCIA DE UNA ÚNICA VOZ MASCULINA DISCERNIBLE DENTRO DE CADA UNO DE LOS ARCHIVOS, SE ADOPTÓ COMO OBJETO DE ESTUDIO LA TOTALIDAD DE LA SEÑAL DE HABLA PRESENTE EN CADA REGISTRO. -

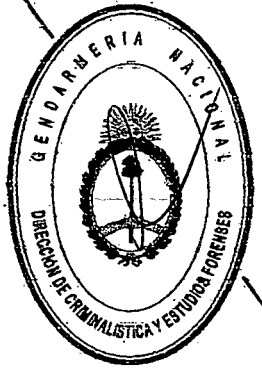
SE EFECTUÓ UN ANÁLISIS DESDE LA PERSPECTIVA FONOAUDIOLÓGICA, MEDIANTE PROCEDIMIENTOS AUDITIVO-PERCEPTUALES Y DE EXPLORACIÓN ACÚSTICA DE LA SEÑAL VOCAL. PARA ELLO, SE SELECCIONARON CINCO (5) ARCHIVOS DE AUDIO, CONFORME AL CRITERIO DE SELECCIÓN ADOPTADO PARA EL ESTUDIO DETALLADO EN EL APARTADO 3.4.5. DEL DESARROLLO DEL INFORME PERICIAL. - - - - - - -

5.A. LAS CARACTERÍSTICAS ACÚSTICO-FONÉTICAS IDENTIFICADAS SE DETALLAN EN EL APARTADO 3.4.5.1. DEL DESARROLLO DEL INFORME PERICIAL. -

5.B. LOS RESULTADOS DEL ANÁLISIS ESPECTRAL SE EXPONEN EN EL APARTADO 3.4.5.2. DEL DESARROLLO DEL INFORME PERICIAL. - - - - - - - -

5.C. EL ANÁLISIS DE LA COHERENCIA AMBIENTAL SE DETALLA EN EL APARTADO 3.4.5.3 Y EXPUESTOS EN LA TABLA A5 DEL ANEXO I. - - - - - -

5.D. SE DETALLARON EVENTOS COMPATIBLES CON RESPIRACIONES DEL INTERLOCUTOR, LAS CUALES SE ENCUENTRAN INDIVIDUALIZADAS Y


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
62
DIR CRIMINALÍSTICA Y EST FORENSES

CONSIGNADAS TEMPORALMENTE EN LOS ARCHIVOS DE TEXTO (.TXT) CORRESPONDIENTES A CADA REGISTRO SONORO, DETALLADO EN EL APARTADO 5.4. DEL INFORME PERICIAL, EXPUESTOS EN TABLA A6 DEL ANEXO I. --

5.E. LA DIVISIÓN ACÚSTICA FORENSE NO POSEE HERRAMIENTA CON ESTANDAR FORENSE PARA LA DETECCIÓN DEEPFAKE TAL COMO FUE DETALLADO EN EL APARTADO 3.4.5.4.1. DEL DESARROLLO DEL INFORME PERICIAL. --

6.A. LOS ANÁLISIS AUDITIVO-PERCEPTUALES EFECTUADOS PERMITIERON IDENTIFICAR Y DOCUMENTAR DISTINTAS CARACTERÍSTICAS DE LAS SEÑALES DE HABLA EXAMINADAS. SIN EMBARGO, LOS RESULTADOS OBTENIDOS NO PERMITEN ESTABLECER DE MANERA CONCLUYENTE SI CORRESPONDEN A EMISIONES VOCALES HUMANAS DE ORIGEN NATURAL. CONFORME LO DETALLADO EN LA CONCLUSIÓN ANTERIOR 5.E. --

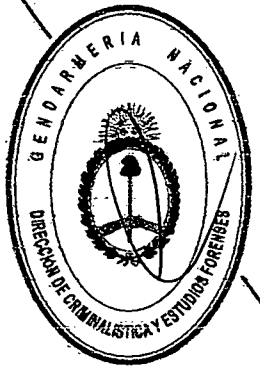
6.B. NO FUE POSIBLE ESTABLECER SI LAS SEÑALES DE HABLA EXAMINADAS CORRESPONDEN A DIEGO ORLANDO SPAGNUOLO, CONFORME A LO EXPUESTO EN EL PUNTO 2 DE LA PRESENTE CONCLUSIÓN. -----------

6.C. LAS CONCLUSIONES ALCANZADAS SE ENCUENTRAN LIMITADAS POR LAS CARACTERÍSTICAS DEL MATERIAL APORTADO, LA AUSENCIA DE INFORMACIÓN CONTEXTUAL Y EL ALCANCE DE LOS PROCEDIMIENTOS APLICADOS. ANTE LA AUSENCIA DE UNA METODOLOGÍA DE DETECCIÓN DE VOZ SINTÉTICA VALIDADA PARA LAS CONDICIONES EXAMINADAS Y LA IMPOSIBILIDAD DE EFECTUAR LA COMPARACIÓN FORENSE DE VOCES MEDIANTE EL SISTEMA DISPONIBLE, NO RESULTA PROCEDENTE ESTABLECER RANGOS CUANTITATIVOS DE CONFIANZA RESPECTO DEL ORIGEN NATURAL O SINTÉTICO DE LAS SEÑALES NI DE SU ATRIBUCIÓN A LA PERSONA DE REFERENCIA. ------------------------------

7. A), B), C) Y D) SI BIEN EXISTEN METODOLOGÍAS Y ALGORITMOS DE DETECCIÓN DE VOCES SINTÉTICAS DESARROLLADOS EN EL ÁMBITO DE LA INVESTIGACIÓN CIENTÍFICA, NO SE DISPONE EN ESTA DIVISIÓN DE HERRAMIENTAS FORENSE NI DE UNA METODOLOGÍA TÉCNICO-CIENTÍFICA VALIDADA PARA SU APLICACIÓN A LAS CONDICIONES PARTICULARES DEL MATERIAL EXAMINADO. POR LO EXPUESTO, NO FUE POSIBLE EMITIR UNA CONCLUSIÓN TÉCNICA. ------------------

8.A. AUNQUE EXISTEN METODOLOGÍAS QUE, BAJO DETERMINADAS CONDICIONES, PERMITEN IDENTIFICAR INDICIOS DE RECOMPRESIÓN, NO SE DISPONE DE UN PROCEDIMIENTO DE APLICACIÓN GENERAL QUE PERMITA ESTABLECER DE MANERA FEHACIENTE LA CANTIDAD DE RECOMPRESIONES EFECTUADAS NI RECONSTRUIR EL HISTORIAL DE CODIFICACIONES Y EXPORTACIONES SUCESIVAS UNICAMENTE A PATIR DE UN ARCHIVO DE AUDIO. EN CONSECUENCIA, NO FUE POSIBLE DETERMINAR SI LOS ARCHIVOS FUERON SOMETIDOS A PROCESOS DE RECOMPRESIÓN NI IDENTIFICAR ALTERACIONES ATRIBUIBLES A DICHOS PROCESOS. --

8.B. EXPUESTO EN EL APARTADO 1 DE LA PRESENTE CONCLUSIÓN. -------


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
63
DIR. CRIMINALÍSTICA Y EST. FORENSES

8.C. Y 8.D. NO ES POSIBLE ESTABLECER LA EXISTENCIA DE ALTERACIONES DERIVADAS DE EXPORTACIONES SUCESIVAS CONFORME LO DETALLA EN EL APARTADO 3.4.3.1 DEL DESARROLLO DEL INFORME PERICIAL. - - -

9.A. SE IDENTIFICARON DISCONTINUIDADES ABRUPTAS EN LA FORMA DE ONDA, COMPATIBLES CON POSIBLES PUNTOS DE CORTE, LAS CUALES FUERON INDIVIDUALIZADAS Y CONSIGNADAS TEMPORALMENTE EN LOS ARCHIVOS DE TEXTO (.TXT), EXPUESTOS EN TABLA A5 DEL ANEXO I, CORRESPONDIENTE A CADA REGISTRO. -

9.B. , 9.C Y 9.D. CONSIDERANDO LAS CARACTERÍSTICAS DEL MATERIAL APORTADO Y LAS HERRAMIENTAS DE ANÁLISIS DISPONIBLES EN ESTA DIVISIÓN, NO FUE POSIBLE RECONSTRUIR LA SECUENCIA CRONOLÓGICA DE EVENTUALES MODIFICACIONES, DETERMINAR LA EXISTENCIA DE SUPERPOSICIÓN DE CAPAS O INSERCIONES DE CONTENIDO SONORO, NI ESTABLECER EL ORIGEN DE LAS DISCONTINUIDADES OBSERVADAS. -

10. NO FUE POSIBLE EFECTUAR LA COMPARACIÓN TÉCNICA SOLICITADA, AL NO OBTENER RESPUESTA AL REQUERIMIENTO DE INDIVIDUALIZAR EL MATERIAL DIFUNDIDO EN MEDIOS DE COMUNICACIÓN U OTRAS VERSIONES CIRCULANTES, POR LO QUE TAMPOCO SE DISPUSO DE LOS ARCHIVOS NECESARIOS PARA EFECTUAR EL COTEJO Y DETERMINAR EVENTUALES DIVERGENCIAS ESTRUCTURALES O MODIFICACIONES ENTRE POSIBLES VERSIONES DEL MATERIAL. - - - - - - - - - - - - - - - - - - -

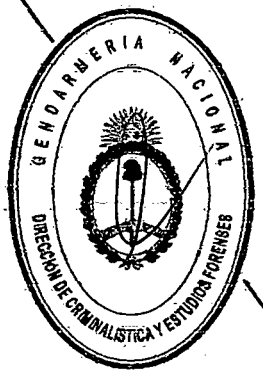
11.A. NO FUE POSIBLE EVALUAR DE MANERA FEHACIENTE LA CONTINUIDAD E INTEGRIDAD DE LA CADENA DE CUSTODIA DEL ELEMENTO DE JUICIO, CONFORME LO DESARROLLADO EN EL APARTADO 3.4.6 DEL PRESENTE INFORME. -

11.B. LOS METADATOS TÉCNICOS EXTRAÍDOS DE LOS ARCHIVOS EXAMINADOS CONTIENEN DIFERENTES CAMPOS DE FECHA, CUYOS VALORES SON SUSCEPTIBLES DE MODIFICACIÓN Y NO CONSTITUYEN, POR SÍ SOLOS, EVIDENCIA SUFICIENTE PARA ESTABLECER DE MANERA FEHACIENTE LA FECHA DE OBTENCIÓN. ASIMISMO, ANTE LA AUSENCIA DE INFORMACIÓN CONTEXTUAL INDEPENDIENTE, Y POR LAS LIMITACIONES EXPUESTAS EN EL APARTADO 3.4.3.1. DEL PRESENTE INFORME PERICIAL, NO FUE POSIBLE DETERMINAR EL LUGAR NI LAS CIRCUNSTANCIAS EN QUE FUERON REALIZADAS. - - - - - - - - - - - - - - - -

11.C. DEL ANÁLISIS EFECTUADO, Y DADAS LAS CARACTERÍSTICAS ACÚSTICAS DEL MATERIAL EXAMINADO, NO FUE POSIBLE IDENTIFICAR DIÁLOGOS, VOCES U OTRAS MANIFESTACIONES VOCALES PRESENTES EN LOS SONIDOS DE FONDO NI, POR CONSIGUIENTE, EFECTUAR PROCEDIMIENTOS DESTINADOS A SU RECUPERACIÓN. - - - - - - - - - - - - - -

11.D. EN EL ARCHIVO DE AUDIO DENOMINADO “DIEGO SOS BUENO.MP4” SE IDENTIFICÓ, MEDIANTE ESCUCHA, UNA MENCIÓN AUDIBLE DEL NOMBRE PROPIO “DIEGO”. -

11.E. SI BIEN SE IDENTIFICARON DISCONTINUIDADES Y OTRAS CARACTERÍSTICAS ACÚSTICAS COMPATIBLES CON POSIBLES OPERACIONES DE EDICIÓN, NO FUE POSIBLE ESTABLECER SI ESTAS


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
64
DIR CRIMINALÍSTICA Y EST FORENSES

INVOLUCRARON LA ELIMINACIÓN DE INTERVENCIONES VOCALES DE UN EVENTUAL INTERLOCUTOR DEL HABLANTE PRINCIPAL. POR LO TANTO, NO SE DISPONE DE ELEMENTOS SUFICIENTES PARA DETERMINAR LAS CAUSAS TÉCNICAS DE LA POTENCIAL AUSENCIA DE SU REGISTRO AUDIBLE.---

11.F. LA TOTALIDAD DE LAS CARACTERÍSTICAS TÉCNICAS Y/O VARIACIONES TÉCNICAS SIGNIFICATIVAS DETECTADAS, MEDIANTE LA METODOLOGÍA DISPONIBLE, SE ENCUENTRAN DETALLADAS EN LA TABLA A5 DEL ANEXO I. NO PUDIENDO ESTABLECER LA EXISTENCIA DE DISTINTOS FORMATOS O MODALIDADES DE CAPTURA DEL AUDIO EN CADA ARCHIVO DE AUDIO. ASIMISMO, LA COMPATIBILIDAD CON UNA GRABACIÓN CONTINUA FUE ABORDADA EN EL PUNTO 4 DE LA PRESENTE CONCLUSIÓN.---------------------------------

11.G. LOS ANÁLISIS DE INTEGRIDAD DE AUDIO PERMITEN IDENTIFICAR INDICIOS TÉCNICOS DE POSIBLES ALTERACIONES, DISCONTINUIDADES O PROCESAMIENTOS SOBRE EL MATERIAL SONORO. SIN EMBARGO, LA AUSENCIA DE INFORMACIÓN CONTEXTUAL INDEPENDIENTE SOBRE EL ORIGEN DE LAS GRABACIONES, LOS DISPOSITIVOS Y CONDICIONES DE CAPTACIÓN, LOS PROCESOS DE TRATAMIENTO POSTERIOR Y LA DISPONIBILIDAD DE LOS ARCHIVOS FUENTE ORIGINALES, SUMADA A LAS LIMITACIONES DE LAS HERRAMIENTAS INFORMÁTICAS FORENSES DISPONIBLES EN ESTA DIVISIÓN, RESTRINGE LA POSIBILIDAD DE ESTABLECER LAS CAUSAS DE LAS CARACTERÍSTICAS ACÚSTICAS DETECTADAS Y RECONSTRUIR LOS EVENTUALES PROCESOS A LOS QUE PUDIERON HABER SIDO SOMETIDOS LOS ARCHIVOS EXAMINADOS.----

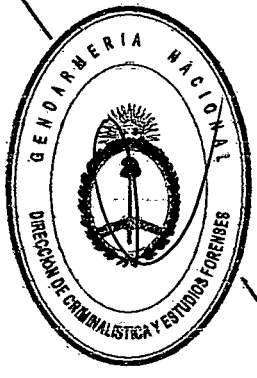
Asimismo, se dio lectura al informe pericial final, a partir de lo cual, conforme orientación telefónica recibida por parte del secretario del juzgado Doctor Alan Vence, las manifestaciones de los peritos de parte se exponen en el ANEXO II del informe pericial.

El presente informe pericial consta de SETENTA Y TRES (73) fojas útiles.–

Es todo cuanto se informa. –

BUENOS AIRES, 11 de noviembre de 2026.

LIC. DIEGO BERMÚDEZ GOLINELLI	**ING. ADRIÁN NOREMBERG**	**ING. MARCELO ANTONIO TOROK**
DNI: 30.234.196	DNI: 21.671.619	DNI: 14.596.326
Perito de parte	Perito de parte	Perito de parte


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
65
DIR. CRIMINALÍSTICA Y EST. FORENSES

LIC. CECILIA BEATRIZ VERA
DNI: 21.524.504
Perito de parte

GABRIEL EDUARDO JOAQUÍN GAETE
Cabo
División Acústica Forense

STEFANÍA MARIANELLA SULIGOY
Alférez
División Acústica Forense

JOHANA MARÍA BELÉN LUGO
Alférez
División Acústica Forense

Los peritos del Ministerio Público Fiscal firman en disidencia en virtud de la negativa a incorporar las manifestaciones en las conclusiones del presente Informe.

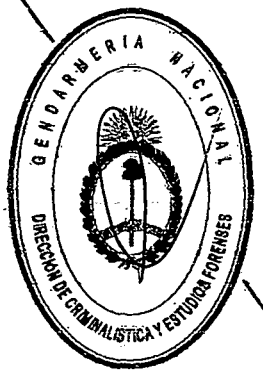
BRUNO MOGLIA
DNI: 31.061.087
Perito de parte

DR. ROBERTO LEONARDO
DNI: 24.577.375
Perito de parte

GALA IGLESIAS BRICKLES
DNI: 35.112.244
Perito de parte

ING. NICOLÁS EDUARDO SANGUINETI
DNI: 31.824.316
Perito de parte

ING. JOSÉ GABRIEL POLLOLA BARÓN
DNI: 26.601.258
Perito de parte


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
66
DIR CRIMINALÍSTICA Y EST FORENSES

ANEXO II

MANIFESTACIONES DE LOS PERITOS DE PARTE

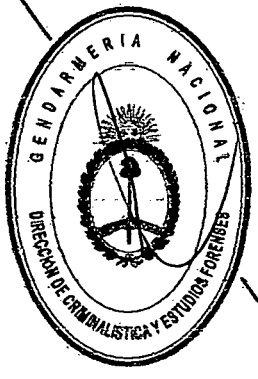
I. La Licenciada Vera y los Ingenieros Golinelli y Torok, manifiestan: que habiendo tomado conocimiento en el día de la fecha de las extracciones aportadas por los peritos oficiales y sin perjuicio de acompañar preliminarmente las mismas, dado que reciben en el presente acto, la copia de traslado de los archivos analizados durante las diligencias periciales, presentarán el pertinente dictamen de parte, una vez recibida copia del dictamen oficial. Acto seguido, dejan constancia que su manifestación vertida en el presente anexo es al solo efecto de realizar sus observaciones, sin avalar en forma alguna las eventuales manifestaciones de los otros peritos de parte.

II. El Ingeniero Noremberg manifiesta que durante las sesiones conjuntas se detectaron inconsistencias en los criterios de factibilidad y/o aptitud de los archivos para cotejo del Protocolo. Vale la pena resaltar que dicho protocolo tiene una emisión original y no cuenta con revisiones a la fecha. El criterio de este perito es que con marcadas evidencias en manipulación de los archivos de audio que afectaron la integridad de los mismos deberían quedar anulados para su comparación. No obstante la presente manifestación y dado que se recibe el material para estudio más profundo se reserva el derecho de emitir un dictamen de parte solo o en conjunto.

III. Los peritos del MPF manifiestan su completa disconformidad con las conclusiones del dictamen pericial debido a que fue cercenado su derecho a manifestar disidencias en las conclusiones del dictamen pericial. De esta manera queda desnaturalizado el debate pericial para el cual han aceptado el cargo por el Ministerio Público Fiscal de la Nación. Sin perjuicio de esto exponen a continuación algunas disidencias, las cuales se vuelcan en este anexo II por indicación de los peritos oficiales:

Sobre los puntos de pericia 2 y 3.

Los Ingenieros José Pollola Barón y Nicolás Sanguineti, y el Sr. Bruno Moglia del Laboratorio de Multimedia Forense de la DATIP, todos peritos de parte del MPF, manifiestan que la verificación de factibilidad realizada en este estudio pericial indicó que **15 archivos de audio son APTOS para la comparación forense de voces**, conforme al Protocolo Único para la Comparación Forense de Voces (Ministerio de Seguridad de la Nación, Resolución Ministerial N.° 526/2024). Esto significa que, en cada uno de esos 15 archivos, la señal vocal identificada presenta una **calidad sonora suficiente** para efectuar una comparación forense con una muestra de voz indubitada (la cual nunca fue aportada


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

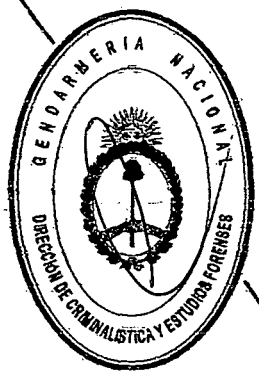
GENDARMERÍA NACIONAL
Foja
67
DIR CRIMINALÍSTICA Y EST FORENSES

por el imputado Diego Orlando Spagnuolo). Por otra parte, en su ítem 4.7, el citado protocolo contempla tres tipos de análisis —**perceptual, semiautomático y automático**— y exige "**realizar al menos dos** tipos de análisis de los mencionados". En ese marco, si **el software Forensia v2.0** no resulta apto para efectuar el análisis automático debido a las limitaciones explicitadas durante el desarrollo, entendemos que este hecho **no justifica, por sí mismo, la cancelación de la comparación**, ya que **sería posible realizar los otros dos tipos de análisis previstos: el perceptual y el semiautomático**. Asimismo, aun si la aplicación de estos dos métodos diera lugar a conclusiones con un menor grado de confianza, esa circunstancia tampoco justificaría prescindir del estudio, toda vez que el punto 6 (c) del Oficio Judicial de fecha 13 de julio de 2026 prevé expresamente que se indiquen "las limitaciones metodológicas del análisis y los rangos de confianza de las conclusiones"

Respecto del método automático, estos peritos del Ministerio Público manifiestan que valoran la prudencia metodológica de los peritos oficiales, pero entienden que sería conveniente someter esta limitación detectada a una verificación empírica antes de renunciar a su factibilidad.

Roberto L. Leonardo, perito de parte del MPF, manifiesta dos motivos por lo que diciente con el análisis y la conclusión de los puntos 2 y 3, en primer lugar, según el dictamen, "la evaluación de factibilidad tiene por objeto determinar si el material reúne condiciones acústicas que permitan su utilización en un eventual análisis comparativo de voces con fines de identificación". Al finalizar la reunión del 2 de septiembre de 2026 se determinó que el archivo identificado como "que tiene lule menem con Karina.mp4" "reunía las condiciones previstas en el Protocolo para ser considerado APTO", y se gestionó en consecuencia "la convocatoria de la persona de referencia para la obtención de muestras indubitadas de su voz". Sin embargo, el 16 de septiembre de 2026, sobre el mismo archivo, los peritos oficiales concluyeron que el cotejo no resultaba técnicamente procedente a pesar de mantener la condición de aptitud. Se establece, de manera contradictoria que la muestra es APTA pero que no puede someterse a ningún método de comparación para el que se evaluó su aptitud.

Hace notar que las características que fundaron la exclusión (atenuación en el rango de baja frecuencia, eventos impulsivos con contenido espectral superior al ancho de banda del habla y discontinuidades abruptas en la forma de onda) son propiedades del mismo


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
68
DIR CRIMINALÍSTICA Y EST FORENSES

archivo dubitado, ya presentes al momento de evaluar la factibilidad. La posterior utilización del programa Forensia v2.0 (sistema automático) también era conocida de antemano y debió considerarse al declarar la aptitud.

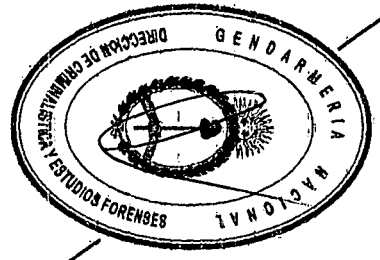
En segundo lugar, manifiesta que carece de fundamentos la afirmada imposibilidad de efectuar el cotejo. El apartado 4.7 del Protocolo (Protocolo Único para la Comparación Forense de Voces - Resolución MS 526/2024) contempla tres tipos de análisis (perceptual, semiautomático y automático) y establece que "Es necesario realizar al menos dos tipos de análisis". El dictamen descarta el análisis automático, pero omite el método perceptual y semiautomático. Este último, según el Protocolo, consiste en realizar el análisis espectrográfico de los segmentos de habla con medición de rasgos como los formantes. La División cuenta con esa capacidad: entre otros, en los estudios referentes al punto pericial 5 del mismo dictamen se efectuó, con el software Praat, el análisis espectrográfico de la estructura armónica, la frecuencia fundamental y los formantes del archivo. La combinación de los análisis perceptual y semiautomático satisface el mínimo de dos métodos que exige el Protocolo. Por lo tanto, la imposibilidad de utilizar Forensia v2.0 no conduce, por sí sola, a la imposibilidad de efectuar el cotejo. A ello se suma que el dictamen reconoce que los hallazgos solo "constituyen indicios acústicos que proporcionan fundamentos razonables para plantear la hipótesis" de una transferencia acústica, y admite que "las observaciones efectuadas **no permiten establecer de manera concluyente dicho mecanismo de producción**". Aun así, esa hipótesis no confirmada se utilizó para excluir por completo el cotejo.

Sobre el punto de pericia 3.

El Ing. Pollola Barón, perito de parte del MPF, manifiesta que, aunque les fue informado que las limitaciones de calibración de FORENSIA justifican no utilizar sus resultados con fines identificatorios, entiende que ello no permite descartar por sí mismo otras técnicas reconocidas de comparación forense, o en todo caso precisar las razones técnicas por las cuales eventualmente tampoco resultan aplicables.

Sobre el punto de pericia 5.

Los Ingenieros José Pollola Barón y Nicolás Sanguineti, peritos de parte del MPF, manifiestan que, si bien el informe señala que no fue posible determinar concluyentemente el origen natural o sintético de las voces examinadas, **el propio análisis identificó características compatibles con el habla humana espontánea, como variaciones de entonación, pausas, cambios en la velocidad de articulación, repeticiones, chasqui-**

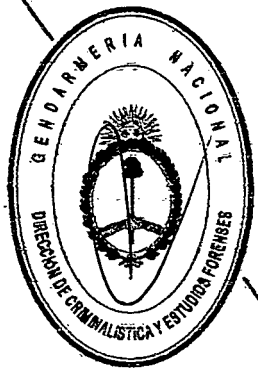

GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

dos linguales y eventos compatibles con respiraciones audibles. Asimismo, entienden que el examen espectrográfico de los cinco archivos seleccionados no evidenció anomalías manifiestas. Por lo tanto, consideran que estos hallazgos deberían valorarse críticamente atribuyendo el peso fáctico en proporciones correctas y coherentes, **destacando que se identificaron características compatibles con una producción vocal natural, sin haberse documentado anomalías específicas atribuibles a síntesis artificial en los análisis realizados**.

Gala Iglesias Brickles, miembro del Laboratorio Multimedia Forense de DATIP y perito de parte del MPF, manifiesta que el informe pericial incurre en una severa incoherencia metodológica: utiliza bibliografía científica para justificar las limitaciones de la percepción auditiva humana, pero omite deliberadamente aplicar los procedimientos instrumentales y auditables que esa misma literatura propone. En efecto, el dictamen cita los trabajos de Yang, Sun, Lyu y Rose (2026) y de García Antuña (2026), pero ignora la tesis central de dichos autores. Puntualmente, Yang et al. (2026) recomienda el uso de un método segmental instrumentalizado diseñado para superar las falencias del oído humano, recomendándolo explícitamente por sobre los sistemas automáticos de detección de deepfakes. Estos últimos operan bajo un esquema de "caja negra" (black box), lo que impide auditar el proceso de análisis y verificar la trazabilidad científica sobre la muestra dubitada.

La pericia sustituyó unilateralmente el análisis estadístico objetivo de jitter y shimmer por una apreciación fonoaudiológica subjetiva, dejando sin respuesta el punto de pericia solicitado y sin justificar la imposibilidad de realizar la medición instrumental. Asimismo, el estudio carece de sustento teórico y reproducibilidad al emplear una grilla de doce categorías sin referencias bibliográficas que acrediten su validez en el ámbito forense, sumado a una total falta de trazabilidad al omitir marcas de tiempo (timestamps) y registros gráficos de respaldo. A ello se suma un dato revelador: la evaluación fonoaudiológica arrojó descripciones prácticamente idénticas en los cinco archivos analizados.

Respecto al examen de las respiraciones efectuado, resulta metodológicamente incompleto y carente de rigor diagnóstico. El informe se limita a consignar un mero registro auditivo de los puntos donde el perito cree percibir inhalaciones, pero omite evaluar la sincronía fisiológica y sintáctica de dichos eventos con el discurso. En acústica forense y detección de síntesis de voz (deepfakes), la sola presencia de un sonido respiratorio no acredita autenticidad. Tal como demuestran Han et al. (2026) y Doan et al. (2023), lo que permite discriminar una emisión humana de una sintética es la correlación temporal y aerodinámica entre la respiración, las pausas y el contexto fonético circundante. Al no

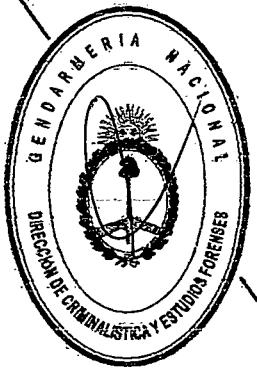

GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
70
DIR CRIMINALÍSTICA Y EST FORENSES

analizar la duración, el espaciado ni la congruencia gramatical de las respiraciones (Layton et al., 2025; Kulangareth et al., 2024), la observación de la contraparte queda reducida a una apreciación subjetiva.

Respecto al análisis de coarticulación, resulta cuestionable la justificación esgrimida para omitirlo, alegando que "no se dispone de un procedimiento de aplicación suficientemente establecido". En el ámbito de la fonética acústica y la lingüística forense, la evaluación instrumental de la coarticulación cuenta con protocolos estandarizados y ampliamente validados desde hace décadas. Específicamente, su medición se realiza de forma objetiva y reproducible mediante el seguimiento de transiciones formánticas (F1–F3) en los límites consonante-vocal y el cálculo de ecuaciones de locus (locus equations, Sussman et al., 1991; Recasens, 2018). Asimismo, los estándares internacionales de la especialidad —como el Best Practice Manual de la ENFSI (2022), también citado en este mismo informe— contemplan expresamente el análisis dinámico de las trayectorias espectrales y formánticas como parte de la metodología forense. Cabe destacar, además, que el propio equipo de peritaje utilizó el software Praat en su informe, herramienta diseñada precisamente para ejecutar este tipo de segmentaciones y mediciones formánticas. En consecuencia, la omisión de este punto pericial no responde a una carencia de métodos en la disciplina, sino a la falta de aplicación técnica por parte de la pericia.

Roberto L. Leonardo, perito de parte del MPF, manifiesta que se advierte una incongruencia en el punto 5 del informe entre los resultados del análisis y la conclusión. Se seleccionó el material, definieron los parámetros de análisis y se aplicaron a los cinco archivos los procedimientos auditivo-perceptuales y espectrográficos, lo que supone que los consideraron idóneos para responder el punto pericial. Todos los resultados apuntan en la misma dirección. En los cinco archivos se relevaron variaciones entonativas, pausas breves de respiración, cambios en la velocidad del habla, repeticiones silábicas (y léxicas en cuatro de ellos), y chasquidos linguales; se documentaron respiraciones audibles; y en el examen espectrográfico de la estructura armónica, la frecuencia fundamental y los formantes "no se identificaron anomalías y/o irregularidades manifiestas en las características consideradas". Sobre esa base, se afirma que "**Los resultados obtenidos permitieron reconocer características habituales de la producción vocal humana, sin que se identificaran anomalías evidentes que fundamentaran una atribución de generación o modificación sintética de las emisiones**", y que "**no se identificaron irregularidades manifiestas en los parámetros acústico-fonéticos evaluados que permitan sustentar una atribución de generación o modificación mediante**


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

sistemas de inteligencia artificial". Sin embargo, al formular la conclusión, no se asumen las consecuencias de los hallazgos. Se recuerdan que los procedimientos se aplicaron "sin constituir un procedimiento específico de detección de habla sintética" y que "los procedimientos empleados no cuentan con capacidad discriminativa validada para establecer el origen natural o sintético de las emisiones", y concluyen que "los resultados tampoco permiten descartar la intervención de dichos sistemas". **Esa reserva no surge de ningún hallazgo concreto del estudio.** Si los procedimientos no eran aptos, no se explica por qué se los aplicó ni por qué se valoraron sus resultados. Si lo eran, como surge de su empleo**, la conclusión debió expresar lo que esos resultados muestran: no se identificó ningún indicio de síntesis, y sí múltiples rasgos característicos de la producción vocal humana. Al destacar la presencia de "características habituales de la producción vocal humana" y la falta de irregularidades que "permitan sustentar" una atribución sintética.**

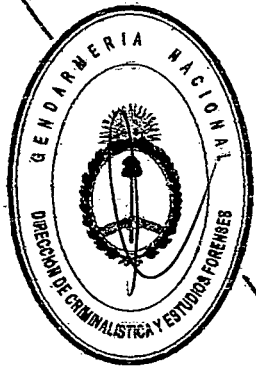
Sobre el *punto de pericia* 6.

Los Ingenieros Nicolás Sanguineti y José Pollola Barón, y el Sr. Bruno Moglia del Laboratorio de Multimedia Forense de la DATIP, todos peritos de parte del MPF, manifiestan que, pese a las limitaciones técnicas mencionadas en los "estudios relativos al análisis de audio sintético" de este punto pericial se describen **numerosas características vocales compatibles con voz humana real de origen no sintético** y no se individualiza **ningún hallazgo atribuible específicamente a generación sintética o modificación mediante inteligencia artificial**. Por este motivo, entendemos que **la conclusión esbozada por los peritos oficiales en este punto se encuentra descontextualizada** con el desarrollo que ellos mismos llevaron a cabo, evitando realizar una valoración que se desprende de la simple lectura estructurada de su análisis.

Sobre el punto de pericia 7.

El Ing. Pollola Barón, perito de parte del MPF, manifiesta que entiende que la ausencia de herramientas de detección de voz sintética validadas y disponibles en la División Acústica Forense constituye una limitación de los recursos y procedimientos aplicados, pero sin embargo esto no debería interpretarse como una imposibilidad técnica general de realizar este tipo de análisis. Si existen metodologías de detección desarrolladas en el ámbito científico, su aplicabilidad, dada la importancia de este aspecto, deberían evaluarse teniendo en cuenta las características y condiciones del material examinado.

Gala Iglesias Brickles, miembro del Laboratorio Multimedia Forense de DATIP y perito de parte del MPF, manifiesta que la bibliografía citada en el propio informe (Yang et al., 2026, ref. 36) señala que la síntesis de voz deja correlatos acústicos medibles a nivel


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

GENDARMERÍA NACIONAL
Foja
72
DIR. CRIMINALÍSTICA Y EST. FORENSES

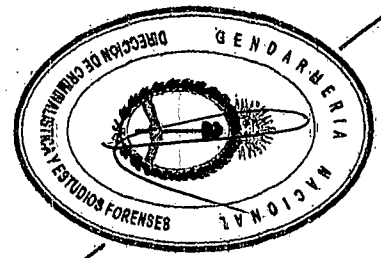
segmental, particularmente en la trayectoria de los formantes vocálicos, las transiciones coarticulatorias y el comportamiento de la frecuencia fundamental (F0). Aun cuando el laboratorio no contara con un software automático para detectar voces sintéticas, el instrumental empleado en la propia pericia —el software Praat— permite relevar, medir y describir objetivamente estos parámetros acústicos básicos. Se advierte que el dictamen no sólo se abstiene de emitir una conclusión técnica, sino que tampoco ofrece observaciones descriptivas sobre lo que se visualiza en los gráficos adjuntos. El informe incorpora vistas espectrográficas a modo meramente ilustrativo, pero omite consignar qué se observa y qué se mide en ellas.

Sobre los puntos de pericia 4 y 9.

Bruno Moglia, del Laboratorio de Multimedia Forense de la DATIP, perito de parte del MPF, manifiesta que el estudio pericial identifica características compatibles con posibles operaciones de edición, pero **no expone una evaluación del contenido** de los fragmentos continuos (aquellos que no muestran en su interior indicios observables de edición) y, en consecuencia, no se examina si estos pasajes conservan enunciados completos o si presentan interrupciones sintácticas. Ese análisis resultaría útil para evaluar, dentro de los límites del material disponible, la compatibilidad de los hallazgos con **operaciones de edición de distinta naturaleza**, a saber: la selección y compilación de fragmentos que conservan información interpretable; la supresión de intervenciones de otros hablantes; y/o la unión de pasajes con el fin de alterar el sentido aparente del discurso.

Sobre el punto de pericia 11.

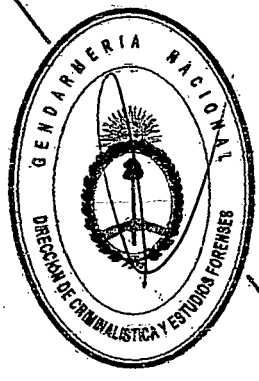
El **Ing. Nicolás Sanguineti**, perito de parte del MPF manifiesta que " en función de lo desarrollado por los peritos oficiales en el apartado 6 de este dictamen y lo informado en este punto de pericia, resulta un **error conceptual** exigir a un dispositivo aportado voluntariamente la misma trazabilidad que a un elemento secuestrado en el marco de un allanamiento donde el material probatorio debe ser recolectado y resguardado inmediatamente con el objetivo de impedir accesos indebidos por parte de los allanados, presumiblemente no interesados en la producción de esa prueba. En cambio, los aportes voluntarios de material probatorio realizados por Secretaría penal en un Juzgado o en el Ministerio Público no prevén un secuestro intempestivo del efecto, coexistiendo en estos procedimientos diversas alternativas en donde la continuidad de la cadena de custodia de esta prueba informática aportada es, por definición, trazable desde **el momento en que ingresó a la esfera de custodia del Estado. En este caso, desde el momento que fuera aportado a la Fiscalía Federal N° 5**; momento en el cual se confeccionó un formulario de cadena de custodia, el cual fue rubricado por los funcionarios intervinientes.


GENDARMERIA NACIONAL
DIRECCION DE CRIMINALISTICA Y ESTUDIOS FORENSES

Por otro lado, no se identificaron metadatos que indiquen manipulaciones posteriores al ingreso del dispositivo aportado a la órbita del proceso penal. Sin entrar en tecnicismos sobre el sistema de archivos del pendrive en cuestión, **manifiesta que no encuentra ningún elemento para poner en duda la cadena de custodia de esta prueba informática** una vez que fuera introducida en forma voluntaria al proceso penal.

Cecilia B Vera
Licenciada en Informática-UNLA
Perito Informático Forense
Perito Registro Profesional L394
COPITEC Registro Profesional Nº 4.476
CPCI Registro Profesional

Marcelo Antonio TOROK
Ingeniero en Informática
Perito Judicial y Policial


GENDARMERÍA NACIONAL
DIRECCIÓN DE CRIMINALÍSTICA Y ESTUDIOS FORENSES