



FACULTAD DE INGENIERÍAS Y CIENCIAS AGROPECUARIAS
ESCUELA DE SONIDO Y ACÚSTICA

CREACIÓN DE UN BANCO DE DATOS DE PERFILES ACÚSTICOS,
MEDIANTE ANÁLISIS DE VOZ Y DE CARACTERÍSTICAS FONÉTICAS DE
REGIONES DEL ECUADOR.

"Trabajo de titulación presentado en conformidad con los requisitos
establecidos para optar por el título de Sonido y Acústica."

Profesor Guía
Jorge Páez Rodríguez,
M.Sc. En Gestión y Evaluación de la
Contaminación Acústica

Autora
María Carolina Monroy Baroja

Año
2015

DECLARACIÓN DEL PROFESOR GUÍA

"Declaro haber dirigido este trabajo a través de reuniones periódicas con el estudiante, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación."

Jorge Páez Rodríguez,
Magister en Gestión y Evaluación de la
Contaminación Acústica
C.I.: 4452061600

DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE

"Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes."

Carolina Monroy Baroja

C.I.: 171189396-4

AGRADECIMIENTOS

Quiero agradecer a todas las personas que creyeron en mí siempre, mi Familia, a mis amigos, a mis perritos, a mis maestros, a mi profesor tutor Jorge Páez, Luis Bravo, Christiam Garzón, profesores en general y a la Universidad por haberme ayudado en todo.

También agradecer a la Policía Nacional, a la Unidad de Criminalística, por ayudarme y darme su confianza, a las personas que me colaboraron con sus voces y al magnífico país que tenemos por recibirme increíble en cada lugar del Ecuador.

DEDICATORIA

Esta tesis está dedicada a mis padres Eduardo y Patricia, a mis hermanas Gabriela y Fernanda, a mi tía Eulalia Buitrón y familia en general, por brindarme su ayuda, su amor, su incondicionalidad y su aliento día tras día, a mis perritos y animalitos por ser alegría y amor y a todos mis amigos por apoyarme, aconsejarme siempre y estar a mi lado.

RESUMEN

El siguiente trabajo se basa en la creación de una base de voces, se logró visitando 15 provincias del Ecuador para realizar las grabaciones de las voces, fue una experiencia bastante enriquecedora poder viajar y conocer un poco más de nuestro hermoso país; y a la vez poder contar con el apoyo total de la Policía Nacional alrededor de todo el Ecuador.

El propósito es crear una base de datos de perfiles acústicos, una vez editadas y analizadas las voces y se obtendrá ciertas características como frecuencia inferior y superior, pitch, intensidad, formantes y los espacios o pausas entre cada sílaba para cada voz de cada región del Ecuador, se realizará varias comparaciones para ver qué resultados nos entregan.

El banco de voces es un trabajo muy útil e importante para mí en lo personal como para la Policía Nacional, ya que tenemos varios registros de voces de todo el Ecuador, con su respectivo tono, volumen y su acento autóctono.

De esta manera se puede ampliar la base de datos de la Policía Nacional y ya se ha pensado ya en realizar proyectos más adelante con esta información, pero analizando diferentes parámetros.

El propósito de este proyecto es encontrar ciertas diferencias y semejanzas entre los hablantes por región, por provincia, por palabra, por género, etc, por medio de tablas y gráficos se podrá ver de una mejor manera los resultados.

ABSTRACT

The following work is based on the creation of a database of voices, this was achieved visiting 15 provinces of Ecuador for recording voices, was quite an enriching experience to travel and learn more of our beautiful country; and also to have the full support of the National Police around the Ecuador.

The purpose is to create a database of acoustic profiles, once edited and analyzed the voices and certain characteristics such as lower and upper frequency, pitch, intensity, formants and spaces or breaks between each syllable for each voice of each region is obtained in Ecuador, several comparisons will be made to see what results we deliver.

The Voice Bank is a very useful and important tool for me personally and for the national police work, since we have several records of voices from around the Ecuador, with its own pitch, volume and your native accent.

This way you can expand the database of the National Police and already thought about making projects in the future with this information, but analyzing different parameters.

The purpose of this project is to find some differences and similarities between speakers by region, by province, by word, by genre, etc., by means of tables and graphs can be seen in a better way the results.

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1 OBJETIVOS.....	2
1.1.1 Objetivo General.....	2
1.1.2 Objetivos Específicos.....	2
1.2 Antecedentes.....	3
1.3 Justificación.....	8
1.4 Alcance	8
2. MARCO TEÓRICO.....	9
2.1 Conceptos Generales.....	9
2.1.1. Acústica Forense	9
2.1.2. Peritaje Acústico	9
2.1.1.1 Fonética	9
2.1.1.2 Fonología	10
2.2 Audio Analógico.....	11
2.2.1 Transductor Electro acústico de entrada, Micrófonos:.....	11
2.2.2 Pre – amplificador:.....	16
2.3 Rango Dinámico	16
2.3.1 Ruido de Fondo	16
2.3.2 Headroom	16
2.3.3 Máximo nivel de distorsión.....	17
2.3.4 Relación señal – ruido	17
2.3.5 Nivel nominal	17
2.4 Audio Digita.....	17
2.4.1 Conversión de audio analógico a digital	17
2.4.2 Interfaces de audio	18
2.4.3 Edición y Procesamiento de Audio	18
2.5 Análisis de Audio	19
2.5.1 Frecuencia de Muestreo	19

2.5.2 Profundidad de Bits (Bit Depth)	19
2.5.3 Formatos de archivos de audio.....	20
2.5.4 Biometría de voz.....	21
2.6 PDS (Procesamiento digital de señales).....	21
2.6.1 Formantes.....	23
2.6.2 Dominio en el tiempo	24
2.6.3 Dominio en la Frecuencias	25
2.6.4 Datos enérgicos de la voz.....	25
2.6.5 Teorema de Fourier	25
2.6.6 Transformada de Fourier	26
2.7 Inteligibilidad de la Palabra	27
2.7.1 Enmascaramiento.....	28
2.7.2 Distorsión.....	29
2.7.3 Salas acústicas y recintos para toma de muestras de voz	30
2.8 Reconocimiento por medio de voz	31
2.8.1 Sistemas de reconocimiento de voz del locutor.....	32
2.8.2 Características de la voz del locutor	33
2.8.3 Naturaleza de la voz	34
2.8.3.1 Rango de Frecuencias de la voz.....	34
2.8.3.2 Clasificación de la voz	35
2.9 Herramientas para reconocimiento de voz.....	36
2.9.1 Analizadores de características de audio	36
2.9.2 Analizadores espectrales.....	37
2.9.3 Medidores de niveles	39
2.10 Características a analizar.....	39
2.10.1 Frecuencia (Hz)	39
2.10.2 Intensidad (dB).....	40
2.10.3 Pitch o tono (Hz)	40
3 METODOLOGÍA.....	42
3.1 Metodología por utilizar.....	42
3.2 Metodología utilizada por la Policial Nacional.....	43

3.3	Uso de las herramientas para reconocimiento de voz de la Policía Nacional.....	43
4	DESARROLLO	45
4.1	Toma de muestras	45
4.2	Banco de Voces.....	47
4.3	Análisis de características de voz.....	48
4.4	Comparación de las características acústicas.....	48
5	ANÁLISIS DE RESULTADOS.....	52
6	PROYECCIONES	106
7	CONCLUSIONES Y RECOMENDACIONES.....	107
7.1	Conclusiones	107
7.2	Recomendaciones.....	110
	REFERENCIAS	112
	ANEXOS	115

ÍNDICE DE FIGURAS

Figura 1: Tracto Vocal	11
Figura 2: Micrófono Omnidireccional	13
Figura 3: Micrófono Bidireccional	13
Figura 4: Micrófono Cardioide	14
Figura 5: Micrófono Supercardioides.....	15
Figura 6: Gráfico de formantes, Intensidad y Pitch con programa PRAAT.....	24
Figura 7: Rango de frecuencia de Voz	35
Figura 8: Analizadores FFT de curvas y de barras.....	37
Figura 9: Espectrograma	38
Figura 10: Analizador espectral por banda de tercio de octava.....	38
Figura 11: Medidor de nivel.....	39
Figura 12: Frecuencia baja y altas	39
Figura 13: Intensidad.....	40
Figura 14: Micrófono Azden SGM- 990	45
Figura 15: Grabadora digital Tascam DR-680.....	46
Figura 16: Toma de muestras	47
Figura 17: Análisis en el PRAAT	49
Figura 18: Mapa del Ecuador	50
Figura 19: Porcentajes de frecuencia inferior (Hz) de cada palabra por género.	55
Figura 20: Porcentajes de frecuencia inferior (Hz) de cada región por género.	56
Figura 21: Porcentajes de frecuencia inferior (Hz) de cada provincia por género.	57
Figura 22: Porcentajes de frecuencia superior (Hz) de cada palabra por género.	58
Figura 23: Porcentajes de frecuencia superior (Hz) de cada región por género.	59
Figura 24: Porcentajes de frecuencia superior (Hz) de cada provincia por género.	60

Figura 25: Porcentajes de intensidad mínima (dB) de cada palabra por género.	63
Figura 26: Porcentajes de intensidad mínima (dB) de cada región por género.	64
Figura 27: Porcentajes de intensidad mínima (dB) de cada provincia por género.	65
Figura 28: Porcentajes de intensidad máxima (dB) de cada palabra por género.	66
Figura 29: Porcentajes de intensidad máxima (dB) de cada región por género.	67
Figura 30: Porcentajes de intensidad máxima (dB) de cada provincia por género.	68
Figura 31: Porcentajes de intensidad (dB) de cada palabra por género.	69
Figura 32: Porcentajes de intensidad (dB) de cada región por género.	70
Figura 33: Porcentajes de intensidad (dB) de cada provincia por género.	71
Figura 34: Porcentajes de pitch mínimo (Hz) de cada palabra por género.	74
Figura 35: Porcentajes de pitch mínimo (Hz) de cada región por género.	75
Figura 36: Porcentajes de pitch mínimo (Hz) de cada provincia por género. ...	76
Figura 37: Porcentajes de pitch máximo (Hz) de cada palabra por género.	77
Figura 38: Porcentajes de pitch máximo (Hz) de cada región por género.	78
Figura 39: Porcentajes de pitch máximo (Hz) de cada provincia por género. ...	79
Figura 40: Porcentajes de pitch (Hz) de cada palabra por género.	80
Figura 41: Porcentajes de pitch (Hz) de cada región por género.	81
Figura 42: Porcentajes de pitch (Hz) de cada provincia por género.	82
Figura 43: Porcentajes de formante F1 de cada palabra por género.	86
Figura 44: Porcentajes de formante F1 de cada región por género.	87
Figura 45: Porcentajes de formante F1 de cada provincia por género.	88
Figura 46: Porcentajes de formante F2 de cada palabra por género.	89
Figura 47: Porcentajes de formante F2 de cada región por género.	90
Figura 48: Porcentajes de formante F2 de cada provincia por género.	91
Figura 49: Porcentajes de formante F3 de cada palabra por género.	92
Figura 50: Porcentajes de formante F3 de cada región por género.	93
Figura 51: Porcentajes de formante F3 de cada provincia por género.	94

Figura 52: Porcentajes de formante F4 de cada palabra por género.	95
Figura 53: Porcentajes de formante F4 de cada región por género.	96
Figura 54: Porcentajes de formante F4 de cada provincia por género.....	97

ÍNDICE DE TABLAS

Tabla 1: Muestras por región por género.	52
Tabla 2: Promedios de Frecuencia inferior y superior de cada palabra por región y género, en Hz.	53
Tabla 3: Promedios de Intensidad mínima, Intensidad máxima e Intensidad decada palabra por región y género, en dB.	61
Tabla 4: Promedios de Pitch mínimo, Pitch máxima y Pitch de cada palabra por región y género, en Hz.	72
Tabla 5: Promedio de las formantes de cada palabra por región y género.	83
Tabla 6: Promedio de las pausas por sílabas de cada palabra por región y género, en segundos.	98
Tabla 7: Anova entre la frecuencia mínima y máxima en función a la región.	100
Tabla 8: Anova entre la frecuencia mínima y máxima en función a la provincia.	100
Tabla 9: Anova entre la frecuencia mínima y máxima en función al género. .	101
Tabla 10: Anova entre la intensidad mínima, máxima e intensidad en función a la región.	101
Tabla 11: Anova entre la intensidad mínima, máxima e intensidad en función a la provincia.	102
Tabla 12: Anova entre la intensidad mínima, máxima e intensidad en función al género.	102
Tabla 13: Anova entre la pitch mínimo, máximo y pitch normal en función a la región.	103
Tabla 14: Anova entre la pitch mínimo, máximo y pitch normal en función a la provincia.	103
Tabla 15: Anova entre la pitch mínimo, máximo y pitch normal en función al género.	104
Tabla 16: Anova entre F1, F2, F3 Y F4 en función a la región.	104
Tabla 17: Anova entre F1, F2, F3 Y F4 en función a la provincia.	105
Tabla 18: Anova entre F1, F2, F3 Y F4 en función a la región.	105
Tabla 19: Frecuencias inferior y superior por región y palabra.	107
Tabla 20: Rangos de frecuencias con respecto a región y género.	107

Tabla 21: Intensidad mínima, máxima y normal por región y palabra.	108
Tabla 22: Rangos de intensidades con relación al género y región.	108
Tabla 23: Pitch normal, inferior y superior con relación al género y región....	109
Tabla 24: Rangos de pitch con relación al género y región.....	109

1. INTRODUCCIÓN

Este trabajo trata acerca de las características de voces del Ecuador para un perfil criminal, para lo cual se tomó muestras en distintas provincias del Ecuador. La creación de la base de datos, la obtención de características y distintos parámetros es lo que se va a obtener en este trabajo.

Se tuvo la oportunidad de trabajar con la Unidad de Criminalística de la Policía Nacional, ya que en el país no hay mucha gente especializada en lo que es acústica forense y me pareció una rama muy interesante para el proyecto.

La acústica forense es una subdivisión de criminalística que apoya en su parte logística a la labor policial, para poder esclarecer hechos analizando las evidencias para luego presentar en los tribunales.

Hoy día la Policía Nacional del Ecuador en la Unidad de Criminalística se puede encontrar varios equipos y personal capacitado para este trabajo de análisis y reconocimiento de voces. La criminalística es una ciencia natural y penal, que investiga, descubre y verifica la existencia de un delito o una conducta presuntamente delictuosa, aplicando sus conocimientos, tecnología y metodología para aportar con evidencias físicas o indicios de un presunto delito.

Para este proyecto se tomó muestras a grupos de 10 personas de cada provincia de diferentes sexos, con diferente tono, dialecto y acento.

Para realizar las grabaciones se utilizó una grabadora de voz Tascam DR-680, un micrófono Azden 990 y la colaboración de la gente de cada provincia.

Se realizará grabaciones de voces de las 3 regiones del Ecuador, que suman en total 15 provincias:

Costa: Guayas, Los Ríos, Manabí, Esmeraldas, El oro

Sierra: Loja, Imbabura, Pichincha, Tungurahua, Azuay, Carchi

Oriente: Sucumbíos, Napo, Orellana, Pastaza

Una vez grabadas todas las voces, se procedió a editar con el programa de edición de audio Cool Edit palabra por palabra y luego se utilizó el software PRAAT para obtener varias características con las cuales se van a trabajar que son: Frecuencia mínima y máxima, Intensidad normal, mínima y máxima, Pitch normal, mínimo y máximo, Formantes y las pausas que existen entre sílaba y sílaba de cada palabra.

1.1. Objetivos

1.1.1 Objetivo General

- Crear un banco de voces, en el cual al analizar sus características básicas se puede llegar a obtener datos fonéticos muy importantes para llegar a conocer un perfil acústico de un hablante dependiendo que región, provincia y género sea.

1.1.2 Objetivos Específicos

- Realizar procesamientos digitales de señales y estudios acerca de los diferentes registros de voz, para conocer el mayor número de características de la voz humana.
- Tomar muestras de voces de las 3 regiones del Ecuador, de la siguiente manera:

Costa: Guayas, Los Ríos, Manabí, Esmeraldas, El Oro

Sierra: Loja, Imbabura, Pichincha, Tungurahua, Azuay, Carchi

Oriente: Sucumbíos, Napo, Orellana, Pastaza

- Crear un banco con 10 registros de voz por provincia, con personas de mismas edades y con las mismas condiciones de grabación.
- Realizar un estudio comparativo con los datos obtenidos.

1.2 Antecedentes

La Acústica Forense es una rama muy compleja y poco conocida en nuestro país dentro de la rama de la Acústica, en la cual se investiga hechos delictivos como: violaciones, desapariciones, violencia intrafamiliar, secuestros y diferentes delitos, en los cuales existen pruebas como grabaciones de voces.

Hay que tomar en cuenta que el análisis que se realiza en estos casos debe contar con personal calificado y con los equipos necesarios para poder obtener la clase de información que estos casos requieren.

En ésta área podemos analizar los diferentes registros de voz con lo cual podemos llegar a conocer varias características de la persona como la edad, el sexo, estrato social, conducta, estado emocional, comportamiento, estado psicológico, región de donde proviene, etc.

Además tenemos registros no vocales como ruido de fondo o diversos sonidos que nos ayudan a ir descubriendo delitos y en caso que se haya utilizado un arma de fuego, podemos reconocer gracias al sonido del disparo, el ambiente acústico de la escena del delito y el tipo de arma.

En nuestro país a mediados del año 1999 se comenzó a aplicar la Acústica Forense para algunos casos, teniendo así mucho campo laboral aún que investigar acerca de la misma.

Para este trabajo de titulación se va a trabajar con la Policía Nacional, específicamente con la Unidad de Criminalística.

La Policía Nacional del Ecuador (2013) señala que:

- **Misión:**

“La Policía Nacional es una institución estatal de carácter civil, armada, técnica, jerarquizada, disciplinada, profesional y altamente especializada, cuya misión es atender la seguridad ciudadana y el orden público, y proteger el libre ejercicio de los derechos y la seguridad de las personas dentro del territorio nacional.”

- **Visión:**

“Policía Nacional Honesta respetuosa de los Derechos Humanos, Efectiva en el cumplimiento de la misión que brinda calidad de servicio a la ciudadanía.”

- **Historia:**

“Breve síntesis histórica de la unidad de criminalística del Ecuador”

La investigación en el Ecuador tiene su génesis en la Policía Nacional, es así que, en la Presidencia de Carlos Arroyo del Río, el 4 de marzo de 1941 se había aprobado el Reglamento de las Oficinas de Seguridad, Identificación y Dactiloscopia existentes en el país, dependientes del Ministerio de Gobierno y Policía; luego en septiembre de 1946, se crea el Servicio Médico Legal en el país, dependiente del Ministerio de Gobierno y adscrito a la Policía Civil Nacional.

A finales de la década de los 50 e inicio de los 60 el gobierno del Dr. Velasco Ibarra dispuso el retorno del Servicio de Investigación Criminal a manos de la Policía Civil Nacional; el 1 de diciembre de 1961 se hace entrega a la institución policial, modernos equipos, instrumentos y material específico para la instalación en Quito de un laboratorio de Investigación Criminal.

En noviembre de 1964 toma el nombre de “Laboratorio Central de Investigación Criminal” y, en febrero de 1975, la denominación de “Laboratorio Central y Peritajes” bajo la dependencia de la Dirección Nacional de Investigaciones, posteriormente la Dirección Nacional de Policía Judicial e Investigaciones.

El Dr. Félix Gonzalo Grijalva Herdoíza, bioquímico farmacéutico fue el primer jefe del Laboratorio de Investigación Criminal, que funcionó en la antigua Comandancia General, ubicado en las calles Cuenca y Mideros, incursionando en las áreas de toxicología, análisis de drogas, balística, documentología y dactiloscopia.

En 1982 se designó como jefe del Laboratorio Central y Peritajes al entonces Capitán de Policía Marco Cuesta, quién con el apoyo de la Dirección Nacional de Tránsito, creó una nueva área denominada: Centro de Investigación de Accidentes Viales.

En 1999 se inicia un proceso de reforma jurídica en el país, tendientes a la aplicación del sistema oral acusatorio, en 2001 se estructura e introduce el Reglamento de Policía Judicial en dónde se define la estructura de la Policía Judicial, dando origen a los Departamentos de Criminalística y Unidades de Apoyo Criminalístico, en concordancia con lo que establece el Reglamento de Policía Judicial publicado en el Registro Oficial No. 368, del 13 de Julio del 2001." (Policía Nacional del Ecuador, 2014)

Avances de la policía científica en el ecuador

Sin lugar a dudas, el avance de la ciencia y de la técnica Criminalística en nuestro país ha tenido un profundo impacto en muchas áreas especializadas de la Criminalística, pudiendo afirmar con meridiana claridad que hemos logrado pasar de la simple lupa al más complejo microscopio (Microscopio de Barrido Electrónico), con el propósito de captar hasta los más pequeños detalles, que en muchas ocasiones se convierten en la clave para orientar la investigación en un proceso judicial penal.

Sobre la base del avance y desarrollo de la tecnología, técnicas e instrumentos criminalísticos el Instituto Tecnológico Superior Policía Nacional desde 1995 ha venido formando a Oficiales, Clases y Policías en las disciplinas de la Criminalística, posicionándose como el único centro de educación superior de la institución policial, responsable de la gestión académica en los distintos ámbitos de la investigación, tanto de campo como de laboratorio.

En nuestros días prevalecen las pruebas aportadas por el hombre de ciencia a través del análisis de los indicios, tanto del estudio de las víctimas, así como los revelados en el lugar de los hechos, valiosa y fundamental fuente de información que permite la reconstrucción del ilícito y la identificación de su autor o autores.

Las modernas técnicas criminalísticas están basadas en logros científicos y son, en su totalidad, valiosos auxiliares del método criminalística, el cual se ha constituido en una estrategia de la investigación.

Significativos han sido los avances de Policía Científica en el Ecuador, sobre todo en los últimos siete años, pues a nuestros laboratorios se les ha entregado tecnología de punta, hecho que posibilita colocarlos en una posición privilegiada en la Región.

La instalación e implementación del Sistema Automatizado de Huellas Dactilares - AFIS a nivel nacional, con un sistema Francés de última generación con 1000 dpis de resolución mayor a cualquier equipo en Latinoamérica, con la ventaja de haber sido implementando en todas las provincias del Ecuador con excepción de Galápagos, tenemos más de 250.000 registros de detenidos, además de 800 registros de cadáveres NN y más de 100 registros biométricos de personas desaparecidas lo que nos ha permitido identificar, más de 100 cadáveres NN en los últimos siete meses. Se han tenido más de 800 hits positivos en la investigación de delitos tales como: robos a domicilios y locales comerciales, así como aquellos relacionados con muertes violentas.

El Sistema Automatizado de Identificación Balística IBIS cuya base de datos y equipos son los más modernos de la región, con tecnología Canadiense, somos el único país en Latinoamérica que se encuentra integrado a la Red Internacional IBIN en la cual se encuentran 14 países la mayoría de Europa y Norteamérica.

Sistema informático Automatizado de Identificación de voces y Rostros AVIS, con tecnología rusa en la comparación de bases de datos, de rostros y voces inició con un registro de todos los sentenciados con más de 10000 registros iniciales se está incrementando las bases de datos y los resultados de identificación de personas y voces serán importantes ya que se ha registrado personas detenidas para investigaciones en todo el país contando con estaciones remotas en todas las provincias.

Contamos con el microscopio de Barrido Electrónico más moderno de Sudamérica cuyos resultados son importantes en la identificación de quien disparo un arma de fuego además de muchos otros usos, habiéndose capacitado a nuestros técnicos en varios países principalmente Bélgica y Holanda, en este sistema se han realizado estudios en muchos casos de connotación nacional y además casos en los cuales se investiga a Policías y militares evidenciando una forma técnica y transparente de trabajo basada en protocolos internacionales.

Contamos con el laboratorio de ADN MITOCONDRIAL más moderno del país el cual está siendo empleado en la extracción de perfiles genéticos de cadáveres NN para ser comparados con familiares de personas desaparecidas.

Se ha implementado el laboratorio de Toxicología que cuenta con equipos con tecnología de punta en el estudio de venenos, tóxicos y sustancias de uso no permitido tanto en casos penales como en casos administrativos para realizar pruebas confirmatorias en el análisis de muestras de servidores policiales para el centro de confianza de la Inspectoría General.

1.3 Justificación

Como sabemos cada persona tiene un registro de voz único que lo podemos comparar con la huella digital, entonces es muy importante tener muestras de voces de un grupo de la población del Ecuador en una base de voces para luego poder compararlas y analizar características básicas como frecuencia, intensidad, pitch, formantes y pausas de las palabras.

Se realiza un estudio a todas las muestras tomadas para luego proceder a realizar gráficos comparativos y así tener una idea más clara de los resultados y conclusiones que deseamos obtener.

1.4 Alcance

Este trabajo se realiza para la Policía Nacional del Ecuador, específicamente al área de Criminalística, para facilitar a ésta Institución a identificar características de la voz humana como frecuencia, intensidad, pitch, formantes y pausas de las palabras y compararlas entre ellas.

Como se conoce en el Ecuador existen 4 regiones, las cuales están conformadas por provincias y en las que se habla con diferente fonética, encontrando distintos tonos, acentos, timbres, fonemas, etc., lo cual hace única a la voz de cada persona.

Por eso es importante que se realice esta base de datos para tener datos relevantes acerca de las diferentes voces del Ecuador, y como éste proyecto pues a un posteriori realizar varios otros proyectos con esta información.

2. MARCO TEÓRICO

2.1. Conceptos Generales

2.1.1. Acústica Forense:

La acústica es una rama de la física que se dedica al estudio de del sonido, ultrasonido e infrasonido en forma de ondas mecánicas que se propagan por distintos medios como: gaseoso, sólido o líquido.

Tiene una aplicación práctica científica dentro de un proceso legal, en el cual se realiza investigaciones especializadas, basadas en muestras de voces las cuales son analizadas detenidamente en un laboratorio, para luego elaborar el peritaje correspondiente.

2.1.2. Peritaje Acústico

Tiene que ver con el esclarecimiento de un hecho legal gracias a la Acústica Forense y a una persona llamada perito forense en este caso, el cual tiene el conocimiento científico, técnico y práctico para dar una información a los tribunales de justicia bajo juramento, con lo cual se convierten en personas muy confiables al momento de validar un hecho.

2.1.1.1. Fonética

Se entiende como fonética al estudio realizado a los sonidos físicos y acústicos producidos por el ser humano, viene del griego foné que quiere decir “voz o sonido”.

2.1.1.2. Fonología

Se entiende como fonología al estudio de una imagen mental o abstracta que el ser humano entiende acerca del lenguaje oral.

Existe la terminología “fonos” que son percepciones reales del sonido, mientras que los “fonemas” son percepciones mentales del sonido, y estos se clasifican en Vocales y Consonantes.

Al emitir vocales no se presenta ningún tipo de obstáculo en el aparato fonador, mientras que si se emite consonantes éstas presentan un obstáculo a lo largo del aparato fonador.

El aparato fonador está compuesto por todos los órganos que participan en la articulación del lenguaje del ser humano, estos son:

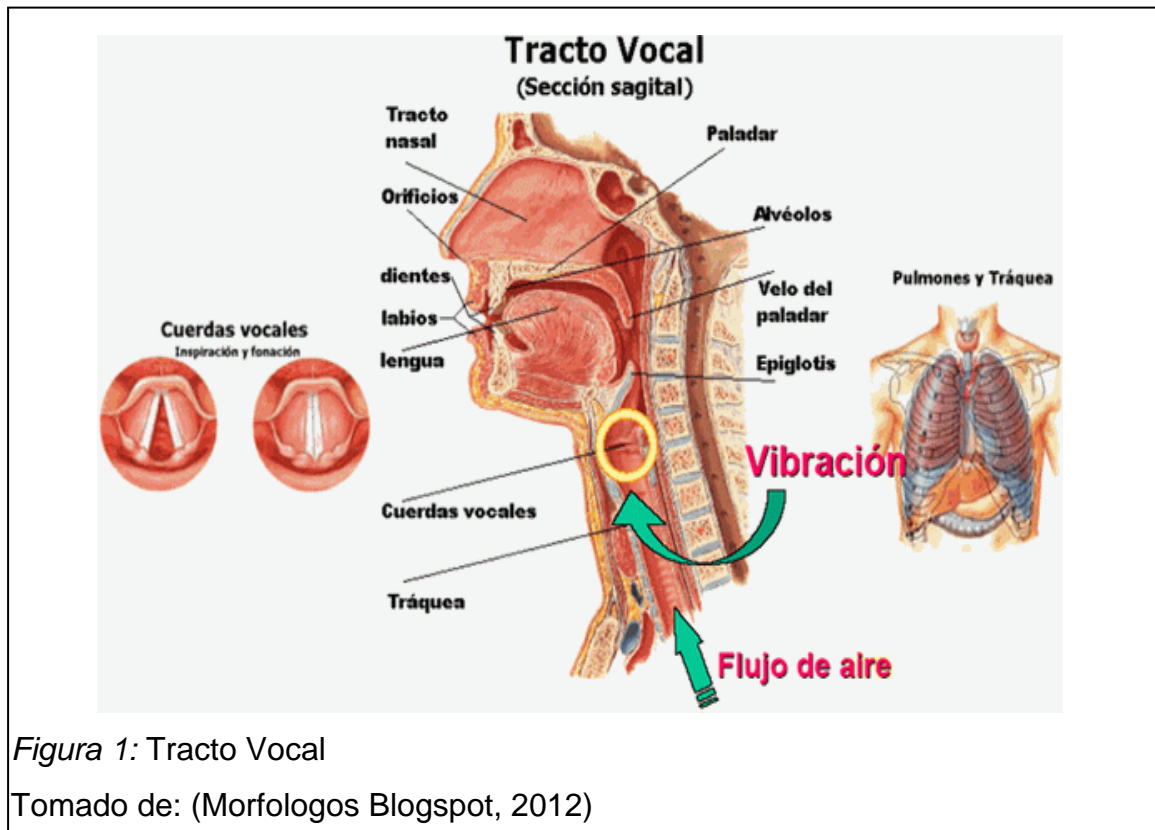
Laringe: Es un órgano tubular donde se produce la voz, de estructura músculo – cartilaginosa, construida por cartílagos en formas de semilunas. Es el órgano que está situado arriba de la tráquea, la cual se comunica con la faringe.

Cavidad Bucal: O conocida como Boca, es la cavidad donde se realizan las funciones como el lenguaje, se articulan las palabras y se direcciona el sonido. Está compuesta por labios, dientes, paladar, lengua, etc.

Cuerdas Vocales: Son aquellas encargadas de producir la voz, son dos pliegues bucales o labios membranosos, no tienen forma de cuerda, y producen vibraciones que a su vez crean los sonidos.

Cavidad Nasal: Sirve como un resonador acústico que se encarga de limpiar el aire que viene del exterior y es aquí donde se generan las letras M, N y Ñ.

Labios: Son aquellos que articulan las palabras, ya que dejan pasar o limitan el flujo de aire que sale de los pulmones hacia afuera. Los labios son el límite exterior del aparato fonador.



2.2. Audio Analógico

Es aquel que se almacena, procesa y reproduce con ayuda de circuitos electrónicos y otros dispositivos de carácter analógico.

Por ejemplo: el cassette, la cinta magnética bobina o el disco de vinilo.

2.2.1. Transductor Electro acústico de entrada, Micrófonos

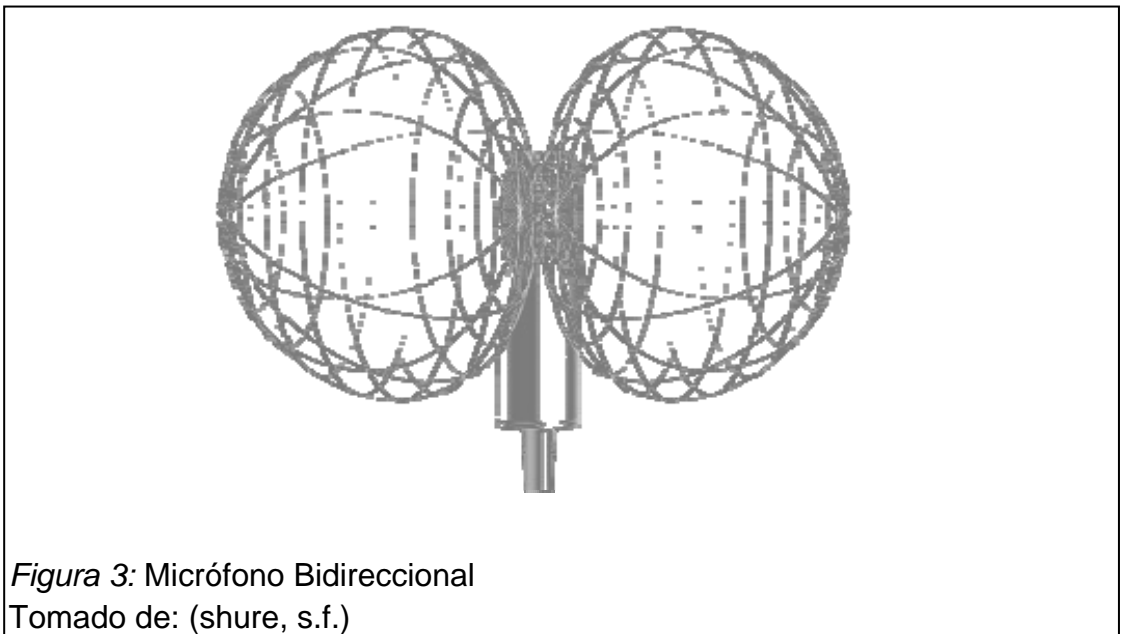
El micrófono es un transductor electro acústico y el primer elemento de la cadena de audio, capta las ondas sonoras transformando potencia acústica en potencia eléctrica, con similares características ondulatorias y con la posibilidad de grabar cualquier elemento, lugar o situación.

Los micrófonos funcionan gracias a dos tipos de transductores. El primero es un transductor mecano acústico llamado diafragma, el cual transforma las variaciones de presión en vibraciones mecánicas y el segundo es un transductor electromecánico el cual transforma vibraciones mecánicas en eléctricas. Entre los dos se transductores se forma un electroacústica.

- **Clasificación por principio de funcionamiento:** Se clasifican en:
 - **Micrófonos Dinámicos de Bobina Móvil:** Estos micrófonos tienen una bobina con espiras de alambre de cobre, el cual se oscila a lo largo del núcleo cilíndrico del imán. El diafragma vibra con variaciones de presión sonora de una onda, impulsando a la bobina dentro de la ranura del imán, generando corriente eléctrica.
 - **Micrófonos de Condensador o Capacitivos:** Estos micrófonos utilizan un campo eléctrico, funcionan por estática, tienen un diafragma muy delgado bañado en oro y una placa posterior metálica la cual suele estar perforada.
Estos micrófonos utilizan una energía extra conocida como Phantom Power o fuente fantasma, la cual suministra 48 voltios para que funcione el micrófono.
- **Clasificación por patrón polar o directividad:** Según la forma en la cual captan el sonido, los micrófonos se clasifican en:
 - **Omnidireccional:** Capta sonido independiente de todas las direcciones, 360°. La desventaja es que también capta sonidos no deseados. Responden mejor a bajas y medias frecuencias 0 Hz a 2kHz).



- **Bidireccional:** Se los conoce también como forma 8, captan el sonido que llega tanto a la parte delantera como trasera, o sea un ángulo de 90° , no captan sonidos por las partes laterales. Captan un 33% de ruido ambiental.

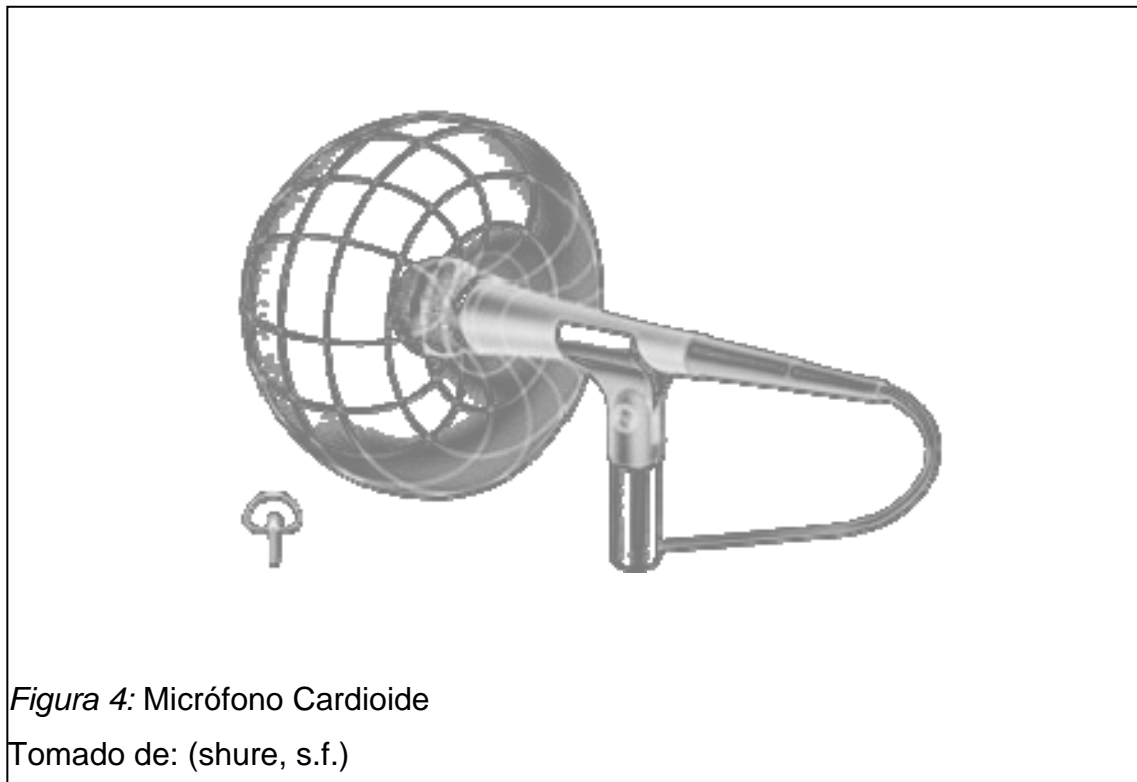


- **Unidireccional:** Captan sonido solamente de una sola dirección, son más direccionales en frecuencias altas que en bajas, ya que la direccionalidad depende de la frecuencia.

Se clasifican en:

- **Cardioides:** Tiene un patrón polar con forma de corazón, teniendo mayor sensibilidad por la parte frontal y mínima en la parte posterior, con un rechazo de 25 dB. Sirve para mejor captación de sonidos agudos en lugares con mucho ruido, tiene un ángulo de captación entre 131° y un ángulo de rechazo máximo de 180° .

El problema de estos micrófonos es que pueden producir retroalimentación por su parte posterior, por lo cual no es muy utilizado en TV, ya que puede captar 33% de ruido ambiental, mientras son buenos micrófonos para captar un sonido cerca de la fuente, en caso del cantante, como micrófonos de mano.



- **Supercardioides:** También captan sonido por la parte frontal comportándose mas direccionales que los cardioides, teniendo un ángulo de captación de 115° y un ángulo máximo de rechazo de 126° . Un problema de estos micrófonos es que al ser más direccional, en su parte posterior se crea un pequeño lóbulo trasero de captación, el cual tiene un rechazo de 12 dB y una captación ambiental de 27%.

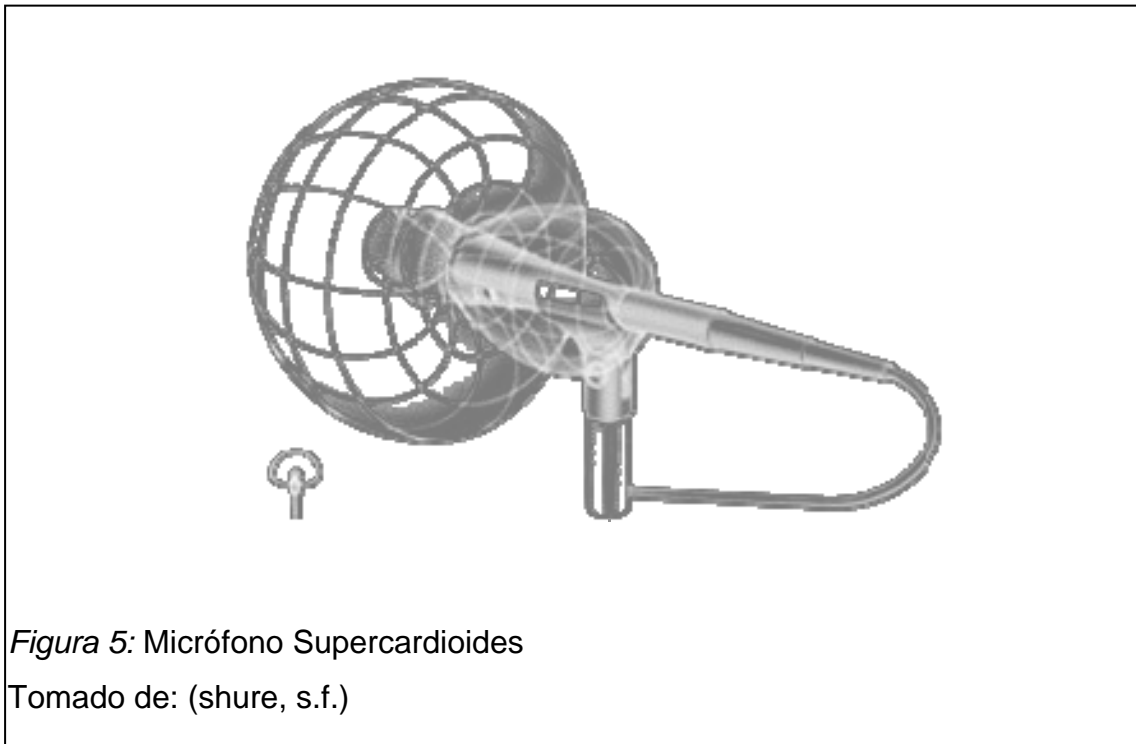


Figura 5: Micrófono Supercardioides

Tomado de: (shure, s.f.)

- **Hipercardioides:** Es el más direccional de los 3, teniendo un ángulo de captación de 105° y un ángulo máximo de rechazo de 110° . Tiene menos captación de manera lateral, pero igual tiene un lóbulo trasero de captación mucho más grande que el de los supercardioides, tienen un rechazo de 6 dB de rechazo por su parte trasera y 25% de captación ambiental.

2.2.2. Pre – amplificador

Es un dispositivo muy importante dentro de la cadena electroacústica, ya que aumenta o amplifica el nivel de la señal actuando sobre la señal de entrada.

Al salir las señales del preamplificador alcanzan los 0 dB, o sea el nivel de línea. Nivelan la tensión eléctrica que llega a las distintas fuentes de audio, para luego de igualarlas, enviarlas al otro equipo como señal de entrada.

La ganancia es la relación entre la salida y entrada, se la expresa en decibelios o decibeles. Hay pre-amplificadores que también regulan tono, balance, refuerzan la sonoridad con volumen bajo, etc.

2.3. Rango Dinámico

Es la relación entre un valor máximo de la señal donde ésta se distorsiona o se recorta y un valor mínimo donde aparece el ruido de fondo del sistema.

Dentro de esto tenemos varios términos que debemos tenerlos claros:

2.3.1. Ruido de Fondo

Es un sonido indeseado que se superpone a la señal útil, ya sea en una medición acústica o al grabar muestras de voces como en este caso, pueden llegar a alterar los resultados del propósito que queremos llegar.

2.3.2. Headroom:

Es la diferencia de nivel entre el nivel máximo (nivel de saturación) y el nivel nominal.

2.3.3. Máximo nivel de distorsión:

Es el máximo nivel que tiene una señal, antes de que llegue a la saturación.

2.3.4. Relación señal – ruido

Es la diferencia entre el nivel nominal y el ruido de fondo.

2.3.5. Nivel nominal

Es un punto óptimo de trabajo que se encuentra dentro del valor máximo y del mínimo.

2.4. Audio Digital

Se trata de codificar una señal eléctrica que representa una onda sonora, almacenándolas en código binario representados de “ceros” y “unos”.

Estas secuencias de valores enteros las obtenemos en el muestreo y la cuantificación digital de la señal eléctrica.

2.4.1. Conversión de audio analógico a digital

Consiste en tomar valores discretos de la amplitud de una señal analógica, en intervalos pequeños de tiempo, ya que mientras más tomas se realicen en un segundo será más confiable la reproducción.

Se le asigna un número binario a cada muestra, ya que el audio digital se trabaja en equipos modernos y computadoras, las cuales tienen como unidad el bit. El bit puede tener solo dos valores “1” o “0”.

Se puede trabajar palabras de 8 0 16 bits para tener una mayor fidelidad, obteniendo así 256 ó 65536 combinaciones distintas con mayor resolución.

2.4.2. Interfaces de audio

Es un dispositivo hardware que transforma la señal análoga en digital, por medio de convertidores A/D y D/A, para pasar así de señal analógica a código binario, se conectan al ordenador y maneja las entradas, las salidas y el procesamiento de las señales.

Existen de dos tipos: las tarjetas de sonido y las interfaces o cajas de hardware externas, las cuales se conectan por medio USB o FireWire. Algunas cuentan con varias entradas para poder grabar más de una señal al mismo tiempo ya sea línea o micrófono y también salidas para conectarlas a sistemas surround o mesa de mezclas.

También tienen puertos MIDI los cuales nos permiten conectar sintetizadores y otros teclados externos, y tienen conexiones S/PDIF. Es muy importante trabajar con interfaces de buena calidad.

Siempre debemos tener en cuenta la frecuencia de muestreo que es la cantidad de muestras tomadas en 1 segundo y la resolución de bit nos da una representación más precisa de la longitud de onda.

2.4.3. Edición y Procesamiento de Audio

La edición de audio consiste en manipular e integrar distintos elementos sonoros, ya sean voces, efectos, pistas musicales o sonidos con diferentes características.

El procesamiento de audio tiene algunas aplicaciones tecnológicas, como por ejemplo en este caso puntual necesitaremos síntesis o reconocimiento del

habla, reproducción en alta resolución, también se puede alterar ciertos parámetros para crear un nuevo sonido de una pista ya existente, variar su contenido armónico, ancho de banda, dinámica, entre otros parámetros.

2.5. Análisis de Audio

Se realiza un análisis espectral ya sea en la pista completa o en parte de ella, se puede ver en gráficos.

2.5.1. Frecuencia de Muestreo

Es el número de muestras que pueden ser tomadas en intervalos de tiempos regulares, para convertirlas de señal continua a señal discreta, para convertir señal de análoga a digital.

Para que el sonido sea más parecido al original, debemos tener una cantidad mayor de frecuencias de muestreo, y que así el sonido en digital será de mayor calidad.

Con el teorema de Nyquist & Shannon podemos deducir que para reproducir una onda exacta, la frecuencia de muestreo debe ser mayor al doble de la máxima frecuencia a ser muestreada. La frecuencia de muestreo más usada es 44.1 Hz.

2.5.2. Profundidad de Bits (Bit Depth)

Es la resolución a la cual se captura la señal de audio relacionado a la amplitud, determinan de qué tamaño serán cada una de las muestras de audio. Podemos trabajar a las siguientes profundidades: 8, 16, 20, 24, 32 y 64 bits.

Determina el rango dinámico en el cual se maneja la señal, ya sean valores máximos o mínimos en dB, entre mayor sea el tamaño de la muestra tendremos una mayor cantidad de información de cada señal.

2.5.3. Formatos de archivos de audio

se tiene varios formatos de audio, cada cual tiene su extensión específica según el archivo y dependen de la comprensión de audio.

Entre los más importantes tenemos: WAV, MP3, MIDI Y OGG, pero de ahí tenemos varios como: Aac, Mp4, Atrac, Ram, Dss, Dvf, Msv, IVS, entre otras varias. Se explicará más detalladamente los formatos más importantes:

- **Wav (Wave Form Audio File):** Se trabaja con Microsoft, tiene excelente calidad de audio, se lo utiliza en el CD, son archivos que tienen pesos grandes entre 20 y 30 Mb, usa la compresión PCM o ADPCM.
- **MP3 (MPEG 1 Layer 3) :** Tiene una gran calidad y un grado muy bueno de compresión, se utiliza para descargar y almacenar música, usa compresión PCM, se usa para videos de internet, en los reproductores, presentan una mínima pérdida de señal, no es tan buena para grabar voces.
- **MIDI (Musical Instrumental Digital Interface):** Es un protocolo de comunicación serial que permite generar sonidos, al compartir su información con dispositivos como computadoras, sintetizadores, controladores, secuenciadores o cualquier dispositivo musical electrónico.

Lo utilizan algunos reproductores de audio, tienen un peso reducido, se utilizan igual en páginas como HTML o composiciones instrumentales.

- **OGG** : Es un archivo abierto y libre compatible con varios códecs de audio, se clasifica en Vorbis y Opus, tienen alta calidad, creados por la fundación Xiph.Org, tiene la misma compresión que el mp3, pero su calidad de reproducción es mejor, puede servir para audio y video, no tiene restricciones con patentes de software.

2.5.4. Biometría de voz

Es un sistema que nos permite validar o identificar las voces de las personas con una certeza absoluta, obviamente se trabaja con sistemas y programas de última tecnología.

Para este trabajo se toma en cuenta los rasgos físicos de la voz, como también las características de la voz como frecuencia, tono, acento y velocidad.

Al unir estos dos aspectos se crea una huella de voz única para cada persona.

Se puede ver que cada día la biometría de la voz nos ayuda en temas de seguridad, como por ejemplo las voces en las claves de tarjetas de crédito, ya que ahora la voz es la clave, en empresas para contabilizar la asistencia en puestos de trabajo, seguridad en sitios web, y obviamente como en este tema de tesis en el reconocimiento de voz para perfiles criminales.

2.6. PDS (Procesamiento digital de señales)

Es la representación, transformación y análisis de la información de las señales digitales de audio, discretizadas en tiempo y amplitud.

Se tienen ventajas y desventajas, entre estas:

- **Ventajas:**
- Las señales al ser digitalizadas pueden ser recuperadas más fácilmente que las analógicas en caso que exista distorsión, ya que las señales analógicas tendrán una forma de onda inaceptable y dañada mientras que se puede recuperar información siendo digitalizado.
- Es mucho más fácil manipular la información, tiene un almacenamiento seguro y es mucho más económico hacerlo con sistemas digitales, ya que con sistemas analógicos se invertía más tiempo, dinero e incluso eran imposibles realizar ciertas tareas.
- Tiene una sensibilidad baja a condiciones ambientales, variación, tolerancia y envejecimiento de sus componentes, lo que hace posible fabricar sistemas con respuesta idénticas y que no varíen con el tiempo.
- Se tiene un rango de error realmente bajo, lo cual permite que tenga una señal de alta calidad y fidelidad, con detección y corrección de errores, proceso que no es compatible con sistemas analógicos.
- Se puede trabajar con varias señales a la vez, ya que los hardware de los sistemas digitales tienen mayor flexibilidad que los analógicos.
- El procesamiento digital de señales realiza tareas muy complejas, trabaja con la información a tiempo real, procesamiento de imagen, multimedia, etc.
- Los sistemas digitales permiten la implementación de algoritmos más sofisticados, mediante operaciones aritméticas o lógicas.

- **Desventajas:**

- Los sistemas digitales requieren dispositivos previos y posteriores al procesamiento como convertidores A/D y D/A, filtros, circuitos, etc.
- Poseen un rango limitado para frecuencias disponibles para el procesamiento.
- Al convertir una señal analógica en digital, se puede producir una distorsión que impide que se pueda reconstruir exactamente la señal analógica original, partiendo de las muestras cuantificadas.
- La sincronización en los sistemas analógico es mucho más fácil que en los sistemas digitales ya que éstos necesitan obtener un grado muy alto de sincronización.
- Se pierde gran cantidad de información por muestreo.
- Los sistemas digitales utilizan grandes anchos de banda, cuando la relación señal – ruido llega a cierto nivel, la calidad decae, en los sistemas analógicos existe una degradación más natural de la señal.
- El rango dinámico de la frecuencia es más limitado.

2.6.1. Formantes

Se las conoce como la frecuencia de resonancia del aparato fonador y tienen una gran cantidad de la información lingüística del habla.

Son características específicas por bandas de frecuencias donde se concentra la mayor parte de energía sonora de un sonido. Tienen 3 parámetros importantes: ancho de banda, centro de frecuencia y energía.

Cuando se modifica la forma del tracto vocal, también se modifican estos parámetros en diferente medida, lo que ocasiona que el filtro y el sonido final cambien.

La laringe produce armónicos que serán reforzados o atenuados por las resonancias o formantes, así los armónicos más cercanos a los valores formánticos se amplificaran más que los alejados.

En el gráfico se puede ver que las formantes son los puntitos rojos que como su nombre lo dice forman una onda.

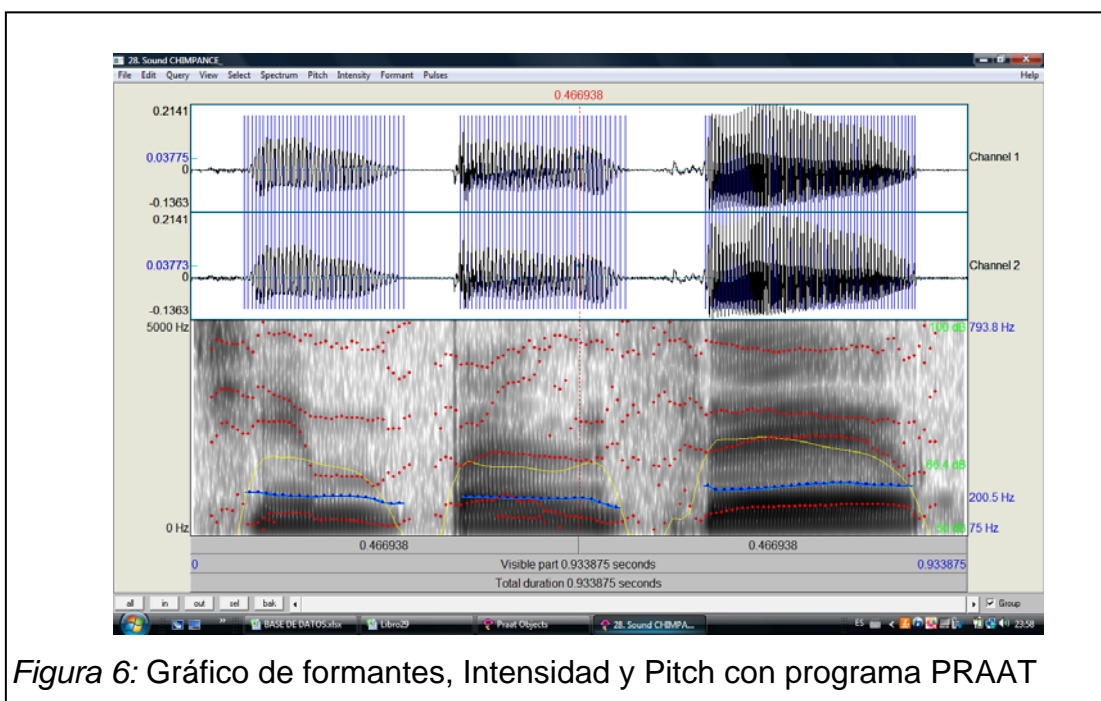


Figura 6: Gráfico de formantes, Intensidad y Pitch con programa PRAAT

2.6.2. Dominio en el tiempo

Es el análisis de la amplitud de una señal representada en el dominio del tiempo en el que fue muestreada.

Existe el dominio en el tiempo discreto y continuo. En el dominio discreto el valor de la señal son puntos discretos del eje temporal y en el continuo el valor de la señal es para todos los números reales.

2.6.3. Dominio en la Frecuencias

Es el análisis de las señales respecto a su frecuencia, mostrando los componentes de una señal dependiendo de la frecuencia en la que oscilan dentro de un rango dinámico, y captura las características espectrales de la

señal de audio. En una forma de onda existen muchas frecuencias además de la frecuencia fundamental

Los armónicos son los componentes de frecuencias individuales del espectro, éstos son enteros simples de la frecuencia fundamental.

2.6.4. Datos energéticos de la voz

Se trata de los anchos de banda donde se sitúa la mayor parte de la densidad de energía de la voz de los hablantes, en los gráficos se puede ver que son las partes más negras o acentuadas.

2.6.5. Teorema de Fourier

También llamada serie de Fourier, es una herramienta matemática que analiza funciones periódicas, descomponiendo esta función en una suma infinita de funciones sinusoidales simples. Es decir que se puede construir una onda periódica compleja si se suma varias ondas simples sinusoidales que tengan amplitud, fases y frecuencia adecuadas.

Mientras más hondas sinusoidales tengamos, la onda se vuelve más cuadrada. Las series de Fourier tienen la forma:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \frac{2n\pi}{T} t + b_n \sin \frac{2n\pi}{T} t \right]$$

Donde a_n y b_n se denominan coeficientes de Fourier de la serie de Fourier de la función $f(x)$.

La serie de Fourier se utiliza mucho en la rama de Ingeniería y Matemáticas, en este caso de se lo usa para acústica y en sistemas de telecomunicaciones.

2.6.6. Transformada de Fourier

Transforma señales entre el dominio del tiempo y el dominio de la frecuencia, tiene también la transformada inversa de Fourier que transforma de cualquiera de los dos dominios al otro, son utilizadas en ingeniería y física.

Son útiles para estudiar señales periódicas, pero las más frecuentes son las no periódicas, para lo cual existe una teoría matemática más ambiciosa con valores complejos y definida en la recta, con una función g , de la cual creó la siguiente fórmula:

$$g(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx$$

Donde f es L^1 , es decir, f tiene que ser una función integrable en el sentido de la integral de Lebesgue. El factor que acompaña a la integral facilita el enunciado de los teoremas referentes a la transformada de Fourier. Esta forma de normalizar la transformada de Fourier es la más común, mas no es universal. En la práctica las variables x y ξ suelen estar asociadas adimensiones como el tiempo (segundos) y frecuencia (Hz) respectivamente, si se utiliza la fórmula alternativa:

$$g(\xi) = \sqrt{\frac{\beta}{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\beta\xi x} dx$$

La constante β cancela las dimensiones asociadas a las variables sosteniendo un exponente adimensional.

La transformada de Fourier goza de una serie de propiedades de continuidad que garantizan que puede extenderse a espacios de funciones mayores e incluso a espacios de funciones generalizadas.

2.7. Inteligibilidad de la Palabra

Es un criterio muy importante en la vida cotidiana. La palabra es el medio de transmisión de información ya se trate de medio de comunicación o medios de difusión.

No sirve de nada que por ejemplo, que una película posea una alta calidad artística y técnicas generales, si al momento de su exhibición se produce una pérdida en la inteligibilidad de los diálogos. El observador se va a esforzar por comprender la obra en vano, pasando así al aburrimiento, irritación, agotamiento y rechazo de la obra.

Existen varios elementos en la cadena electroacústica de audio que puede afectar la inteligibilidad. Además algunos de estos aspectos técnicos son de gran importancia como: la potencia de la emisión de la voz, el nivel de ruido ambiente (que puede enmascarar las voces) y las características acústicas de los locales.

Un gran porcentaje de los problemas de inteligibilidad en los medios de difusión se debe a deficiencias en las fuentes sonoras, al tener un incorrecto dominio de la voz pueden darse las siguientes deficiencias:

- La incorrecta producción de las vocales al no pronunciarlas con el énfasis necesario, generando una proyección débil de la voz y el ruido la puede enmascarar.
- Demasiada unión de palabras.

- Al articular incorrectamente las consonantes, lo cual ocasiona la inteligibilidad de la palabra en general y deficiencia en la misma que ninguna tecnología la puede salvar.
- El descenso de la intensidad al finalizar frases, etc.

Estas deformaciones puede que no tengan gran transcendencia en una conversación en la cual también funcionan mecanismos psicoacústicos de la audición para ayudar a la comprensión. Estas pérdidas de información en las cadenas electroacústicas, también ocasionan un descenso en la inteligibilidad de la palabra para el espectador.

La lengua que hablan en diversos pueblos o regiones de un país, degenera la lengua ya sea por factores físicos o socioculturales. En la representación dramática, la forma en que se escriben los libretos son muy importantes ya que unos son para ser leídos y no para ser actuados y viceversa. No se trata de que los actores griten deformando la intención dramática del dialogo o que lo hagan de una forma ajena a la idiosincrasia del personaje, se trata de optimizar cada vinculo con la inteligibilidad, para así lograr transmisión de estados de ánimo e ideas a través de una elemento natural de la banda sonora que es la voz humana.

2.7.1. Enmascaramiento

Es un fenómeno que es producido cuando el oído está expuesto a escuchar dos o más sonidos al mismo tiempo, el un sonido se interpone al otro, es decir lo enmascara. Por ejemplo: Una conversación en la calle con ruido y además una bocina de un automóvil, obviamente la voz humana no se va a entender o se va a enmascarar ya sea por el ruido de tráfico y otras personas, y más aun si al momento de la conversación el automóvil toca su bocina.

El enmascaramiento de un sonido depende mucho de la intensidad y el espectro del sonido en más carante. Un ruido que cubra el espectro de

frecuencia de la voz y este 10 dB por encima de la misma, impide totalmente entender un diálogo. Si el ruido en cambio está 10 dB por debajo, existe una cierta molestia pero se entienden algo las palabras.

En ocasiones hay ruidos fuera del espectro de la voz, entonces se escucharan los dos sonidos sin que interfiera el uno con el otro.

Se puede dar en pistas de bandas sonoras, ya que se debe tener cuidado al colocar varios elementos sonoros a actuar en el mismo tiempo, ya que obviamente se enmascaran los que tienen espectros en frecuencias similares, dificultando la mezcla y el producto final. En caso que existan arreglos orquestales mal hechos, en los que existen elementos con diferentes funciones que entran en la misma zona del espectro.

En cine es una lucha diaria ya que hay que evitar que los diálogos se enmascaren con el sonido ambiente, los sonidos que tengan espectros similares a la voz humana tenderán a enmascarar las voces.

Cuando se tiene varias fuentes sonoras, puede ser que la señal deseada se enmascare, aquí entra un papel muy importante que es la audición humana o efecto psicoacústico, ya que subconscientemente se concentra sobre la fuente sonora deseada y filtra los sonidos no deseados.

2.7.2. Distorsión

La distorsión afecta a una señal de audio en fase, frecuencia y amplitud. Hay varias clases de distorsión, las más importantes son:

- **Distorsión armónica:** Se da cuando una señal de entrada se ve afectada en su timbre en la etapa de salida, entre las distorsiones armónicas más conocidas tenemos:

- **Distorsión por saturación:** Se presenta cuando se sobrealimenta el rango dinámico de un equipo, recortando los picos y valles de la señal original, esta distorsión cambia el timbre de la señal.
- **Distorsión total armónica (THD):** Se presenta por la circuitería de un sistema, afectando a la forma de onda resultante, ya que introduce frecuencias armónicas, que son frecuencias múltiplos de la señal de entrada. Quiere decir que el THD es un porcentaje de la relación que existe entre la energía de las frecuencias armónicas no deseables con respecto a la de la señal original.
Existen otros tipos de distorsión como la distorsión por fase y la que es producida por micrófonos de condensador que es generada cuando las placas se ponen en contacto, produciendo recorte de forma triangular en los picos y valles de la onda.
- **Distorsión por cruce:** Se produce por la no linealidad que presentan los dispositivos electrónicos, al no encontrarse polarizados, afectando directamente a la curva de transferencia la cual altera la relación de los voltajes de entrada con los voltajes de salida de los distintos dispositivos.

2.7.3. Salas acústicas y recintos para toma de muestras de voz

Los fenómenos acústicos en los locales son tan antiguos como el sonido. El hombre desde las cavernas tuvo que arreglárselas con los misteriosos ecos y persistencia de los sonidos en sus recintos.

La explicación científica de la reverberación y de la reflexión data del siglo pasado, es ahí donde el hombre aprende empíricamente a aplicar ciertas reglas para evitar las influencias de locales con características acústicas nocivas a determinadas informaciones sonoras.

Una sala de locución, deben tener las siguientes características:

- Se debe tratar de aislar los ruidos y vibraciones provenientes de exteriores, utilizando paredes, tumbados, ventanas y puertas preparadas para el efecto.
- Se debe determinar el aislamiento y acondicionamiento acústico de las salas de locución, sabiendo elegir materiales adecuados para recubrir la parte interior y con sus respectivas dimensiones para obtener un sonido de buena calidad.
- En fin, se debe tener una buena acústica evitará algo no deseado como las reverberaciones, reflexiones, ruidos y resonancias del sonido.

2.8. Reconocimiento por medio de voz:

En si el reconocimiento de voz se trata como su nombre lo dice, de reconocer señales en este caso voces.

Existen dos formas para reconocer al locutor:

El locutor debe pronunciar una frase que es dependiente en el sistema, mientras que la otra es donde el locutor debe pronunciar una frase cualquiera que sea independiente del sistema, obviamente la implementación de esta última es más compleja y tiene un menor porcentaje de exactitud al momento del reconocimiento.

Para el reconocimiento de voces se utiliza características físicas y de comportamiento de la voz de cada individuo para crear modelos biométricos que se utilizarán para tareas posteriores de comparación y reconocimiento de voces.

En ocasiones hay factores internos de cada individuo como: estado emocional, edad, características físicas, patologías del aparato fonador, cantidad de palabras habladas, y factores externos también como: ruido ambiental, ruido eléctrico, microfonía, reverberación, decodificación, distorsión, etc. y estos pueden afectar los resultados obtenidos con este método de análisis.

2.8.1. Sistemas de reconocimiento de voz del locutor

Luego de obtener modelos matemáticos que representan características distintas señales de voz, se determina si una o varias muestras de voz tienen algún tipo de semejanza o pertenecen a un modelo de voz previamente analizado.

A este conjunto de procesos se los llama sistema de reconocimiento automático de locutor y se pueden encontrar estos dos tipos de análisis:

- **Identificación:** En estos sistemas se tiene como finalidad reconocer si la voz de una persona se encuentra dentro de un conjunto X de locutores que han sido previamente modelados y registrados dentro del sistema. A este tipo de identificación se lo llama CONJUNTO CERRADO. En la salida de estos sistemas se pueden hallar X posibles resultados que correspondan a cada locutor.

Pueden existir casos en que la voz analizada no se encuentre dentro del conjunto de locutores, en este caso se denomina identificación del CONJUNTO ABIERTO, ya que la voz no pertenece al conjunto de X locutores.

- **Verificación:** En este sistema solo se tienen dos entradas: la voz a verificar y una voz de evidencia.

En estos sistemas podemos obtener dos posibles salidas, aceptación o rechazo de las voces, esto quiere decir si pertenecen o no a la misma persona, dependiendo si superan o no un determinado umbral de similitud.

Para casos judiciales se utiliza análisis de similitud en forma estadística, se denomina Likelihood Ratios o relaciones de verosimilitud, también utilizados en otros tipos de análisis forense como para investigar ADN.

2.8.2. Características de la voz del locutor

Para reconocer la identidad de un locutor es necesario analizar las características socio-psicológicas o llamadas de alto nivel y las características fisiológicas o también llamadas de bajo nivel.

La identidad del locutor se la puede distinguir con las características de alto nivel, las cuales se crean a partir del entorno o estado de ánimo en el que se desenvuelve la persona analizada. Algunas de las características pueden ser: el ritmo, el acento, las pausas, la duración de la palabras, la semántica, el tipo de personalidad, influencia parental, las pronunciaciones, el estatus socio económico, la modulación, el nivel de educación y la entonación.

En cambio la parte anatómica del locutor se puede distinguir gracias a las características de bajo nivel, poniendo un interés especial en el análisis de los órganos de fonación con los que se guardan relación con las características físicas como: sexo, edad, peso, altura, etc.

Se utilizan dos tipos de características en el análisis de señales de voz:

- Características acústicas: Cuantifican las cualidades físicas de un determinado sonido, estos valores son obtenidos desde la representación digital de la voz.

Este parámetro acústico se lo puede clasificar en parámetros cuasi periódicos y parámetros de perturbación:

- Parámetros cuasi periódico: Constan de la Frecuencia fundamental, formante, envolvente Espectral.

- **Parámetros de perturbación:** Es la relación armónico-ruido, Jitter y Shimmer.
- **Características de representación:** Representan el comportamiento dinámico de una señal, mas no alguna cualidad física de la voz. Son conjuntos de valores calculados a partir de una representación grafica de la voz.
Algunos de los métodos más usados actualmente para la extracción de características de representación son: L.P.C. (Código de predicción lineal) o M.F.C.C. (Coeficientes de Frecuencia MelCesptrum).

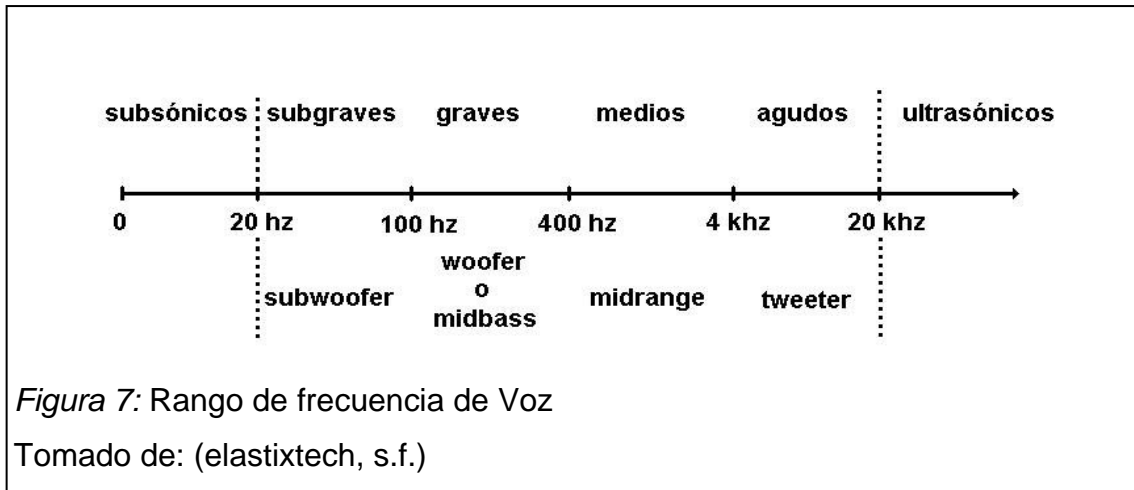
2.8.3. Naturaleza de la voz

La voz humana es un tipo de señal que fluctúa en el transcurso del tiempo, formando así una onda estacionaria.

Si analizamos la onda en periodos cortos de tiempo en milisegundos, en esos segmentos existen características cuasi estacionarias, también hay en segmentos ruidosos que se los denomina sonidos sonoros o sordos.

2.8.3.1. Rango de Frecuencias de la voz

Existen varias frecuencias entre las cuales se pueden generar ondas mecánicas longitudinales, las cuales se reducen a los límites de frecuencia que pueden estimular el oído humano para que el cerebro las perciba como una onda acústica. Los rango de audible es de 20 Hz a 20 KHz, mientras que el rango de la voz más o menos va de 250 y 3.000 Hz, algunos fonemas se encuentran situados entre los 4.000 y los 8.000 Hz.



2.8.3.2. Clasificación de la voz

La voz se clasifica en masculina y femenina.

- **Voz Femenina:** Dentro de las voces femeninas tenemos :
 - **Soprano:** Es la voz más aguda y se divide en 3 tipos:
 - Ligera: Posee un timbre muy fuerte, tiene mucha movilidad, es muy brillante, es poco potente.
 - Dramático: Se trata de la voz más grave dentro de los sopranos, teniendo una mayor potencia, volumen y amplitud.
 - Lírica: Es similar a la ligera, es muy expresiva y más ligera, pero a la vez tiene su parte dramática.
 - **Mezzosopranos:** Es una voz intermedia entre el soprano y el contralto, el timbre es cálido, es muy expresiva y a la vez puede ser ligera.
 - **Contralto:** Es la voz más grave de las mujeres, tiene un timbre grave, su agilidad es media, es una voz escasa.

- **Voz Masculina:** Dentro de las voces masculinas tenemos:
- **Contratenor:** Es la voz más aguda en hombres, no hay que confundirla con los falsetes. El tono de voz del contratenor recuerda a los Castrati, es un hombre con voz de soprano infantil, así ya haya pasado la pubertad.
Su voz se usa para ópera, e igual cantan en iglesias ya que en cierta época se prohibió cantar a mujeres.
- **Tenores:** Se clasifican en:
 - **Lírico:** Posee una gran agilidad, tiene un timbre flexible, seductor y brillante. Tiene registro cálido y suave.
 - **Dramático:** Es un poco más grave que el anterior, tiene el timbre más potente y una voz muy enérgica
- **Barítonos:** Es la voz intermedia entre el tenor y el bajo. Cuenta con una agilidad media y un timbre opaco.
- **Bajo:** Es la voz menos ágil y la más grave.

2.9. Herramientas para reconocimiento de voz

A continuación se va a presentar algunas herramientas para reconocer voces:

2.9.1. Analizadores de características de audio

Todo archivo de audio tiene determinadas características, como por ejemplo: nombre, bit rate, el tamaño del archivo, su duración, fecha de creación, extensión del archivo, frecuencia de muestreo entre otras. Esto podemos encontrar en la parte que no indica las propiedades del archivo.

2.9.2. Analizadores espectrales

Como se conoce al audio se lo representa como voltaje versus tiempo, para conocer cómo se comporta la onda en el dominio versus la frecuencia se utiliza un analizador espectral, el cual nos ayuda con una pantalla en la cual se puede observar las ondas y sus características.

Estos analizadores espectrales trabajan con la transformada de Fourier, los que se basan en algo ritmos matemáticos descomponiendo la señal en sus armónicos.

Hay algunos tipos de analizadores espectrales:

- **Analizadores FFT:** También llamado analizador de señal dinámica, este dispositivo analiza la amplitud y la frecuencia, para lo cual utiliza anchos de banda paralelos y en tiempos instantáneos. En frecuencias inferiores de 500 Hz se utiliza interpolación para tratar de completar, mientras que en frecuencias altas se realiza un análisis de hasta centésimo de octava.

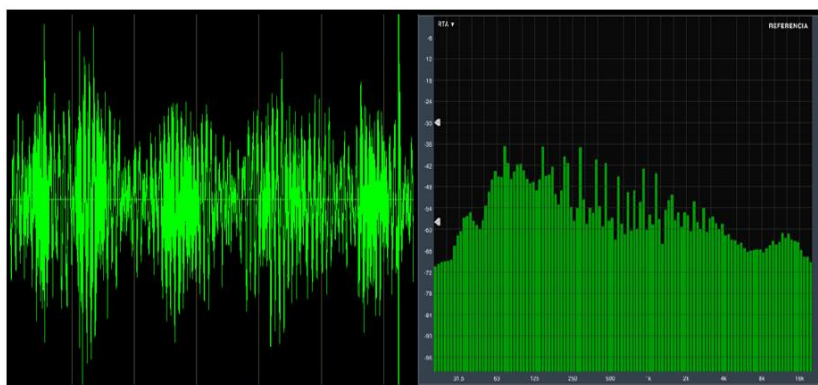


Figura 8: Analizadores FFT de curvas y de barras.
Tomado de: (Pepeferrersonido, 2011)

- **Espectrogramas:** Con este gráfico se puede ver la densidad del espectro de una señal, trabajando en 3D. Se puede ver el tiempo, frecuencia y amplitud. Trabaja bastante con una codificación de colores fuertes para sus gráficos.

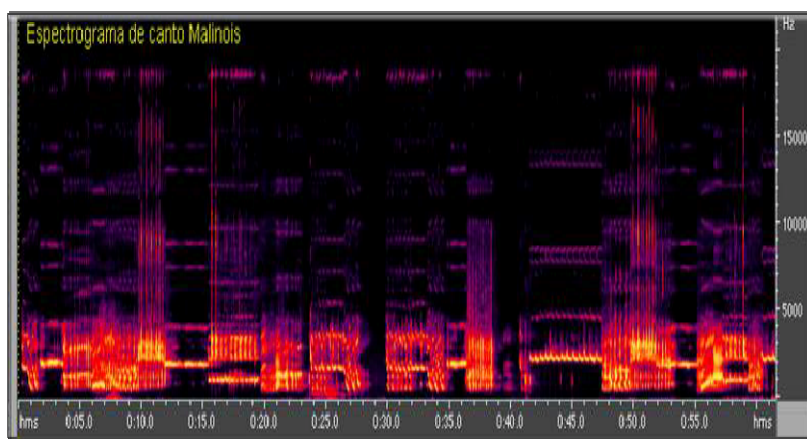


Figura 9: Espectrograma
Tomado de: (cced, s.f.)

- **Analizadores por banda de octava:** Es similar al analizador FFT, se tiene mejor análisis para las frecuencias bajas ya que se lo hace de manera logarítmica, y no se utiliza interpolación para gráfica lineal. Estos analizadores utilizan anchos de banda variables dependiendo de una frecuencia central, ya que a menor frecuencia central, el ancho de banda será menor.

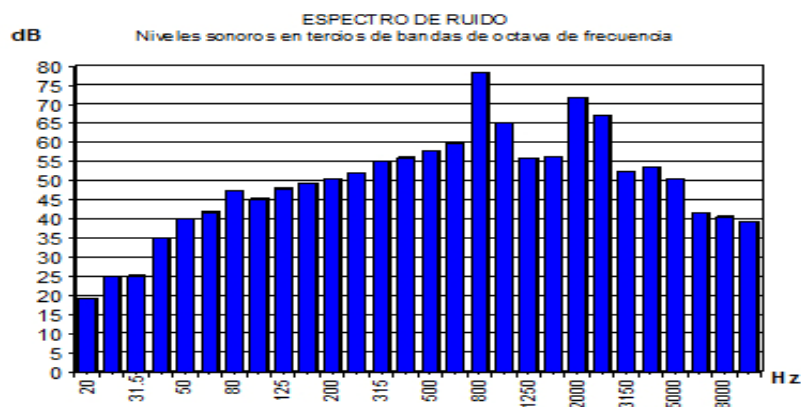


Figura 10: Analizador espectral por banda de tercio de octava.

Tomado de: (bizkaia, s.f.)

2.9.3. Medidores de niveles

También existen los VU meters o LED meters, miden el nivel de amplitud eléctrica de una señal en función al tiempo, éste puede ser peak o RMS, teniendo un tiempo de respuesta configurable en slow o fast.

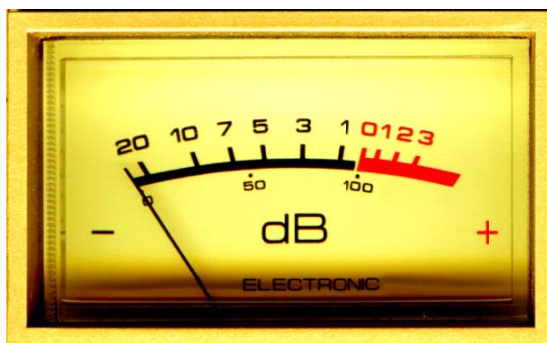


Figura 11: Medidor de nivel
Tomado de: (escuela evolutiva, s.f.)

2.10. Características a analizar

A continuación se conocerá la definición de las características que se necesitarán para analizar:

2.10.1. Frecuencia (Hz)

Es la cantidad de repeticiones que tiene una onda en un tiempo definido.

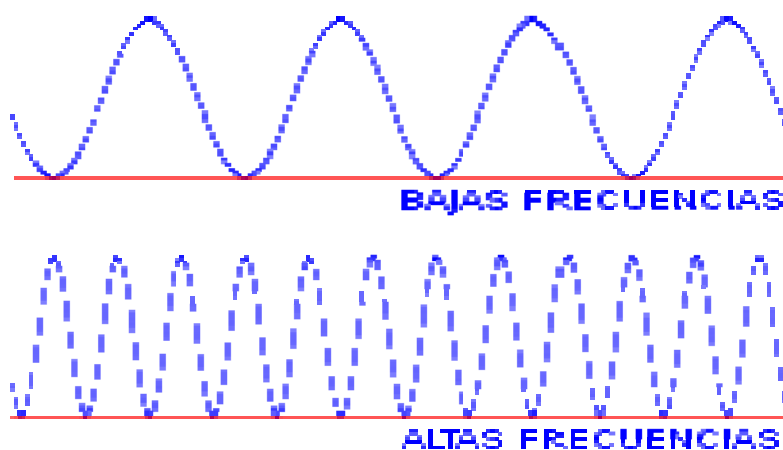
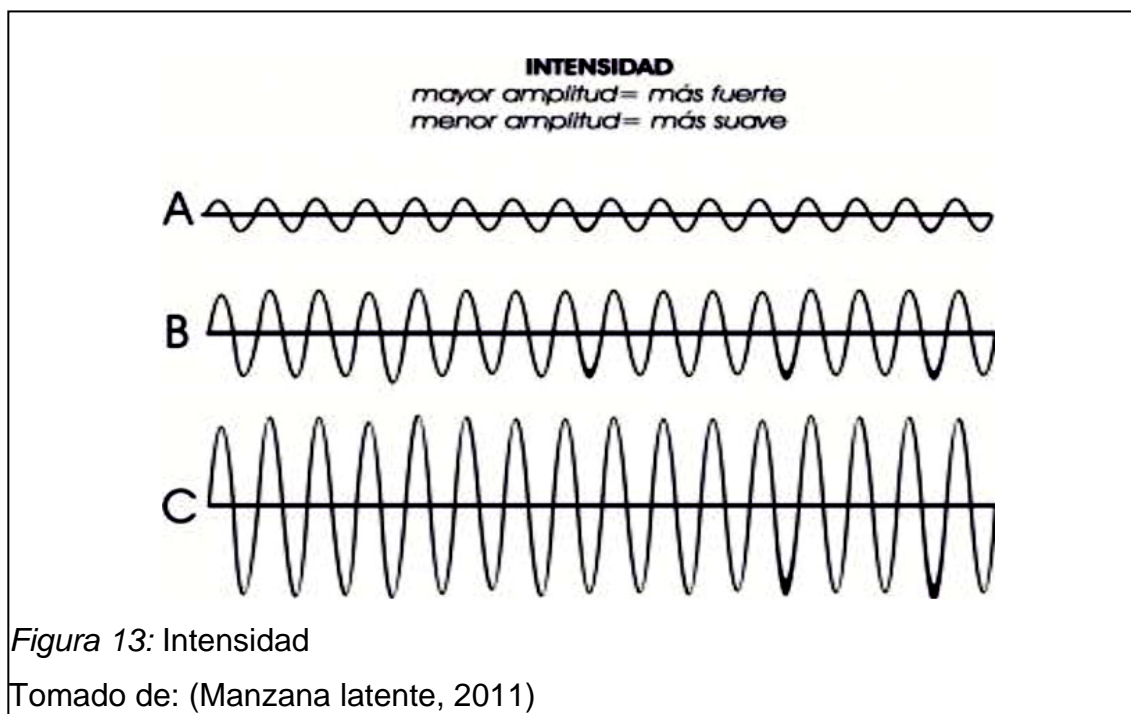


Figura 12: Frecuencia baja y altas

2.10.2. Intensidad (dB)

Es la fuerza o el volumen con el cual se pronuncia una palabra, teniendo sonidos fuertes y suaves, y expresando estados de ánimos.

Con la intensidad de la voz se puede expresar estos estados de ánimo: el miedo, la agresividad, el cansancio, la debilidad, la felicidad, la enfermedad, el nerviosismo, la tristeza, etc.; como también puede describir tamaños y distancias, y con tonos agudos o graves, se refuerza la lejanía y la proximidad.



2.10.3. Pitch o tono (Hz)

Es una característica fundamental en la comunicación, se puede definir como la impresión que nos causa una frecuencia de vibración a la que una onda determinada se manifiesta. Se tiene tono agudo y grave, que dependen de la cantidad de movimiento producidas por las cuerdas vocales al emitir una palabra.

Mientras el número de vibraciones sea mayor más aguda será la voz, o sea tendrá el tono más alto, y mientras existan menor número de vibraciones tendrá un sonido grave y un tono más bajo. El tono nos da mucha información sobre las emociones de una persona y lo que quiere expresar.

El tono tiene un papel muy importante al construir ambientes sonoros y escenarios, se dice que los tonos agudos tienen que ver con colores claros,

con luminosidad, con brillos, etc., mientras que los tonos graves tienen que ver con colores oscuros.

Como se mencionó anteriormente se pueden describir efectos de proximidad y de lejanía. Se dice que los tonos agudos están ligados con representar sentimientos como alegría, sorpresa, optimismo pero también nerviosismo, miedo y tensión, y los tonos graves nos representan las sensaciones de tristeza, depresión, pesimismo, melancolía, etc.

3. METODOLOGÍA

3.1. Metodología por utilizar:

En este proyecto se utilizará una metodología descriptiva, ya que vamos a llegar a conocer ciertas características de la voz humana analizando las muestras de voces tomadas.

Para realizar este trabajo hasta el momento vamos a seguir los siguientes pasos:

- Crear un banco de datos fonéticos de 15 diferentes provincias del Ecuador.
- Editar las palabras a ser analizadas con el programa de edición de audio Cool Edit.
- Utilizando el programa de análisis fonético PRAAT, se analiza las voces según provincia, región, género, palabra.
- Una vez analizadas las voces y teniendo ya valores de diferentes características como: intensidad, pitch, formantes, frecuencias, etc., se procede a hacer estudios estadísticos.
- Se utiliza un programa estadístico llamado SPSS, el cual nos ayuda a realizar tablas comparativas de las palabras con respecto a las formantes, intensidad, pitch, frecuencias superior e inferior, y pausas entre sílabas.
- Se realiza gráficos comparativos para obtener una mejor visualización de los promedios obtenidos.

3.2. Metodología utilizada por la Policial Nacional

La metodología que se utiliza en el área técnico-científica de la Policía Nacional del Ecuador, obedece a un proceso en dónde el examinador analiza las grabaciones de audio, a fin de identificar los acontecimientos y a los hablantes o locutores.

El análisis de voz con fines criminalísticas y/o forenses, se enfoca en un procedimiento en el cual, una voz desconocida es cotejada con una voz conocida.

Son los métodos semiautomático y automático, los utilizados en identificación de voces, considerados también como objetivos, dentro de la escala comparativa en la identificación por voz / tarea de reconocer la identidad de la persona de entre un grupo de muestras de voz, mediante la comparación de los parámetros deducidos de estas muestras.

1 - N = 1 contra la base de datos

1 – 1 = 1 contra un conocido.

Para el propósito legal, la intervención de un examinador (perito) es crucial puesto que deduce del análisis acústico las características de la frecuencia del tono principal, a través de la segmentación de la señal de voz, y comparación de los formantes (geometría de órganos articulatorios), a través de la identificación, posición y dinámica de la conducta de formantes.

3.3. Uso de las herramientas para reconocimiento de voz de la Policía Nacional

Es un trabajo bastantes grato, poder trabajar con la Unidad de Criminalística de la Policía Nacional, como es obvio ellos manejan un tipo de información bastante reservada, por lo cual no pueden entregar mayor detalle de las herramientas y equipos, sin embargo me permito compartir los datos generales:

- DSP Sona Graph 5000 de Key Elemetrics
- IKAR LAB de Speech Techonology Center
- Así mismo, en el sistema AVIS+F, utilizamos varias herramientas.
- Voice Grid.
- SIS II.
- Sound Cleaner.
- Import Utility.
- Multi Passt Port.
- Examiner.

4. DESARROLLO

A continuación se va a explicar cómo se realizó el proyecto detenidamente:

4.1. Toma de muestras

La toma de muestras se realizó en 15 provincias del Ecuador, a las cuales se viajó para poder obtener estas grabaciones:

Costa: Guayas, Los Ríos, Manabí, Esmeraldas, El oro

Sierra: Loja, Imbabura, Pichincha, Tungurahua, Azuay, Carchi

Oriente: Sucumbíos, Napo, Orellana, Pastaza

Se grabó con una grabadora Tascam DR- 680 y un micrófono Azden SGM-990, con una profundidad de bits de 24 bits y una frecuencia de muestreo de 44.100 Hz, los hablantes tuvieron que leer ciertas palabras, el ambiente ideal era una sala acondicionada especialmente para estas grabaciones sin ruido, pero al realizarlas se tuvo varios inconvenientes con el lugar donde se grabaron las diferentes voces, ya que no en todos los lugares había silencio total, las instalaciones no fueron lo suficientemente adecuadas para las grabaciones porque existía ruido.

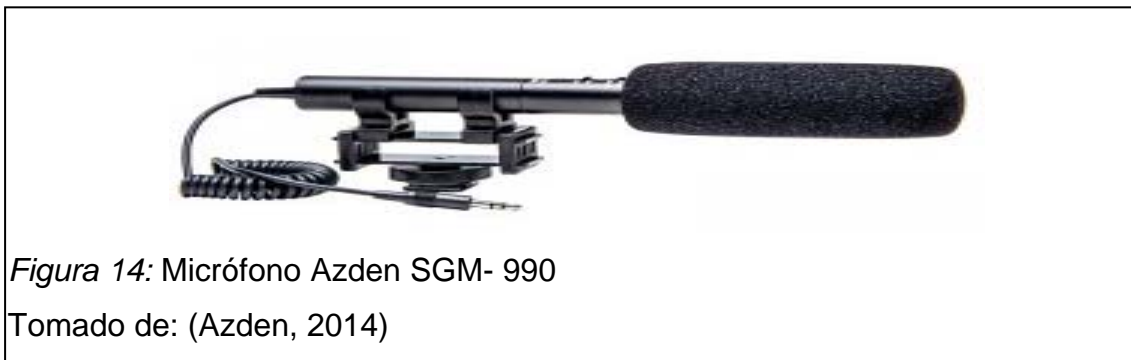




Figura 15: Grabadora digital Tascam DR-680

Tomado de: (Tascam, s.f.)

Las palabras que los hablantes grabaron fueron:

PALABRAS	
Ahora	Encebollado
Arcoíris	Ferrocarril
Chimpancé	Fútbol
Cocodrilos	Jaula
Drenaje	Loro

Se usó estas palabras ya que para reconocer las voces de presuntos delincuentes, en la Policía Nacional les hacen leer palabras de este tipo, que sean palabras conocidas, fáciles de leer y pronunciar y que se sepa que tienen un acento especial dependiendo de su pronunciación en cada provincia. Además lo más característico de cada provincia es como pronuncian consonantes como la r, rr, ll, s, c.

Tuvo mucho que ver su forma de pronunciar y leer las palabras, ya que no todas las personas leían exactamente a la primera la palabra real, por lo cual se realizaron 2 muestras de cada persona.

Fue complicado conseguir la gente en las diferentes provincias, ya que no querían colaborar ya sea por vergüenza a leer o porque no sabían leer bien, pero al final se logró el objetivo.

Además como se indicó anteriormente, el lugar influyó mucho en el resultado final.



Figura 16: Toma de muestras

4.2. Banco de Voces

Una vez grabada las voces, se procedió a editar, cortando palabra por palabra y luego realizamos el banco de voces.

El banco de voces o base de datos la realizamos palabra por palabra por cada hablante.

4.3. Análisis de características de voz

Se pudo determinar las siguientes características:

- Frecuencia Mínima y Máxima
- Intensidad mínima, máxima y promedio
- Pitch mínimo, máximo y promedio
- Las 4 formantes: F1, F2, F3 Y F4
- Espacios temporales o pausas de las palabras

Las características que vamos a obtener son las que el programa Praat nos brinda, ya que al ser un programa gratuito tiene algunas restricciones, y bajo estas características vamos a trabajar.

La definición de estas palabras se encuentran en el marco teórico.

4.4. Comparación de las características acústicas

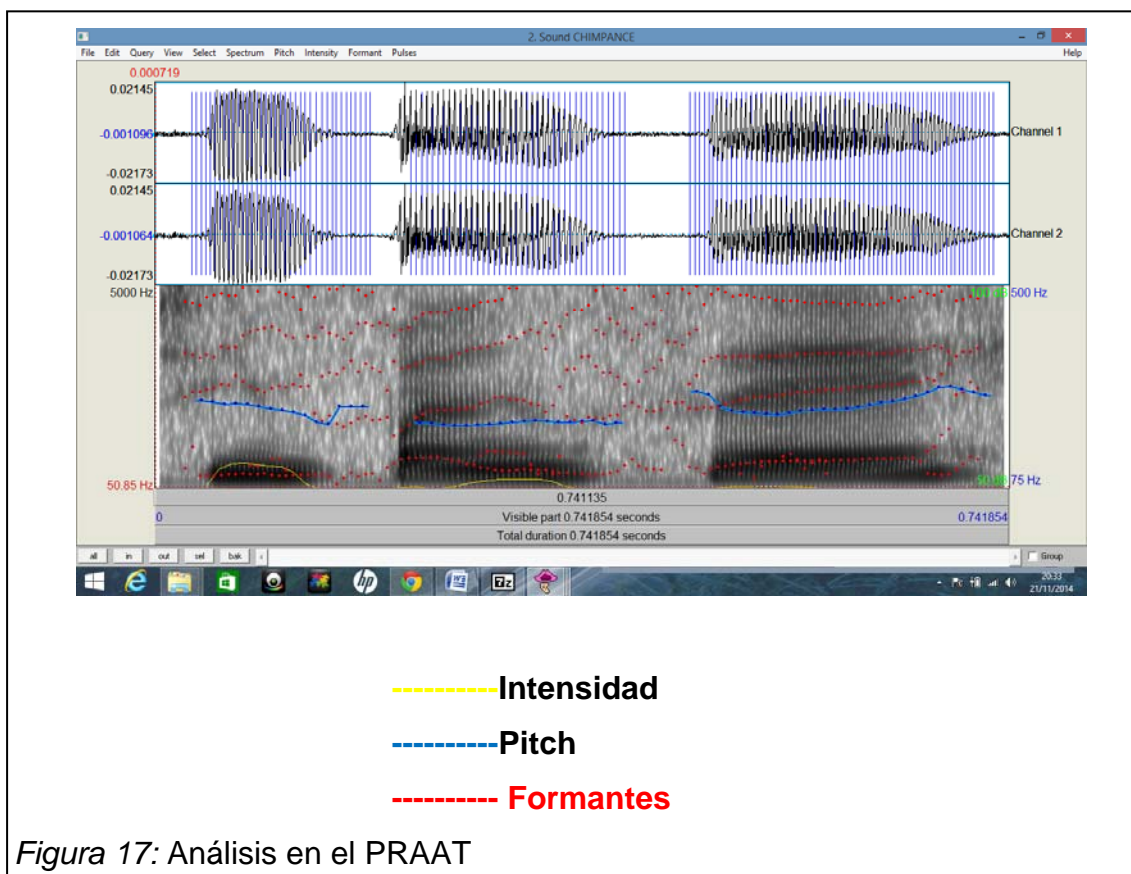
Se va a utilizar un software gratuito llamado Praat, el cual analiza científicamente cada palabra, obteniendo las características mencionadas anteriormente.

En el programa Praat se puede grabar voces con cualquier tipo de archivo de audio y ver los espectrogramas. Obviamente como es una versión gratuita tiene ciertas restricciones al momento de analizar las voces, y obviamente se lo puede automatizar para utilizarlo en investigaciones de alto nivel para calcular jitter, shimmer, etc.

Para realizar la comparación de características de voces, primero se procede a abrir el software Praat, nos vamos a "Open (Abrir)", damos clic en "Readfrom file...", ahí se va abriendo voz por voz y palabra por palabra para ser analizada.

Una vez con la palabra cargada, se procede a cortar los bordes o excesos de la palabra tanto al comienzo como al final de la palabra, luego procedemos a tomar la frecuencia menor y la mayor, la intensidad menor, mayor y promedio, el pitch menor, mayor y promedio, luego vemos las formantes F1, F2, F3 y F4 y además vemos los silencios o espacios que existen dentro de la palabra.

A continuación se pueden ver gráficos de cada una de las características:



4.5. Análisis estadísticos por regiones

En el análisis de voces se comparan las voces por provincias, es decir un ejemplo vemos las diferencias y paralelismos entre Sierra - Costa, Costa - Oriente y Oriente - Sierra.

A continuación se podrá ver de mejor manera, el mapa del Ecuador, para darnos cuenta que provincias no mas se pudo visitar, existen provincias que en

sus fronteras se encuentra un dialecto totalmente diferente al de la ciudad, ya que al vivir en medio de 2 provincias tienden a coger el acento de las dos provincias vecinas y lo mezclan, teniendo así un dialecto mixto.



Fue una experiencia muy enriquecedora, poder conocer además de nuevos acentos, nuevas culturas, nuevas personas, nuevos lugares y además poder contar con su ayuda fue muy importante y gratificante.

Por cuestiones económicas no se pudo viajar a tomar las muestras en las Islas Galápagos, hubiera sido realmente estupendo poder hacerlo.

Se puede reconocer auditivamente las diferencias entre los acentos de las diferentes provincias, es fácil reconocer de donde es una persona solamente

por su acento, pero en este trabajo se reconocerá objetivamente, es decir con las grabaciones y el software antes mencionado, es un reconocimiento mas real.

Se realizarán un análisis descriptivo indicando tablas y gráficos para ver los resultados de las comparaciones entre regiones, provincias, géneros, y se incluirá un análisis inferencial con varias hipótesis para podernos dar cuenta mejor de las conclusiones, todo esto se encuentra en la parte de Análisis de resultados.

Se harán tablas para ver el ANOVA, que es un análisis de la varianza, ésta es una herramienta estadística muy potente y de gran utilidad tanto en la industria, para el control de procesos, como en el laboratorio de análisis, para el control de métodos analíticos, para comparación de múltiples poblaciones, etc.

Para aceptar las hipótesis en el análisis ANOVA el valor de la significancia es aquel que nos da la respuesta, si es un valor menor a 0,05 se rechaza la hipótesis y si el valor es mayor a 0,05 quiere decir que se acepta la hipótesis.

5. Análisis de resultados

Al momento de analizar las voces se pudo ver que existían grabaciones en malas condiciones ya que existía mucho ruido o se pronunció muy seguidas las palabras por el hablante, al principio se comenzó a editar las palabras en el Cool Edit normalizándolas y usando un filtro llamado noise reduction, y se pudo observar que se perdían características propias de las voces, por lo cual se volvió a editar las voces pero en su estado natural, es decir sin filtrar nada para así poder entregar la base de datos sin pérdida de información.

A continuación se presentan tablas y gráficos para ver ciertos parámetros muy importantes que fueron analizados detenidamente.

Tabla 1: Muestras por región por género.

				Recuento
Región	Costa	Genero	Masculino	230
			Femenino	140
	Sierra	Genero	Masculino	190
			Femenino	180
	Oriente	Genero	Masculino	280
			Femenino	210

Tabla 2: Promedios de Frecuencia inferior y superior de cada palabra por región y género, en Hz.

					Frec_Inf	Frec_Sup
					Media	Media
Palabra	Ahora	Región	Costa	Masculino	147,61	1858,43
				Femenino	196,21	2261,29
			Sierra	Masculino	166,93	2242,42
				Femenino	180,59	2878,61
			Oriente	Masculino	129,14	3690,86
				Femenino	138,89	4033,38
	Arcoíris	Región	Costa	Masculino	118,66	4215,87
				Femenino	114,05	4388,36
			Sierra	Masculino	84,30	4405,47
				Femenino	80,51	4142,63
			Oriente	Masculino	72,52	4512,36
				Femenino	79,87	4750,95
	Chimpancé	Región	Costa	Masculino	55,18	4482,04
				Femenino	53,93	4373,63
			Sierra	Masculino	63,04	4213,26
				Femenino	69,29	4094,17
			Oriente	Masculino	76,66	4319,83
				Femenino	83,26	4744,38
	Cocodrilos	Región	Costa	Masculino	85,49	4421,96
				Femenino	95,41	4689,93
			Sierra	Masculino	64,95	4686,00
				Femenino	64,22	4862,94
			Oriente	Masculino	79,32	4844,04
				Femenino	82,38	4685,48
	Drenaje	Región	Costa	Masculino	92,75	4505,61
				Femenino	155,83	4626,93
			Sierra	Masculino	91,04	4388,89
				Femenino	113,48	4315,33
			Oriente	Masculino	120,31	4115,29
				Femenino	160,80	4410,90
Encebollado	Región	Costa	Masculino	105,10	4430,13	
			Femenino	134,98	4361,79	
		Sierra	Masculino	124,39	3762,42	
			Femenino	139,72	4189,83	
		Oriente	Masculino	116,43	4586,68	
			Femenino	104,52	4815,14	
Ferrocarril	Región	Costa	Masculino	89,58	4561,65	
			Femenino	120,98	4742,43	
		Sierra	Masculino	89,13	4295,84	

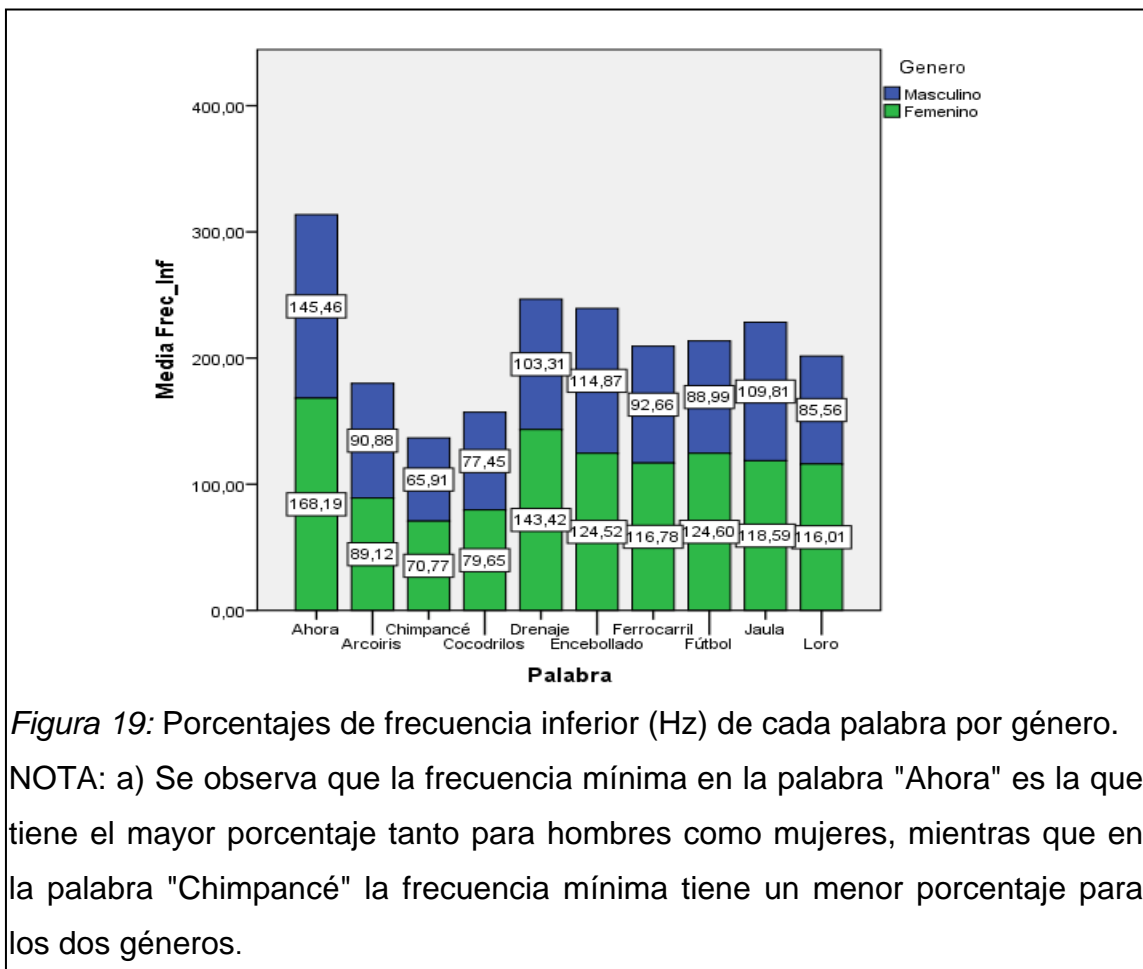
				Femenino	110,64	4376,94
			Oriente	Masculino	97,60	4430,50
				Femenino	119,24	4722,20
	Fútbol	Región	Costa	Masculino	91,16	3911,26
				Femenino	118,37	4147,86
			Sierra	Masculino	89,13	2960,58
				Femenino	127,60	3767,56
			Oriente	Masculino	87,12	4030,60
				Femenino	126,17	4336,90
	Jaula	Región	Costa	Masculino	121,99	4667,74
				Femenino	140,53	4729,13
			Sierra	Masculino	104,44	4760,32
				Femenino	103,37	4740,17
			Oriente	Masculino	103,44	4743,75
				Femenino	115,95	4697,10
	Loro	Región	Costa	Masculino	90,89	4902,48
				Femenino	120,76	4848,31
			Sierra	Masculino	84,34	4922,89
				Femenino	112,46	4893,78
			Oriente	Masculino	82,02	4057,57
				Femenino	116,12	4437,29

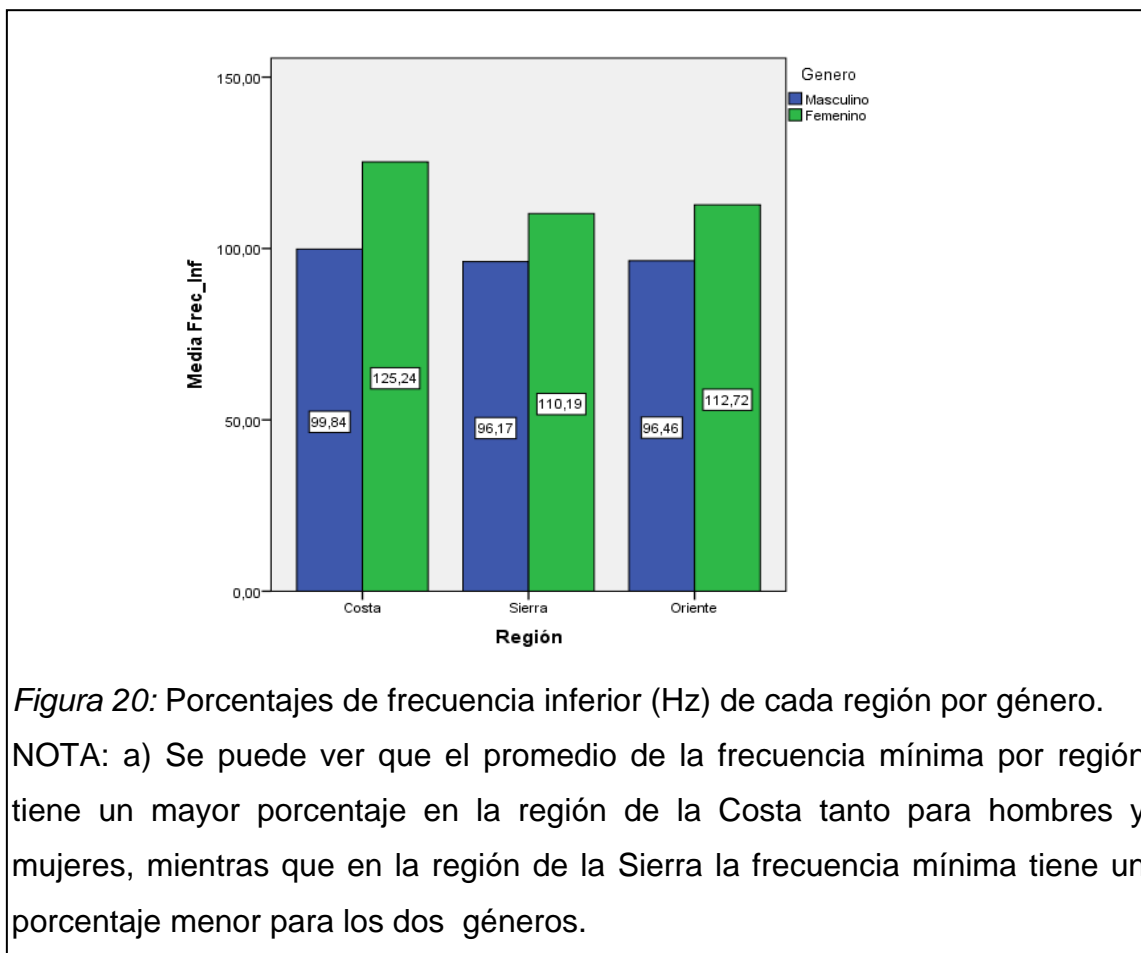
NOTA: a) Se puede ver que existe una diferencia bastante variable entre la frecuencia inferior y superior de las palabras, dependiendo de la región que sean.

b) Se aprecia que los mayores promedios de frecuencia mínima se presenta en los hablantes de género femenino y tienen el siguiente orden: 7 muestras de la Costa, 2 muestras de la Sierra y 1 muestra del Oriente.

c) Se puede ver que los promedios de frecuencia máxima se presenta en los hablantes de género femenino y muestras van en este orden: 4 muestras de la Costa, 4 muestras del Oriente y 2 muestras de la Sierra, lo que demuestra que tienen una frecuencia mínima y máxima similar en costa y oriente.

d) Se puede decir que las mujeres tienen la frecuencia mínima y máxima superior a los hombres.





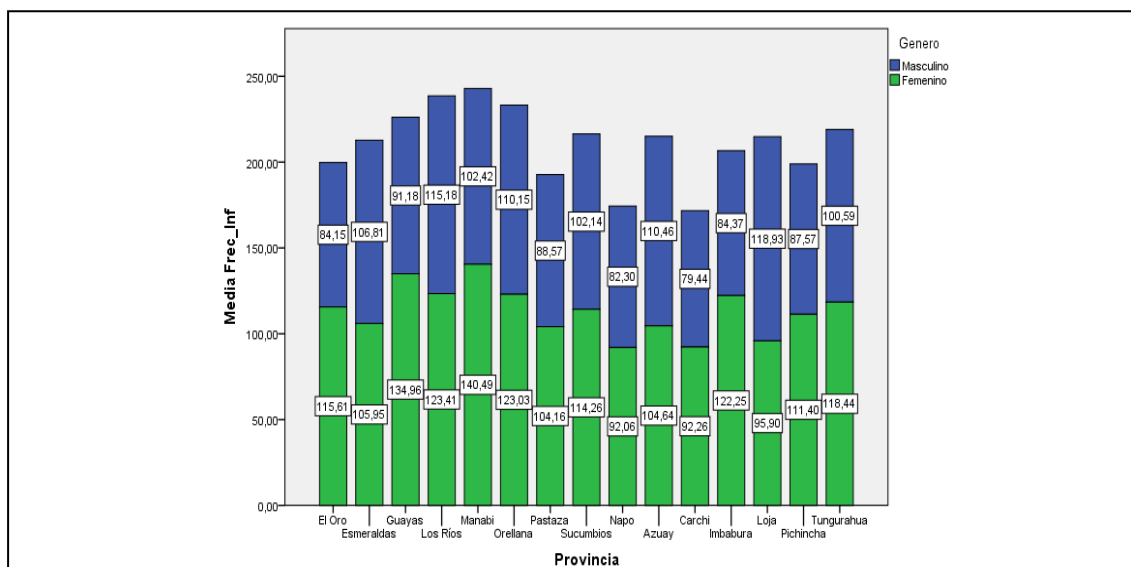


Figura 21: Porcentajes de frecuencia inferior (Hz) de cada provincia por género.

NOTA: a) Se puede ver que el promedio de la frecuencia inferior por provincia tiene un mayor porcentaje en la provincia de Manabí para mujeres y en la provincia de Loja en cambio el promedio de la frecuencia mínima tiene un mayor porcentaje para los hombres.

b) Se puede ver que el promedio de de la frecuencia inferior por provincia tiene un menor porcentaje en la provincia del Napo tanto para hombres como para mujeres.

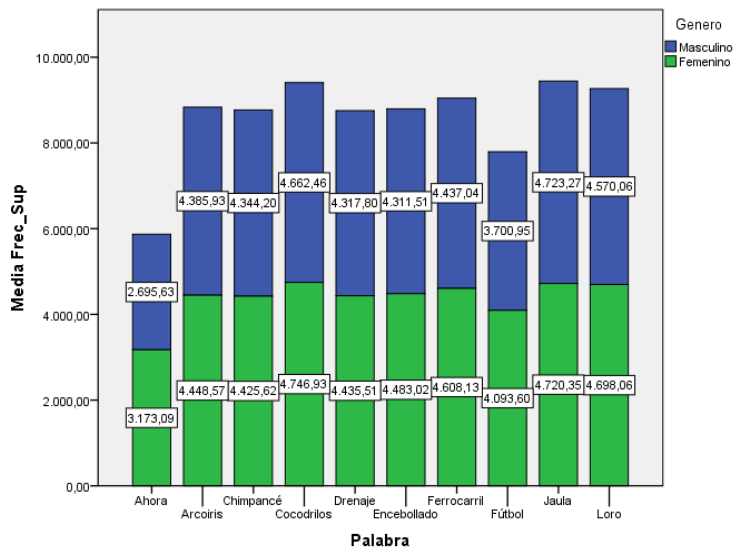


Figura 22: Porcentajes de frecuencia superior (Hz) de cada palabra por género.

NOTA: a) Se puede ver que el promedio de la frecuencia máxima en la palabra "Cocodrilos" es la que tiene el mayor porcentaje para las mujeres, en cambio para los hombres el promedio máximo se encuentra en la palabra "Jaula".

b) Mientras que el promedio de la frecuencia máxima de en la palabra "Ahora" tiene un menor porcentaje para los dos géneros.

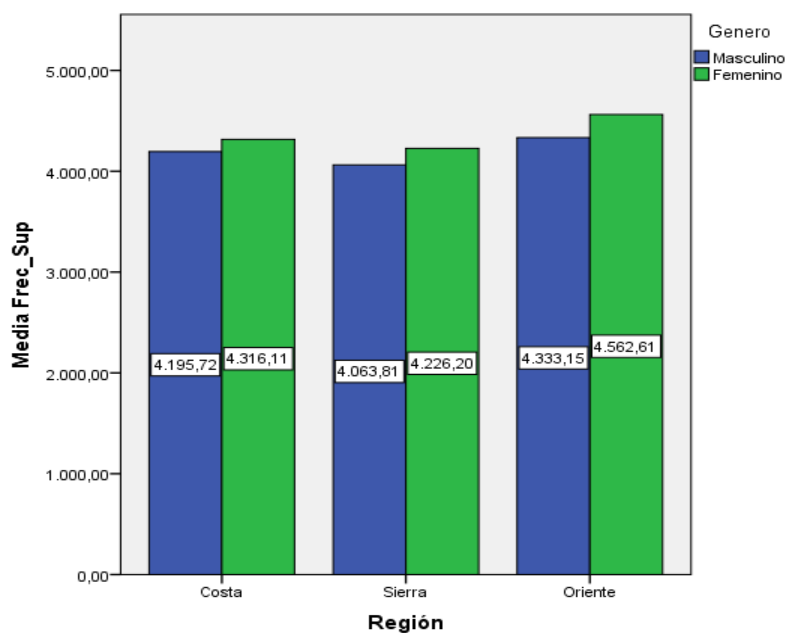


Figura 23: Porcentajes de frecuencia superior (Hz) de cada región por género.

NOTA: a) Se puede ver que el promedio de la frecuencia superior por región tiene un mayor porcentaje en la región del Oriente tanto para hombres y mujeres, mientras en la región de la Sierra la frecuencia superior tiene un porcentaje menor para los dos géneros.

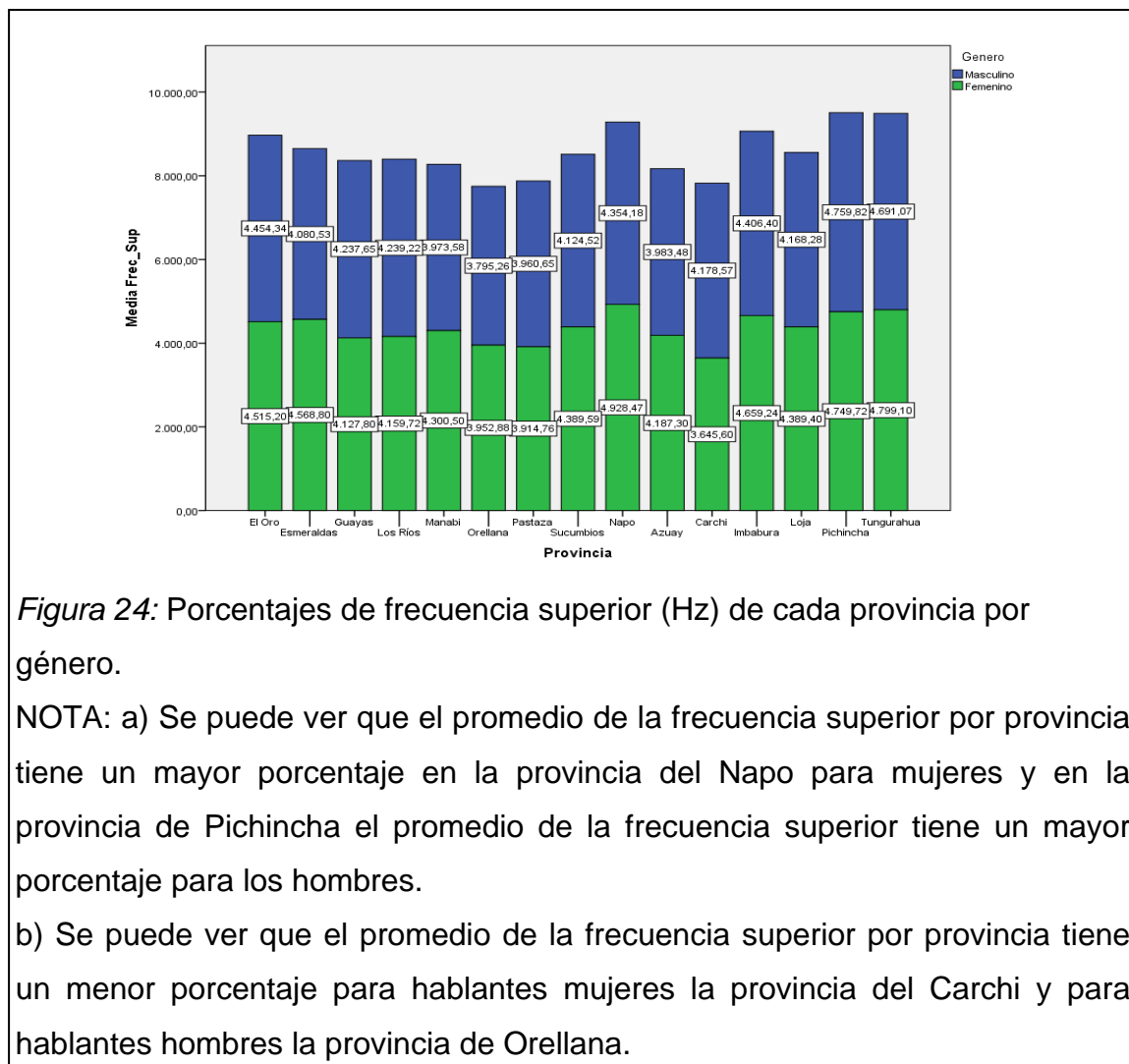


Tabla 3: Promedios de Intensidad mínima, Intensidad máxima e Intensidad de cada palabra por región y género, en dB.

Palabra	Ahora	Costa	Masculino	Int_min	Int_max	Intensidad
				Media	Media	Media
	Ahora	Costa	Masculino	44,99	59,3	56,0
			Femenino	47,91	61,7	58,7
		Sierra	Masculino	46,61	59,0	55,8
			Femenino	47,83	59,2	56,3
		Oriente	Masculino	46,43	60,3	56,9
			Femenino	49,72	62,2	59,3
	Arcoíris	Costa	Masculino	26,66	59,4	55,2
			Femenino	29,58	59,1	55,2
		Sierra	Masculino	25,78	57,8	53,8
			Femenino	27,71	58,5	54,1
		Oriente	Masculino	28,12	58,9	54,1
			Femenino	27,33	61,4	56,8
	Chimpancé	Costa	Masculino	27,70	58,9	52,8
			Femenino	27,77	60,8	54,4
		Sierra	Masculino	27,86	56,9	50,3
			Femenino	29,38	57,3	52,3
		Oriente	Masculino	29,72	58,7	52,7
			Femenino	28,53	60,7	55,3
	Cocodrilos	Costa	Masculino	31,25	59,6	54,4
			Femenino	28,69	61,2	55,4
		Sierra	Masculino	27,78	59,7	54,1
			Femenino	28,20	58,8	53,6
		Oriente	Masculino	30,57	59,3	54,1
			Femenino	28,26	61,1	56,0
	Drenaje	Costa	Masculino	38,22	58,0	54,2
			Femenino	38,02	59,2	54,3
		Sierra	Masculino	31,64	56,6	52,4
			Femenino	32,22	57,4	53,1
		Oriente	Masculino	33,04	57,4	52,8
			Femenino	34,52	60,0	55,8
Encebollado	Costa	Masculino	31,90	60,1	55,7	
		Femenino	31,01	56,6	55,6	
	Sierra	Masculino	31,59	57,8	53,7	
		Femenino	34,74	58,3	54,6	
	Oriente	Masculino	31,76	59,3	54,1	
		Femenino	33,66	60,8	56,5	
Ferrocarril	Costa	Masculino	28,81	59,1	54,0	
		Femenino	28,55	61,2	55,9	

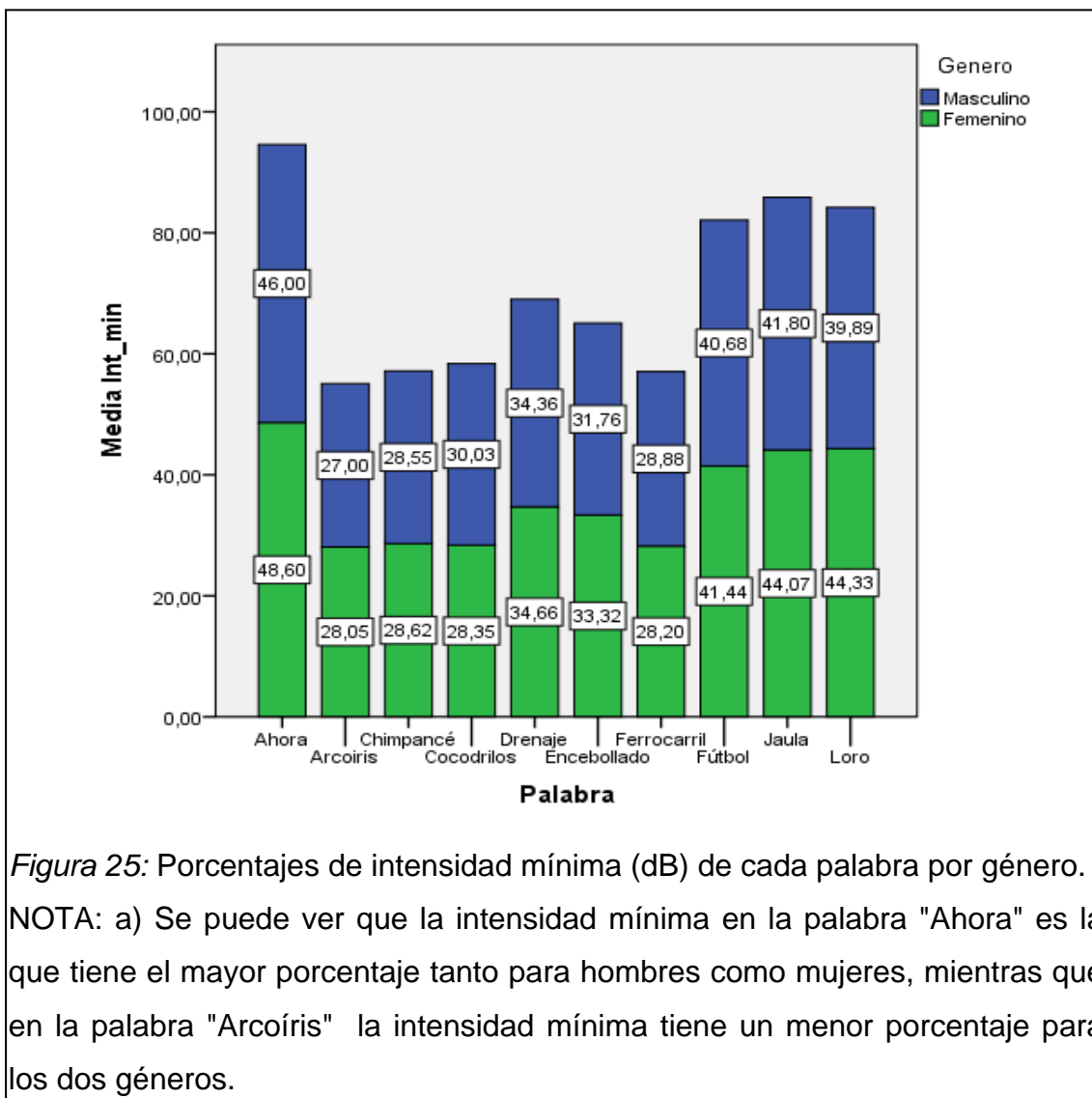
		Sierra	Masculino	28,11	57,9	53,1
			Femenino	27,32	56,5	52,0
		Oriente	Masculino	29,44	58,2	52,2
			Femenino	28,72	61,3	56,4
	Fútbol	Costa	Masculino	40,90	61,0	56,8
			Femenino	42,42	62,7	58,1
		Sierra	Masculino	42,20	58,9	54,9
			Femenino	41,34	60,5	66,6
		Oriente	Masculino	39,48	60,2	55,4
			Femenino	40,88	62,9	58,2
	Jaula	Costa	Masculino	42,34	61,1	57,5
			Femenino	43,90	60,9	57,1
		Sierra	Masculino	42,30	60,6	56,7
			Femenino	44,79	60,2	56,8
		Oriente	Masculino	41,01	59,9	56,3
			Femenino	43,57	58,7	55,1
	Loro	Costa	Masculino	44,31	60,2	57,7
			Femenino	45,57	61,8	58,6
Sierra		Masculino	43,87	60,3	57,0	
		Femenino	46,20	60,7	64,4	
Oriente		Masculino	33,56	48,5	45,6	
		Femenino	41,95	56,6	53,8	

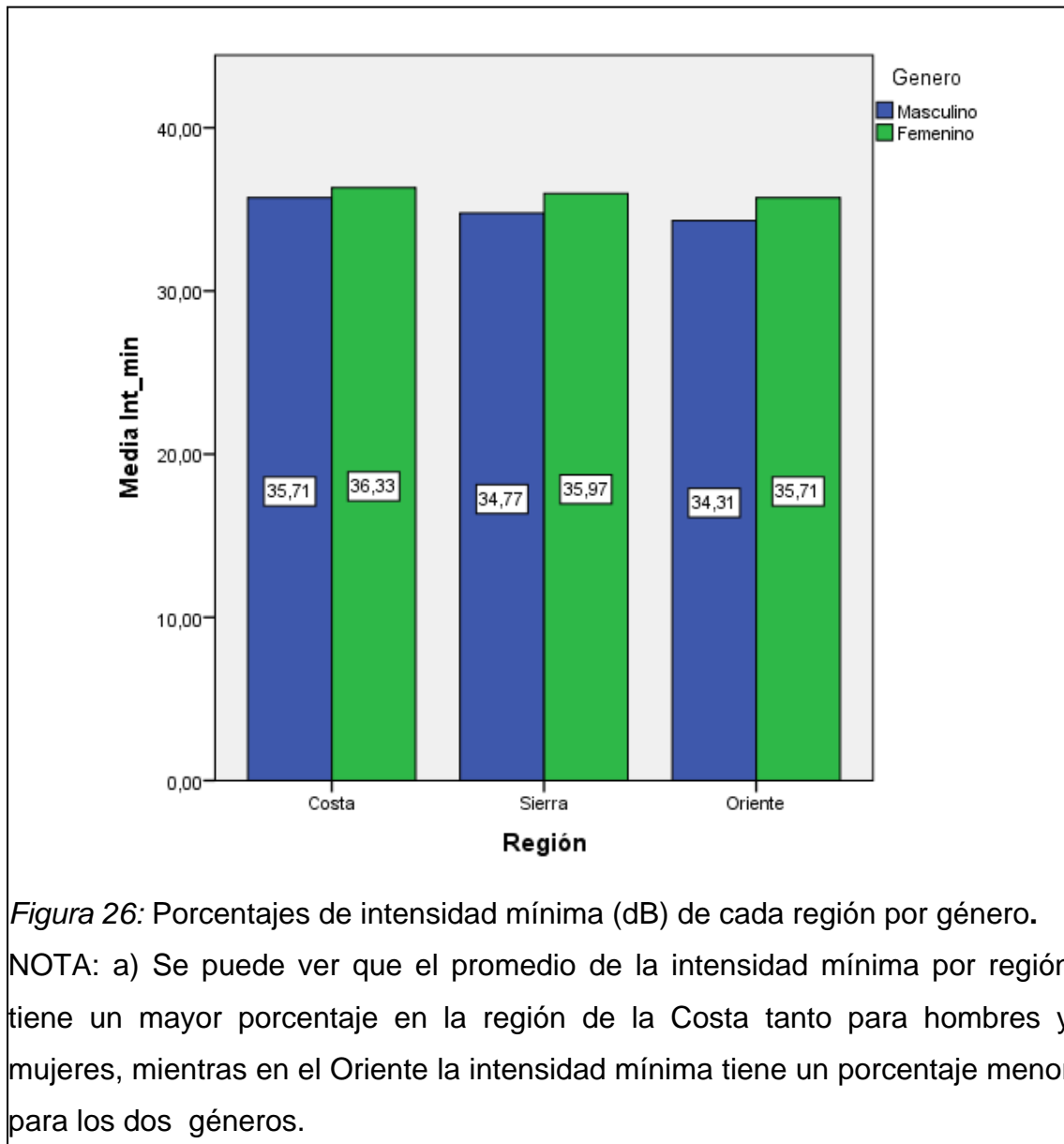
NOTA: a) Se puede ver que existe una diferencia de dB dependiendo de la palabra que se pronuncie, relacionado directamente con la región a la que pertenezcan.

b) Se puede ver que los mayores promedios de la Intensidad mínima se presentan en los hablantes de género femenino en el siguiente orden: 3 muestras Sierra, 2 muestras Costa, y 2 muestras Oriente, también existen valores altos para ciertas palabras en hablantes hombres en el siguiente orden: 2 muestras Costa y 1 muestra Oriente.

c) Se puede ver que los mayores promedios de la Intensidad máxima se presentan en los hablantes de género femenino en el siguiente orden: 6 muestras Oriente y 3 muestras Costa, también existe un valor alto para un hablante de la Costa.

d) Se puede ver que los mayores promedios de la Intensidad se presentan en los hablantes de género femenino en el siguiente orden: 7 muestras Oriente y 1 muestra Sierra, también existen valores altos para ciertas palabras en hablantes hombres en el siguiente orden: 1 muestras Costa y 1 muestra Oriente.





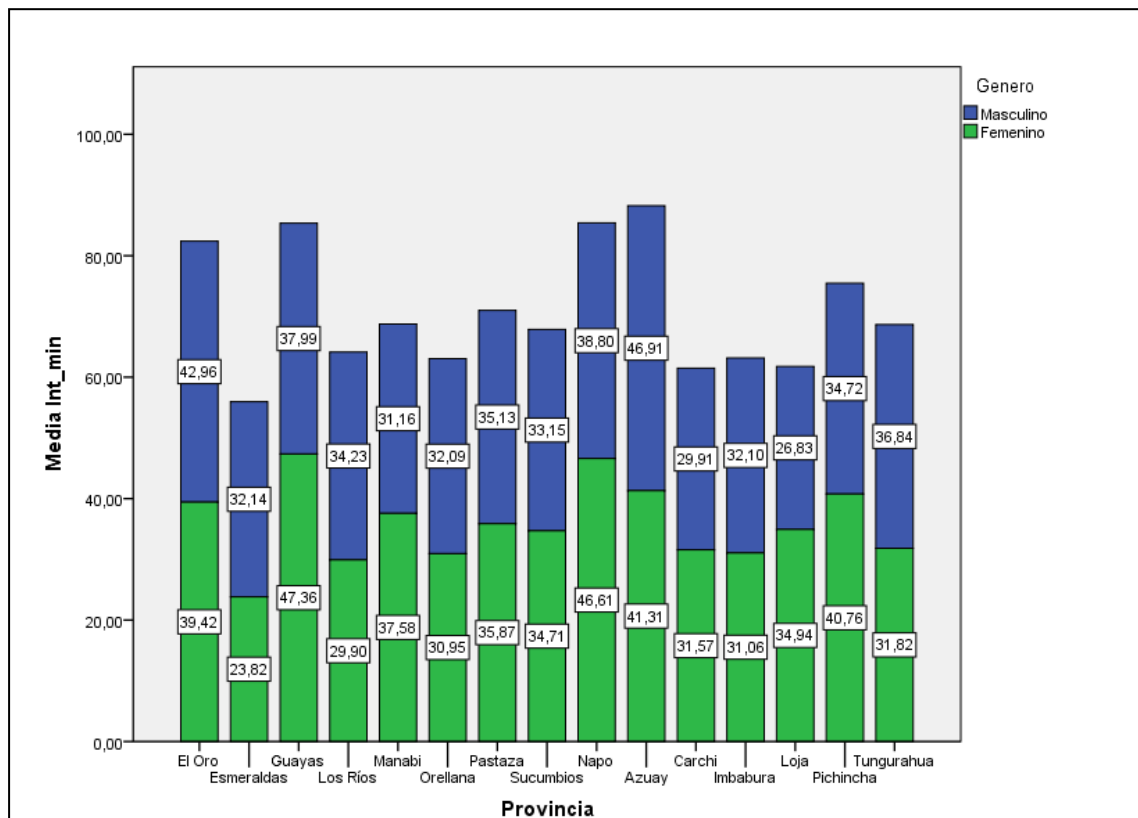


Figura 27: Porcentajes de intensidad mínima (dB) de cada provincia por género.

NOTA: a) Se puede ver que el promedio de la intensidad mínima por provincia tiene un mayor porcentaje en la provincia del Guayas para mujeres y en la provincia de Azuay el promedio de la intensidad mínima tiene un mayor porcentaje para los hombres.

b) Se puede ver que el promedio de la intensidad mínima por provincia tiene un menor porcentaje para hablantes mujeres la provincia de Esmeraldas y para hablantes hombres la provincia de Loja.

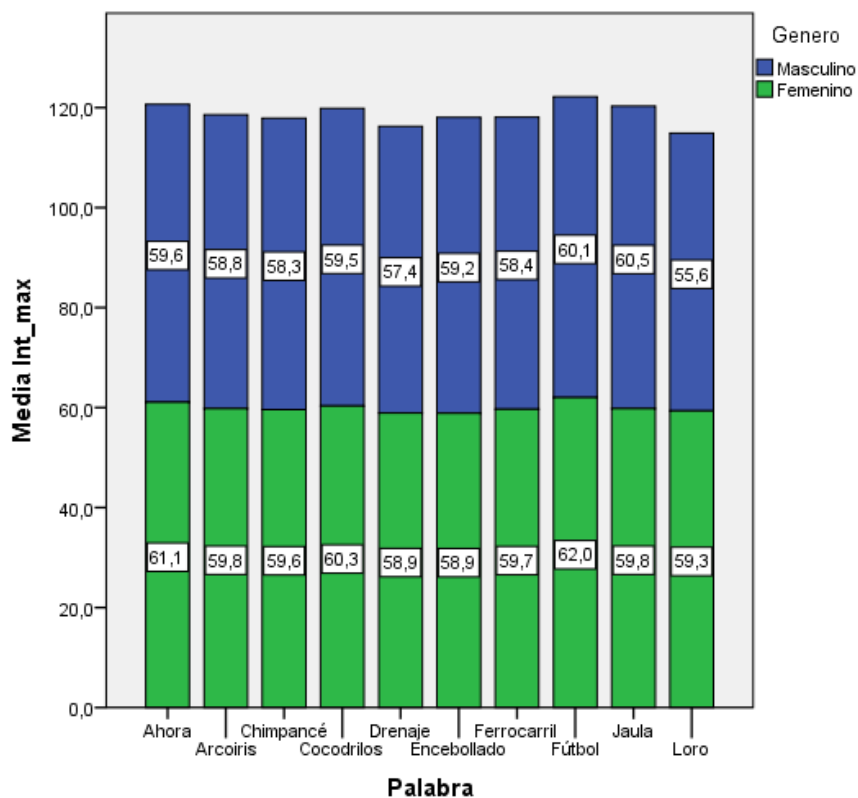
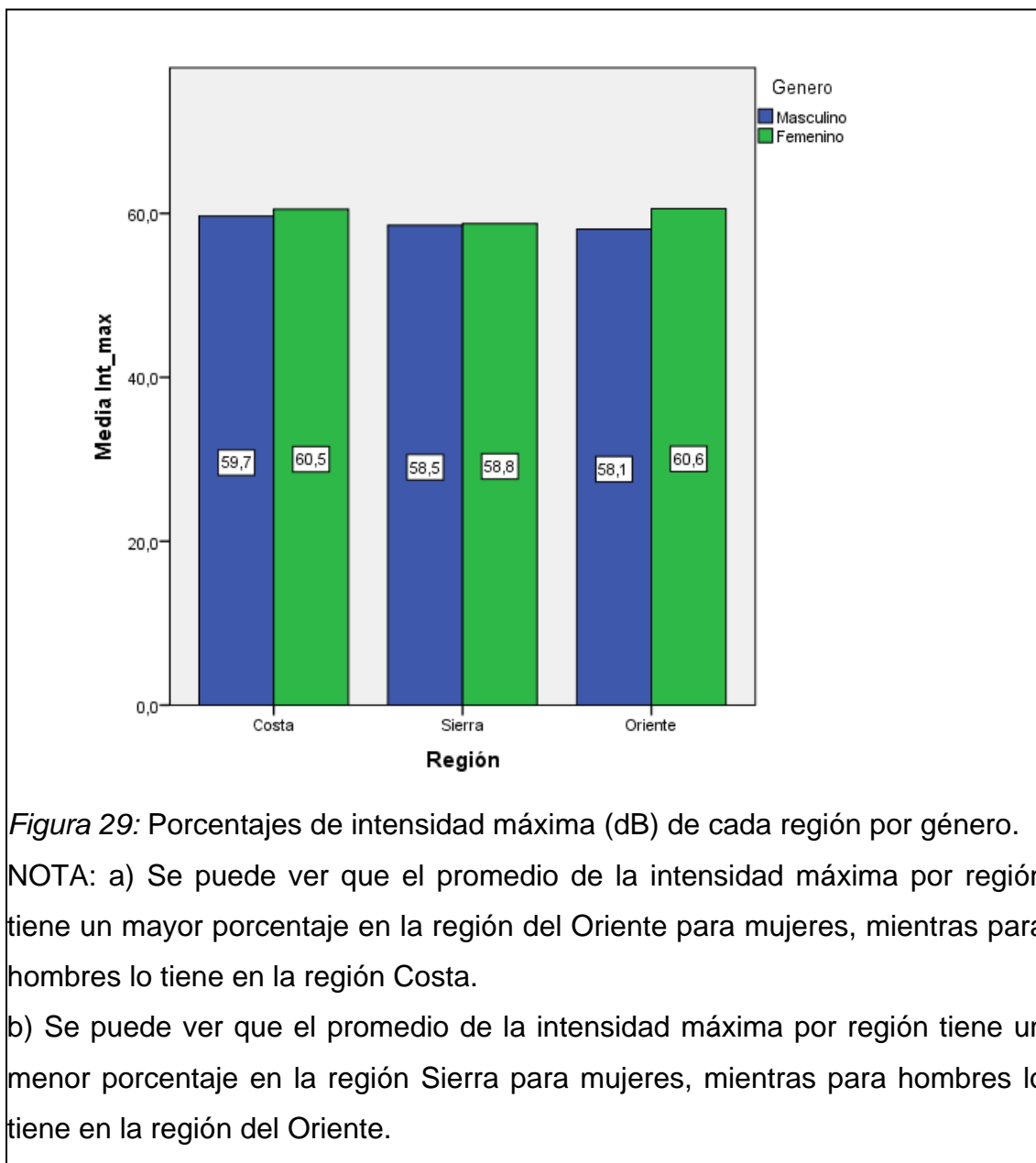


Figura 28: Porcentajes de intensidad máxima (dB) de cada palabra por género.

NOTA: a) Se puede ver que la intensidad máxima en la palabra "Fútbol" es la que tiene el mayor porcentaje para hablantes mujeres, mientras que en la palabra "Jaula" la intensidad máxima tiene un mayor porcentaje para los hombres.

b) Se puede ver también que la intensidad máxima en la palabra "Drenaje" y "Encebollado" es la que tiene el menor porcentaje para hablantes mujeres, mientras que la palabra "Drenaje" tiene el menor porcentaje de intensidad máxima para los hablantes hombres.



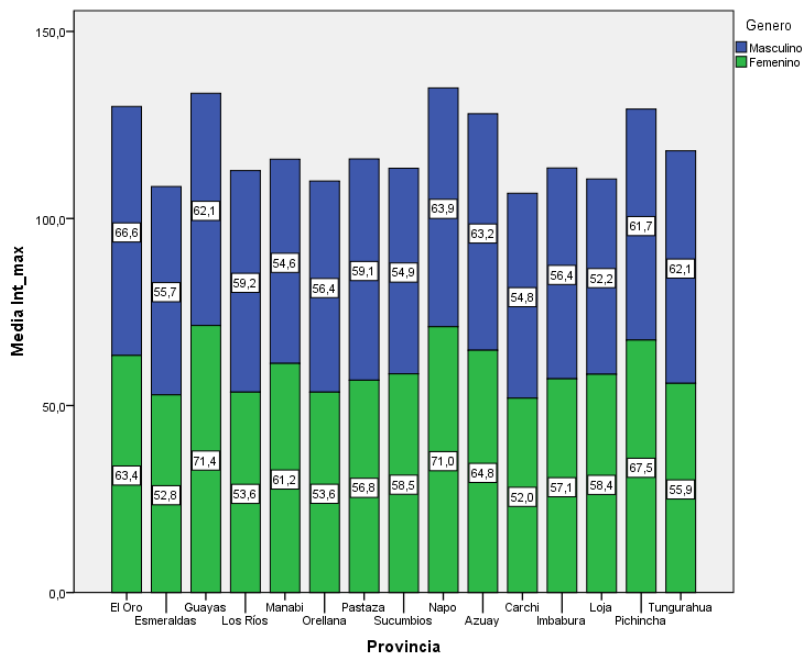


Figura 30: Porcentajes de intensidad máxima (dB) de cada provincia por género.

NOTA: a) Se puede ver que el promedio de la intensidad máxima por provincia tiene un mayor porcentaje en la provincia del Guayas para mujeres y en la provincia del Oro para los hombres.

b) Se puede ver que el promedio de la intensidad máxima por provincia tiene un menor porcentaje en la provincia de Carchi para mujeres, mientras que Loja tiene el menor porcentaje en hablantes hombres.

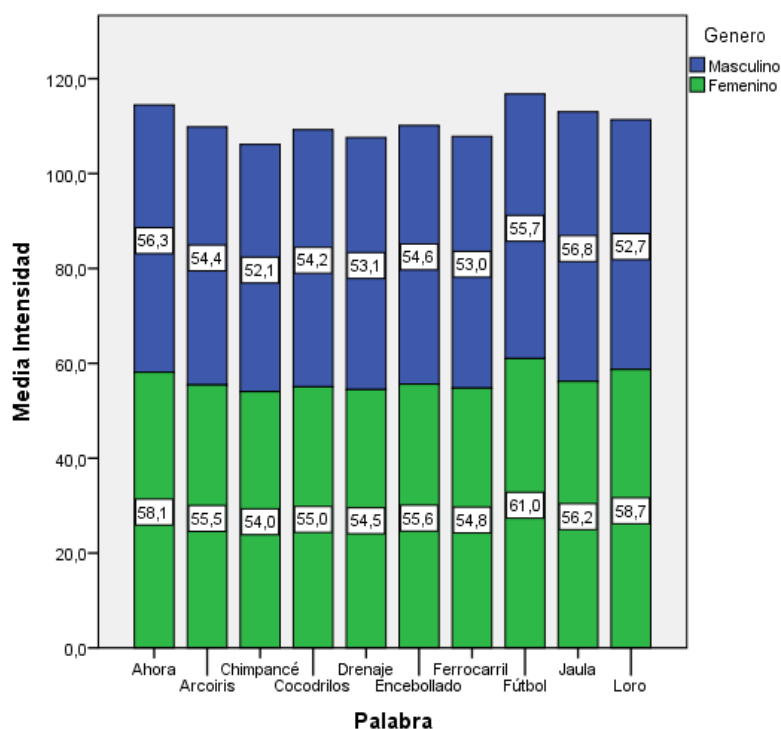


Figura 31: Porcentajes de intensidad (dB) de cada palabra por género.

NOTA: a) Se puede ver que la intensidad en la palabra "Fútbol" es la que tiene el mayor porcentaje para hablantes mujeres, mientras que en la palabra "Jaula" la intensidad máxima tiene un mayor porcentaje para los hombres.

b) Se puede ver también que la intensidad máxima en la palabra "Chimpancé" es la que tiene el menor porcentaje para hablantes de los dos géneros.

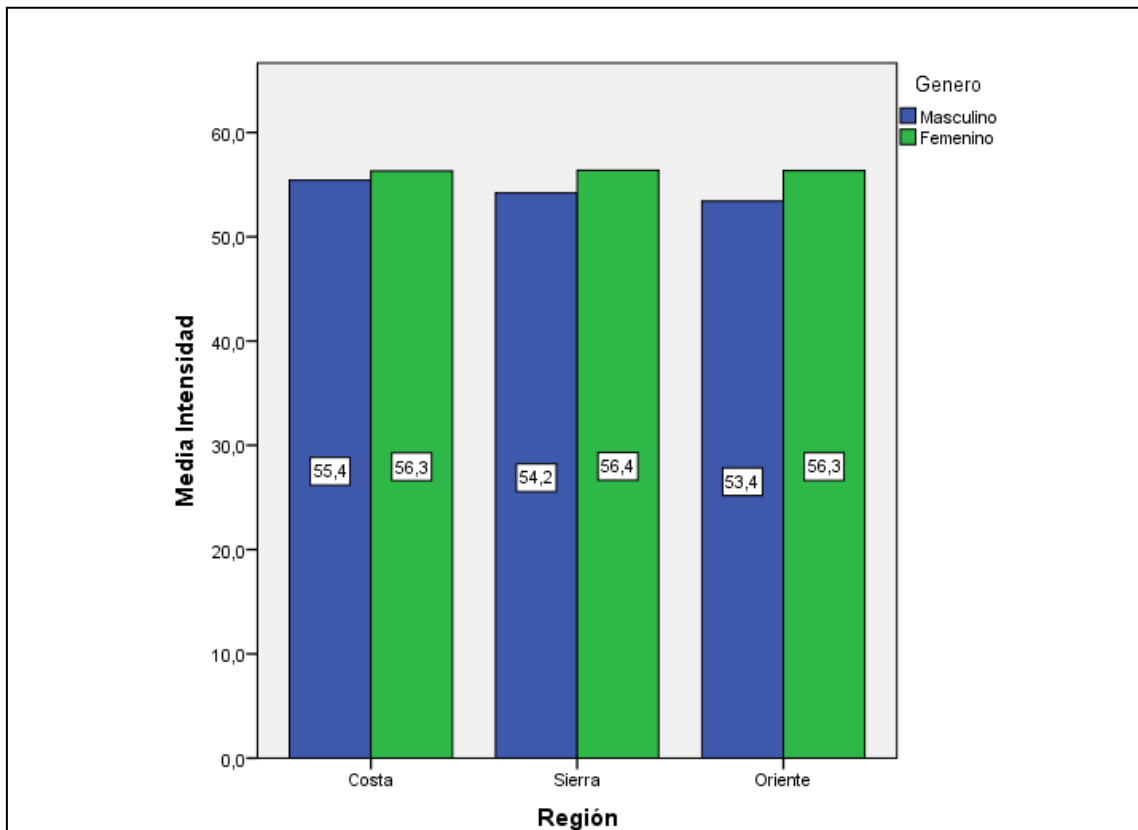


Figura 32: Porcentajes de intensidad (dB) de cada región por género.

NOTA: a) Se puede ver que el promedio de la intensidad por región tiene un mayor porcentaje en la región de la Sierra pero en este caso los valores son muy parecidos en las 3 regiones para las mujeres, mientras que para hombres el máximo nivel de intensidad lo tienen en la región Costa.

b) Se puede ver que el promedio de la intensidad máxima por región tiene un menor porcentaje en la región del Oriente, tanto para mujeres y para hombres.

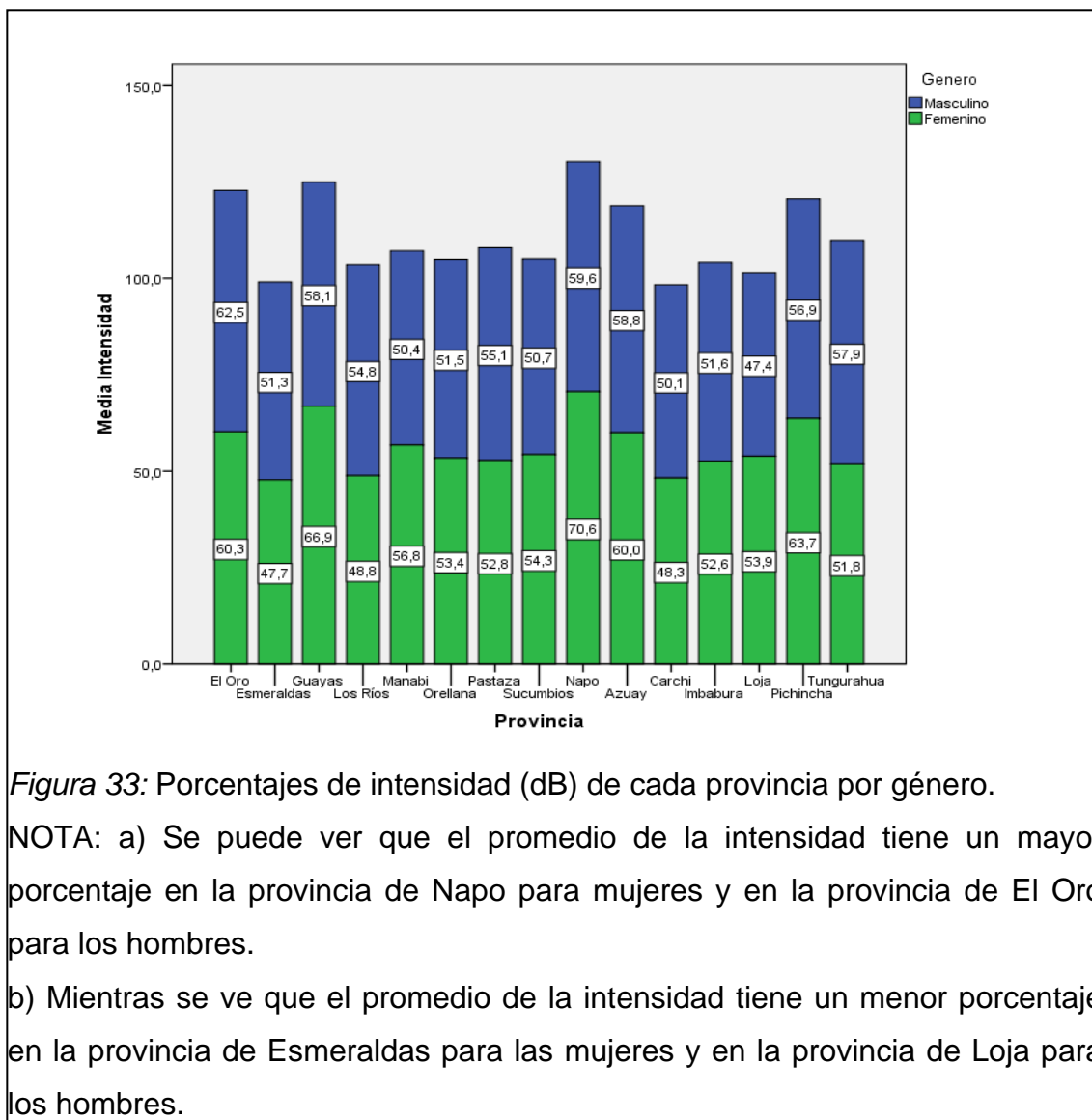


Figura 33: Porcentajes de intensidad (dB) de cada provincia por género.

NOTA: a) Se puede ver que el promedio de la intensidad tiene un mayor porcentaje en la provincia de Napo para mujeres y en la provincia de El Oro para los hombres.

b) Mientras se ve que el promedio de la intensidad tiene un menor porcentaje en la provincia de Esmeraldas para las mujeres y en la provincia de Loja para los hombres.

Tabla 4: Promedios de Pitch mínimo, Pitch máxima y Pitch de cada palabra por región y género, en Hz.

				Pitch_min	Pitch_max	Pitch
				Media	Media	Media
Palabra	Ahora	Costa	Masculino	109,18	163,40	132,56
			Femenino	174,97	272,18	223,44
		Sierra	Masculino	109,14	152,51	133,28
			Femenino	148,87	249,63	213,42
		Oriente	Masculino	109,68	177,78	137,91
			Femenino	175,53	263,80	244,14
	Arcoíris	Costa	Masculino	102,08	161,22	133,40
			Femenino	137,08	265,55	221,88
		Sierra	Masculino	106,02	215,84	135,96
			Femenino	142,42	279,80	217,42
		Oriente	Masculino	108,97	199,82	138,64
			Femenino	153,23	271,08	222,53
	Chimpancé	Costa	Masculino	102,82	172,53	126,92
			Femenino	153,63	277,06	217,27
		Sierra	Masculino	105,13	202,23	132,01
			Femenino	164,39	280,01	216,93
		Oriente	Masculino	102,22	187,41	131,92
			Femenino	177,58	267,31	220,66
	Cocodrilos	Costa	Masculino	91,51	151,63	121,47
			Femenino	120,85	264,12	195,58
		Sierra	Masculino	94,12	186,61	126,42
			Femenino	135,96	296,20	197,62
		Oriente	Masculino	96,01	213,35	129,81
			Femenino	143,87	240,47	195,69
	Drenaje	Costa	Masculino	104,28	154,27	128,79
			Femenino	166,02	269,71	217,32
		Sierra	Masculino	102,29	160,35	130,67
			Femenino	163,23	253,84	211,33
Oriente		Masculino	107,00	172,77	133,07	
		Femenino	168,00	270,35	218,10	
Encebollado	Costa	Masculino	102,44	153,81	126,18	
		Femenino	141,05	267,70	213,67	

		Sierra	Masculino	102,11	155,41	126,32	
			Femenino	141,06	252,98	206,79	
		Oriente	Masculino	104,06	163,48	130,13	
			Femenino	152,10	258,33	209,40	
		Ferrocarril	Costa	Masculino	97,00	174,47	131,06
				Femenino	167,83	288,80	217,29
	Sierra		Masculino	100,99	157,68	129,07	
			Femenino	162,53	256,13	201,47	
	Oriente		Masculino	103,97	166,55	131,50	
			Femenino	158,00	264,41	213,70	
	Fútbol	Costa	Masculino	114,57	154,71	139,15	
			Femenino	201,42	281,24	236,85	
		Sierra	Masculino	112,68	153,06	135,96	
			Femenino	167,78	260,10	224,09	
		Oriente	Masculino	111,04	167,40	144,34	
			Femenino	187,61	260,21	231,39	
	Jaula	Costa	Masculino	117,38	175,46	139,98	
			Femenino	175,46	271,77	225,12	
		Sierra	Masculino	111,92	153,72	133,69	
			Femenino	182,69	250,48	223,88	
		Oriente	Masculino	118,29	171,84	142,01	
			Femenino	176,50	244,37	214,85	
	Loro	Costa	Masculino	109,04	157,66	133,14	
			Femenino	186,66	278,97	226,97	
Sierra		Masculino	105,21	170,93	132,42		
		Femenino	155,69	257,52	218,83		
Oriente		Masculino	89,69	131,85	112,97		
		Femenino	166,96	234,62	199,33		

NOTA: a) Se puede ver que existe una diferencia de Hz dependiendo de la palabra que se pronuncie, relacionado directamente con la región a la que pertenezcan.

b) Se puede ver que los mayores promedios del pitch mínimo se presentan en los hablantes de género femenino en el siguiente orden: 6 muestras del Oriente, 3 muestras de la Costa y 1 muestra de la Sierra.

c) Se aprecia que los mayores promedios del pitch máximo se presentan en los hablantes de género femenino en el siguiente orden: 6 muestras de la Costa, 3 muestras de la Sierra y 1 muestra del Oriente.

d) Se observa que los mayores promedios del pitch se presentan en los hablantes de género femenino en el siguiente orden: 5 muestras de la Costa, 4 muestras del Oriente y 1 muestra del Sierra.

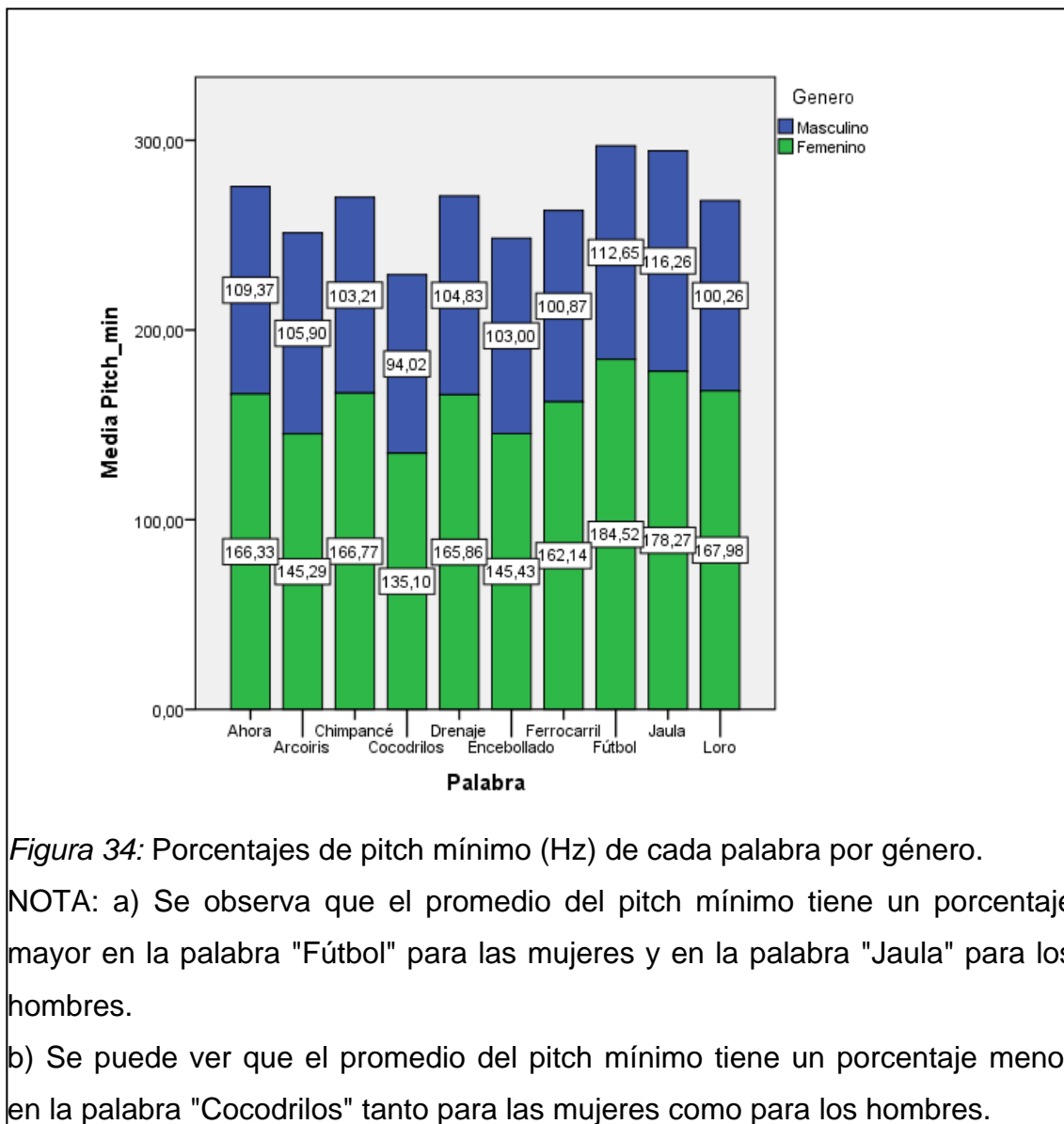


Figura 34: Porcentajes de pitch mínimo (Hz) de cada palabra por género.

NOTA: a) Se observa que el promedio del pitch mínimo tiene un porcentaje mayor en la palabra "Fútbol" para las mujeres y en la palabra "Jaula" para los hombres.

b) Se puede ver que el promedio del pitch mínimo tiene un porcentaje menor en la palabra "Cocodrilos" tanto para las mujeres como para los hombres.

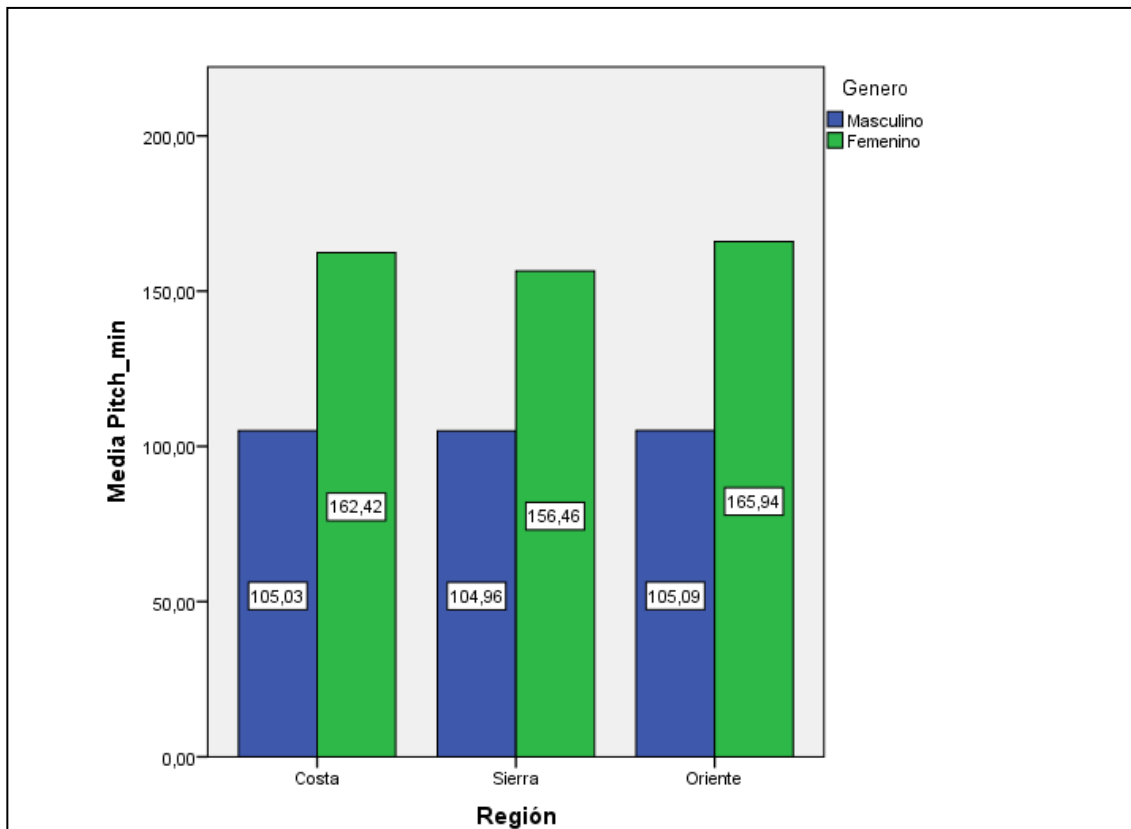


Figura 35: Porcentajes de pitch mínimo (Hz) de cada región por género.

NOTA: a) Se puede ver que el promedio del pitch mínimo por región tiene un mayor porcentaje tanto para hombres como para mujeres en la provincia del Oriente.

b) Se aprecia que el promedio del pitch mínimo por región tiene un menor porcentaje tanto para hombres como para mujeres en la provincia de la Sierra.

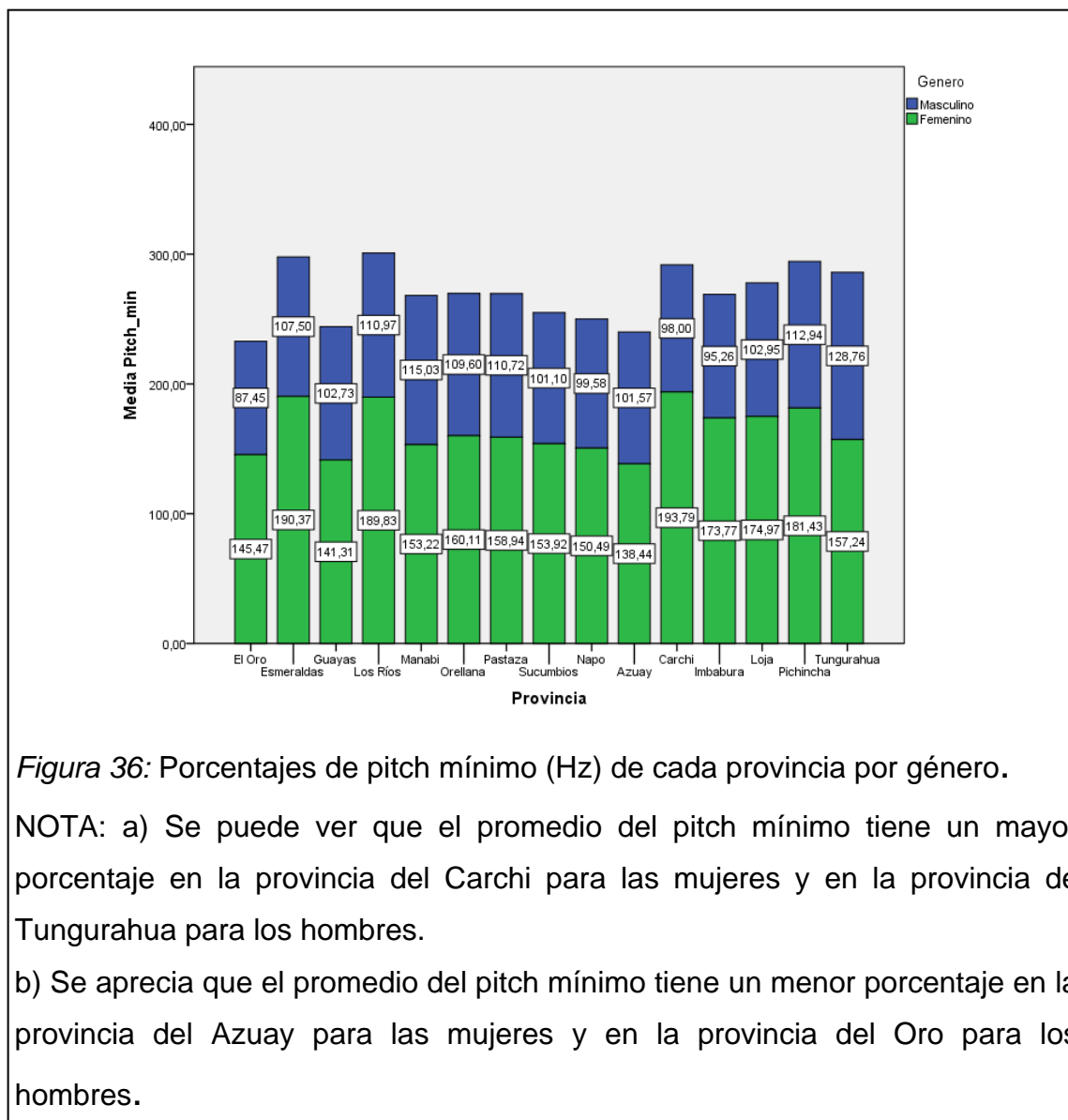


Figura 36: Porcentajes de pitch mínimo (Hz) de cada provincia por género.

NOTA: a) Se puede ver que el promedio del pitch mínimo tiene un mayor porcentaje en la provincia del Carchi para las mujeres y en la provincia de Tungurahua para los hombres.

b) Se aprecia que el promedio del pitch mínimo tiene un menor porcentaje en la provincia del Azuay para las mujeres y en la provincia del Oro para los hombres.

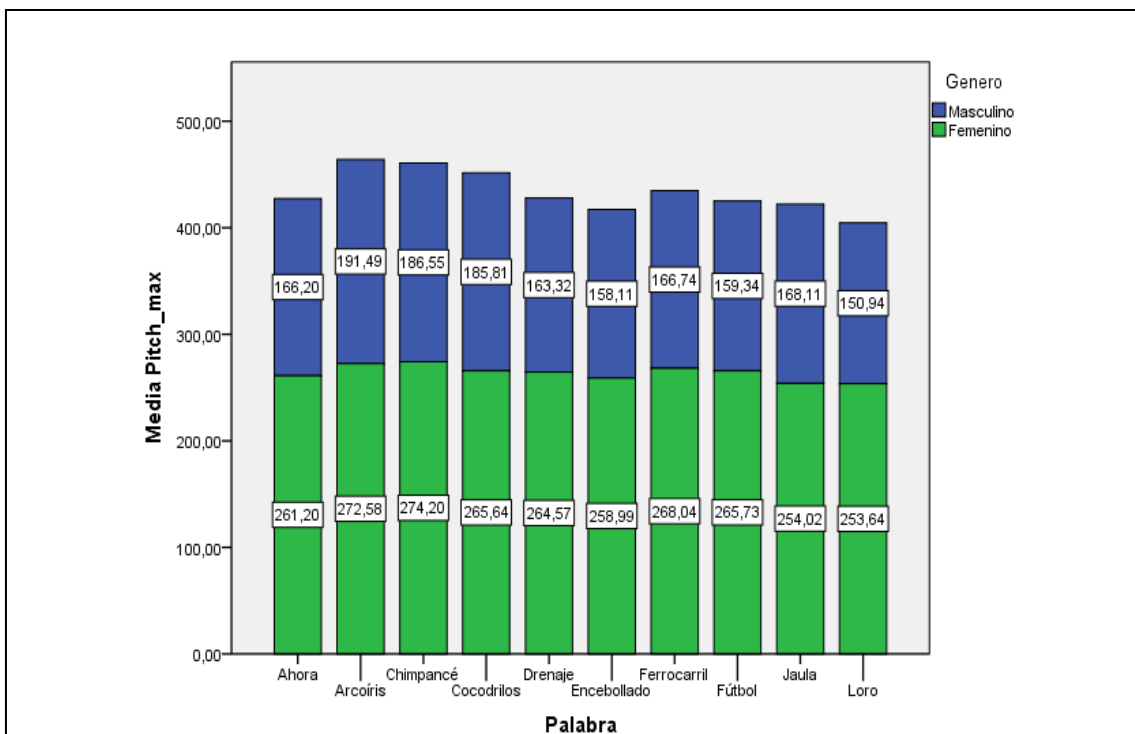


Figura 37: Porcentajes de pitch máximo (HZ) de cada palabra por género.

NOTA: a) Se puede ver que en el promedio del pitch máximo el mayor porcentaje tienen las mujeres en la palabra "Chimpancé", mientras que en hombres se encuentra en la palabra "Arcoiris".

b) Se puede ver que en el promedio del pitch máximo el menor porcentaje se encuentra en la palabra "Loro" tanto para hombres como mujeres.

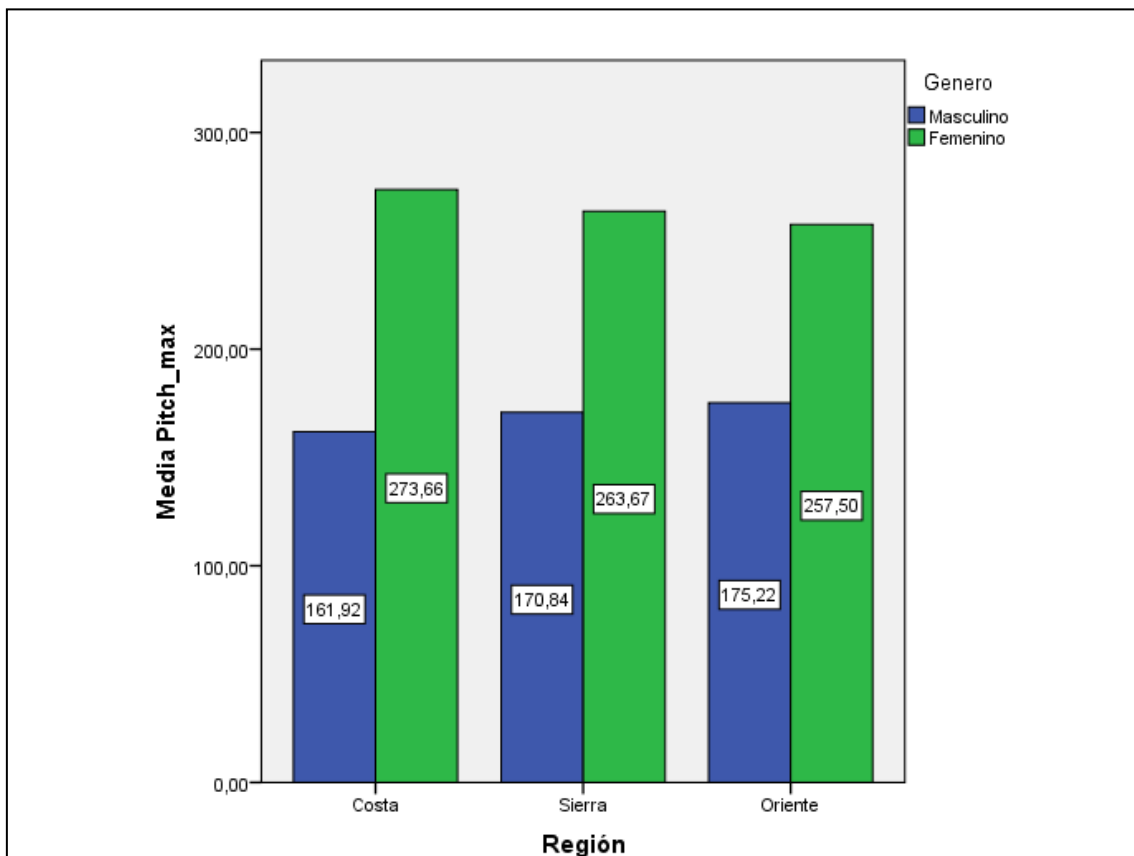


Figura 38: Porcentajes de pitch máximo (Hz) de cada región por género.

NOTA: a) Se puede ver que el promedio del pitch máximo por región tiene un mayor porcentaje para mujeres en la región del Oriente y para hombres en la región de la Costa.

b) Se puede ver que el promedio del pitch máximo por región tiene un menor porcentaje en la región Costa para las mujeres, y en la región del Oriente para los hombres.

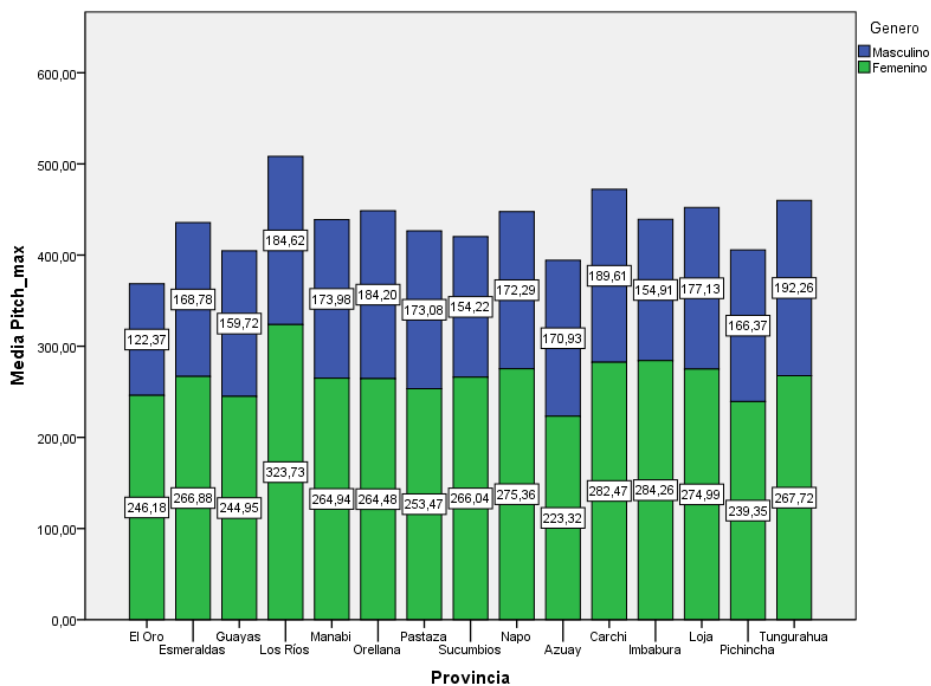


Figura 39: Porcentajes de pitch máximo (Hz) de cada provincia por género.

NOTA: a) Se puede ver que el promedio del pitch máximo por provincia tiene un mayor porcentaje en la provincia de los Ríos para mujeres y en la provincia del Tungurahua para los hombres.

b) Se puede ver que el promedio del pitch máximo por provincia tiene un menor porcentaje en la provincia del Azuay para mujeres, mientras que en la provincia del Oro tiene el menor porcentaje en hablantes hombres.

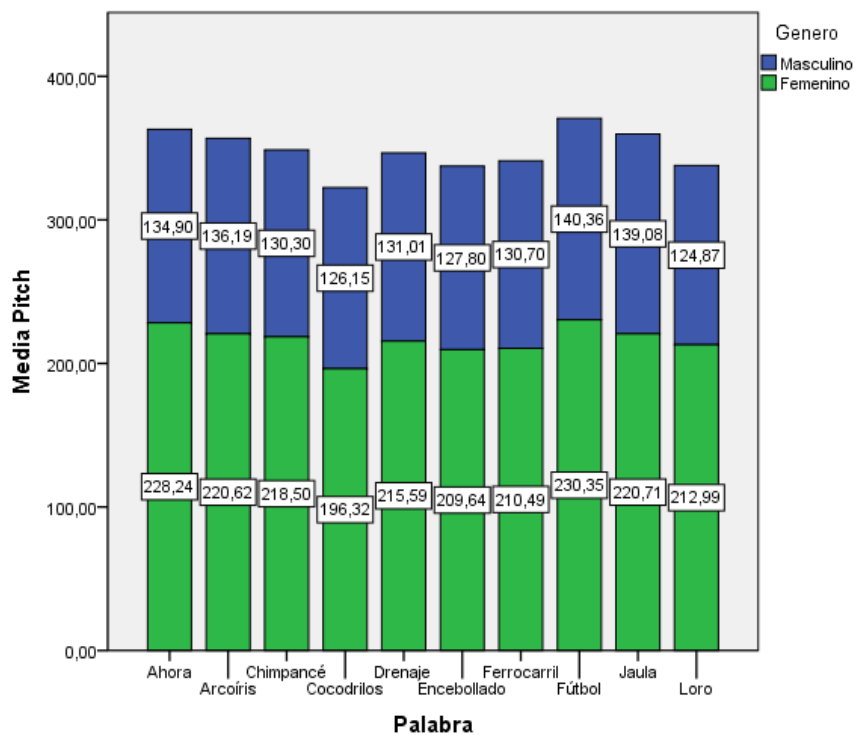
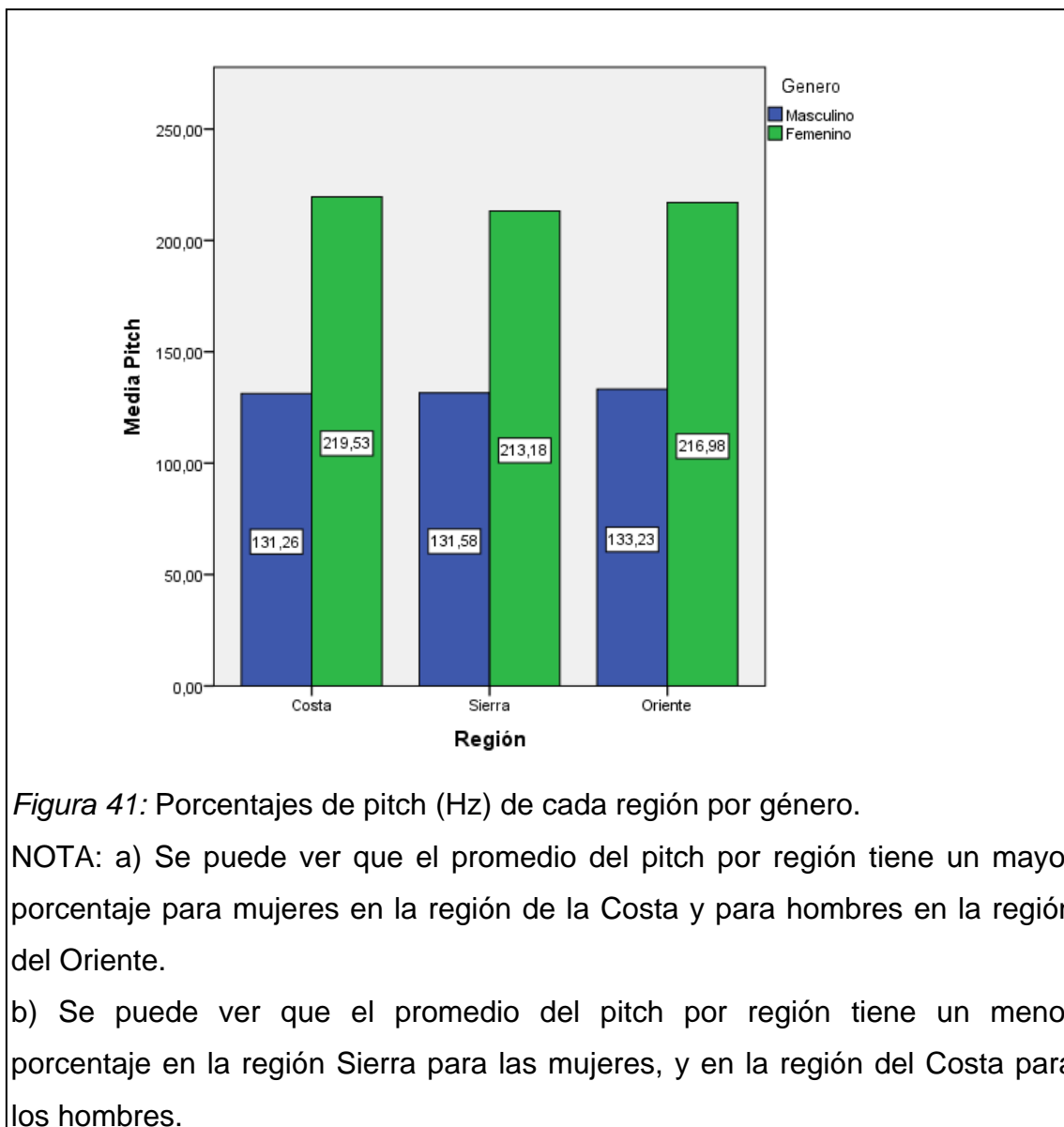


Figura 40: Porcentajes de pitch (Hz) de cada palabra por género.

NOTA: a) Se puede ver que en el promedio del pitch el mayor porcentaje que tiene la palabra "Fútbol" tanto para mujeres como para hombres.

b) Se puede ver que en el promedio del pitch máximo el menor porcentaje se encuentra en la palabra "Cocodrilos" tanto para hombres como mujeres.



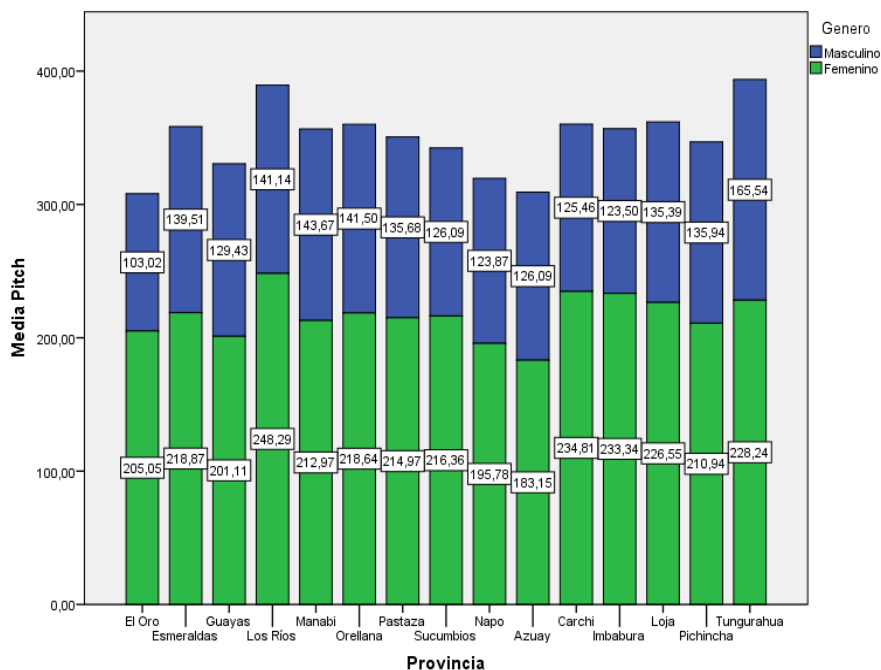


Figura 42: Porcentajes de pitch (Hz) de cada provincia por género.

NOTA: a) Se puede ver que el promedio del pitch por provincia tiene un mayor porcentaje en la provincia de los Ríos para mujeres y en la provincia del Tungurahua para los hombres.

b) Se puede ver que el promedio del pitch por provincia tiene un menor porcentaje en la provincia del Azuay para mujeres, mientras que en la provincia del Oro tiene el menor porcentaje en hablantes hombres.

Tabla 5: Promedio de las formantes de cada palabra por región y género.

				F1	F2	F3	F4	
				Media	Media	Media	Media	
Palabra	Ahora	Costa	Masculino	556,88	1194,88	2608,52	3834,23	
			Femenino	639,86	1319,72	2642,50	4008,23	
		Sierra	Masculino	561,93	1244,40	2725,37	3942,44	
			Femenino	665,96	1400,75	2845,76	4089,29	
		Oriente	Masculino	573,09	1212,54	2620,72	3868,22	
			Femenino	688,75	1352,53	2749,93	4074,02	
		Arcoíris	Costa	Masculino	468,99	1790,65	2803,17	3922,33
				Femenino	525,21	1878,86	2883,47	4144,88
			Sierra	Masculino	508,79	1851,51	2871,34	4060,55
				Femenino	556,19	2051,76	3032,55	4276,57
			Oriente	Masculino	515,30	1837,79	2868,14	4053,39
				Femenino	548,73	2016,84	2986,17	4304,49
	Chimpancé	Costa	Masculino	590,78	1869,74	2913,25	4067,38	
			Femenino	535,35	1883,38	2980,76	4184,35	
		Sierra	Masculino	605,20	1928,72	3033,23	4180,15	
			Femenino	685,59	2014,80	3088,77	4269,66	
		Oriente	Masculino	540,91	1849,45	2896,26	4110,75	
			Femenino	583,62	1932,23	3014,80	4271,97	
	Cocodrilos	Costa	Masculino	483,42	1553,08	2743,07	3902,50	
			Femenino	513,44	1678,96	2940,47	4143,66	
		Sierra	Masculino	544,55	1723,68	2932,14	4057,19	
			Femenino	569,18	1804,19	3055,94	4194,68	
			Oriente	Masculino	463,56	1552,69	2735,22	3969,67
				Femenino	509,43	1734,20	2977,46	4184,84
Drenaje		Costa	Masculino	497,88	1761,25	2766,94	4024,64	

			Femenino	544,81	1790,76	2875,94	4101,40
		Sierra	Masculino	537,13	1853,21	2900,80	4109,40
			Femenino	604,47	2012,52	3025,47	4230,24
		Oriente	Masculino	509,26	1795,89	2743,00	4032,12
			Femenino	592,45	1954,86	2916,10	4258,19
	Encebollado	Costa	Masculino	500,59	1582,75	2814,29	3971,45
			Femenino	537,23	12297,29	2879,79	4121,69
		Sierra	Masculino	511,47	1580,26	2886,35	4040,22
			Femenino	545,36	1703,82	2977,33	4180,99
		Oriente	Masculino	485,20	1540,80	2827,46	4020,31
			Femenino	537,40	1676,64	2887,53	4239,17
	Ferrocarril	Costa	Masculino	481,41	1609,16	2645,89	3922,88
			Femenino	524,94	1763,87	2774,86	4081,25
		Sierra	Masculino	482,29	1719,38	2705,75	3991,82
			Femenino	501,53	1803,06	2743,90	3932,64
		Oriente	Masculino	465,47	1721,34	2748,72	3952,85
			Femenino	510,60	1906,65	2798,48	4176,20
	Fútbol	Costa	Masculino	424,78	1259,77	2751,76	3841,29
			Femenino	441,10	1306,18	2861,78	4045,50
		Sierra	Masculino	437,32	1293,47	2806,66	3858,78
			Femenino	473,50	1394,88	2958,67	4070,93
		Oriente	Masculino	464,83	1337,97	2825,89	3948,48
			Femenino	436,77	1321,91	2908,34	4127,37
	Jaula	Costa	Masculino	534,84	1260,18	2621,32	3839,93
			Femenino	617,85	1423,40	2707,93	4002,89
		Sierra	Masculino	538,25	1321,77	2675,65	3908,37
			Femenino	658,68	1515,32	2880,48	4131,52

		Oriente	Masculino	539,15	1311,77	2603,70	3869,66
			Femenino	591,38	1383,11	2666,15	3868,55
	Loro	Costa	Masculino	441,41	1283,99	2669,07	3798,26
			Femenino	501,56	1368,48	2728,31	4030,23
		Sierra	Masculino	492,32	1352,23	2765,50	3925,20
			Femenino	507,04	1457,38	2884,67	4103,94
		Oriente	Masculino	384,56	1055,04	2185,80	3207,61
			Femenino	471,51	1273,83	2559,62	3732,98

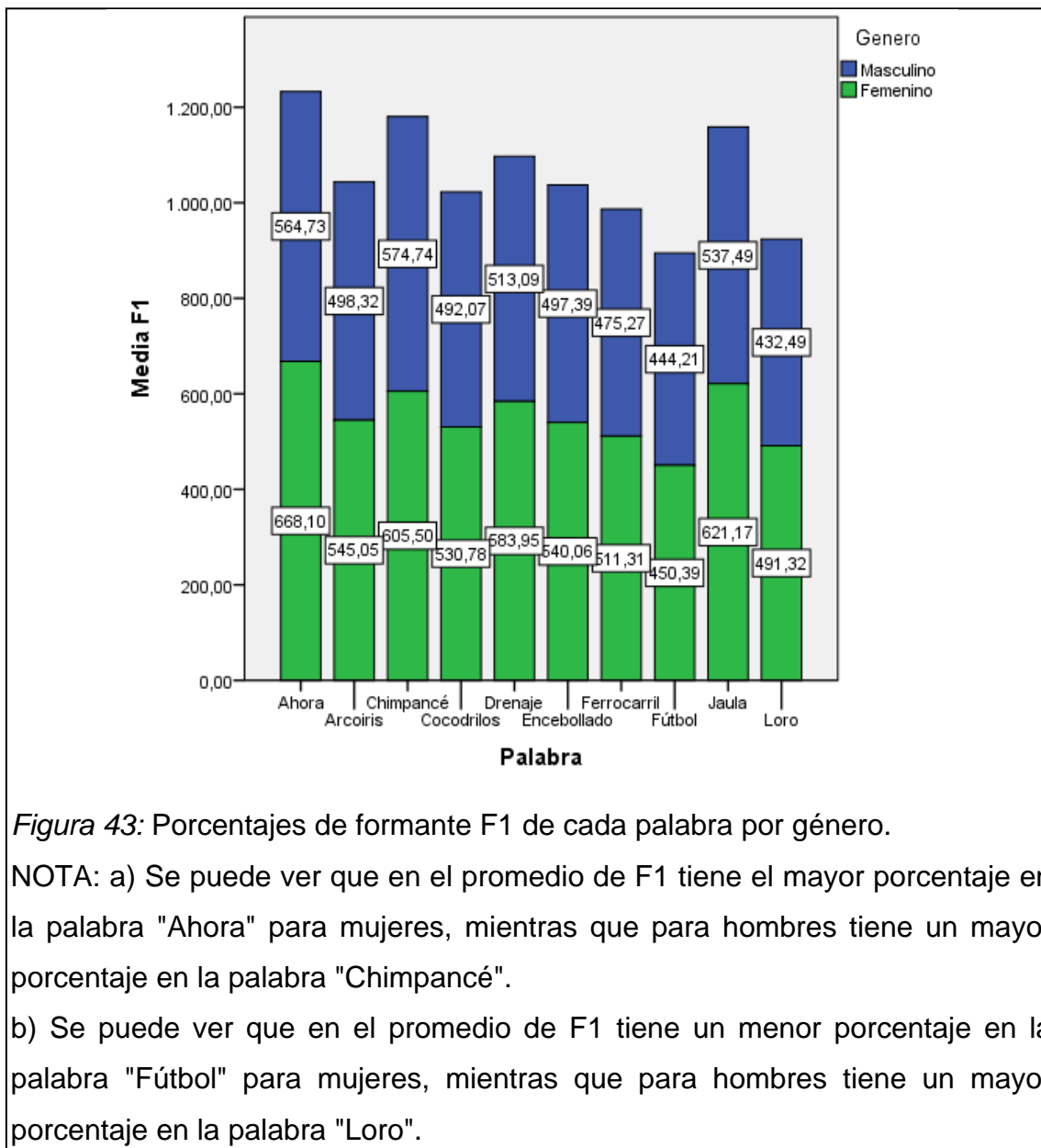
NOTA: a) Se puede ver que las formantes F1 tienen un rango de 380 Hz hasta 660 Hz, y se puede ver que los mayores promedios de F1 se presentan en los hablantes de género femenino en este orden: 8 muestras de la Sierra y 2 muestras del Oriente.

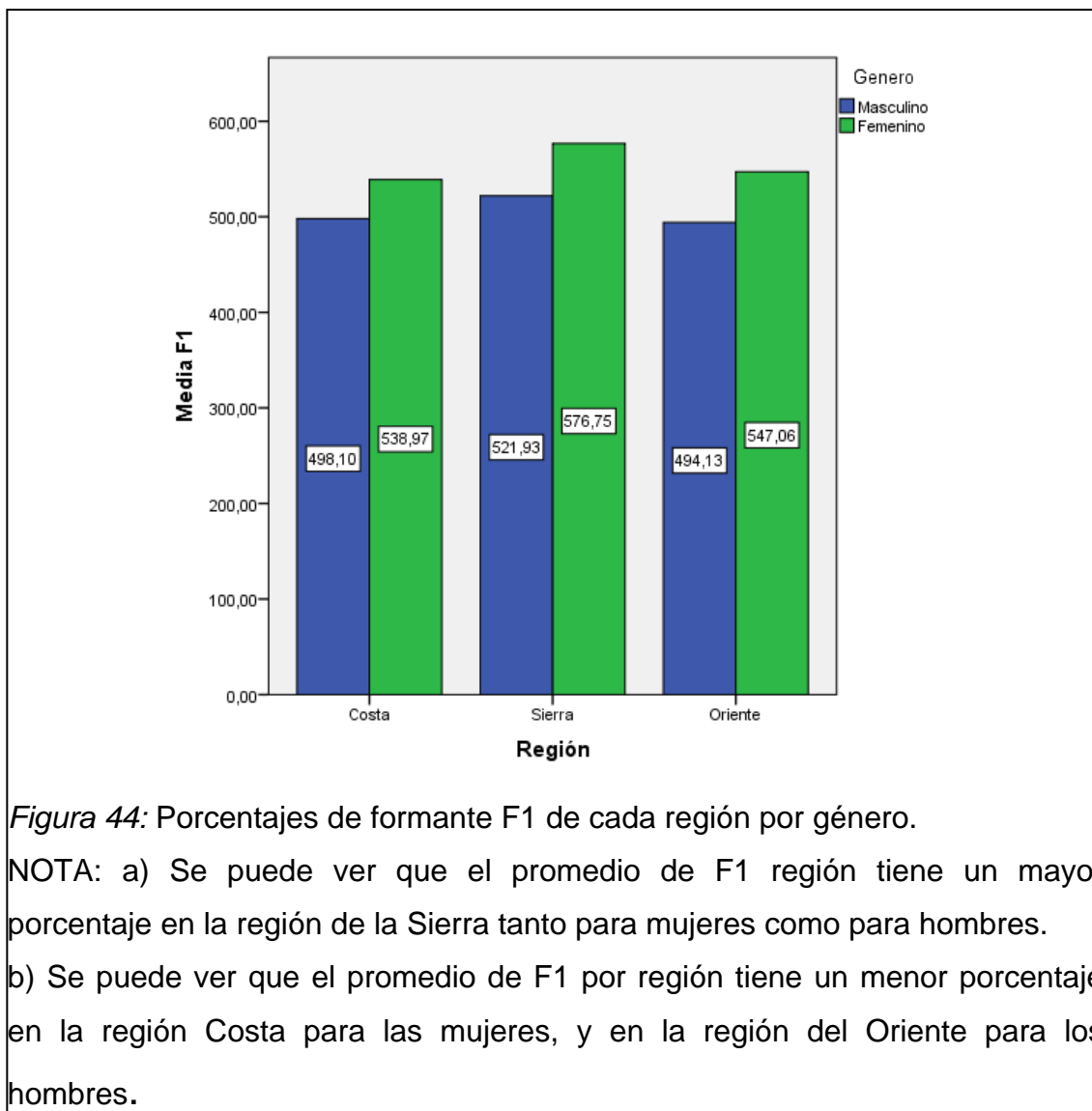
b) Se puede ver que las formantes F2 tienen un rango de 1194 Hz hasta 2014 Hz, y se puede ver que los mayores promedios de F2 se presentan en los hablantes de género femenino en este orden: 9 muestras de la Sierra y 1 muestra del Oriente.

c) Se puede ver que las formantes F3 tienen un rango de 2600 Hz hasta 3055 Hz, y se puede ver que los mayores promedios de F3 se presentan en los hablantes de género femenino en este orden: 9 muestras de la Sierra y 1 muestra del Oriente.

d) Se puede ver que las formantes F4 tienen un rango de 3800 Hz hasta 4300 Hz, y se puede ver que los mayores promedios de F4 se presentan en los hablantes de género femenino en este orden: 6 muestras de la Sierra y 4 muestras del Oriente.

e) Los mayores promedios de la tabla total de las 4 formantes presenta gran porcentaje para las hablantes mujeres en un 95% en las mujeres de la Sierra y un 5% las del Oriente.





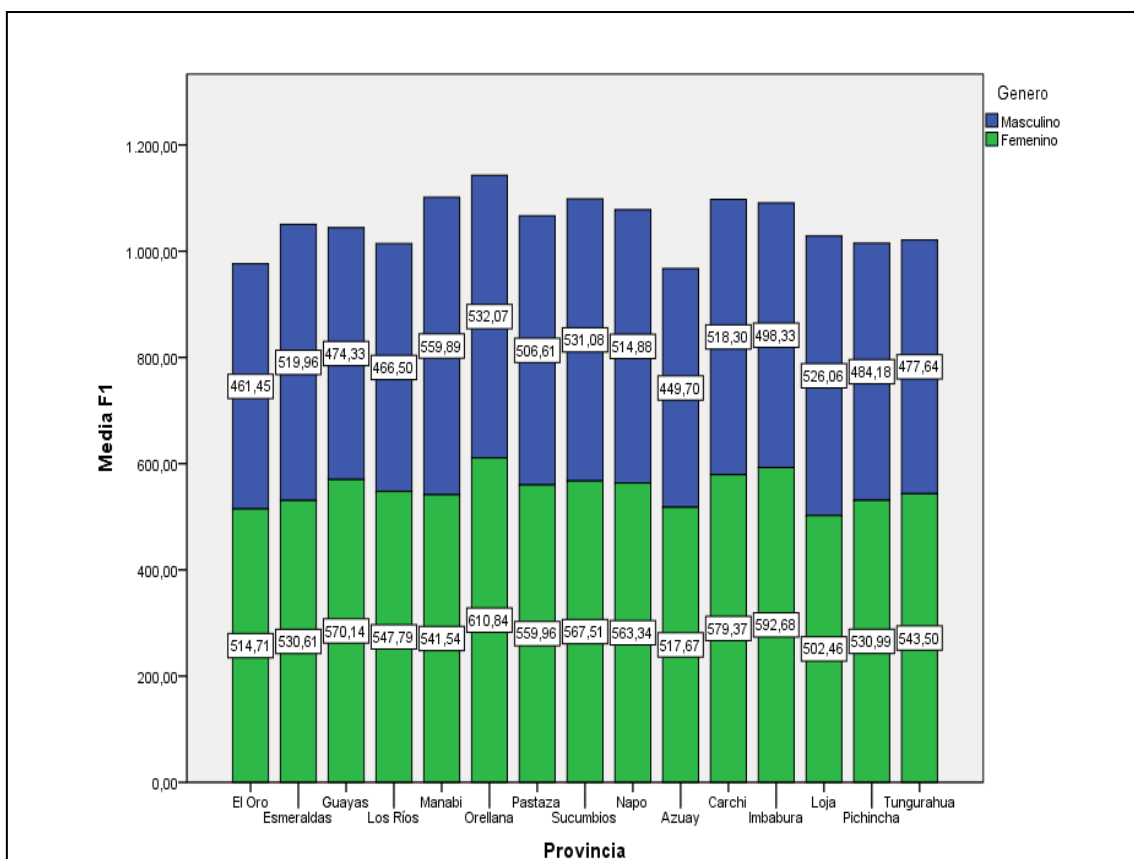


Figura 45: Porcentajes de formante F1 de cada provincia por género.

NOTA: a) Se puede ver que el promedio de F1 por provincia tiene un mayor porcentaje en la provincia de Orellana para mujeres y en la provincia de Manabí para los hombres.

b) Se puede observar que el promedio de F1 por provincia tiene un menor porcentaje en la provincia del Loja para mujeres, mientras que en la provincia de Azuay tiene el menor porcentaje en hablantes hombres.

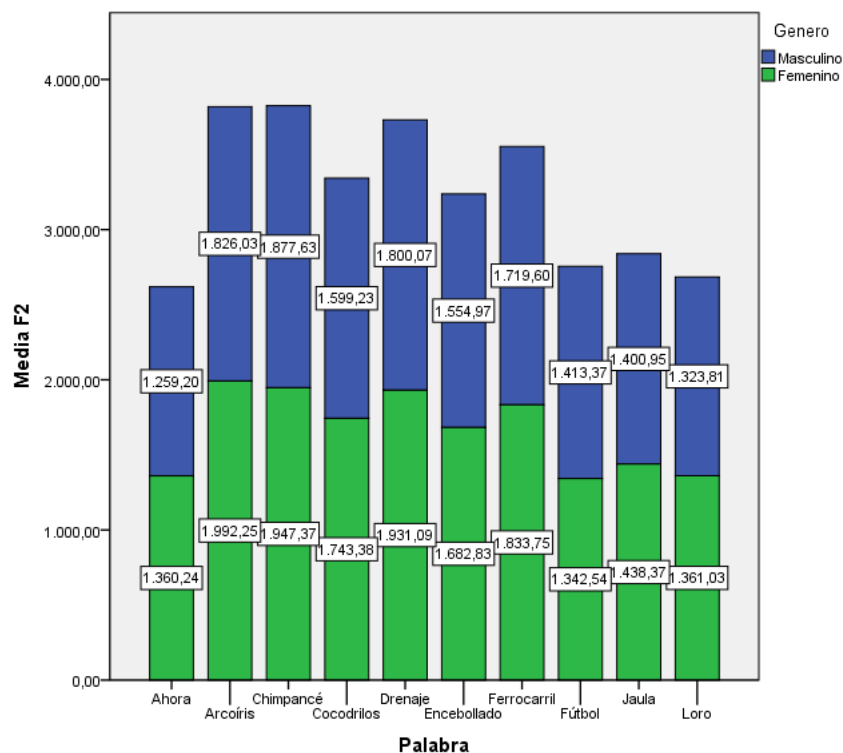
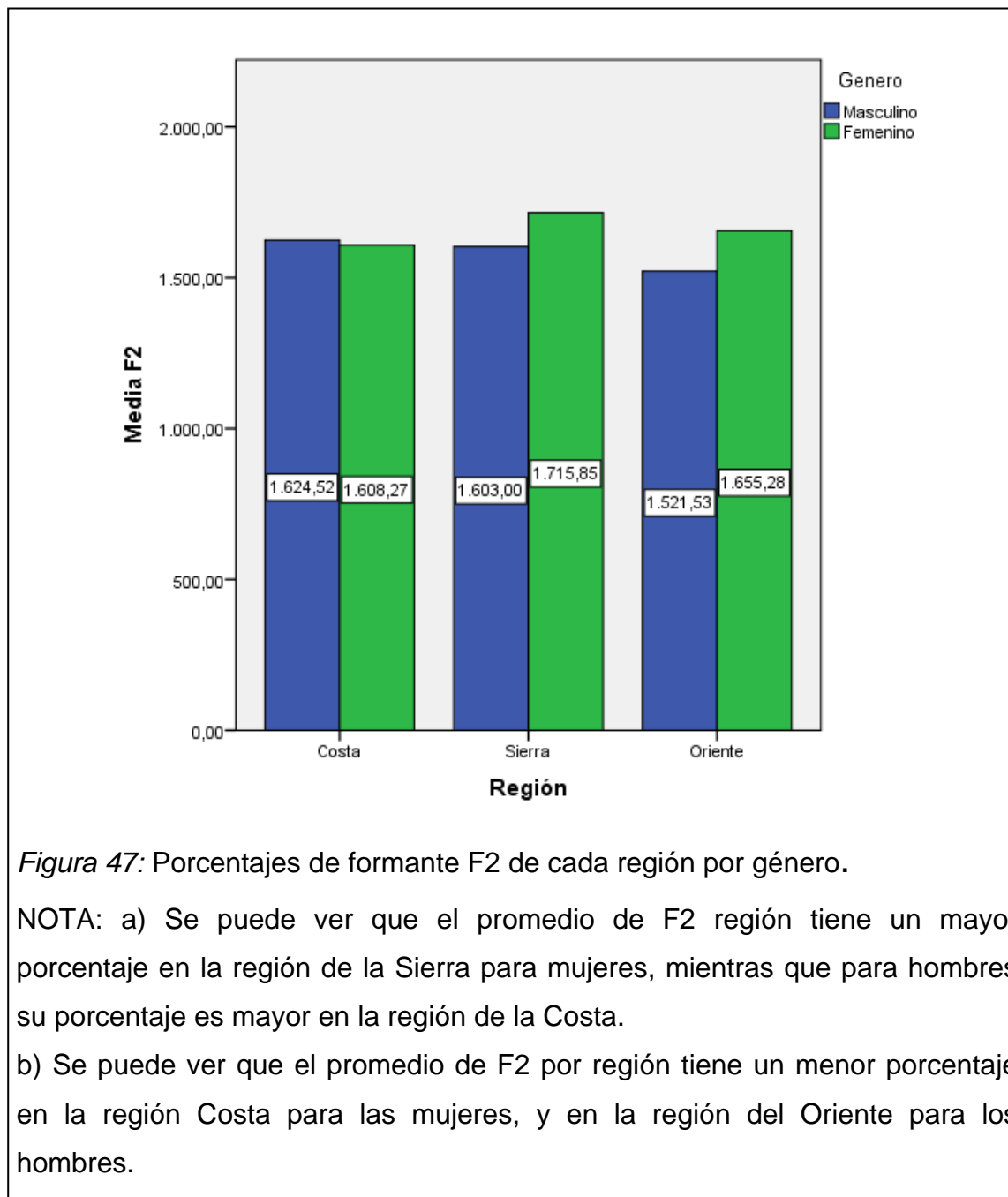
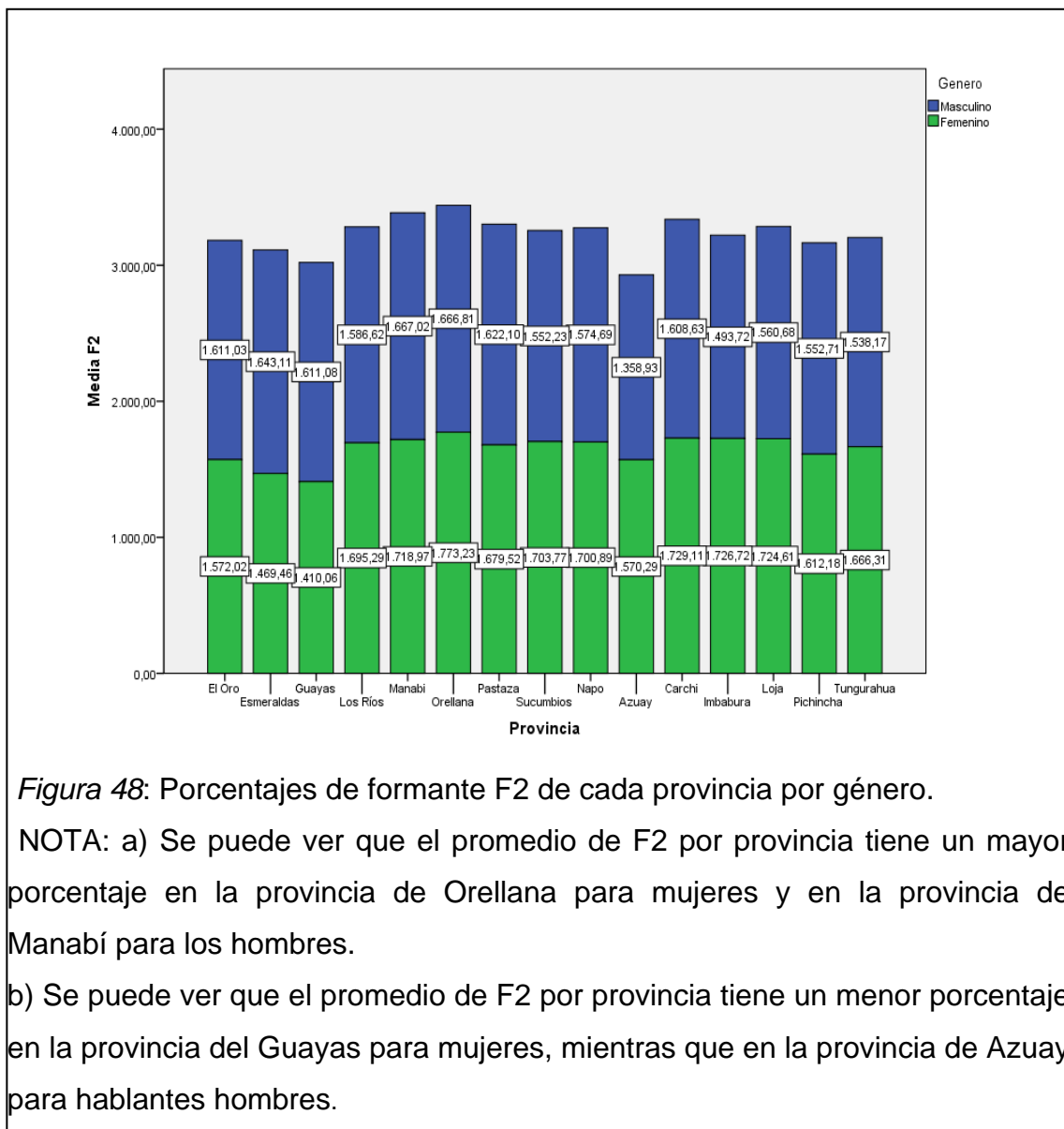


Figura 46: Porcentajes de formante F2 de cada palabra por género.

NOTA: a) Se puede ver que en el promedio de F2 tiene el mayor porcentaje en la palabra "Arcoíris" para mujeres, mientras que para hombres tiene un mayor porcentaje en la palabra "Chimpancé".

b) Se puede ver que en el promedio de F2 tiene un menor porcentaje en la palabra "Fútbol" para mujeres, mientras que para hombres tiene un mayor porcentaje en la palabra "Ahora".





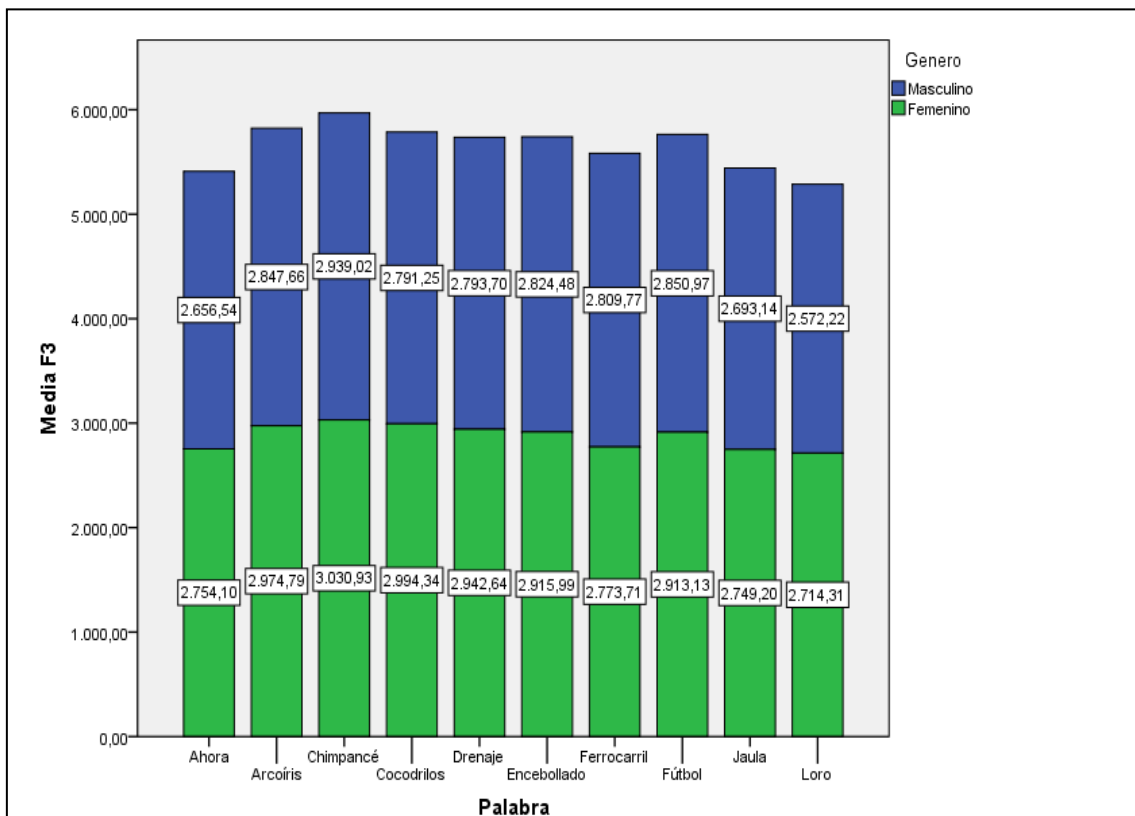
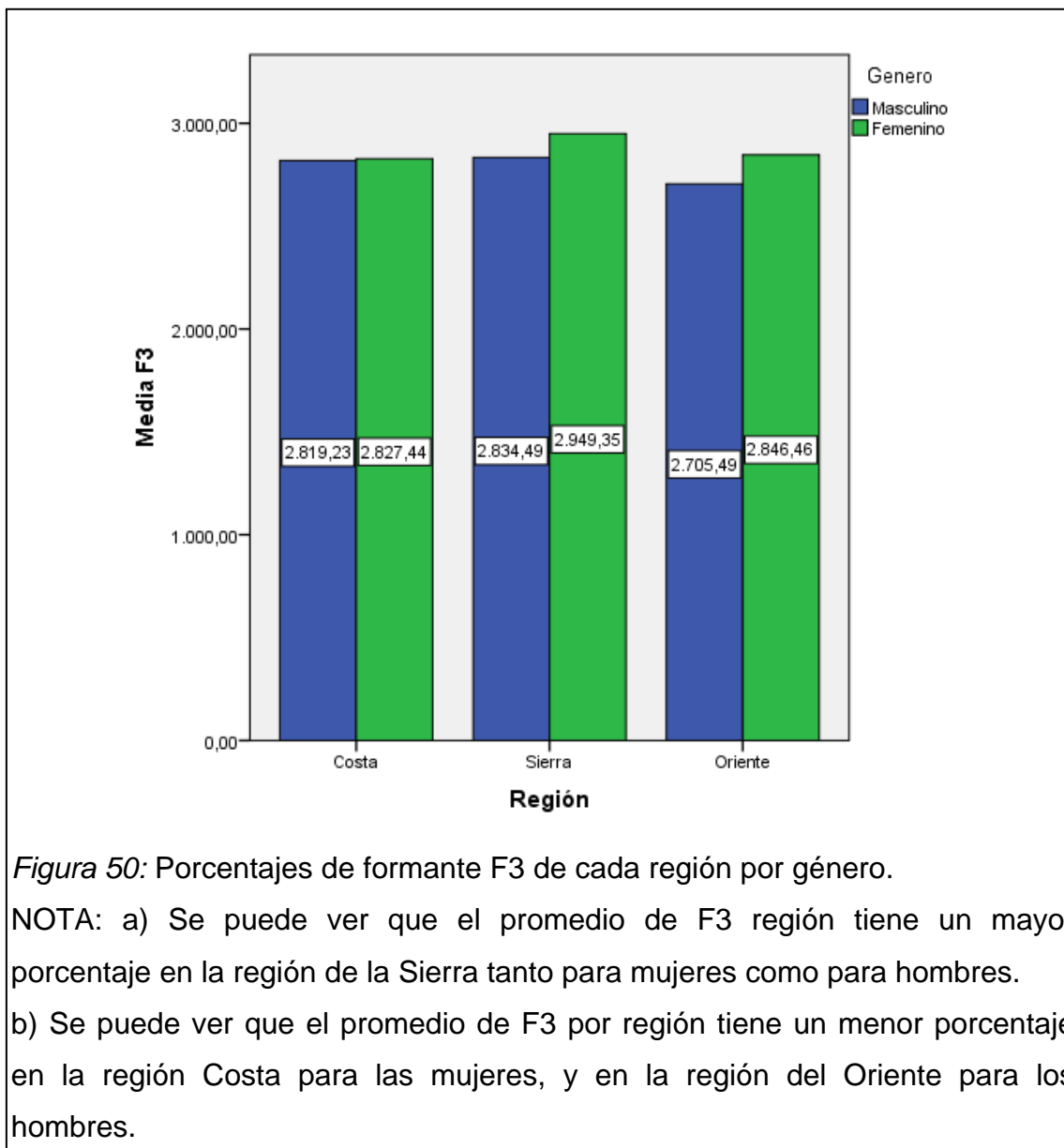


Figura 49: Porcentajes de formante F3 de cada palabra por género.

NOTA: a) Se puede ver que en el promedio de F3 tiene el mayor porcentaje en la palabra "Chimpancé" tanto para mujeres como para hombres.

b) Se puede ver que en el promedio de la F3 tiene un menor porcentaje en la palabra "Loro" tanto para mujeres como para hombres.



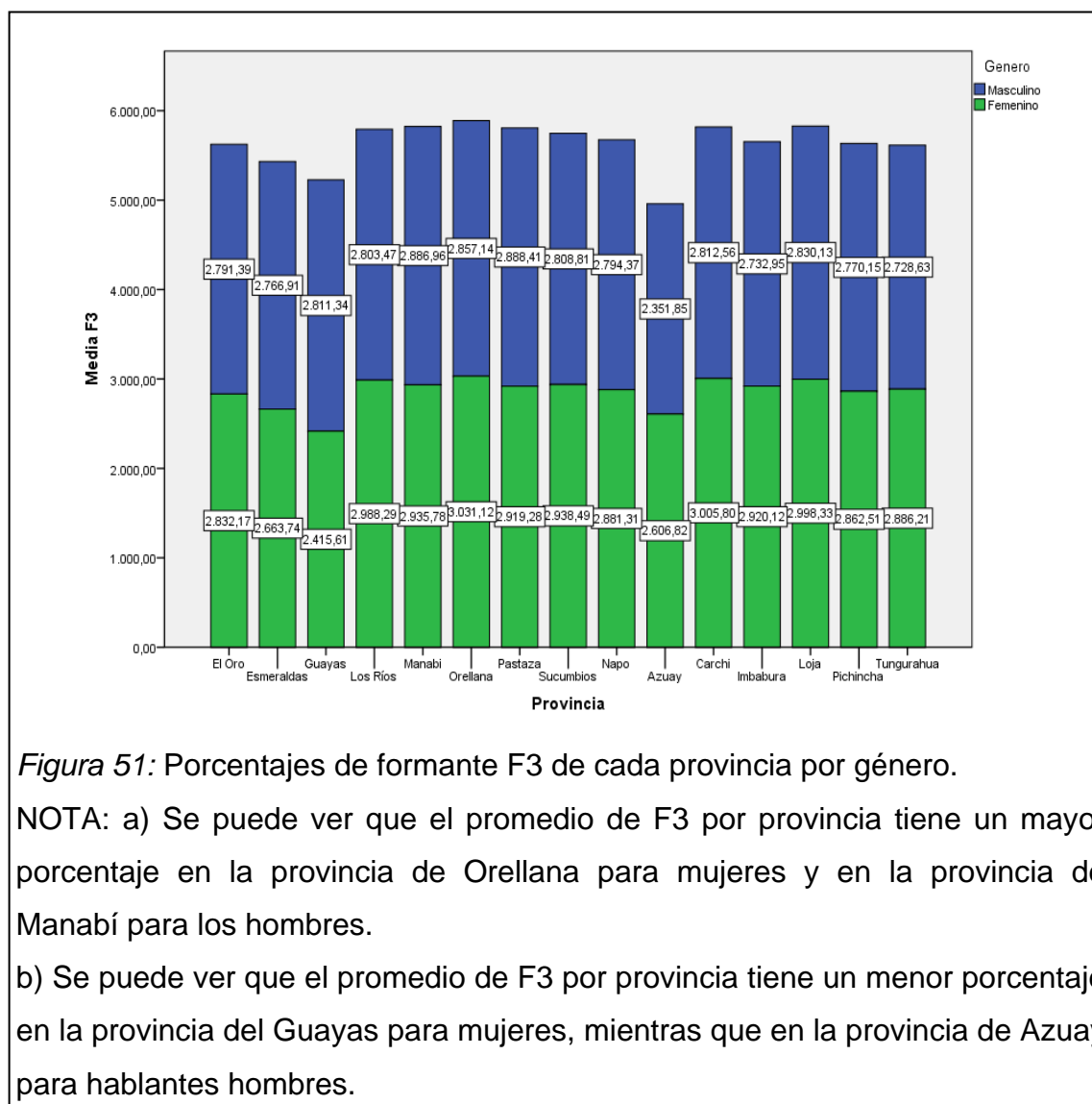
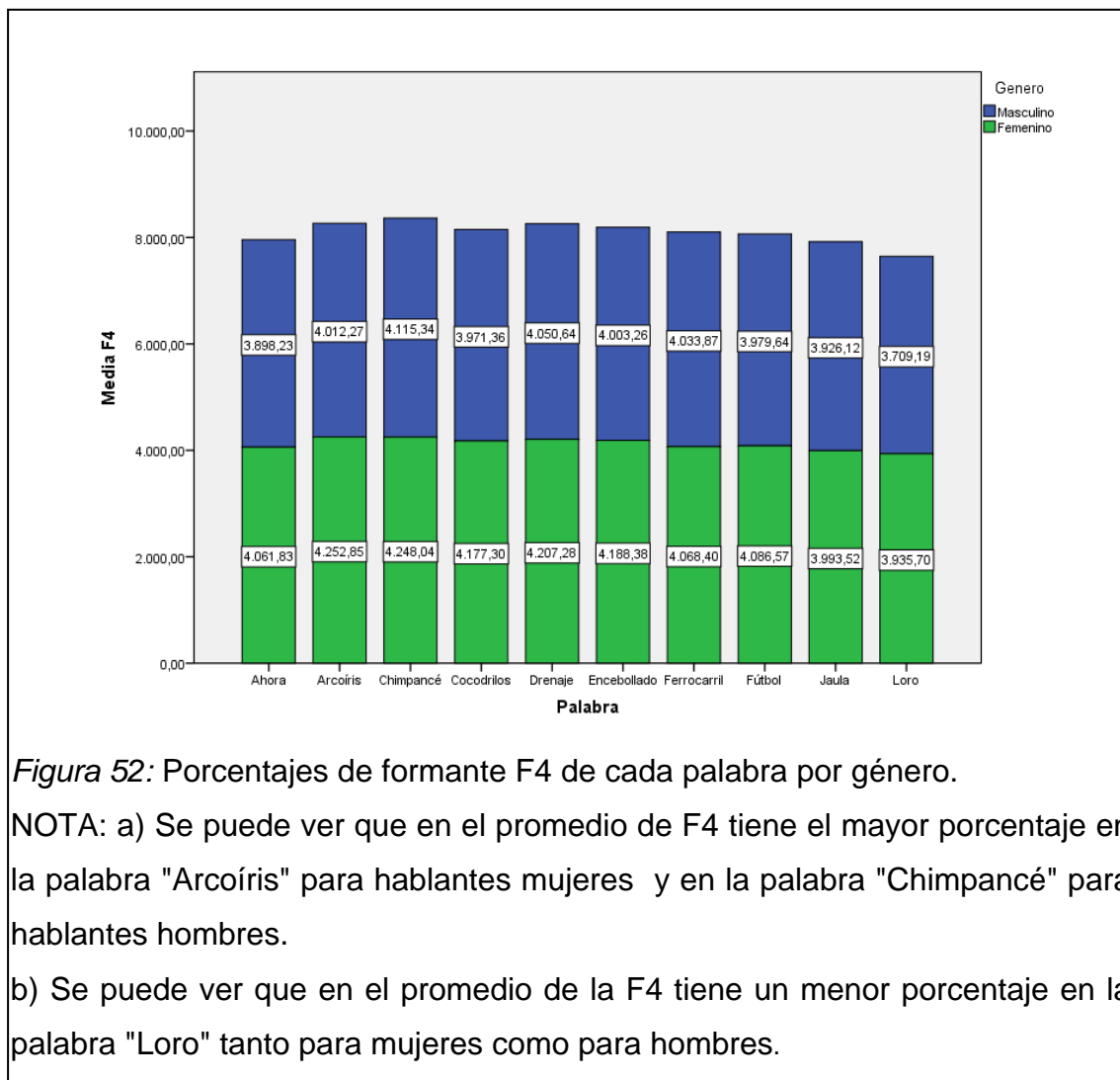


Figura 51: Porcentajes de formante F3 de cada provincia por género.

NOTA: a) Se puede ver que el promedio de F3 por provincia tiene un mayor porcentaje en la provincia de Orellana para mujeres y en la provincia de Manabí para los hombres.

b) Se puede ver que el promedio de F3 por provincia tiene un menor porcentaje en la provincia del Guayas para mujeres, mientras que en la provincia de Azuay para hablantes hombres.



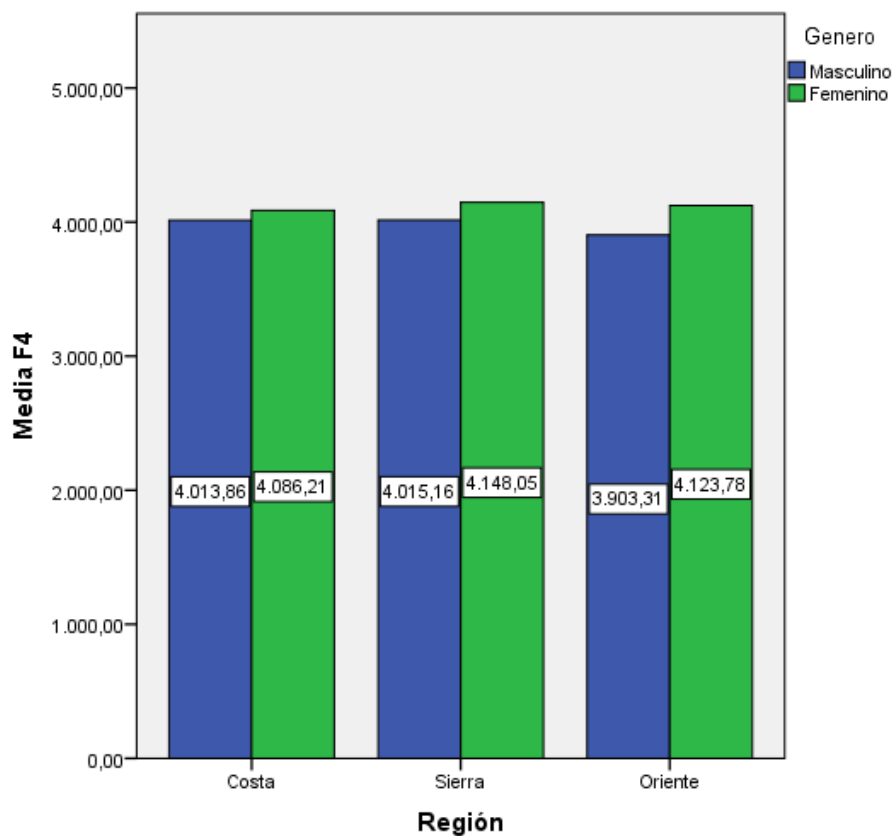


Figura 53: Porcentajes de formante F4 de cada región por género.

NOTA: a) Se puede apreciar que el promedio de F4 región tiene un mayor porcentaje en la región de la Sierra tanto para mujeres como para hombres.
 b) Se puede ver que el promedio de F4 por región tiene un menor porcentaje en la región Costa para las mujeres, y en la región del Oriente para los hombres.

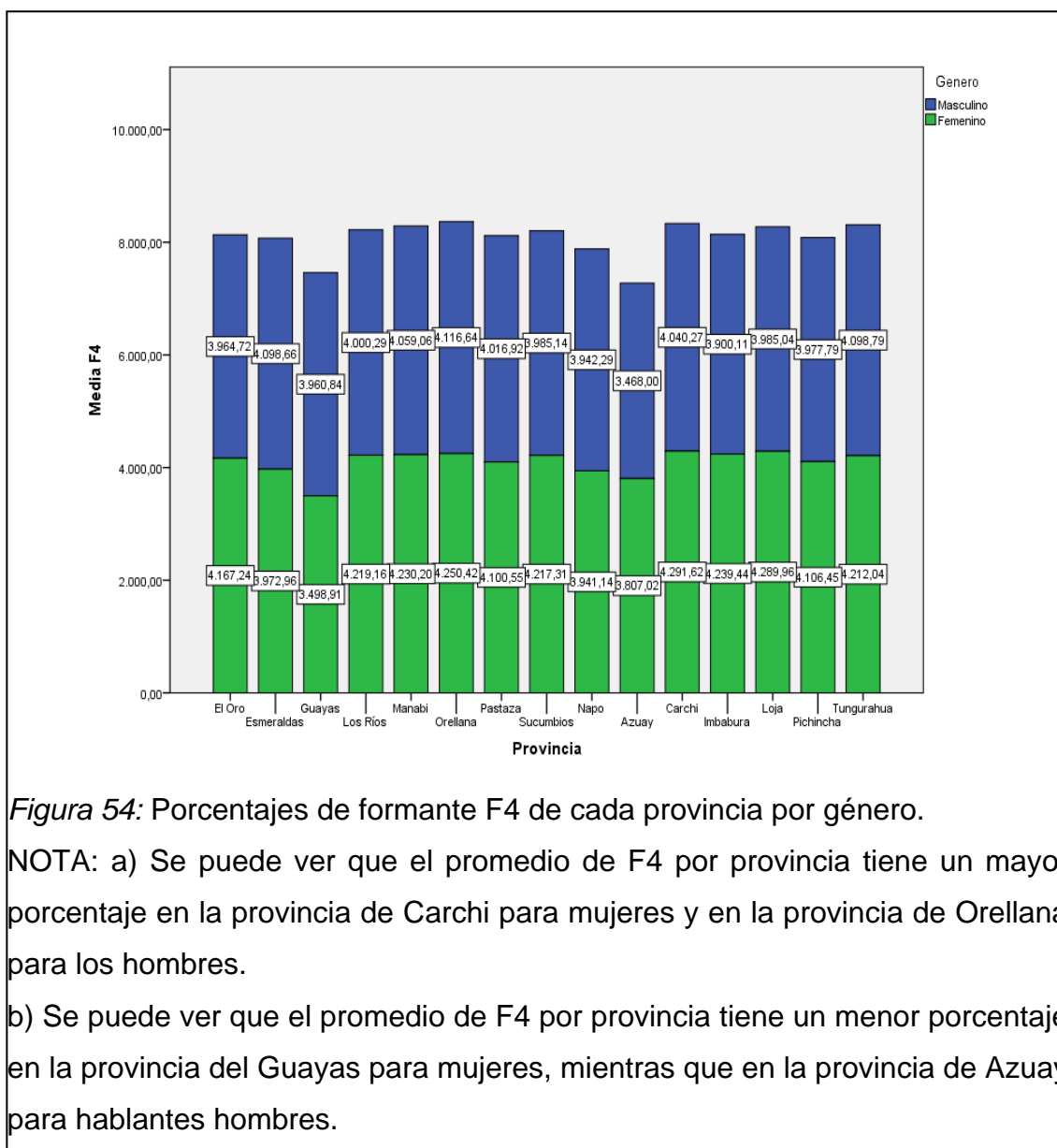


Figura 54: Porcentajes de formante F4 de cada provincia por género.

NOTA: a) Se puede ver que el promedio de F4 por provincia tiene un mayor porcentaje en la provincia de Carchi para mujeres y en la provincia de Orellana para los hombres.

b) Se puede ver que el promedio de F4 por provincia tiene un menor porcentaje en la provincia del Guayas para mujeres, mientras que en la provincia de Azuay para hablantes hombres.

Tabla 6: Promedio de las pausas por sílabas de cada palabra por región y género, en segundos.

Palabra	Ahora			Pausa1	Pausa2	Pausa3	
				Media	Media	Media	
		Costa	Masculino	,02			
			Femenino	,01			
		Sierra	Masculino	,01			
			Femenino	,02			
		Oriente	Masculino	,01			
			Femenino	,02			
		Arcoíris	Costa	Masculino	,09	,09	
				Femenino	,07	,07	
			Sierra	Masculino	,09	,09	
				Femenino	,10	,10	
			Oriente	Masculino	,12	,12	
				Femenino	,10	,10	
	Chimpancé	Costa	Masculino	,07	,07		
			Femenino	,06	,06		
		Sierra	Masculino	,08	,08		
			Femenino	,12	,12		
		Oriente	Masculino	,08	,08		
			Femenino	,08	,08		
	Cocodrilos	Costa	Masculino	,06	,02		
			Femenino	,06	,05		
		Sierra	Masculino	,06	,02		
			Femenino	,05	,01		
		Oriente	Masculino	,06	,04		
			Femenino	,07	,03		
Drenaje	Costa	Masculino	,06	,02			
		Femenino	,08	,04			
	Sierra	Masculino	,13	,06			
		Femenino	,10	,06			
	Oriente	Masculino	,08	,04			
Encebollado	Costa	Sierra	Masculino	,07	,01		
			Femenino	,07	,02		
		Oriente	Masculino	,08	,04		
			Femenino	,07	,05		

		Oriente	Masculino	,08	,00	
			Femenino	,07	,02	
	Ferrocarril	Costa	Masculino	,04	,07	,02
			Femenino	,03	,07	,02
		Sierra	Masculino	,03	,08	,01
			Femenino	,02	,07	,02
		Oriente	Masculino	,01	,09	,02
			Femenino	,02	,08	,07
	Fútbol	Costa	Masculino	,01		
			Femenino	,00		
		Sierra	Masculino	,00		
			Femenino	,01		
		Oriente	Masculino	,01		
			Femenino	,03		
	Jaula	Costa	Masculino	,01		
			Femenino	0,00		
		Sierra	Masculino	,01		
			Femenino	0,00		
		Oriente	Masculino	,00		
			Femenino	0,00		
Loro	Costa	Masculino	,00			
		Femenino	0,00			
	Sierra	Masculino	,00			
		Femenino	,00			
	Oriente	Masculino	,00			
		Femenino	,00			

NOTA: a) Se puede ver que las pausas dependen mucho del número de sílabas que esté compuesta la palabra a ser analizada.

b) Las pausas tienen una duración de pocos milisegundos, dependiendo la palabra.

c) Las pausas dependen mucho de la procedencia del hablante, es decir de la región y provincia que sean.

Tabla 7: Anova entre la frecuencia mínima y máxima en función a la región.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Frec_Inf	Entre grupos	10034,109	2	5017,055	1,965	,141
	Dentro de grupos	3132979,313	1227	2553,365		
	Total	3143013,422	1229			
Frec_Sup	Entre grupos	18637410,735	2	9318705,367	8,832	,000
	Dentro de grupos	1293589966,192	1226	1055130,478		
	Total	1312227376,927	1228			

NOTA: a) Se acepta la hipótesis que la frecuencia inferior tiene la misma media en las distintas regiones ya que la significancia es de 0.141 quiere decir mayor a 0.05, pero la frecuencia superior no tiene la misma media en distintas regiones ya que la significancia es 0, menor a 0.05.

Tabla 8: Anova entre la frecuencia mínima y máxima en función a la provincia.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Frec_Inf	Entre grupos	132653,59	14	9475,26	3,82	,000
	Dentro de grupos	3010359,83	1215	2477,66		
	Total	3143013,42	1229			
Frec_Sup	Entre grupos	95673194,32	14	6833799,59	6,82	,000
	Dentro de grupos	1216554182,61	1214	1002103,94		
	Total	1312227376,93	1228			

NOTA: a) Se rechaza la hipótesis que la frecuencia inferior y la superior tiene la misma media en distintas provincias ya que sus valores son 0.

Tabla 9: Anova entre la frecuencia mínima y máxima en función al género.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Frec_Inf	Entre grupos	94270,491	1	94270,491	37,971	,000
	Dentro de grupos	3048742,931	1228	2482,690		
	Total	3143013,422	1229			
Frec_Sup	Entre grupos	8505988,373	1	8505988,373	8,005	,005
	Dentro de grupos	1303721388,554	1227	1062527,619		
	Total	1312227376,927	1228			

NOTA: a) Se rechaza la hipótesis que la frecuencia inferior y la superior tiene la misma media en los distintos géneros, ya que sus valores son menores a 0,05.

Tabla 10: Anova entre la intensidad mínima, máxima e intensidad en función a la región.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Int_min	Entre grupos	223,828	2	111,914	,756	,470
	Dentro de grupos	181724,248	1227	148,105		
	Total	181948,077	1229			
Int_max	Entre grupos	339,771	2	169,885	1,714	,181
	Dentro de grupos	121588,703	1227	99,094		
	Total	121928,474	1229			
Intensidad	Entre grupos	253,929	2	126,965	,899	,407
	Dentro de grupos	173276,038	1227	141,219		
	Total	173529,967	1229			

NOTA: a) Se acepta la hipótesis que la intensidad mínima, máxima e intensidad tiene la misma media en distintas regiones, ya que los valores de la significancia es mayor a 0.05.

Tabla 11: Anova entre la intensidad mínima, máxima e intensidad en función a la provincia.

ANOVA						
		Suma de cuadrados	Gl	Media cuadrática	F	Sig.
Int_min	Entre grupos	27021,164	14	1930,083	15,136	,000
	Dentro de grupos	154926,913	1215	127,512		
	Total	181948,077	1229			
Int_max	Entre grupos	23306,327	14	1664,738	20,509	,000
	Dentro de grupos	98622,147	1215	81,170		
	Total	121928,474	1229			
Intensidad	Entre grupos	25226,430	14	1801,888	14,762	,000
	Dentro de grupos	148303,537	1215	122,061		
	Total	173529,967	1229			

NOTA: a) Se rechaza la hipótesis que la intensidad mínima, máxima e intensidad tiene la misma media en distintas provincias, ya que sus valores de significancia son menores a 0.05.

Tabla 12: Anova entre la intensidad mínima, máxima e intensidad en función al género.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Int_min	Entre grupos	343,586	1	343,586	2,323	,128
	Dentro de grupos	181604,490	1228	147,886		
	Total	181948,077	1229			
Int_max	Entre grupos	444,783	1	444,783	4,496	,034
	Dentro de grupos	121483,691	1228	98,928		
	Total	121928,474	1229			
Intensidad	Entre grupos	1265,119	1	1265,119	9,018	,003
	Dentro de grupos	172264,848	1228	140,281		
	Total	173529,967	1229			

NOTA: a) Se acepta la hipótesis que la intensidad mínima tiene la misma media en los distintos géneros ya que su valor es mayor a 0,05.

a) Se rechaza la hipótesis que la intensidad máxima e intensidad tienen la misma media en distintos géneros, ya que su significancia es menor a 0,05.

Tabla 13: Anova entre la pitch mínimo, máximo y pitch normal en función a la región.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Pitch_min	Entre grupos	4275,441	2	2137,721	,957	,384
	Dentro de grupos	2740767,569	1227	2233,714		
	Total	2745043,010	1229			
Pitch_max	Entre grupos	25804,664	2	12902,332	1,791	,167
	Dentro de grupos	8840255,876	1227	7204,773		
	Total	8866060,541	1229			
Pitch	Entre grupos	8489,012	2	4244,506	1,523	,218
	Dentro de grupos	3419236,437	1227	2786,664		
	Total	3427725,448	1229			

NOTA: a) Se acepta la hipótesis que el pitch mínimo, máximo y pitch normal tiene la misma media en distintas regiones, ya que sus valores de significancia es mayor a 0,05.

Tabla 14: Anova entre la pitch mínimo, máximo y pitch normal en función a la provincia.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Pitch_min	Entre grupos	188098,150	14	13435,582	6,384	,000
	Dentro de grupos	2556944,860	1215	2104,481		
	Total	2745043,010	1229			
Pitch_max	Entre grupos	426045,966	14	30431,855	4,381	,000
	Dentro de grupos	8440014,575	1215	6946,514		
	Total	8866060,541	1229			
Pitch	Entre grupos	369905,252	14	26421,804	10,498	,000
	Dentro de grupos	3057820,196	1215	2516,724		
	Total	3427725,448	1229			

NOTA: a) Se rechaza la hipótesis que el pitch mínimo, máximo y pitch normal tiene la misma media en distintas provincias, ya que su valor de significancia es menor a 0,05.

Tabla 15: Anova entre la pitch mínimo, máximo y pitch normal en función al género.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
Pitch_min	Entre grupos	971489,835	1	971489,835	672,655	,000
	Dentro de grupos	1773553,175	1228	1444,262		
	Total	2745043,010	1229			
Pitch_max	Entre grupos	2676603,671	1	2676603,671	531,043	,000
	Dentro de grupos	6189456,869	1228	5040,274		
	Total	8866060,541	1229			
Pitch	Entre grupos	2139664,280	1	2139664,280	2039,894	,000
	Dentro de grupos	1288061,168	1228	1048,910		
	Total	3427725,448	1229			

NOTA: a) Se rechaza la hipótesis que el pitch mínimo, máximo y pitch normal tiene la misma media en distintos géneros.

Tabla 16: Anova entre F1, F2, F3 Y F4 en función a la región.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
F1	Entre grupos	287065,669	2	143532,835	12,431	,000
	Dentro de grupos	14167432,200	1227	11546,400		
	Total	14454497,870	1229			
F2	Entre grupos	5314380842,811	2	2657190421,405	1,133	,322
	Dentro de grupos	2877064226969,150	1227	2344795621,002		
	Total	2882378607811,960	1229			
F3	Entre grupos	3783079,379	2	1891539,690	15,799	,000
	Dentro de grupos	146898779,162	1227	119721,906		
	Total	150681858,541	1229			
F4	Entre grupos	2014369,855	2	1007184,927	5,617	,004
	Dentro de grupos	220029077,860	1227	179322,802		
	Total	222043447,714	1229			

NOTA: a) Se rechaza la hipótesis que la F1, F3 y F4 tienen la misma media en distintas regiones ya que su significancia es menor a 0,05, mientras se acepta la hipótesis que F2 tiene la misma media en distintas regiones ya que su significancia es mayor a 0,05.

Tabla 17: Anova entre F1, F2, F3 Y F4 en función a la provincia.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
F1	Entre grupos	851973,751	14	60855,268	5,436	,000
	Dentro de grupos	13602524,118	1215	11195,493		
	Total	14454497,870	1229			
F2	Entre grupos	45228693656,560	14	3230620975,469	1,384	,153
	Dentro de grupos	2837149914155,400	1215	2335102810,004		
	Total	2882378607811,960	1229			
F3	Entre grupos	17798555,225	14	1271325,373	11,624	,000
	Dentro de grupos	132883303,316	1215	109368,974		
	Total	150681858,541	1229			
F4	Entre grupos	26993287,660	14	1928091,976	12,010	,000
	Dentro de grupos	195050160,054	1215	160535,111		
	Total	222043447,714	1229			

NOTA: a) Se rechaza la hipótesis que la F1, F3 y F4 tienen la misma media en distintas provincias ya que su significancia es menor a 0.05 y se acepta la hipótesis que F2 tiene la misma media en distintas provincias, ya que su significancia es mayor a 0,05.

Tabla 18: Anova entre F1, F2, F3 Y F4 en función a la región.

ANOVA						
		Suma de cuadrados	gl	Media cuadrática	F	Sig.
F1	Entre grupos	816491,461	1	816491,461	73,519	,000
	Dentro de grupos	13638006,409	1228	11105,868		
	Total	14454497,870	1229			
F2	Entre grupos	3344997621,061	1	3344997621,061	1,427	,233
	Dentro de grupos	2879033610190,900	1228	2344489910,579		
	Total	2882378607811,960	1229			
F3	Entre grupos	4921769,899	1	4921769,899	41,465	,000
	Dentro de grupos	145760088,642	1228	118697,141		
	Total	150681858,541	1229			
F4	Entre grupos	10605852,882	1	10605852,882	61,597	,000
	Dentro de grupos	211437594,832	1228	172180,452		
	Total	222043447,714	1229			

NOTA: a) Se rechaza la hipótesis que la F1, F3 y F4 tienen la misma media en distintos géneros, ya que su significancia es menor a 0,05 mientras se acepta la hipótesis que F2 tiene la misma media en distintos géneros, ya que tiene una significancia mayor a 0,05. Ver anexo Base de Datos.

6. PROYECCIONES:

Para un futuro se podría realizar una toma de voces bajo ciertos parámetros:

- Sería ideal tener una sala totalmente adecuada para poder grabar voces bajo las mismas condiciones y también es muy importante tener los mismos equipos para esto.
- La base de datos que se puede obtener nos sirve para varios proyectos venideros y a la Policía Nacional también, ya que ayuda a la hora de tener varias voces grabadas en cada provincia.
- Se puede también hacer proyectos a un futuro grabando a un numero X de personas de las mismas edades y bajo las mismas condiciones de grabación, para poder analizar las voces y ver sus resultados.
- Con un programa más profesional como los que manejan en la policía Nacional se podría sacar más características de las voces, ya no solo de la palabra sino de oraciones completas.

7. CONCLUSIONES Y RECOMENDACIONES:

7.1. Conclusiones:

1. Se concluye que en la Costa y Sierra tanto los hombres como las mujeres tienen la frecuencia inferior en la palabra "Chimpancé" mientras que las frecuencia superior se encuentra en la palabra "Loro", y en el Oriente para los ambos géneros la frecuencia inferior se encuentra en la palabra "Arcoíris"; la frecuencia superior se encuentra en la palabra "Cocodrilos" para los hombres y para las mujeres en la palabra "Encebollado".

Tabla 19: Frecuencias inferior y superior por región y palabra.

Región	Género	Freq. Inf. (Hz)	Palabra	Freq. Sup. (Hz)	Palabra
Costa	Masculino	55,18	Chimpancé	4902,48	Loro
	Femenino	53,93	Chimpancé	4848,31	Loro
Sierra	Masculino	63,04	Chimpancé	4922,89	Loro
	Femenino	69,29	Chimpancé	4893,78	Loro
Oriente	Masculino	72,52	Arcoíris	4844,04	Cocodrilos
	Femenino	79,87	Arcoíris	4815,14	Encebollado

2. Se concluye en la tabla que los rangos de frecuencia inferior y superior son mayores para las mujeres en las 3 regiones del Ecuador.

Tabla 20: Rangos de frecuencias con respecto a región y género.

Región	Género	Freq. Inf. (Hz)	Freq. Sup. (Hz)
Costa	Masculino	55,18 - 147,61	1858,43 - 4902,48
	Femenino	53,93 - 196,21	2261,29 - 4848,31
Sierra	Masculino	63,04 - 166,93	2242,42 - 4922,89
	Femenino	64,22 - 180,59	2878,61 - 4893,78
Oriente	Masculino	72,52 - 129,14	3960,86 - 4844,04
	Femenino	79,87 - 160,80	4033,38 - 4815,14

3. Se aprecia que para los hombres la intensidad mínima en las 3 regiones se da en la palabra "Arcoíris" y para las mujeres en la Costa se da en "Chimpancé", en la Sierra y Oriente en "Ferrocarril"; la intensidad máxima para los hombres en la Costa y Sierra tiene la palabra "Jaula" mientras que en el Oriente en la palabra "Loro" y para las mujeres en Costa y Oriente en la palabra "Fútbol" y en la palabra "Loro" en la Sierra; y para el parámetro de Intensidad para los hombres para la Costa y Sierra se da en la palabra "Loro" y en la palabra "Ahora" para el Oriente, mientras que para mujeres se da en la palabra "Ahora" para la Costa y Oriente y para la Sierra en la palabra "Fútbol".

Tabla 21: Intensidad mínima, máxima y normal por región y palabra.

Región	Género	Int. Min. (dB)	Palabra	Int. Max. (dB)	Palabra	Int. (dB)	Palabra
Costa	Masculino	26,66	Arcoíris	61,1	Jaula	57,7	Loro
	Femenino	27,77	Chimpancé	62,7	Fútbol	58,7	Ahora
Sierra	Masculino	25,78	Arcoíris	60,6	Jaula	56,7	Loro
	Femenino	27,32	Ferrocarril	60,7	Loro	66,6	Fútbol
Oriente	Masculino	28,12	Arcoíris	60,3	Ahora	59,6	Ahora
	Femenino	27,33	Ferrocarril	62,9	Fútbol	59,3	Ahora

4. Se observa en la tabla que los rangos de intensidad inferior, superior y normal son mayores para las mujeres en las 3 regiones del Ecuador, la diferencia no es mucha pero si se analiza auditivamente la diferencia de dB es mínima, no es notoria, pero si hay niveles iguales se le debe sumar 3dB al valor anterior, mientras que para los niveles que difieran entre 15 dB o más, la suma sería aproximadamente igual al más alto de los dos.

Tabla 22: Rangos de intensidades con relación al género y región.

Región	Género	Int. Inf. (Hz)	Int. Sup. (Hz)	Int. (Hz)
Costa	Masculino	26,66 - 44,99	58 - 61,1	52,8 - 57,7
	Femenino	27,77 - 47,91	56,6 - 62,7	54,3 - 58,6
Sierra	Masculino	25,78 - 46,41	56,6 - 60,6	50,3 - 57
	Femenino	27,32 - 49,72	56,5 - 60,7	52 - 66,6
Oriente	Masculino	28,12 - 46,43	48,5 - 60,3	45,6 - 56,9
	Femenino	27,32 - 49,72	56,6 - 62,2	53,8 - 59,3

5. Se puede concluir que el pitch mínimo se encuentra con mayor porcentaje en la palabra "Cocodrilos" tanto para hombres como para mujeres para la Costa y la Sierra y para el Oriente en hombres el pitch mínimo se encuentra en la palabra "Loro" y en mujeres en la palabra "Cocodrilos"; el pitch máximo se encuentra para los hombres en la Costa en la palabra "Jaula", en la Sierra en la palabra "Arcoíris" y en el Oriente "Cocodrilos", mientras que para las mujeres en la Costa tiene el pitch máximo en la palabra "Ferrocarril", para la Sierra en la palabra "Cocodrilos" y en el Oriente en la palabra "Arcoíris"; y para los hombres el pitch se encuentra en la Costa en la palabra "Jaula", para la Sierra en la palabra "Arcoíris y Fútbol" y para el Oriente en la palabra "Fútbol" , mientras que para las mujeres se encuentra en la palabra "Ahora".

Tabla 23: Pitch normal, inferior y superior con relación al género y región.

Región	Género	Pitch. Min. (Hz)	Palabra	Pitch. Max. (Hz)	Palabra	Pitch. (Hz)	Palabra
Costa	Masculino	91,51	Cocodrilos	175,46	Jaula	139,98	Jaula
	Femenino	120,85	Cocodrilos	288,8	Ferrocarril	236,85	Fútbol
Sierra	Masculino	94,12	Cocodrilos	215,84	Arcoíris	135,96	Arcoíris/ Fútbol
	Femenino	135,96	Cocodrilos	296,2	Cocodrilos	224,09	Fútbol
Oriente	Masculino	89,69	Loro	213,35	Cocodrilos	144,34	Fútbol
	Femenino	143,87	Cocodrilos	271,08	Arcoíris	244,14	Ahora

6. Se concluye que en los niveles de pitch tanto para mínimo, máximo y normal en el género femenino se tiene valores mucho mayores que los de los hombres.

Tabla 24: Rangos de pitch con relación al género y región.

Región	Género	Pitch. Min. (Hz)	Pitch.Max. (Hz)	Pitch. Min. (Hz)
Costa	Masculino	91,51 - 117,38	151,63 - 175,40	121,47 - 139,98
	Femenino	120,85 - 201,42	265,55 - 228,80	195,58 - 236,85
Sierra	Masculino	94,12 - 112,68	152,51 - 215,84	126,32 - 135,96
	Femenino	120,85 - 182,69	249,63 - 296,20	197,62 - 224,09
Oriente	Masculino	96,01 - 118,29	131,85 - 213,35	112,97 - 144,34
	Femenino	143,87 - 264,41	240,47 - 271,08	195,69 - 244,14

7. Se observa que los espacios o pausas de cada palabra dependen del número de sílabas que tiene cada una de ellas, es decir de cuan largo es la palabra, de los acentos y de la dificultad que éstas dan al pronunciar una palabra dependiendo la región que se encuentren.
8. Comparando a los hablantes se puede observar que los resultados de los espacios o pausas de las palabras son independientes de género y región.
9. Se puede ver que la comparación entre las mismas personas de una región o provincia tienen el mismo silencio o pausa en el espacio temporal de la palabra.
10. Esta base de datos puede ayudar a la Policía Nacional para su trabajo en el área de reconocimiento de voces y acústica forense, para poder distinguir voces en función de género, provincia y región.
11. La herramienta PRAAT que es un sistema gratuito en este caso nos ayudó a distinguir diferentes características necesarias para caracterizar una voz, como: La frecuencia mínima y máxima, Pitch, Intensidad y formantes.

7.2. Recomendaciones

- Las grabaciones deben tener exactamente las mismas condiciones de equipo y de sala para que sean ideales para este tipo de análisis.
- Se recomienda para próximos trabajos que se realicen de este tipo, tener en cuenta que el software sea mucho más avanzado que el utilizado.

- Para que este trabajo sea de mayor ayuda se debería crear un base de voces más extensa, y con mejor calidad de grabación, para de esta manera poder analizar y obtener más características de la voz.

REFERENCIAS

- 4 ujaen. (s.f.). *Transformada de Fourier*. Recuperado el 13 de Febrero de 2015 de http://www4.ujaen.es/~jalmira/transformada_fourier_almira.pdf
- Arellano, A. (2007). *Técnicas de Grabación y Reproducción del Sonido*. Recuperado el 12 de Mayo de 2015 de <http://tesis.ipn.mx/xmlui/bitstream/handle/123456789/6201/ICE105.pdf?sequence=1>
- Arte Sonoro. (2013.). *Rango dinámico ¿Qué es y para qué sirve?*. Recuperado el 25 de Agosto de 2014 de <http://www.artesonoro.com.mx/2013/06/rango-dinamico-que-es-y-para-que-sirve.html>
- Astormastering. (s.f.). *Criterios de Selección de Micrófonos*. Recuperado el 18 de Marzo de 2015 de http://www.astormastering.com.ar/Clase_11_Criterios_de_Seleccion_de_Microfonos.pdf
- Biblioteca Cenace. (s.f.). *Diseño de Salas de Locución y Acondicionamiento Acústico*. Recuperado el 15 de Octubre de 2014 de http://biblioteca.cenace.org.ec/jspui/bitstream/123456789/1008/12/Mora_Luis_Enrique.PDF
- Es scribd. (s.f.). *Espectro de la Voz Humana y la Música*. Recuperado el 3 de Abril de 2015 de <http://es.scribd.com/doc/213944701/Espectro-de-La-Voz-Humana-y-La-Musica>
- Elastixtech. (s.f.). *Transmisión de la Voz*. Recuperado el 15 de Marzo de 2015 de <http://elastixtech.com/fundamentos-de-telefonía/transmision-de-la-voz/>
- Educación. (2008). *Diseño de Materiales Multimedia*. Recuperado el 9 de Septiembre de 2014 de [Formatos de archivos de audio. http://www.ite.educacion.es/formacion/materiales/107/cd/audio/audio0102.html](http://www.ite.educacion.es/formacion/materiales/107/cd/audio/audio0102.html)
- Escuela de cine online nucine. (s.f.). *Fundamentos de Micrófonos*. Recuperado el 15 de Octubre de 2014 de <http://www.escueladecineonline.nucine.com/tecno-audio2.htm>

- Elai upm. (s.f.). *Introducción al Procesamiento Digital de Señales*. Recuperado el 19 de Mayo de 2015 de http://www.elai.upm.es/webantigua/spain/Publicaciones/pub01/intro_proc_sdi_g.pdf
- lie fing edu uy. (s.f.). *Formantes*. Recuperado el 29 de Abril de 2015 de <http://ie.fing.edu.uy/investigacion/grupos/gmm/audio/seminario/seminariosviejos/2003/charlas/charla1/voz11.htm>
- García, D. (s.f.). *La Frecuencia Fundamental de la Voz y sus Efectos en Reconocimiento de Habla Continua*. Recuperado el 2 de Diciembre de 2014 de <http://www.sepln.org/revistaSEPLN/revista/26/garcia.pdf>
- Gómez, E. (2009). *Representación de Señales de Audio*. Recuperado el 12 de Septiembre de 2014 de <http://www.dtic.upf.edu/~egomez/teaching/sintesi/SPS1/Tema3representacion.pdf>
- Harris, C. (1991). *Handbook of Acoustical Measurements and Noise Control*. Tercera Edición, Acoustical Society of America (A.S.A.).
- Introsis blogspot. (s.f.). *Introducción a Sistemas de Audio*. Recuperado el 11 de Febrero de 2015 de <http://introsis.blogspot.com/2007/06/rango-dinmico-nivel-nominal-snr.html>
- ICKROM. (s.f.). *Clasificación de los Micrófonos*. Recuperado el 20 de Enero de 2015 de http://www.ickrom.com.mx/info/informacion_tecnica/clasificacion_de_microfonos.php
- Labrada, J. (1995). *El registro sonoro*, Primera Edición, Editorial Voluntad S.A.
- Martín, J. (2006). *Analizadores de Espectros. Tipos [Ii]*. El analizador FFT. Revista española de electrónica, ISSN 0482-6396, Nº 619
- Manzero A. (2010). *Memoria de Testigos*. Recuperado el 18 de Diciembre de 2014 de <http://www.casadellibro.com/libro-memoria-de-testigos-obtencion-y-valoracion-de-la-prueba-testific-al/9788436823516/1658772>
- Miyara. F. (1999). *Acústica y Sistemas de Sonido*. Tercera Edición, Editora de la Universidad Nacional de Rosario
- Papoulis, A. (1985). *Sistemas digitales y analógicos, transformadas de Fourier, estimación espectral*. Recuperado el 23 de Septiembre de 2014 de .

- Pino, P., Díaz, J., Jiménez, C y Rothman, H. (2004). *Revista Ingeniería*. Recuperado el 24 de Mayo de 2015 de Identificación de algunos parámetros espectrales que determinan la calidad de la voz UC, Vol. 11
- Quilis, A. (1988). *Fonética Acústica de la Lengua Española*. Recuperado el 10 de Abril de 201 de <http://www.casadellibro.com/libro-fonetica-acustica-de-la-lengua-espanola/9788424901318/436297>
- Román, D. (1992). *Manual Praat*. Recuperado el 10 de Julio de 2014 de http://domingo-roman.net/manual_analisis_acustico.html
- Shure. (s.f.). *Microfonos: Patrones Polares / Direccionalidad*. Recuperado el 16 de Agosto de 2014 de http://www.shure.es/asistencia_descargas/contenido-educativo/microfonos/microphone_polar_patterns
- Taller de audio. (s.f.). *Audio Digital*. Recuperado el 15 de Noviembre de 2014 de http://cmm.cenart.gob.mx/tallerdeaudio/cursos/cursoardour/Teoria_y_tecnicas/Audiodigital.html
- Tartamudez pro (s.f.). *Formante*. Recuperado el 17 de Marzo de 2015 de <http://www.tartamudez.pro/2009/05/formante.html>
- Unad edu. (s.f.). *Ventajas del Procesado Digital de Señales*. Recuperado el 25 de Octubre de http://datateca.unad.edu.co/contenidos/299004/Ventajas_PDS.pdf
- Unimúsica Perú. (s.f.). *Academia Unisono*. Recuperado el 2 de Julio de 2014 de Fundamentos Básicos On-Line. de http://www.unimusicaperu.com/produccion_musical_profundidad_bits.htm
- UV. (s.f.). *Introducción al Procesado Digital de Señales*. Recuperado el 18 de Diciembre de 2014 de http://www.uv.es/soriae/tema_1_pds.pdf
- Uva. (s.f.). *Micrófonos*. Recuperado el 13 de Mayo de 2015 de http://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_04_05/io8/public_html/microfonos.htm
- Yio. (s.f.). *Profundidad de Bits y Frecuencia de Muestreo*. Recuperado el 26 de Abril de 2015 de <http://www.yio.com.ar/contenido.php?artid=57>

ANEXOS