



VIGILADA MINEDUCACIÓN

Tesis de Maestría en Ingeniería Electrónica

SISTEMA DE MEDICIÓN Y BIOREALIMENTACIÓN
PARA EL MANEJO DE LA ANSIEDAD-ESTADO EN
ESTUDIANTES DEL ÁREA DE LA SALUD

Jesús Ernesto Moreno Rizo

Tutor:

Dr. Juan López López

Co-Tutor:

Dra. Laura Andrea León

Co-Tutor:

Msc. Sandra Liliana Cancino

Una tesis presentada en cumplimiento de los requisitos para el grado de
Maestría en Ingeniería Electrónica

Julio 2020

*“Inteligencia más carácter, ese es el objetivo
de la verdadera educación.”*

Martin Luther King Jr.

Agradecimientos

Es para mí importante destacar que este trabajo no hubiese sido posible sin el apoyo de un gran grupo de personas, familiares y amigos. En primer lugar, me gustaría agradecer a mis papás Ernesto Moreno y Claudia Patricia Rizo de Moreno, ya que su apoyo, su sabiduría y amor incondicional han hecho de mí la persona que soy. Adicionalmente, me gustaría agradecer a mis directores de tesis, que me han guiado a lo largo de este arduo pero agradable proceso enseñándome a ver el lado positivo de la vida y a ser mejor cada día.

Abstract

Hoy en día, el estrés se ha convertido en un hecho habitual de la vida del ser humano, ya que es común al presentarse en todos los entornos laborales en mayor o menor medida, y ciertamente, se ha establecido que las condiciones laborales tienen un impacto significativo en la salud de los empleados [1]. La exposición a condiciones laborales adversas que genera este tipo de respuestas de forma crónica, puede desencadenar el comúnmente conocido “burnout” [2], el cual puede acarrear consecuencias devastadoras para el trabajador, su familia, el ambiente laboral y la organización, [3]. No obstante, el deterioro hacia esta condición extrema, comprende un tiempo considerable durante el cual el sujeto ha experimentado la ansiedad-estado. Los efectos de la ansiedad-estado resultan de particular importancia cuando la población afectada involucra al personal médico asistencial. El estudio reportado por Richardsen y Burke [4] que relaciona el estrés en los profesionales de la salud con su nivel de desempeño, reveló que las consecuencias de la prevalencia de altos niveles de estrés se revierten en un desempeño profesional inadecuado hacia sus pacientes. Similarmente, Daniel S. Tawfik, y colegas [5] reportaron que, el agotamiento y el bajo bienestar han sido reconocidos como riesgos laborales comunes entre los profesionales de la salud. Por lo tanto, el proveer herramientas tecnológicas que faciliten la identificación y adecuado manejo de la ansiedad-estado se constituye como una estrategia de promoción y prevención ante esta creciente problemática. En particular, el introducir estas herramientas a los estudiantes del área de la salud tiene el potencial de aumentar los niveles de autocuidado tanto durante sus etapas de formación como de su ejercicio profesional.

Mediante la integración del diseño de software y el procesamiento de la señal de voz fue posible generar una herramienta de biorealimentación de la ansiedad-estado a partir de la caracterización de la señal de voz en el dominio del tiempo y frecuencia. Este desarrollo fue validado mediante una comparación entre los resultados obtenidos de dos tipos de pruebas psicométricas autoevaluativas ampliamente usadas (SAM e IDARE) antes y después de la biorealimentación. Se obtuvieron diferencias estadísticamente significativas entre las mediciones previas y posteriores a la biorealimentación mostrando una tendencia a la disminución asociada a dicha intervención. Lo anterior da lugar a afirmar que:

1. Es posible identificar de manera objetiva la ansiedad-estado de manera menos intru-

siva que los sistemas de medición alámbricos comerciales actuales, sin incurrir en la subjetividad de las herramientas autoevaluativas tradicionales.

2. La biorealimentación a partir de la medición de características no lingüísticas de la señal de voz se constituye como una estrategia viable para el manejo de esta condición.
3. Para una adecuada identificación de la ansiedad-estado se debe prestar particular atención a aquellas características no lingüísticas de la señal de voz que describan la variación de la banda de frecuencias del hablante sin dejar de lado aquellas que aporten información del contexto fonético del mismo. En particular las características de tono, centroides, asimetría y pendiente espectral acompañadas de la sonoridad integrada y MFCC resultan de particular interés.

Keywords: Ansiedad-estado, SAM, IDARE, Biorealimentación, Bioseñales del estrés, Sistema para el autocuidado.

Índice de cuadros

2.1. Características No Lingüísticas de La Voz Utilizadas en la Detección de Estrés	12
3.1. Base de datos generada	19
4.1. Resultados pruebas estadísticas entre grupos	49
4.2. Resultados pruebas estadísticas intra sujetos (LB vs MPI) - Etapa TN . . .	52
4.3. Resultados pruebas estadísticas intra sujetos (LB vs MPI) - Etapa P	53

Índice de figuras

2.1. Representación del mecanismo de producción del habla humana	10
2.2. Representación gráfica del SAM	15
3.1. Etapas principales de la investigación	19
3.2. Generación de registro del participante	21
3.3. Interfaz de usuario	23
3.4. Esquema general de procesamiento de audio	24
3.5. Ilustración de la superposición de ventana	25
3.6. Esquema de eliminación de silencios y segmentación del discurso	27
3.7. Señal de voz segmentada	28
3.8. Cálculo de los MFCC	33
3.9. Esquema general del cálculo de la Sonoridad Integrada	35
3.10. Pipeline general para la extracción de las características de voz	36
4.1. Resultado del filtrado	41
4.2. Resultado de eliminación de silencios	42
4.3. Segmentación del discurso	43
4.4. Arquitectura de la estructura de los datos obtenidos para el registro de TN y P	44
4.5. Variación porcentual Texto Neutro	45
4.6. Variación porcentual Palabras	47

4.7. Comparación entre grupos del tono para registro durante la etapa de línea base (LB-TN)	54
4.8. Comparación entre grupos del tono para registro durante la etapa de medición post intervención (MPI-TN)	55
4.11. Comparación intra sujeto G1 del tono para registro LB y MPI de TN	55
4.9. Comparación entre grupos del tono para registro durante la etapa de línea base (LB-P)	56
4.12. Comparación intra sujeto G1 del tono para registro LB y MPI de Palabras	56
4.10. Comparación entre grupos del tono para registro durante la etapa de medición post intervención (MPI-P)	57
4.13. Comparación intra sujeto G2 del tono para registro LB y MPI de TN	58
4.14. Comparación intra sujeto G2 del tono para registro LB y MPI de Palabras	58
4.15. Resultados test psicométrico SAM dimensión Valencia	59
4.16. Resultados test psicométrico SAM dimensión Dominancia	59
4.17. Resultados test psicométrico SAM dimensión Activación	60
4.18. Gráfica de caja y bigote para la prueba psicométrica IDARE	61
4.19. Gráfica de caja y bigote para la prueba psicométrica IDARE corregida	63

Índice general

Agradecimientos	II
Índice de cuadros	VI
Índice de figuras	VII
1. Introducción	1
1.1. Antecedentes	1
1.2. Ansiedad-Estado en el Sistema de Salud	2
1.3. Ansiedad-Estado en la Población Estudiantil	4
1.4. Motivación para la investigación	4
1.5. Objetivos	5
1.5.1. Objetivo General	5
1.5.2. Objetivos Específicos	5
1.6. Estructura del Documento	5
2. Marco Teórico	7
2.1. Efectos de la Ansiedad-Estado	7
2.2. Medición de Ansiedad-Estado	8
2.3. Generación del Habla	10
2.4. Efectos de la Ansiedad-Estado Sobre el Habla	11
2.5. Biorealimentación	12

2.6. Pruebas Psicométricas	13
2.7. Inducción de Ansiedad-Estado	14
3. Metodología	17
3.1. Tipo de Intervención de la Investigación	17
3.2. Selección de Participantes	17
3.3. Criterios de Inclusión	18
3.4. Criterios de Exclusión	18
3.5. Base de Datos Generada	18
3.6. Procedimientos y Protocolo	18
3.6.1. Convocatoria a participar	19
3.6.2. Distribución de la aplicación experimental	19
3.6.3. Registro de sujeto experimental	20
3.6.4. Registro de línea base	20
3.6.5. Inducción de ansiedad-estado	20
3.6.6. Registro psicométrico post-inducción	20
3.6.7. Intervención	21
3.6.8. Registro psicométrico final	22
3.7. Desarrollo de Interfaz Experimental	22
3.8. Pre-procesamiento y Procesamiento de la Señal de Voz	24
3.8.1. Generalidades	24
3.8.2. Preprocesamiento	26
3.8.3. Procesamiento para la extracción de las características de la señal de voz	28
3.9. Análisis Estadístico	36
3.9.1. Verificación de los criterios estadísticos	36
3.9.2. Variables Paramétricas (con distribución normal)	37
3.9.3. Variables No Paramétricas (con distribución diferente a la normal) . .	37
3.9.4. Pruebas entre grupos	37

3.9.5. Pruebas intra sujetos	38
3.9.6. Definición de las hipótesis	38
4. Resultados - Discusión y Contribución	41
4.1. Resultados del Preprocesamiento	41
4.1.1. Filtrado de las señales	41
4.1.2. Eliminación de silencios y segmentación del discurso	42
4.1.3. Procesamiento de la señal de voz	43
4.1.4. Resultados de las pruebas psicométricas	57
4.2. Aportes más Relevantes de la Presente Investigación	64
4.2.1. Diseño e implementación de sistema de biorealimentación para el manejo de la ansiedad-estado	64
4.2.2. Validación de la metodología experimental mediante resultados de pruebas psicométricas	64
4.2.3. Identificación de variables no lingüísticas de la señal de voz relevantes para la identificación de la ansiedad-estado	65
5. Conclusión y Trabajos Futuros	66
5.1. Conclusión	66
5.2. Trabajos Futuros	67
Glosario	67
Referencias	69

Capítulo 1

Introducción

1.1. Antecedentes

El estrés supone un hecho habitual de la vida del ser humano, ya que cualquier individuo, con mayor o menor frecuencia, lo ha experimentado a lo largo de su vida [6]. El estrés es una respuesta inespecífica del organismo ante una diversidad de factores o exigencias del ambiente. Se trata de un proceso adaptativo y de emergencia, siendo imprescindible para la supervivencia de la persona. En todo caso, el estrés surge de la relación entre la persona y su ambiente. La medida en la que el sujeto percibe las demandas ambientales como un peligro para su bienestar y si exceden o igualan sus recursos para enfrentarse a ellas [7]. No obstante, la vivencia del estrés tanto positivo como negativo va a depender de la valoración que realiza el individuo de las demandas de la situación y de sus propias capacidades para hacer frente a las mismas [6].

Es común experimentar este tipo de respuesta en el entorno laboral y ciertamente, se ha establecido que las condiciones laborales tienen un impacto significativo en la salud de los empleados [1] y en particular, aquellos entornos laborales que generan en los empleados un estado emocional alterado, también conocido como ansiedad. Dado que el termino *estrés* ha sido usado libremente para abarcar un amplio conjunto de conceptos es necesario aclarar que para el alcance de este documento este termino hará referencia a la ansiedad-estado. Ahora bien, la ansiedad ha sido definida por Weinberg y Gould [8], como un estado emocional ne-

gativo caracterizado por el nerviosismo, preocupación y aprensión, asociados con el arousal o activación del cuerpo. Sin embargo es Spielberger [9], quien expresa que la ansiedad puede ser clasificada como; Ansiedad rasgo: que es la predisposición de una persona, como rasgo de su personalidad, a percibir ciertas situaciones como amenazadoras o no, respondiendo a ellas con niveles variados de ansiedad; y Ansiedad estado, que es un estado emocional inmediato, caracterizado por un sentimiento de miedo, aprensión y tensión, acompañado o asociado con la activación del sistema nervioso autónomo. Cuando un individuo es expuesto a condiciones laborales adversas y éste no es capaz de regular la ansiedad-estado de forma adecuada se habla del estrés crónico, que en el contexto laboral se conoce comúnmente como “burnout” [2]. El burnout, se caracteriza por un agotamiento abrumador, actitudes negativas o falta de compromiso e insatisfacción con el desempeño laboral [10, 11] y puede conllevar consecuencias devastadoras para el trabajador, su familia, el ambiente laboral y la organización [3]. No obstante, el deterioro hacia esta condición extrema, comprende un tiempo considerable durante el cual el sujeto haya padecido la ansiedad-estado.

1.2. Ansiedad-Estado en el Sistema de Salud

Los efectos de la ansiedad-estado resultan de particular importancia cuando la población afectada involucra al personal médico asistencial. El estudio reportado por Richardsen y Burke [4] que relaciona el estrés en los profesionales de la salud con su nivel de desempeño, reveló que las consecuencias de la prevalencia de altos niveles de estrés en el personal médico se revierten en un desempeño profesional inadecuado hacia sus pacientes. Similarmente, Tawfik y colegas [12] reportó que, el agotamiento y el bajo bienestar han sido reconocidos como riesgos laborales comunes entre los profesionales de la salud. En particular para los médicos, la prevalencia del agotamiento se estima en más del 50 %, la fatiga excesiva en el 45 %, y una tasa de suicidio 3 a 5 veces mayor que en la población general [12]. Por otra parte, investigaciones longitudinales enfocadas en estudiar la prevalencia y los antecedentes de la ansiedad-estado en psicólogos clínicos han reportado que cerca del 40 % de los psicólogos informaron altos niveles de agotamiento emocional, el 10 % informó reacciones de despersonalización, y el

19% informó una sensación reducida de logro personal a lo largo del tiempo [13]. El aumento de la ansiedad en los profesionales de la salud ha sido asociado con eventos de seguridad del paciente, incluyendo errores médicos [5, 14, 15]. El bajo bienestar en los profesionales de la salud, como por ejemplo en el dominio de la fatiga, altos niveles de estrés, depresión, mala calidad de vida, entre otros, se han relacionado con la reducción de la seguridad del paciente. A pesar de que la relación entre el grado de seguridad de las unidades médicas y la seguridad de los pacientes bajo el cuidado de un personal médico cada vez más afectado por esta condición, se ha venido estudiando en los últimos años, no se ha llegado a un consenso general y continúa siendo un tema controversial [16].

Aun así, según el artículo de opinión publicado por el diario El Espectador [17], un estudio adelantado por la Facultad de Enfermería y Rehabilitación de la Universidad de La Sabana alertó sobre los frecuentes problemas de seguridad que corren los pacientes en Colombia. “El estudio estableció que en Colombia cada año mueren alrededor de 180.000 personas hospitalizadas por errores médicos siendo los niños y los adultos mayores de 75 años de edad las poblaciones más comúnmente afectadas”. Ante esto, Beatriz Pérez (directora del programa de Enfermería de la Universidad de La Sabana) comentó en medios de comunicación: “Por lo menos el 80% de estos accidentes se podrían evitar. Es inconcebible que el error humano de los médicos y del personal de la salud esté matando más a los pacientes que las mismas enfermedades. . . A causa de estos accidentes, se calcula que al menos 23 millones de personas mueren cada año en Latinoamérica.” Similarmente, el estudio realizado en Colombia por los doctores e investigadores, Olga Paredes y Pablo Sanabria en el 2008, titulado “Prevalencia del síndrome de Burnout en residentes de especialidades médico-quirúrgicas, su relación con el bienestar psicológico y con variables sociodemográficas y laborales” [18] , muestra que entre 135 médicos residentes el 43% puntuaron en la categoría media y alta para esta condición, y el 57% en baja; en cuanto al bienestar psicológico, ninguno de los encuestados expresó niveles altos.

1.3. Ansiedad-Estado en la Población Estudiantil

En el artículo de revisión de literatura realizado por McConville y colegas [19], en el que se tuvieron en cuenta 19 estudios científicos los cuales involucraban un total de 1815 estudiantes de pregrado o posgrado del área de la salud tales como medicina, psicología, trabajo social, enfermería, terapia ocupacional, podología, dietética y fisioterapia, demostró que: “El estrés, la ansiedad y la depresión pueden tener un efecto negativo en la salud y en la capacidad del estudiante para estudiar y desempeñarse en la práctica clínica”. Por lo tanto, es ideal introducir herramientas que provean a los estudiantes estrategias para el adecuado manejo de la ansiedad-estado y aumentar el autocuidado durante sus etapas de formación profesional, ya que esto constituye una estrategia de promoción y prevención ante esta creciente problemática. Así mismo, el adecuado manejo de la ansiedad-estado y aumento del autocuidado por parte de los estudiantes, tienen el potencial de mejorar el bienestar psicológico de los estudiantes [20,21] y facilitar potencialmente un enfoque más centrado en el paciente [22].

1.4. Motivación para la investigación

Por lo tanto, el desarrollo e introducción de herramientas tecnológicas que aporten al adecuado manejo de estados alterados como el estrés o ansiedad-estado a lo largo del proceso formativo de los futuros profesionales de la salud, constituyen un mecanismo de prevención para mantener y mejorar la salud de los empleados en general, combatir la disminución en la productividad laboral y la disminución de la frecuencia de eventos adversos causados por errores humanos. Entre las intervenciones no invasivas que se han empleado para abordar esta problemática se encuentra la biorealimentación la cual consiste en proveerle al individuo información sobre sus diversas respuestas fisiológicas para que aprenda a modificarlas activamente [23]. El entrenamiento de biorealimentación puede modificar los síntomas de ansiedad-estado al reducir la excitación psicofisiológica desadaptativa [23]. La biorretroalimentación también es una herramienta terapéutica para facilitar el aprendizaje de la autorregulación de

funciones autónomas para mejorar la salud. Debido a los avances tecnológicos y el creciente interés en las terapias alternativas, la biorealimentación permanece en la atención de los investigadores sobre posibles aplicaciones del método en medicina [24]. Tomando en cuenta el panorama anteriormente descrito, resulta de interés responder la pregunta ¿Cuál es el efecto que tiene un sistema de medición y biorealimentación para el manejo de la ansiedad-estado, en una población estudiantil del área de la salud a nivel de pregrado?

1.5. Objetivos

Con base en lo anterior, este proyecto tiene como objetivo evaluar el desempeño del sistema de biorealimentación propuesto para el manejo de la ansiedad-estado. Dicho sistema emplea un micrófono de un computador para medir el nivel de agitación que presenta el usuario y le presenta al mismo una biorealimentación visual a forma de pauta para promover la autorregulación del nivel de ansiedad-estado. Para dicho fin, el sistema mide las características no lingüísticas de la señal de voz, de forma no invasiva y en tiempo real.

1.5.1. Objetivo General

Estudiar el efecto de un sistema de biorealimentación en el manejo de la ansiedad-estado, en una población estudiantil del área de la salud.

1.5.2. Objetivos Específicos

- Diseñar un sistema de biorealimentación para el manejo de la ansiedad-estado.
- Implementar un protocolo de experimentación que integre el sistema de biorealimentación para el manejo de la ansiedad-estado.
- Validar si el sistema de biorealimentación promueve la reducción de los niveles de la ansiedad-estado.

1.6. Estructura del Documento

Este documento está compuesto de cuatro secciones. En el marco teórico se presenta el marco conceptual del trabajo y de investigaciones relevantes, presentando un apartado de definiciones técnicas para las variables no lingüísticas de la señal de voz a ser extraídas

y analizadas. A continuación, se presenta la metodología empleada para dar respuesta a la pregunta de investigación, y cumplir los objetivos propuestos. Seguidamente, se darán a conocer los resultados obtenidos del proceso experimental, con sus correspondientes análisis determinando así las contribuciones del trabajo al área de la medición y adecuado manejo de la ansiedad-estado. Finalmente, se presenta una síntesis final de los resultados obtenidos y las razones por las cuales se presentaron algunas variaciones a los resultados esperados. Se concluye con algunas recomendaciones para trabajos futuros a los cuales da paso los resultados del presente trabajo. Dado que se demuestra que, la biorealimentación en base al procesamiento de la señal de voz se constituye como una estrategia viable para el adecuado manejo de la ansiedad-estado en la población de estudiantes del área de la salud.

Capítulo 2

Marco Teórico

2.1. Efectos de la Ansiedad-Estado

Se ha sugerido que, los mecanismos biológicos resultantes de la ansiedad-estado pueden deteriorar la salud física. Una de las hipótesis es que el sistema nervioso autónomo (SNA) y el eje hipotálamo-hipófisis-suprarrenal (HHA) se debilitan. Esto da como resultado la sobre-activación de las funciones vitales (por ejemplo, frecuencia cardíaca y presión arterial) generando daño al metabolismo y al sistema inmunológico [25]. El estrés psicológico se asocia con una variedad de trastornos crónicos de salud y enfermedades cardiovasculares importantes [12]. En general la respuesta fisiológica del cuerpo humano se caracteriza por cambios repentinos en el sistema nervioso autónomo que preparan al cuerpo para las respuestas de “lucha o huida” ante estímulos externos [26]. Según Melamed y colegas [27], los mecanismos potenciales que relacionan el agotamiento con las enfermedades cardiovasculares incluyen sus asociaciones con componentes del síndrome metabólico, la alteración del eje HHA, inflamación, trastornos del sueño, inmunidad reducida, cambios en la coagulación sanguínea, cambios en la fibrinólisis y la adopción de conductas de salud deficientes; como el consumo de tabaco y el sedentarismo.

Además, se ha demostrado que la ansiedad-estado en el ámbito laboral está asociada con la incidencia de eventos cardiovasculares, y esta relación está mediada por los efectos directos de mecanismos neuroendocrinos y los efectos indirectos de conductas no saludables,

en particular mala alimentación y baja actividad física [28]. El agotamiento también puede desencadenar el presentismo [29] y el ausentismo [30–34]. Mientras el ausentismo significa ausencia del trabajador, el presentismo representa un fenómeno en el cual las personas van a trabajar incluso cuando están enfermas, lo que lleva a una pérdida de productividad, gestando un círculo vicioso, como consecuencia de los problemas de salud causados por el agotamiento. Es posible que los trabajadores no alcancen el rendimiento deseable en el trabajo, lo que a su vez puede llevar a niveles crecientes de agotamiento emocional [35]. La salud debilitada del trabajador junto con su capacidad funcional disminuida puede llevar al ausentismo, una gran causa de preocupación para el trabajador y las organizaciones dadas las consecuencias económicas para ambas partes. Estos problemas pueden representar el comienzo de un proceso de declive social que involucra la pérdida de empleos e incluso la exclusión permanente del mercado laboral de las personas afectadas. Para las organizaciones, el ausentismo significa una pérdida de mano de obra, gastos adicionales asociados con trabajadores temporales y una disminución de la productividad [30].

2.2. Medición de Ansiedad-Estado

Los métodos convencionales para estimar el estrés, utilizan algunas variables fisiológicas como la frecuencia cardíaca y la variabilidad de la frecuencia cardíaca en los dominios de tiempo y frecuencia. Otras variables fisiológicas, de medición no invasiva, evaluadas frecuentemente son: presión arterial, electroencefalografía, pulsioximetría y respuesta galvánica de la piel. El análisis de la información resultante de la medición de estas variables fisiológicas se gesta mediante la aplicación de técnicas de procesamiento en tiempo, frecuencia, y tiempo-frecuencia en conjunto con técnicas de aprendizaje automático, obteniendo niveles de precisión variados. Muchos síntomas fisiológicos de la ansiedad-estado, como los que se mencionaron previamente, pueden medirse empleando una multitud de sensores e instrumentos, como, por ejemplo, mediante analizadores químicos, sistemas para la medición de la conductancia de la piel, electrocardiogramas, electroencefalogramas, entre otros. Sin embargo, dichos métodos y sistemas son intrínsecamente intrusivos, ya que requieren la ubicación de

sensores en sitios anatómicamente apropiados, los cuales limitan la movilidad del sujeto para garantizar una interacción directa y adecuada entre usuarios y dispositivos. La implementación de estas técnicas de sensado es adecuada para un ambiente de laboratorio controlado, donde el participante puede realizar una tarea específica mientras se encuentra conectado a los equipos de monitorización. Hoy en día existe un número de dispositivos alámbricos de bioalimentación comerciales entre los que se destacan el ProComp Infiniti - 8 Channel Biofeedback & Neurofeedback System v6.0 (Infiniti-8) y el FlexComp Infiniti - 10 Channel System (FlexComp), entre otros. Sin embargo, ningún dispositivo comercial encontrado cuenta con la capacidad de hacer la captura de las variables fisiológicas de forma mínimamente invasiva, es decir sin estar en contacto directo con el cuerpo del sujetos. La adquisición de estas variables en un contexto no controlado y cambiante como en el que se suelen desempeñar los profesionales de la salud sigue siendo un problema por resolver. Por lo tanto, vale la pena explorar métodos menos intrusivos para estimar la ansiedad-estado. Los investigadores han reconocido ampliamente que la producción vocal humana está influenciada por la ansiedad-estado [36]. Este hecho plantea la voz humana como una fuente potencial para la detección no intrusiva de un nivel elevado de ansiedad-estado.

La detección de la ansiedad-estado mediante el habla es de gran interés, ya que el habla se considera una modalidad no invasiva, menos intrusiva y menos costosa en comparación con otros métodos para la evaluación del estrés [37]. Aunque no existe una única definición acordada sobre el habla bajo estrés, se puede referir que “el estrés es una variabilidad observable en ciertas características del habla debido a una combinación de respuesta inconsciente a estresores y / o control consciente” [38]. En la detección automática de la ansiedad-estado se involucra a varias disciplinas, como la informática, la lingüística, la medicina y psicología ya que es únicamente a partir de una mirada interdisciplinaria que se posible abarcar la multidimensionalidad de este proceso psicofisiológico [39].

La importante contribución del habla humana en la evaluación del estrés ha sido probada por varios estudios [40]. El discurso se ve afectado por las emociones del hablante, ya que

las emociones afectan la tensión muscular, que a su vez afecta el patrón de interacción de flujo de aire en el tracto vocal [41]. Aunque en algunos casos no se observa ningún efecto, hay muchos casos en los que la alteración del habla bajo estrés es significativa y se percibe fácilmente. El nivel de cambio en la producción del habla depende de la intensidad y el tipo de emoción expresada (por ejemplo, ira o miedo) y / o las condiciones ambientales en las que se encuentra el hablante por ejemplo, el efecto Lombard, que consiste en la tendencia involuntaria a incrementar el esfuerzo vocal cuando se habla en un lugar ruidoso para mejorar la audibilidad de la voz [42].

2.3. Generación del Habla

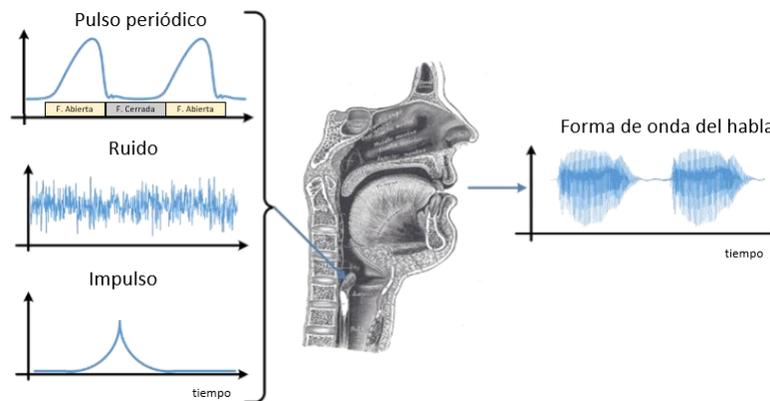


Figura 2.1: Representación del mecanismo de producción del habla humana

Se puede considerar como un sistema con 3 entradas posibles (pulsos periódicos, ruido o impulso). Estas son moduladas en el tracto vocal por los articuladores (sistema) que producen la forma de onda.

En la figura (2.1) se muestra una representación general de la producción del habla. Aquí, la producción del habla típica ocurre cuando una fuente de excitación (impulso), excita la estructura resonante del tracto vocal, lo que resulta en una forma de onda de habla de salida [43]. Si consideramos una vocal, la figura (2.1) muestra el ciclo glótico resultante de la apertura y cierre de los pliegues vocales (fase cerrada, fase abierta) que excita el tracto vocal configurado correspondiente. El flujo de determinadas configuraciones aporta componentes de resonancia que se suman a las componentes frecuenciales derivadas del ciclo glótico. De esta

forma, es posible clasificar la señal mecánica producida por el tracto bucal en sonidos sonoros (componentes frecuenciales producidos por el involucramiento de los pliegues vocales) y sonidos sordos (componentes frecuenciales producidos por resonancias en las cavidades de la parte superior del tracto vocal). Los articuladores en el tracto vocal deben posicionarse continuamente para producir un habla fluida en el tiempo. Estas reconfiguraciones tardan como mínimo 30-50 milisegundos en realizarse [44], por lo tanto, se puede estudiar la señal de voz como una señal semiestacionaria para segmentos de tiempo adecuados. Las características vocales se pueden clasificar en tres componentes: excitación del habla (fuente), tracto vocal (sistema de filtro) y señal del habla (salida).

2.4. Efectos de la Ansiedad-Estado Sobre el Habla

Durante condiciones de estrés, el primer componente (fuente) se ve afectado por el aumento de la tensión del músculo de los pliegues vocales, el segundo componente (filtro) se ve afectado por el cambio en la posición de los articuladores del tracto vocal y el tercer componente (salida) se ve afectado debido a su vinculación con los otros dos componentes. La identificación del habla bajo estrés se basa en la identificación de una serie de características acústicas. Estas incluyen los componentes prosódicos (características suprasegmentales) tales como el tono (f_0), intensidad, duración, entre otras y los componentes espectrales, tales como los coeficientes cepstrales de frecuencia de Mel (MFCC), análisis de coeficientes predictivos lineales (LPC), la identificación de formantes, estadísticas espectrales, entre otros. Las características prosódicas pueden ser más relevantes para identificar el nivel de excitación, mientras que las características espectrales pueden ser más relevantes para identificar los estados de activación emocional [41]. En general, el discurso de estrés produce un tono más agudo f_0 , mayores amplitudes en la forma de onda glótica y más pulso glótico asimétrico en comparación con la condición neutra. Estos cambios tienen impacto en la intensidad del espectro; provocando un espectro de frecuencias con mayor número de componentes en las altas frecuencias. La tabla 2.1, resume las características del habla y el número de estudios asociados a dicha característica según el artículo de revisión de literatura efectuado por

Giannakakis [36].

Cuadro 2.1: Características No Lingüísticas de La Voz Utilizadas en la Detección de Estrés

Características	Número de estudios que la reporta
Tono (f_0)	11
Formantes 1 y 2	4
Centroide Espectral	3
Dispersión Espectral	2
Asimetría Espectral	2
Curtosis Espectral	1
Entropía Espectral	2
Pendiente Espectral	4
MFCC	4
LPCC	4
Sonoridad Integrada	4
Jitters o fluctuaciones	2
Duración del Segmento de Discurso	2

2.5. Biorealimentación

Adicionalmente a la medición de la ansiedad-estado, es importante considerar qué se puede hacer para que esta información le resulte útil a quien está experimentado dicho fenómeno. Aproximadamente desde el año 2000, la literatura científica ha venido reportando acerca de la efectividad de los sistemas de biorealimentación para el manejo de la ansiedad-estado. En general, la biorealimentación es una intervención eficaz que se puede utilizar para reducir la ansiedad-estado y mejora el rendimiento. Esto es especialmente cierto entre los profesionales, cuyo desempeño laboral requiere una gestión adecuada de la ansiedad-estado [26]. La biorealimentación es una técnica en la que los individuos aprenden a modificar sus respuestas fisiológicas con el fin de mejorar la salud física, mental y emocional [45]. Según la Asociación de Psicofisiología Aplicada y Biofeedback (AAPB), la Alianza Internacional de Certificación de Biofeedback (BCIA) y la Sociedad Internacional de Neurofeedback e Investigación (ISNR), el término biorealimentación puede ser definido como el “proceso que permite a una persona aprender a cambiar la actividad fisiológica con el fin de mejorar la salud y el rendimiento. Esto implica el uso de instrumentos que de forma rápida y precisa le presenten información

de su estado metabólico actual y de esta forma se logre una realimentación al usuario. La presentación de esta información, a menudo junto con cambios en el pensamiento, las emociones y el comportamiento, resultan en los cambios fisiológicos deseados. Con el tiempo, estos cambios pueden perdurar sin el uso continuo del instrumento.” (Aprobado el 18 de mayo de 2008 por la AAPB, el BCIA y el ISNR) [46]. La biorealimentación clínica se puede usar para controlar los síntomas de enfermedades, así como para mejorar la salud y el bienestar general a través de la capacitación en el manejo de la ansiedad-estado [47]. Las investigaciones han demostrado que las intervenciones de biorealimentación son eficaces para tratar una variedad de afecciones médicas [45].

2.6. Pruebas Psicométricas

Si bien es cierto, que cambios en las variables fisiológicas mencionadas pueden dar una idea del estado de activación de una persona y esto puede asociarse con la respuesta de ansiedad-estado, la inespecificidad de esta, junto con la dependencia que tiene con la valoración subjetiva de las situaciones, hace que sea impreciso medir niveles de ansiedad-estado valorando únicamente elementos fisiológicos es por esto que se debe recurrir a instrumentos psicométricos para validar los resultados. En cuanto a las pruebas psicométricas, se encontró a lo largo de la búsqueda bibliográfica que las pruebas más ampliamente utilizadas para la medición y caracterización de la ansiedad-estado fueron el Inventario De la Ansiedad Rasgo-Estado (IDARE) y el Self Assesment Manikin (SAM). La prueba IDARE es una medida comúnmente utilizada para caracterizar el rasgo y estado de la ansiedad [36]. Se puede utilizar en entornos clínicos para diagnosticar la ansiedad y para distinguirla de los síndromes depresivos. En su versión más popular (Y), tiene 20 elementos para evaluar la ansiedad por rasgo y 20 para la ansiedad-estado. Todos los artículos se califican en una escala de 4 puntos (por ejemplo, de “nunca” a “siempre”). Las puntuaciones más altas indican mayor ansiedad. Los coeficientes de consistencia interna para la escala han oscilado entre .86 y .95; los coeficientes de fiabilidad test-retest han oscilado entre .65 y .75 en un intervalo de 2 meses [9].

Por otra parte, el SAM, es un test psicométrico el cual emplea una evaluación pictórica no

verbal, mide el nivel de placer o valencia emocional, la activación y la dominancia asociados con cómo se siente una persona en determinado momento. El método SAM mide la respuesta emocional basado en un modelo dimensional de emociones visualmente representado por cara, lo que lo hace un método efectivo, rápido, y es independiente de la cultura en la que se aplique. Este test emplea las dimensiones de valencia activación y dominancia (VAD) (2.1) [48].

- Valencia: mide lo agradable, positivo, negativo o neutral.
- Activación: mide la intensidad o fuerza de la emoción.
- Dominancia: mide cómo es la emoción controladora/ dominante o sumisa.

Durante la aplicación de la prueba SAM, el participante debe elegir una de las cinco figuras mostradas para cada una de las afirmaciones expositivas asociadas a las dimensiones VAD según se puede observar en la figura (2.2). Posteriormente, se calcula en una escala de 9 puntos dónde 5 de los 9 puntos significa un estado neutro para cada situación de evaluación, dependiendo del tipo de test que se esté aplicando, se realiza la calificación del test. Para esta investigación se define que la valencia se mide como 9 lo más alegre y 1 lo más triste, en activación 1 lo más pasivo y 9 lo más activo y en dominancia se mide como 1 lo más dominado y 9 lo más dominante [48].

2.7. Inducción de Ansiedad-Estado

Aunque la ansiedad-estado puede ocurrir en muchos aspectos de la vida real, en la mayoría de los estudios reportados en la literatura científica, esta condición es inducida a los participantes en entornos controlados. Según el artículo de revisión de literatura efectuado por Giannakakis y colegas [36], las tareas de aritmética mental parecen ser los factores estresantes más utilizados para provocar estrés relacionado con la realización de labores que requieren eficacia a nivel mental y procedimental. En particular, para este estudio se les solicitó a los participantes que intentaran solucionar un juego aritmético considerado desafiante bajo un límite de tiempo de 10 minutos. El sudoku es un juego aritmético japonés cuyo objetivo es rellenar una cuadrícula de 9×9 regiones divididas en subcuadrículas de 3×3 celdas (81 celdas

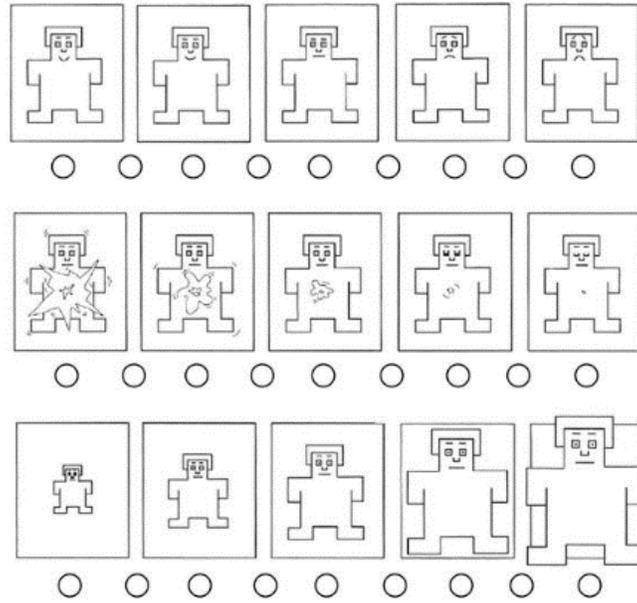


Figura 2.2: Representación gráfica del SAM

en total) con las cifras del 1 al 9 partiendo de algunos números ya dispuestos en algunas de las celdas (21 en este caso en particular). Los números deben ser asignados de forma tal que no se repitan en una misma fila, columna o subcuadrícula. El sudoku planteado tiene una única solución, ya que como fue demostrado por el matemático Gary McGuire, este tipo de esquemas numéricos cuentan con una única solución posible si se asignan 17 cifras o más inicialmente. Como se menciona en Giannakakis y colegas [36], se ha encontrado que este tipo de actividad contra reloj induce cambios considerables en la concentración de hormona adrenocorticotropa (ACTH), cortisol (suero y saliva), hormona del crecimiento, prolactina, así como aumentos en la frecuencia cardíaca, todos factores asociados a estados emocionales alterados como la ansiedad-estado.

La mayoría de los estudios mencionados en esta sección han sido realizados en laboratorio o ambientes bien controlados. Los estresores inductores fueron generalmente intensos para lograr una cantidad prominente y medible de ansiedad-estado. Sin embargo, en condiciones de la vida real, los estresores suelen ser estímulos complejos que involucran muchos aspectos de la personalidad o pueden ocurrir múltiples estresores debido a la complejidad de la forma

de vida. La mayoría de los investigadores admite que no existe un protocolo experimental adoptado y comúnmente utilizado y la estimación del nivel de ansiedad-estado inducido como parámetro comparativo o verdad fundamental de contraste, subyace en la subjetividad [36]. Por lo tanto, es una práctica común entre los investigadores que al enfrentar la dificultad de establecer la verdad base se hace necesario una medición de referencia para la calibración del protocolo y la determinación de condiciones basales. Se debe mencionar que, debido a las restricciones asociadas a la presente coyuntura de salubridad a nivel mundial, esta investigación no se llevó a cabo en un ambiente de laboratorio. Como también es importante aclarar que, la incontrolabilidad asociada a la condición de pandemia, pudo dejar a las personas en estados de estrés permanente. Esta consideración y otras prácticas comunes establecidas en la literatura científica han sido tenidas en cuenta para el desarrollo de la metodología investigativa del presente trabajo la cual se describe a continuación.

Capítulo 3

Metodología

3.1. Tipo de Intervención de la Investigación

El diseño experimental propuesto para esta investigación es un modelo mixto (inter e intra participantes) [49] en el cual se aplicará un protocolo experimental para la aplicación de pruebas psicométricas y adquisición de señales de voz.

El protocolo experimental, descrito más adelante, se desarrolló en una sesión de una duración aproximada de 45 minutos. Debido a la contingencia de salud pública a nivel mundial, y en particular a la suspensión de actividades presenciales dentro de las instalaciones de la Escuela Colombiana de Ingeniería Julio Garavito a partir del 24 marzo del 2020, según lo señala la comunicación institucional No.6 del 23 de marzo de 2020, acatando a la cuarentena nacional declarada por el Gobierno Colombiano, el desarrollo de la presente investigación se realizó de forma virtual. Para superar los obstáculos que supone dicha situación, se desarrolló una aplicación de escritorio permitiendo la realización del protocolo experimental de forma remota.

3.2. Selección de Participantes

En esta investigación se utilizó una muestra aleatoria sobre la base de personas que deseen participar voluntariamente en el proyecto. Estas fueron seleccionadas teniendo en cuenta su estado de salud, condiciones físicas y perfil académico. La decisión de participar en el presente

estudio fue completamente voluntaria y no tuvo ninguna consecuencia desfavorable el elegir no tomar parte en ello.

3.3. Criterios de Inclusión

- Estudiante de pregrado o posgrado del área de la salud.
- Estudiante que haya cursado 20 % de los créditos del programa académico de pregrado.
- Participantes sanos de entre 18 y 35 años de edad (rango de edades que abarca el 76.8 % de la población universitaria [50])

3.4. Criterios de Exclusión

- Haber sido diagnosticado con alguna enfermedad neurológica como depresión, desorden de estrés postraumático, o cualquier otra.
- Haber sido diagnosticado con alguna enfermedad cardiovascular como hipertensión, hipotensión, o cualquier otra.
- No haber cursado 20 % de los créditos del programa académico de pregrado.
- Personas en estado de embarazo.

3.5. Base de Datos Generada

Se obtuvo una base de datos de un total de 8 participantes, de los cuales 6 fueron mujeres y 2 fueron hombres. Estos participantes se asignaron a dos grupos experimentales (G1), control y experimental (G2) como se ilustra en la tabla (3.1). El desbalance en el sexo de los participantes puede atribuirse a la distribución demográfica de las universidades en Colombia. Según el reporte del DANE del 2018, aproximadamente el 70 % de los estudiantes en el área de la salud son mujeres [51].

3.6. Procedimientos y Protocolo

El desarrollo de la presente investigación se realizó mediante la ejecución de las etapas ilustradas en el diagrama de la figura (3.1).

Cuadro 3.1: Base de datos generada

Grupos	Participantes Femeninos	Participantes Masculinos	Control sobre Biorealimentación
G1	2	1	No
G2	4	1	Si



Figura 3.1: Etapas principales de la investigación

3.6.1. Convocatoria a participar

Se realizó el envío de un correo electrónico de convocatoria para estudiantes interesados en participar en el desarrollo de esta investigación como sujetos de prueba. Adjunto al correo se les hizo entrega del formulario “Información para el participante y Consentimiento informado” y se les indicó a los interesados que podían manifestar cualquier duda o inquietud por ese medio. Así mismo, se les indicó a los sujetos, que podrían confirmar su decisión al enviar a vuelta de correo el documento debidamente diligenciado con su información personal (Nombre, No. CC, firma y fecha de diligenciamiento).

3.6.2. Distribución de la aplicación experimental

Se le envió un segundo correo electrónico con el link de acceso al repositorio virtual de la aplicación correspondiente a aquellos estudiantes que enviaron el formulario de “Información para el participante y Consentimiento informado” debidamente diligenciado. La asignación de los participantes a grupo control (G1) y grupo experimental (G2) se realizó de forma aleatoria según el orden de recepción de los correos confirmatorios anteriormente descritos. A través del repositorio virtual los participantes pudieron acceder a un documento de instrucciones para la descarga, instalación y envío de los resultados de la aplicación experimental y el ejecutable

de instalación correspondiente a su grupo experimental.

3.6.3. Registro de sujeto experimental

Una vez terminado el proceso de instalación y ejecución inicial de la aplicación se les solicitó a los participantes ingresar algunos datos personales como nombre, edad, y género para generar el registro correspondiente. Estos datos se mantuvieron confidenciales en todo momento y solo fueron conocidos por los investigadores.

3.6.4. Registro de línea base

En esta etapa se le suministró al participante la prueba psicométrica SAM con el objeto de registrar el estado inicial de los participantes. Debido a que el desarrollo de la investigación no se realizó en un mismo entorno experimental, se tomó la precaución de promover un estado neutro en los participantes al solicitarles que participaran en un ejercicio de respiración durante un minuto. A continuación, se les solicitó realizar la segunda prueba psicométrica (IDARE) para completar el perfil psicológico de línea base. Posteriormente se les solicitó seguir las instrucciones suministradas por la interfaz gráfica de usuario para obtener los registros de voz correspondientes. Los registros de voz constan de la grabación de la lectura de un texto neutro, es decir un texto expositivo carente de contenido emocional, y el registro individual de 10 palabras neutras. Para la extracción de las características de tono velocidad y volumen, se implementó una ventana de Hamming de duración de 30 milisegundos y con un traslape del 50%. Las generalidades del procesamiento de la señal de voz empleado se presentan en el diagrama de la figura (3.2).

3.6.5. Inducción de ansiedad-estado

Con el objetivo de inducir ansiedad-estado en los participantes, se les solicitó realizar una actividad aritmética desafiante que consistió en resolver un sudoku, bajo un límite de tiempo de 10 minutos.

3.6.6. Registro psicométrico post-inducción

Una vez culminada la fase de inducción, se les solicitó a los participantes realizar nuevamente las pruebas psicométricas descritas en la etapa de registro de línea base, para evidenciar

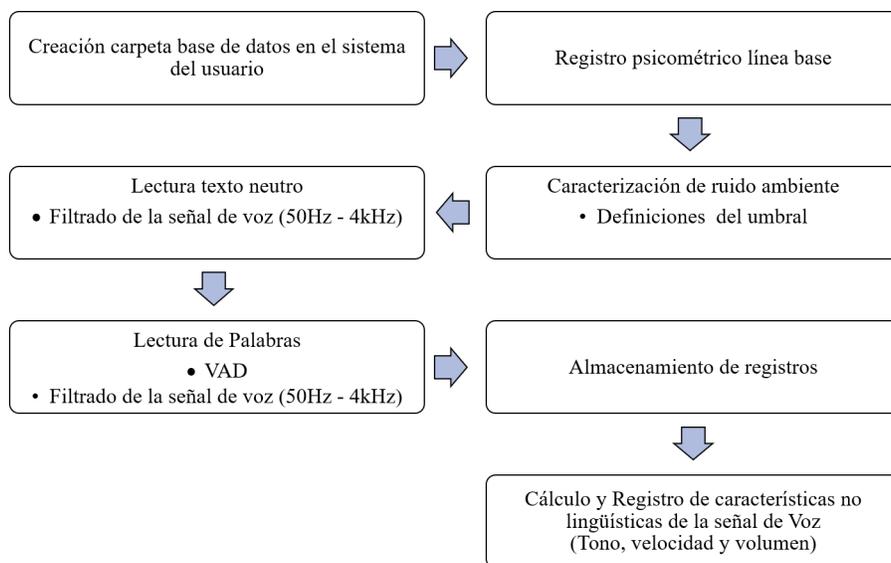


Figura 3.2: Generación de registro del participante

el efecto de la etapa de inducción. Vale la pena mencionar, que para este registro no se les solicitó a los participantes que realizaran el ejercicio de respiración mencionado anteriormente.

3.6.7. Intervención

Inicialmente, se les solicitó a los participantes (tanto del grupo control como al experimental) que leyeran nuevamente el texto neutro. Esto con el objeto de generar un segundo registro de la lectura para así poder hacer un análisis comparativo posteriormente. Se les solicitó a los participantes intentar imitar la pronunciación de cada una de las 10 palabras previamente registradas en la etapa de línea base, teniendo en cuenta las características de tono, velocidad y volumen. Para este fin, se les informó que la aplicación haría la reproducción del registro capturado durante la etapa de línea base antes de iniciar la captura del nuevo registro. Una vez terminada la captura y procesamiento del nuevo registro se les presentó la biorealimentación asociada a su desempeño. Dicha biorealimentación se reportó en un recuadro en la parte inferior de la aplicación mediante una descripción gráfica (a través del encendido de una serie de indicadores) junto con una descripción escrita de cómo modificar la pronunciación para lograr reproducir el desempeño de la etapa previa. Por ejemplo, si el participante realiza una pronunciación más rápida, más grave y con una sonoridad menor al

registro inicial la aplicación le indica que debe pronunciar la palabra más lento, más agudo y más alto (figura 3.3 - f). Es a partir de este momento que la aplicación adopta un comportamiento determinado según el grupo experimental al cual pertenezca el participante. Para los participantes del grupo control, la biorealimentación se genera de forma aleatoria; es decir, no está relacionada con el desempeño del participante. Sin embargo, los criterios de éxito son los mismos para ambos grupos. Por otra parte, la biorealimentación del grupo experimental está directamente relacionada con las características extraídas del nuevo registro.

3.6.8. Registro psicométrico final

Una vez terminada la etapa de intervención, se le solicitó al participante realizar por última vez las pruebas psicométricas anteriormente descritas para evidenciar el efecto de la intervención en el estado emocional del participante.

3.7. Desarrollo de Interfaz Experimental

Para la aplicación del protocolo experimental de forma remota, fue necesario diseñar e implementar una interfaz experimental que integrara cada una de las etapas anteriormente descritas. El desarrollo de la interfaz se realizó mediante la plataforma de diseño de aplicaciones de Matlab *App Designer* en la cual se aplicó un método de diseño de software que supone un nivel de conocimiento para la manipulación de herramientas informáticas básico. Por lo tanto, fueron requerimientos de diseño generar una interfaz fluida e intuitiva pero que a la vez garantizara la integridad y secuencialidad del protocolo experimental. Adicionalmente, se tuvo en cuenta en el diseño la generación automática de registros tipo *datalogger* asociados a la interacción del usuario, ventanas de dialogo emergentes ante una manipulación inadecuada de la interfaz y la generación automática de archivos *.wav* con los registros de voz generados a lo largo del protocolo experimental. Con la intención de facilitar la distribución de esta herramienta, se generó un ejecutable de instalación semiautomático, el cual fue acompañado por una guía de instalación gráfica la cual ilustraba el paso a paso del proceso.

La figura (3.3) ilustra los elementos principales de la interfaz gráfica. En la figura 3.3 - a se observa la ventana para el registro de la información general del participante. Esta cuenta

con una restricción la cual obliga al usuario a confirmar el diligenciamiento del documento de consentimiento informado antes de iniciar con el protocolo experimental. Una vez finalizado el registro de la información, se le indicó al usuario que se dirigiera a la pestaña de pruebas psicométricas donde el usuario tendría que diligenciar las pruebas SAM e IDARE (figuras 3.3 b y c respectivamente). A continuación, se le solicitó al usuario proceder con los registros de voz de línea base (figura 3.3 - e). Posteriormente, se le solicitó al usuario proceder a la actividad (figura 3.3 - d) donde se le retó a resolver el sudoku. El usuario tuvo la libertad de presentar tantos intentos de solución como deseara durante el tiempo de solución estipulado y para cada solución propuesta la interfaz proporcionaba si la solución era correcta o no mediante una ventana de dialogo emergente. Una vez cumplido el tiempo límite o haber solucionado el sudoku y nuevamente habiendo diligenciado las pruebas psicométricas SAM e IDARE, se le presentó al usuario la etapa de biorealimentación (figura 3.3 - f) para la cual fue necesario incorporar las técnicas de pre-procesamiento y procesamiento descritas a continuación.

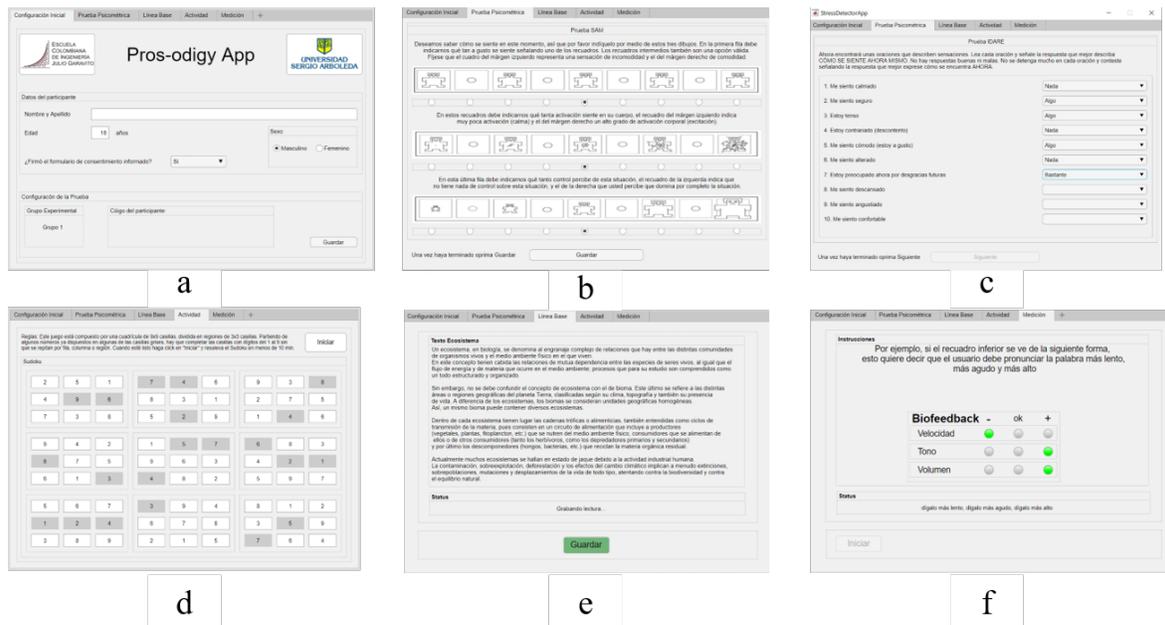


Figura 3.3: Interfaz de usuario

3.8. Pre-procesamiento y Procesamiento de la Señal de Voz

La metodología seguida en este trabajo comprende secuencialmente los aspectos ilustrados en la figura (3.4):

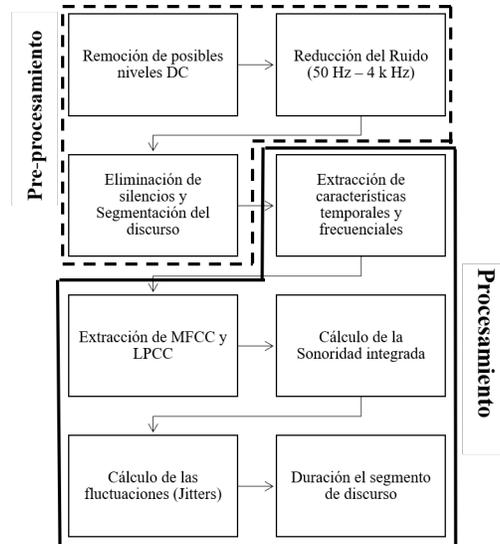


Figura 3.4: Esquema general de procesamiento de audio

3.8.1. Generalidades

Para la adquisición del audio se empleó el micrófono integrado del computador en el cual se ejecutó la interfaz de usuario, con una frecuencia de muestreo de 16 k Hz y una resolución de 16 bits. Cuando las características de audio evolucionan con el tiempo, a menudo es útil dividir el procesamiento en regiones de igual tamaño o regiones de características similares. Este acto de división de esta manera generalmente se llama segmentación. Esta estrategia es tan importante y se usa con tanta frecuencia, que sorprendentemente hay poca discusión al respecto en la literatura de investigación de audio. De hecho, la mayoría de los trabajos de investigación no lo mencionan en absoluto, considerándolo una técnica de implementación que quizás no sea digna de discusión en un trabajo de investigación. Para la extracción de características no lingüísticas de la señal de audio es conveniente dividir el registro en marcos

o ventanas de tiempo más pequeños. Sin embargo, durante este proceso de división puede suceder que la característica de interés no esté contenida en su totalidad por una ventana de análisis y por lo tanto en términos prácticos aparenta estar oculta. La solución al problema de pérdida de características es superponer ventanas de tiempo, lo que significa que, en lugar de una segmentación directa del vector de audio en partes secuenciales, cada nueva ventana contiene una parte de la ventana anterior y parte de la ventana siguiente. La superposición garantiza que las características de audio que exceden la longitud de la ventana estipulada, serán contenidas por completo en la ventana posterior superpuesta. El grado de superposición (generalmente expresado como un porcentaje) describe la cantidad de la información de la ventana anterior que se repite en la ventana siguiente. Tanto para el preprocesamiento como para el procesamiento, se implementó una ventana tipo Hamming (según el estándar del estado del arte [52] con una longitud de 30 milisegundos y una superposición del 50%. La figura (3.5) ilustra el proceso para una superposición del 50%, lo cual es muy común. El esquema a del la figura (3.5) muestra una gráfica de una forma de onda en el tiempo, la cual es dividida en una secuencia de cuadros superpuestos, numerados del 0 al 6 como se puede observar en el esquema b de la figura (3.5).

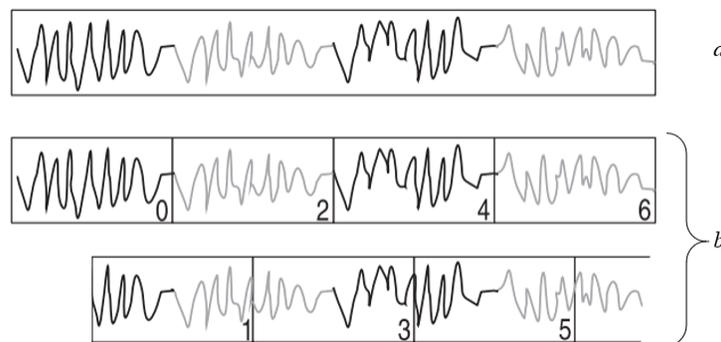


Figura 3.5: Ilustración de la superposición de ventana

3.8.2. Preprocesamiento

Remoción de niveles DC

Para eliminar el componente de DC introducido por los componentes análogos del sistema de adquisición (tarjeta de audio y micrófono) basta con tomar un promedio móvil a largo de la señal y restar el promedio de la señal. De esta forma, se logra nivelar la señal de audio con el voltaje de referencia y se asegura que las señales sean comparables.

Reducción del Ruido

Se aplica un filtrado pasa banda con frecuencias de corte de 50 Hz y 4 kHz a la señal de audio. Esta banda de frecuencias fue elegida teniendo en cuenta el rango de frecuencias fisiológico para la voz humana de tal forma que se atenúen los componentes frecuenciales adicionales a la voz, mientras que se conserve la información de interés. Esta etapa de filtrado se realizó mediante la aplicación de un filtro Butterworth con una atenuación para las frecuencias de corte 60 dB.

Eliminación de silencios y Segmentación del Discurso

Las señales de voz generalmente contienen muchos segmentos de silencio. Por lo tanto, en el análisis de voz es necesario aplicar primero un método de eliminación de silencio, para detectar segmentos de voz "limpios". El método implementado se basa en la detección y diferenciación de la señal de voz de los segmentos de señal que contengan silencios o ruido ambiental. El algoritmo se describe a continuación y se ilustra en el diagrama de las figuras 3.6.

1. La señal de audio se convierte en una representación de tiempo-frecuencia utilizando la ventana y la superposición especificada.
2. Se calcula para cada ventana la energía a corto plazo (STE - ecuación 3.1) y la dispersión espectral (ecuación 3.8).

$$STE = \sum_{i=1}^l |x_i^2| \quad (3.1)$$

3. Se crean histogramas tanto para la distribución de energía como para la distribución espectral.
4. Para cada histograma, se determina un umbral de acuerdo con

$$U = \frac{W * (M_1 + M_2)}{W + 1} \quad (3.2)$$

Donde W es un factor de ponderación determinado empíricamente, M_1 y M_2 son los máximos locales primero y segundo, respectivamente.

5. Tanto la dispersión espectral como la energía a corto plazo se suavizan al pasar a través de filtros de mediana móvil de longitud de cinco muestras, con el fin de atenuar las componentes de alta frecuencia de la señal resultante.
6. Se generan las máscaras de segmentación comparando la energía y la dispersión espectral con sus respectivos umbrales. Para que una ventana se declare como discurso, tanto la energía como la dispersión espectral deben estar por encima de sus respectivos umbrales.
7. Las regiones declaradas como habla se fusionan si la distancia entre ellas es menor que una distancia mínima.

La aplicación de este proceso a la señal de voz se ilustra en la figura (3.7)

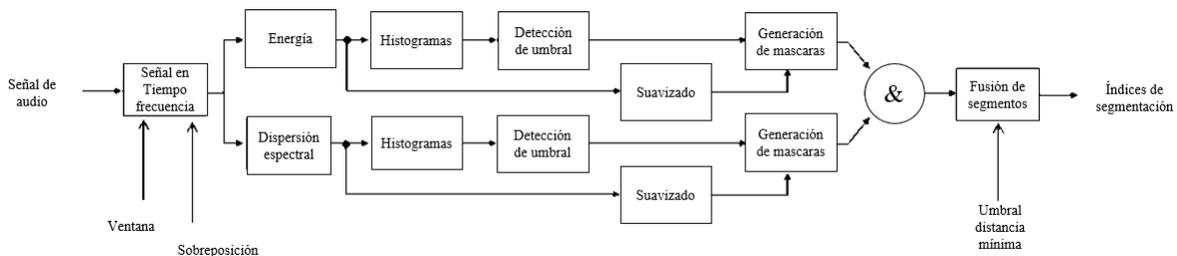


Figura 3.6: Esquema de eliminación de silencios y segmentación del discurso

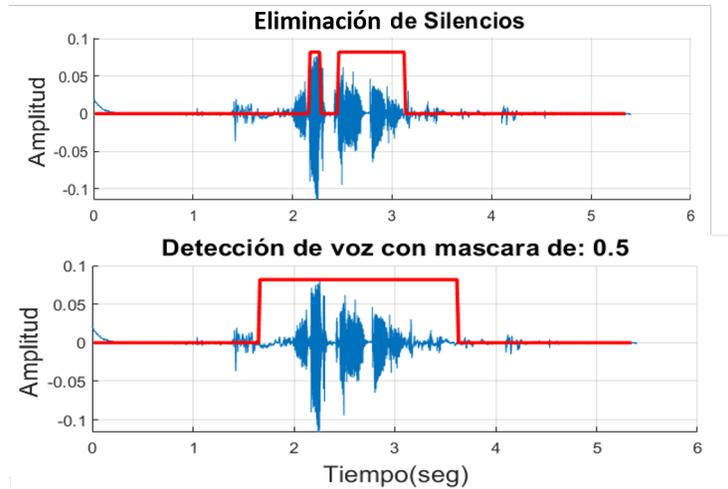


Figura 3.7: Señal de voz segmentada

3.8.3. Procesamiento para la extracción de las características de la señal de voz

Esta sección tiene como objetivo exponer las estrategias de extraer las características no lingüísticas de la señal de voz pre-procesada. En el dominio del tiempo, las características que son extraídas están relacionadas con la amplitud o la energía de la señal. Por otra parte, para el dominio de la frecuencia, las características extraídas se relacionan con los momentos del espectro frecuencial.

Tecnología

La extracción de características se gestó en el entorno de programación de Matlab 2020-a, con la adición de la toolbox para el procesamiento de audio *AudioToolbox*.

Tono (f_0) y Formantes (F_1, F)

Las formantes son componentes frecuenciales distintivas de la señal acústica producida por el habla o el canto. La información que los humanos requieren para distinguir entre los sonidos del habla se puede representar de manera puramente cuantitativa especificando picos en el espectro de frecuencias. Las formantes son producidas por resonancia de tracto vocal. La frecuencia fundamental f_0 o tono puede cuantificarse como una frecuencia, pero el tono

no es una propiedad física puramente objetiva; es un atributo psicoacústico subjetivo del sonido. Históricamente, el estudio del tono y la percepción del tono han sido un problema central en la psicoacústica, y han sido fundamentales para formular y probar teorías de representación, procesamiento y percepción del sonido en el sistema auditivo [52]. Para el alcance de este documento, se define el tono o frecuencia fundamental (f_0), como la frecuencia de mayor energía y está relacionada con el número de vibraciones por segundo producidas por los pliegues vocales. Similarmente las formantes F_1 , y F_2 hacen referencia a las frecuencias de resonancia armónicas de segundo y tercer orden respectivamente. El tono del orador es uno de los indicadores más estudiados del estrés emocional que generalmente aumenta durante condiciones de estrés [36]. Los hablantes estresados tienden a tensionar de forma involuntaria los músculos del tracto vocal, lo cual a su vez agudiza la frecuencia fundamental del hablante [38].

Centroide espectral

El centroide espectral representa el “centro de gravedad” del espectro. Se usa como una indicación de brillo y comúnmente es empleado en análisis de música y clasificación de género. Se define como la suma ponderada en frecuencia normalizada por la suma no ponderada de frecuencias, definido en la ecuación 3.3.

$$Centroide = \frac{\sum_{k=b_1}^{b_2} f_k s_k}{\sum_{k=b_1}^{b_2} s_k} \quad (3.3)$$

Donde f_k es la frecuencia en Hz correspondiente a segmento k-ésimo, s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral. Sus valores típicamente pertenecen al rango de 80 Hz a 180 Hz. A medida que su valor aumenta se interpreta que hay una mayor presencia de frecuencias agudas.

Dispersión espectral

La dispersión espectral representa el “ancho de banda instantáneo” del espectro. Se utiliza como una indicación del dominio de un tono. Por ejemplo, la propagación aumenta a medida que los tonos divergen y disminuye a medida que convergen los tonos y está definido por la ecuación 3.4.

$$Disp = \sqrt{\frac{\sum_{k=b_1}^{b_2} (f_k - \mu_1)^3 s_k}{\sum_{k=b_1}^{b_2} s_k}} \quad (3.4)$$

Donde f_k es la frecuencia en Hz correspondiente a segmento k-ésimo, s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral, μ_1 es el centroide espectral, calculado según ecuación 3.3.

Asimetría espectral

La asimetría espectral mide la simetría alrededor del centroide. En fonética, la asimetría espectral a menudo se denomina inclinación espectral y se usa con otros momentos espectrales para distinguir el lugar de la articulación. Para señales armónicas, indica la fuerza relativa de armónicos superiores e inferiores. Por ejemplo, en la señal de cuatro tonos, hay un sesgo positivo cuando el tono inferior es dominante y un sesgo negativo cuando el tono superior es dominante. Está definido por la ecuación 3.5.

$$Asmt = \frac{\sum_{k=b_1}^{b_2} (f_k - \mu_1)^3 s_k}{(\mu_2)^3 \sum_{k=b_1}^{b_2} s_k} \quad (3.5)$$

Donde f_k es la frecuencia en Hz correspondiente a segmento k-ésimo, s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral, μ_1 es el centroide espectral, calculado según ecuación 3.3 y μ_2 es la dispersión espectral, calculada según ecuación 3.4.

Curtosis espectral

La curtosis espectral mide la planitud, o no Gaussianidad, del espectro alrededor de su centroide. Se usa para indicar el pico de un espectro. Por ejemplo, a medida que aumenta el ruido blanco en la señal de voz, la curtosis disminuye, lo que indica un espectro menos picudo. Está definido por la ecuación 3.6.

$$Curtosis = \frac{\sum_{k=b_1}^{b_2} (f_k - \mu_1)^4 s_k}{(\mu_2)^2 \sum_{k=b_1}^{b_2} s_k} \quad (3.6)$$

Donde f_k es la frecuencia en Hz correspondiente a segmento k-ésimo, s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral, μ_1 es el centroide espectral, calculado según ecuación 3.3 y μ_2 es la dispersión espectral, calculada según ecuación 3.4.

Entropía espectral

La entropía espectral es una medida del “desorden”, las regiones del habla sonora tienen una entropía más baja en comparación con las regiones del habla sorda. Está definido por la ecuación 3.7.

$$Entrp = \frac{-\sum_{k=b_1}^{b_2} s_k \log(s_k)}{\log(b_2 - b_1)} \quad (3.7)$$

Donde s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral.

Pendiente espectral

La pendiente espectral se ha utilizado ampliamente en el análisis del habla, particularmente en el modelado del estrés del hablante [53]. La pendiente está directamente relacionada con las características resonantes de los pliegues vocales y también se ha aplicado a la identificación del hablante. La pendiente espectral es un aspecto socialmente importante del timbre. Se ha demostrado que la discriminación de pendiente espectral ocurre en el desarrollo de la primera infancia [54]. La pendiente espectral es más pronunciada cuando la energía en los

formantes inferiores es mucho mayor que la energía en los formantes superiores. Está definido por la ecuación 3.8.

$$Pendiente = \frac{\sum_{k=b_1}^{b_2} (f_k - \mu_f)(s_k - \mu_s)}{\sum_{k=b_1}^{b_2} (f_k - \mu_f)^2} \quad (3.8)$$

Donde f_k es la frecuencia en Hz correspondiente a segmento k-ésimo, s_k es el valor espectral en el segmento k-ésimo, b_1 y b_2 son los bordes de la banda de frecuencias sobre los cuales se calcula el centroide espectral, μ_f es la frecuencia promedio del espectro y μ_s es el valor espectral promedio.

Coeficiente cepstral de Frecuencia Mel (MFCC)

El cepstrum de frecuencia de Mel (MFC) es una representación del espectro de potencia de un sonido. Esta representación se basa en una transformación de coseno lineal de un espectro de potencia logarítmica en una escala de frecuencia melódicas (Mel) no lineal. Se le añade una “C” a la sigla para hacer referencia que se toman los coeficientes del cepstrum de frecuencia Mel. Para modelar los MFCC se define una matriz de frecuencias de Mel equidistantes que se convertirán en frecuencias centrales de funciones de dispersión triangulares (es decir, un banco de filtros de Mel). Esto está relacionado con el espectro de potencia del segmento de voz, luego la contribución de cada filtro se suma para producir un vector de potencia. La conversión de intensidad-volumen se realiza transformando el vector de potencia al dominio logarítmico, antes de que se tome una transformada discreta del coseno para producir los MFCC finales [55].

Los MFCC se derivan comúnmente de la siguiente manera:

1. Se calcula la transformada de Fourier de un segmento de longitud conveniente de la señal.
2. Se mapea las potencias del espectro obtenido anteriormente en la escala melódica perceptual (Mel), usando ventanas superpuestas triangulares.
3. Se calcula el logaritmo de las potencias de cada una de las frecuencias Mel.

4. Se aplica la transformada discreta del coseno a las potencias de los logaritmos de las frecuencias Mel.
5. Se define a las amplitudes del espectro resultante como los MFCC.

La figura 3.8 ilustra una representación gráfica a este proceso.

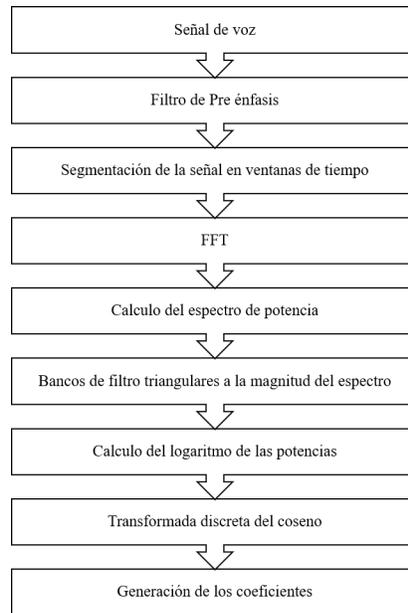


Figura 3.8: Cálculo de los MFCC

Cálculo de coeficientes de predicción lineal (LPCC)

El análisis LPC se define como un modelo autorregresivo. La respuesta espectral de la señal se modela como el producto de un filtro de todos los polos y la respuesta espectral de una señal de error. Esta es una representación natural del paradigma del filtro fuente. La señal de error $E(z)$, puede, por lo tanto, considerarse como la fuente de excitación y el filtro de todos los polos como el filtro del tracto vocal. De manera similar al análisis de espectrograma y Cepstrum para señales de voz, el análisis LPC se implementa cada 10 a 30 ms. La justificación de esto es que se supone que las propiedades estadísticas del aspecto del filtro de una señal de voz son aproximadamente estacionarias en estas escalas de tiempo. Los LPCC se obtienen a partir de la minimización de la suma del error cuadrático $E(z)$. Se inicia

por modelar la señal como una combinación lineal de señales y errores anteriores (ecuación 3.9).

$$s[n] = e[n] + \sum_{j=1}^p a_j s[n-j] \quad (3.9)$$

Donde $s[n]$ es la señal y su transformada Z es $S(z)$, $e[n]$ es el error de la señal y su transformada Z es $E(z)$ y $a_{(1...P)}$ son los parámetros de los filtros de longitud p constituido únicamente por solo polos.

Tomando la transformada Z de la ecuación 3.9 se obtiene:

$$S(Z) = \frac{1}{1 - \sum_{j=1}^p a_j z^{-j}} E(Z) \quad (3.10)$$

El cálculo de los LPCC se realizó secuencialmente empleando la función de Matlab `lpc` la cual implementa un filtro FIR de diferente orden según lo que se estime; en este caso se escogió un filtro de orden 8 siguiendo la regla general para determinar el orden del polinomio de predicción la cual estipula que el grado se obtiene como dos veces el número esperado de formantes más dos. Un filtro de mayor orden podría retornarnos datos poco relevantes para el análisis de voz. Dicha función utiliza un método de autocorrelación del modelo autorregresivo el cual permite encontrar los coeficientes del filtro derivando la señal e igualándola a cero para encontrar las ecuaciones de Yule-Walker la cual se resuelve con el algoritmo de Levinson-Durbin [56].

Sonoridad Integrada

La sonoridad es una medida subjetiva de la intensidad con la que un sonido es percibido por el oído humano. Es decir, la sonoridad es el atributo que nos permite ordenar sonidos en una escala del de mayor intensidad al de menor intensidad. Para el alcance de este documento, se define la sonoridad integrada de la señal de audio, según el estándar ITU-R BS.1770-4. El esquema general del algoritmo se presenta en el diagrama de la figura 3.9.

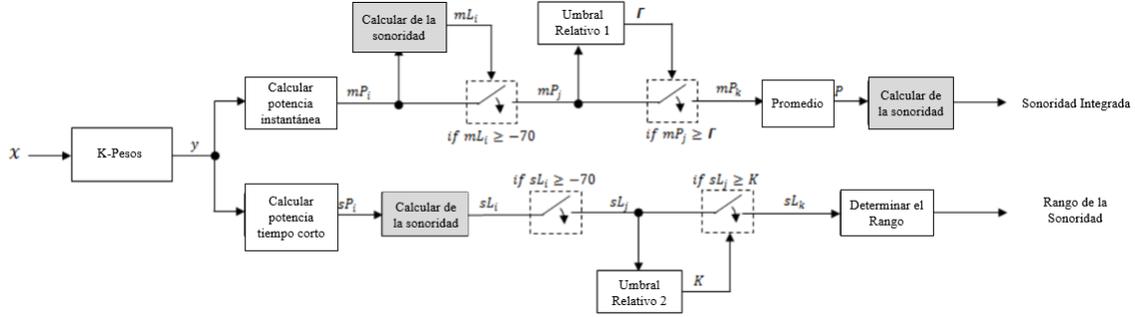


Figura 3.9: Esquema general del cálculo de la Sonoridad Integrada

Jitters o fluctuaciones

La fluctuación, es la diferencia absoluta promedio entre intervalos consecutivos, dividida por el intervalo promedio (un intervalo es el tiempo entre dos puntos consecutivos). Como el jitter se usa a menudo como una medida de la calidad de la voz, los intervalos a menudo se consideran períodos glóticos. Por esta razón, para su cálculo se debe determinar los límites de la posible duración del intervalo o la posible diferencia en la duración de los intervalos consecutivos. La fluctuación de fase local se define como la diferencia de segundo orden absoluta del proceso puntual y está dado como un valor entre 0 y 2. Está definido por la ecuación 3.11.

$$jitter = \sum_{i=1}^N \frac{|T_i - T_{i-1}|}{N - 1} \quad (3.11)$$

Donde T_i es la duración del i -ésimo intervalo y N es el número de intervalos.

Duración del segmento de discurso

Longitud del segmento del discurso en segundos. Se calcula a través de la ecuación 3.12:

$$Drtn = \frac{M}{F_s} \quad (3.12)$$

Donde M es el número de muestras del segmento de discurso y F_s es la frecuencia de muestreo.

El diagrama de la figura 3.10 ilustra la secuencia de extracción de las características no

lingüísticas de la señal de voz:

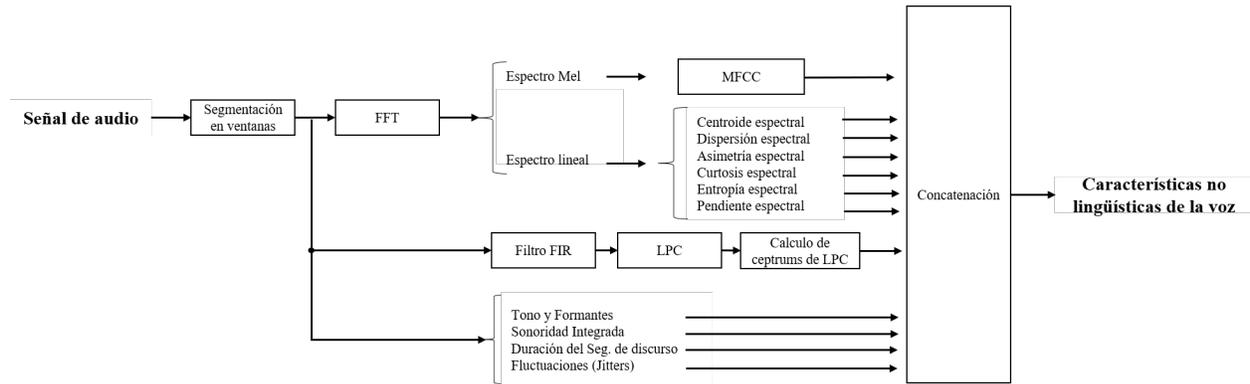


Figura 3.10: Pipeline general para la extracción de las características de voz

3.9. Análisis Estadístico

3.9.1. Verificación de los criterios estadísticos

Prueba de bondad de ajuste a distribución Normal-Lilliefors

La prueba de Lilliefors es una prueba de bondad de ajuste usada para comprobar si un conjunto de observaciones se ajusta a una distribución normal. Esta prueba es una mejora sobre la prueba de Kolomogorov-Smirnov (K-S), corrigiendo la K-S para valores pequeños en las colas de las distribuciones de probabilidad. A diferencia de la prueba K-S, Lilliefors se puede usar cuando no se conoce la media o la desviación estándar de la población. Esencialmente, la prueba de Lilliefors es una prueba K-S que le permite estimar estos parámetros a partir de su muestra. La definición estadística de la prueba de Lilliefors está dada por la ecuación 3.13

$$D = \max_x |F(\hat{x}) - G(x)| \quad (3.13)$$

Donde \hat{F}_x es la función de distribución acumulada empírica (CDF) de los datos de la muestra y $G(x)$ es el CDF de la distribución hipotética con parámetros estimados iguales a los parámetros de la muestra [57]. De esta forma, la prueba de Lilliefors establece las siguientes hipótesis:

- **H0:** Las observaciones de la muestra x proceden de una distribución normal.
- **H1:** Las observaciones de la muestra x no proceden de una distribución normal.

Por lo tanto, es necesario emplear dos tipos de prueba estadística según el resultado de la prueba de Lilliefors.

3.9.2. Variables Paramétricas (con distribución normal)

Se emplearon las pruebas estadísticas de prueba t de Student para dos muestras independientes para las comparaciones intrasujeto y una prueba de permutación de muestras basada en los estadísticos t para las comparaciones entre grupos. Se optó por emplear la prueba de permutación de muestra independiente basada en el estadístico t ya que esta prueba realiza la comparación estadística simultáneamente de múltiples variables [58].

3.9.3. Variables No Paramétricas (con distribución diferente a la normal)

Se empleó una prueba estadística de suma de rangos Wilcoxon para las distribuciones no paramétricas tanto en las comparaciones intrasujetos e intergrupos.

3.9.4. Pruebas entre grupos

Test de permutaciones [58]: test de significancia estadística para el estudio de diferencias entre grupos. La distribución del estadístico estudiado (media, mediana...) se obtiene calculando el valor de dicho estadístico para todas las posibles reorganizaciones de las observaciones en los distintos grupos. Dado que implica calcular todas las posibles situaciones, se trata de un test exacto.

En general su aplicación sigue la secuencia descrita a continuación:

- En primer lugar, se calcula la diferencia entre las medias de los dos grupos, lo que se conoce como diferencia observada.

- Todas las observaciones se combinan juntas sin tener en cuenta el grupo al que pertenecían.
- Se calculan todas las posibles permutaciones en las que las observaciones pueden ser distribuidas en dos grupos de tamaño distintos.
- Para cada permutación, se calcula la diferencia entre medias. El conjunto de valores calculados forma la distribución exacta de las posibles diferencias siendo cierta la hipótesis nula. A esta distribución se le conoce como “permutation distribution of the mean difference”.
- El p-value de dos colas se calcula como la proporción de permutaciones en las que, el valor absoluto de la diferencia calculada, es mayor o igual al valor absoluto de la diferencia observada.

3.9.5. Pruebas intra sujetos

Para las variables con una distribución no normal se emplea la prueba no paramétrica de rango de Wilcoxon. Esta prueba compara las medianas de las muestras. La **hipótesis nula** es que los datos en x y y son muestras de distribuciones continuas con medianas iguales, frente a la **alternativa** de que no lo son. El resultado h es 1 si la prueba rechaza la hipótesis nula al nivel de significancia del 5%, o 0 en caso contrario. Como alternativa, en el caso en que las variables tengan una distribución normal se aplica la prueba de *dos muestras t student*. La **hipótesis nula** es que los datos en los vectores x y y provienen de muestras aleatorias independientes de distribuciones normales con medias iguales y varianzas iguales pero desconocidas, La **hipótesis alternativa** es que los datos en x e y provienen de poblaciones con medias distintas. El resultado h es 1 si la prueba rechaza la hipótesis nula al nivel de significancia del 5%, o 0 en caso contrario.

3.9.6. Definición de las hipótesis

Nota: Interpretar “estadístico de prueba” como un término intercambiable por “media” o “mediana” según sea el caso (paramétrica o no paramétrica) teniendo en cuenta el resultado

de la prueba de normalidad y por lo tanto del tipo de prueba estadística aplicada.

Entre grupos (Pre intervención - LB)

H0: No existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición pre intervención de los grupos G1 y G2.

$$H_0 : \mu_{G1} = \mu_{G2} \quad (3.14)$$

H1: Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición pre intervención de los grupos G1 y G2.

$$H_1 : \mu_{G1} \neq \mu_{G2} \quad (3.15)$$

Entre grupos (Post intervención - MPI)

H0: No Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición post intervención de los grupos G1 y G2.

$$H_0 : \mu_{G1} = \mu_{G2} \quad (3.16)$$

H1: Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición post intervención de los grupos G1 y G2.

$$H_1 : \mu_{G1} \neq \mu_{G2} \quad (3.17)$$

Intra sujetos G1 (Pre vs Post-Intervención)

H0: No existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz de para la medición pre intervención y post intervención de un sujeto experimental de Grupo 1.

$$H_0 : \mu_{Sbj_iPre} = \mu_{Sbj_iPost} \quad (3.18)$$

H1: Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición pre intervención y post intervención de un sujeto experimental de Grupo 1.

$$H_1 : \mu_{Sbj_iPre} \neq \mu_{Sbj_iPost} \quad (3.19)$$

Intra sujetos G2 (Pre vs Pos-Intervención)

H0: No Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz de para la medición pre intervención y post intervención de un sujeto experimental de Grupo 2.

$$H_0 : \mu_{Sbj_iPre} = \mu_{Sbj_iPost} \quad (3.20)$$

H1: Existe una diferencia estadísticamente significativa entre los estadísticos de prueba de las características no lingüísticas de la señal de voz para la medición pre intervención y post intervención de un sujeto experimental de Grupo 2.

$$H_1 : \mu_{Sbj_iPre} \neq \mu_{Sbj_iPost} \quad (3.21)$$

Capítulo 4

Resultados - Discusión y Contribución

4.1. Resultados del Preprocesamiento

4.1.1. Filtrado de las señales

Para la etapa de procesamiento de la señal de voz, como se indicó en la metodología, se realizó una etapa de eliminación de nivel DC y filtrado la cual se ilustra en la siguiente figura 4.1 para la que se obtuvo una razón señal a ruido de 8.9 dB lo cual es una mejora considerable respecto de su valore inicial (-10.8 dB).

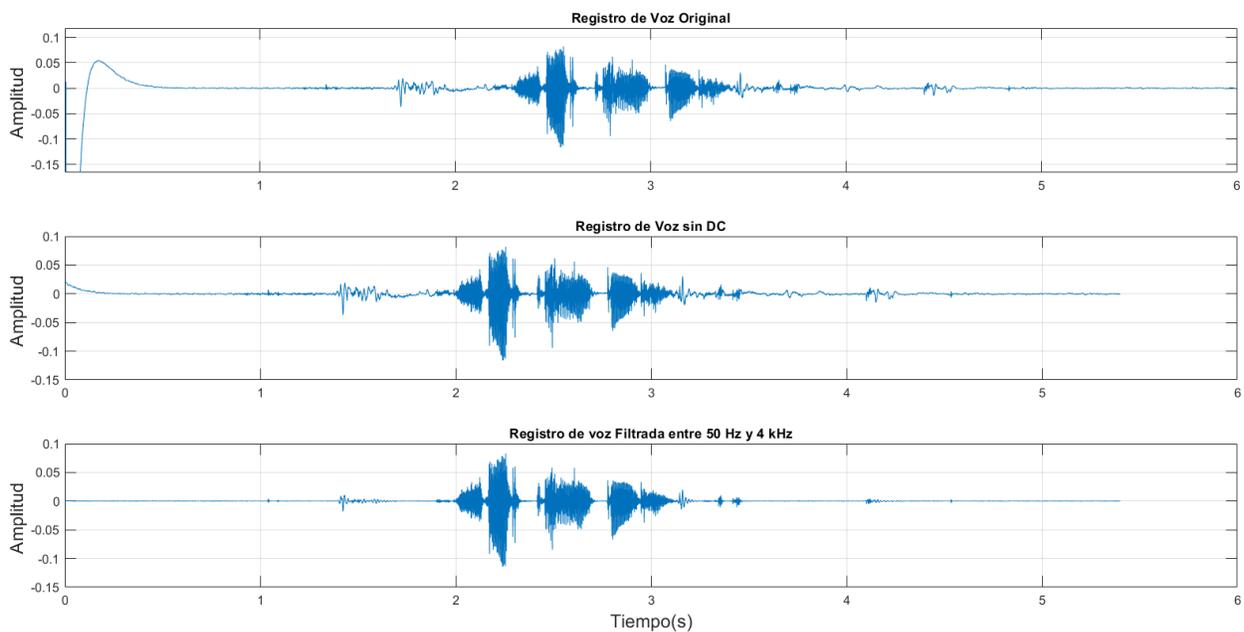


Figura 4.1: Resultado del filtrado

4.1.2. Eliminación de silencios y segmentación del discurso

La figura 4.2 ilustra el resultado del proceso de eliminación de silencios. Se puede apreciar que para este caso particular el algoritmo de detección de voz identificó dos segmentos de señal que son potencialmente voz, sin embargo, al evaluar los umbrales de clasificación anteriormente descritos el programa solo conserva el segundo segmento (delimitado por las fronteras rojas).

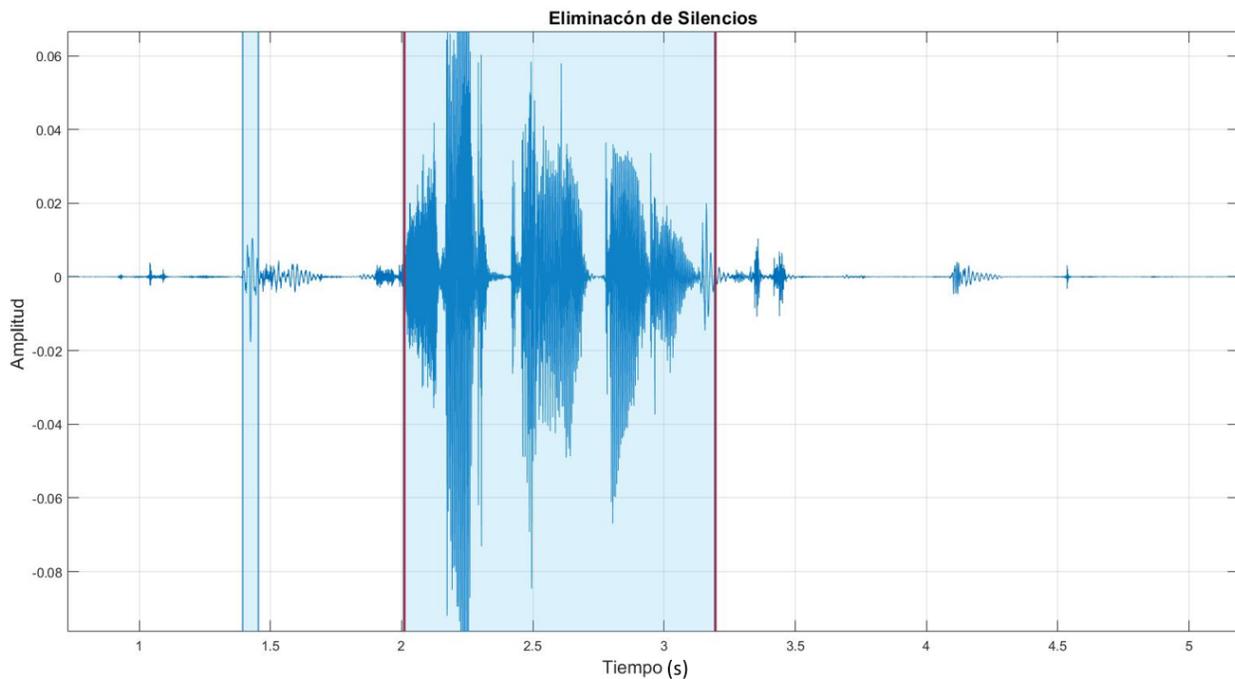


Figura 4.2: Resultado de eliminación de silencios

Por otra parte, la figura 4.3 ilustra la segmentación del discurso generado a partir de la lectura del texto neutro. En contraste con el proceso de eliminación de silencios, en este caso el sistema separa el registro de audio en sub segmentos según los umbrales de clasificación de voz y de duración. Por lo tanto, se logra agrupar unidades fonéticas que no necesariamente corresponden a unidades gramaticales. A modo de ilustración, para el texto: - . . . para un ecosistema. . . – puede verse segmentado en 2 unidades fonéticas de la forma “paraun” y “ecosistema”. Esta clasificación se considera acertada ya que el objetivo de este tipo de procesamiento no es identificar las unidades gramaticales sino las características fonéticas de

los segmentos de discurso descritas en tiempo y frecuencia. Aún más, esta segmentación es deseable ya que permite segmentar el registro de tal manera que los períodos transitorios de reacomodación del aparato sonoro queden agrupados y por lo tanto sea posible hacer una exploración de sus dinámicas a partir de las características no lingüísticas de la señal de voz.

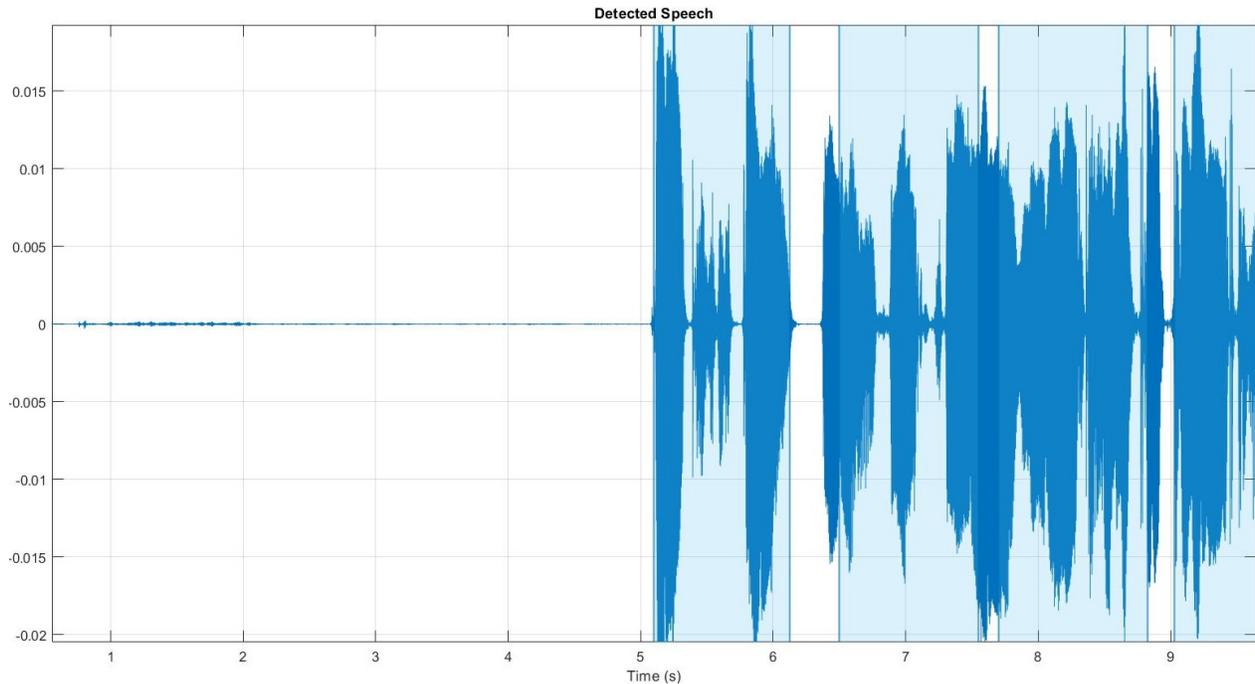


Figura 4.3: Segmentación del discurso

4.1.3. Procesamiento de la señal de voz

Del procesamiento de las señales de voz se obtuvieron dos estructuras de datos. Una para los resultados del procesamiento de la señal de voz de la lectura del texto neutro (TN) pre y post intervención y otra para los resultados del procesamiento de la señal de voz del conjunto de palabras individuales (P). La figura 4.4 presentan la arquitectura de las estructuras de los datos obtenidos.

Es importante destacar que el conjunto de características extraídas para el registro de TN y palabras es ligeramente diferente. La diferencia radica en el intercambio de las variables fluctuaciones y duración del segmento de discurso promedio extraídas para el registro TN por la característica de LPCC extraída para el registro de palabras. Este intercambio se

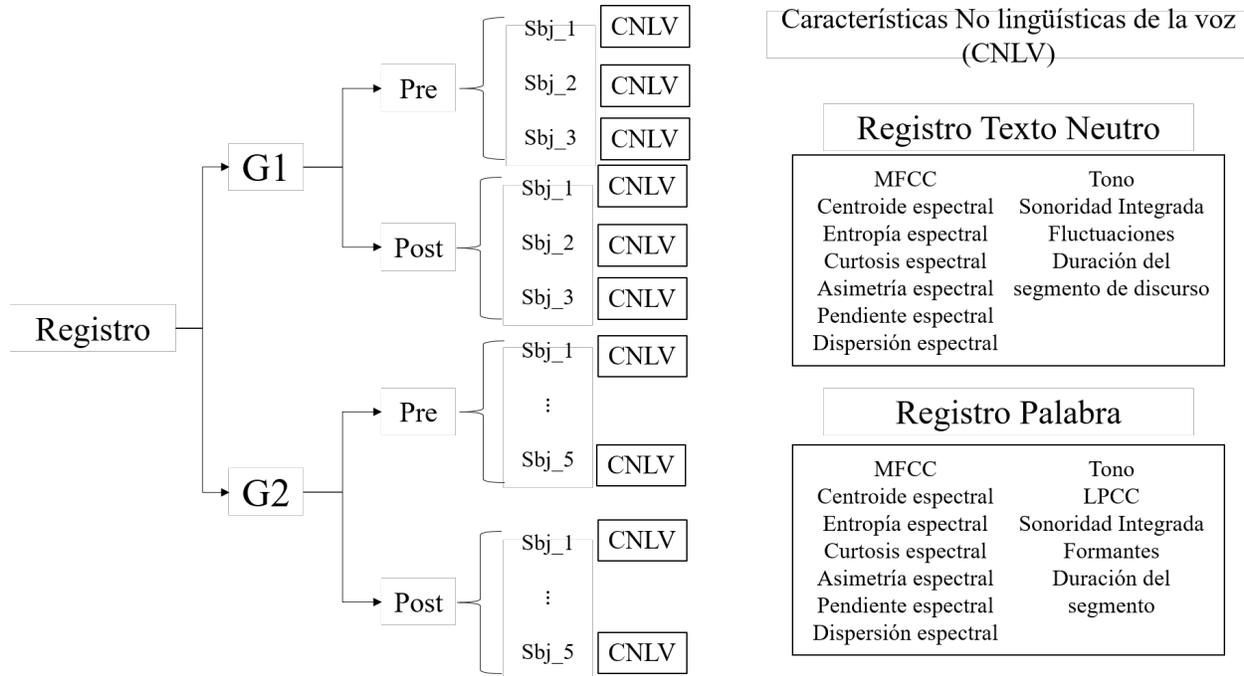


Figura 4.4: Arquitectura de la estructura de los datos obtenidos para el registro de TN y P

aplica, no porque las variables sean intercambiables, sino por interpretación apropiada de las variables a partir de su definición teórica anteriormente descrita. No tendría sentido hablar de una comparación de modelos autorregresivos para unidades fonéticas tan disímiles a las que se obtiene del proceso de segmentación del registro de TN, pero si cobra sentido hacer este tipo de comparación para el registro de palabras. Similarmente, no tiene sentido hablar de fluctuaciones entre segmentos de discurso o duración de segmento de discurso promedio cuando no existe un discurso en el sentido estricto como en el caso de los registros de las palabras.

Variación porcentual de la señal de voz

Las figuras 4.5 y 4.6 presentan una visión más global del comportamiento de las variables extraídas tanto para los registros de texto neutro como para los registros de palabras. Estas gráficas de barras representan la diferencia observable entre los eventos de línea base (LB) y la medición post intervención (MPI) en términos porcentuales.

En cuanto a la variación porcentual obtenida para los registros de lectura del texto neutro se puede observar que tanto para el grupo experimental como para el grupo control se obtu-

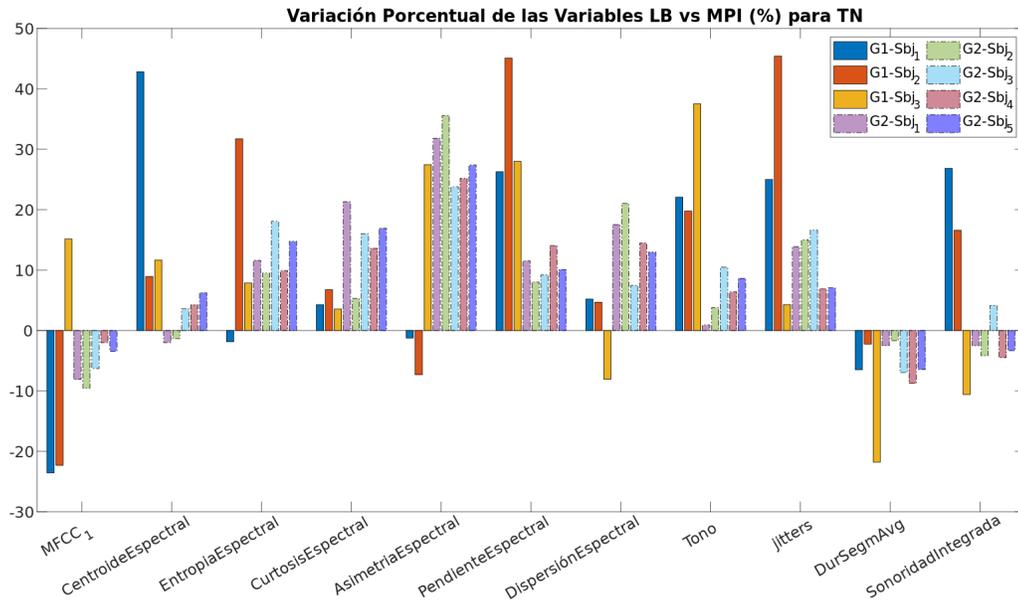


Figura 4.5: Variación porcentual Textos Neutros

vieron variaciones significativas. En el caso del grupo control, se destacan las variaciones de las variables espectrales, las cuales en su mayoría presentan un incremento de hasta un 45%. El aumento generalizado de este conjunto de variables, refleja un desplazamiento de la banda de frecuencias. Similarmente, el grupo experimental también presenta un rango de variación porcentual elevado, aunque ligeramente menor al del grupo control teniendo una variación porcentual máxima de 37%. Vale la pena mencionar, que el comportamiento de la variación porcentual es poco menos uniforme para los sujetos del grupo control comparada con los del grupo experimental. Esto es claramente apreciable, dado que el grupo experimental presenta 6 variables para las cuales uno o más participantes tuvieron variaciones opuestas al resto del grupo (i.e. Sonoridad Integrada) mientras que el grupo control presenta un total de 8 variables para las cuales se observa este mismo comportamiento a pesar de ser un grupo de menor tamaño. Esta mayor uniformidad del grupo experimental también se puede apreciar a partir de las diferencias porcentuales entre participantes del mismo grupo. Tomando una perspectiva más puntual sobre el análisis de las variables, se observa un desplazamiento de banda de frecuencias a partir del aumento de los centroides espectrales indicando una banda con frecuencias más altas. El ancho de banda instantáneo se mantiene constante. Del mismo

modo, el aumento en la asimetría señala una mayor dominancia del tono fundamental en comparación a las formantes. El aumento de la entropía sugiere un aumento en la duración de las secciones de habla sorda lo cual también se ve evidenciado en el aumento de la pendiente espectral. Esto junto con el aumento de las fluctuaciones y la sonoridad en adición a la disminución de la duración del segmento de discurso, pone en manifiesto un comportamiento representativo de la ansiedad-estado. Por otra parte, las variaciones de los MFCC no presentan un patrón claro el cual ayude a informar si el oyente hipotético podría distinguir entre los 2 registros. Esta observación es destacada por Villarreal y colegas [59] en el desarrollo de un polígrafo basado en señales de voz. En este estudio se destaca cómo ciertas características de la voz pueden ser “fingidas” o controladas conscientemente para enmascarar el estado mental del sujeto. De ahí, la importancia de emplear un amplio rango de características. Al contrastar estas observaciones con el comportamiento observado para los sujetos del grupo experimental, se aprecian algunas similitudes. Como, por ejemplo, el aumento porcentual de los centroides espectrales, aunque este se dé en menor medida, lo cual conlleva a afirmar que el desplazamiento de banda de frecuencias es de una menor magnitud. Lo sutil de este cambio puede deberse a que el grupo experimental tiene un número mayor de sujetos que aumentar la “inercia” o resistencia al cambio de las variables. Esta observación ratifica que el aumento de la dispersión espectral es mayor para este grupo. La variable con el mayor cambio porcentual es la asimetría espectral. Como se indicó anteriormente, esto indica una mayor dominancia del tono fundamental y resalta la relevancia que tiene el comportamiento del tono como una variable no lingüística para la identificación de la ansiedad-estado en comparación a las formantes. Por otra parte, el ligero aumento de la entropía sugiere un aumento en la duración de las secciones de habla sorda, el cual no fue tan notorio como para el grupo control y nuevamente esto se ve respaldado por el comportamiento de la pendiente espectral. Para el grupo experimental, el comportamiento de las fluctuaciones y la duración del segmento de discurso promedio siguen el comportamiento observado en el grupo control. Sin embargo, la sonoridad presenta un comportamiento opuesto. Nuevamente, se pone en manifiesto que para

una adecuada caracterización de la ansiedad-estado mediante el análisis de la señal de voz hace falta tomar una mirada global. Una observación parcial y hasta cierto punto intuitiva, como la disminución de la sonoridad integrada, podría conducir erróneamente a concluir que el nivel de ansiedad-estado ha disminuido cuando no es así, como se puede observar tomando en cuenta el resto de variables. En particular la variación porcentual máxima para Variables Congruentes (variable para la cual todos los sujetos presenta una variación positiva o negativa) entre participantes del grupo experimental es de 13% mientras que para el grupo control es 39% de diferencia. Sin embargo, ambos grupos presentan variaciones máximas, tanto positivas como negativas, con magnitudes comparables a diferencia del comportamiento observado para los registros de voz de palabras los cuales se presentan en la figura 4.6.

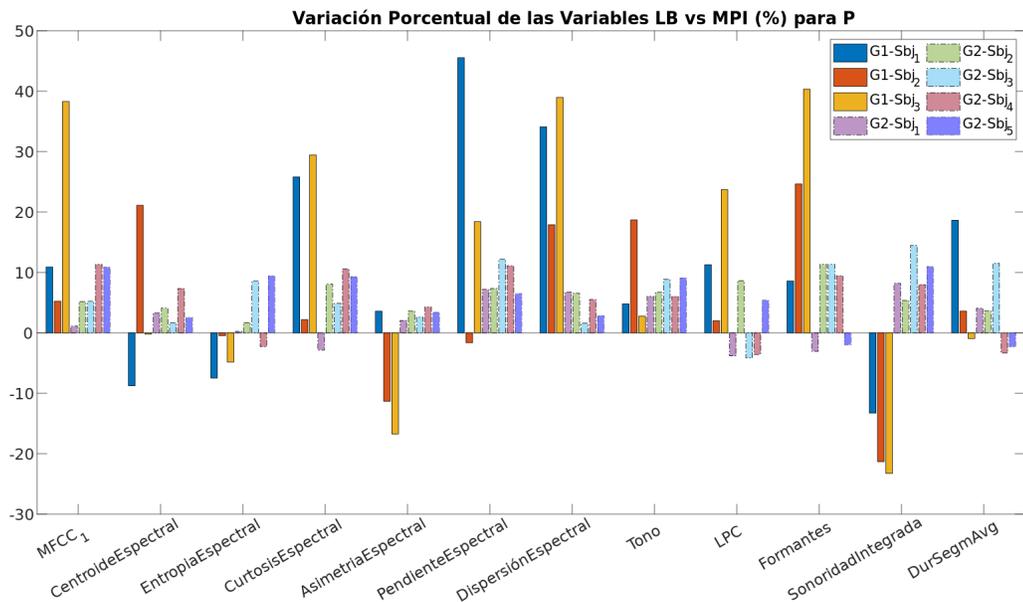


Figura 4.6: Variación porcentual Palabras

Para el caso de los registros de palabras, el rango de variación porcentual para el grupo experimental es considerablemente menor, tan solo de 14% mientras que para el grupo control es de 73%. Sin embargo, se observa un mayor número de variables congruentes para el grupo control 16 comparado con las 12 variables del grupo experimental. Es de destacar que, en la mayoría de estas variables no congruentes, se involucra el sujeto 1 del grupo experimental. Como se verá más adelante, este sujeto también reportó unas puntuaciones atípicas para las

pruebas psicométricas. Una variable de particular interés, como se mencionó anteriormente es la variable del tono (f_0), dado que ha sido identificada en numerosos estudios como una de las características no lingüísticas de la señal de voz más sensible a la ansiedad-estado. Se observa que esta variable presenta un comportamiento congruente y positivo para los dos grupos y para los dos tipos de registros lo cual concuerda con lo reportado en la literatura. Sin embargo, la variación porcentual es menor para el grupo experimental tanto en el registro de lectura de texto neutro como para el de las palabras 8 % y 9 % respectivamente, comparado con el 38 % del texto neutro y el 17 % de palabras observado en el grupo control. Esta diferencia en magnitudes también se ve reflejada para los centroides espectrales. Para el grupo control esta variable es incongruente y sumado al comportamiento de la asimetría espectral sugiere que hay sesgo negativo asociado a dominancia de las formantes superiores. Incluso el aumento de la formante la cual es mayor al del tono comprueba un desplazamiento de la banda de frecuencias hacia las frecuencias superiores aumentando el desprecio de la banda espectral. Adicionalmente la disminución de la entropía asociada a una mayor presencia de regiones de habla sonora sugiere una articulación vocal más lenta, lo cual puede deberse al aumento de la tensión muscular del tracto vocal en condición de ansiedad-estado. En contraste, el grupo experimental, a pesar de presentar en general un aumento en las variables anteriormente mencionadas, la magnitud del cambio respecto a los registros de la línea base son considerablemente menores. Esto sugiere que la bioalimentación se constituye como un factor relevante para el manejo del efecto de la ansiedad-estado en el control motor asociado a la producción de voz y por extensión para el sistema nervioso autónomo.

Resultados de pruebas estadísticas de la señal de voz

En cuanto a la comprobación de las hipótesis estadísticas planteadas en la metodología, la tabla 4.1 presentan los resultados para la determinación de la existencia de diferencias estadísticamente significativas entre los grupos.

A partir de los resultados presentados en la tabla 4.1, es posible afirmar que no existe una diferencia estadísticamente significativa entre los grupos control y experimental para los

Cuadro 4.1: Resultados pruebas estadísticas entre grupos

Tipo de Registro	Entre grupos línea base (LB)						Entre grupos post intervención (MPI)					
	Palabras			Texto Neutro			Palabras			Texto Neutro		
Variables	Hipótesis	P-Valor	Valor Medio	Hipótesis	P-Valor	Valor Medio	Hipótesis	P-Valor	Valor Medio	Hipótesis	P-Valor	Valor Medio
MFCC	H1	0.095	-10.2	H1	0.084	-5.8	H1	0.039*	-18.1	H1	0.073	-12.4
Centroide Espectral	H0	0.048*	21.1	H0	0.016*	25.4	H1	0.052*	7.5	H0	0.025*	23.4
Entropía Espectral	H0	0.016*	3.5	H0	0.046*	4.1	H0	0.450	4.2	H0	0.025*	3.8
Curstosis Espectral	H0	0.041*	3.6	H0	0.025*	3.5	H0	0.067	3.4	H0	0.014*	3.4
Asimetría Espectral	H0	0.05 *	1.2	H0	0.038*	1.1	H1	0.027*	3.5	H0	0.033*	1.2
Pendiente Espectral	H0	0.013*	-1.8	H0	0.002*	-1.3	H1	0.039 *	-0.2	H0	0.018*	-1.1
Dispersión Espectral	H0	0.03*	14.6	H0	0.019*	15.2	H1	0.329	6.2	H0	0.049*	14.5
Tono	H0	0.005*	134	H0	0.037*	126	H1	0.015*	142	H0	0.042*	203
LPCC	H0	0.036*	12.5				H1	0.950	5.1			
Sonoridad Integrada	H1	0.051*	9.2	H0	0.029*	7.8	H1	0.029*	9.8	H1	0.574	14.5
Formantes	H1	0.085	165				H1	0.885	174			
Duración del segmento	H0	0.011*	3.2				H1	0.055	2.1			
Fluctuaciones				H0	0.026*	4				H1	0.519	13
Duración del segmento de discurso				H0	0.05*	4.2				H1	0.420	2.7

H0: No existe una diferencia estadísticamente significativa.

H1: Existe una diferencia estadísticamente significativa

registros de LB (texto neutro y palabras) ni para el registro MPI de TN (H0 con P-valores ≤ 0.05 para la mayoría de las características). Por otra parte, sí se observa una diferencia estadísticamente significativa entre los grupos control y experimental para los registros de Palabras en la MPI (H1 con P-valores ≤ 0.05 para la mayoría de las características). Este resultado es de particular relevancia ya que es en este evento en el cual se emplea la biorealimentación como estrategia para el manejo de la ansiedad-estado y por lo tanto es en esta etapa en la que se espera observar unas diferencias entre los registros. En base á lo anterior, es posible afirmar que:

1. Para una comparación entre grupos, las características no lingüísticas extraídas durante la etapas de línea base (texto neutro y palabras) y la etapa de MPI de texto neutro, son estadísticamente equivalentes. Por lo tanto, se puede afirmar con un nivel de confianza estadística de $\alpha = 0,05$ que la condición psicofisiológica de los participantes al inicio de la prueba y previo a la intervención mediante la biorealimentación fue equivalente. Se puede ver que la tabla 4.1 reporta una diferencia estadística para las variables de MFCC , sonoridad integrada y formantes. Sin embargo, como se explicó anteriormente en la sección de resultados y análisis de la variación porcentual este comportamiento no contrarresta la evidencia asociada a las demás variables al analizarla tomando una perspectiva global de la señal.
2. Para una comparación entre grupos, las características no lingüísticas extraídas durante

las etapas de MPI de palabras, existe una diferencia estadísticamente significativa entre los grupos control y experimental. Por lo tanto, se puede afirmar con un nivel de confianza estadística de $\alpha = 0,05$ que la condición psicofisiológica de los participantes cambió como resultado del proceso de biorealimentación. No obstante, se podría llegar a plantear objeciones para la afirmación anterior, con base en el comportamiento de las variables entropía y curtosis espectral según se presentan en la tabla 4.1, dado que no se cuenta con la suficiente evidencia estadística como para rechazar la hipótesis nula. Sin embargo, como se mencionó previamente en la sección de análisis de las variaciones porcentuales para el registro de palabras, estas variables están relacionadas con la presencia de habla sorda y la planitud del espectro de frecuencias respectivamente. Dado que el conjunto de palabras que se le presentó a los dos grupos fue el mismo, este comportamiento es de esperarse y no se opone a la evidencia estadística del resto de variables.

3. Para una adecuada identificación de la ansiedad-estado a partir de las características no lingüísticas de la señal de voz se debe prestar particular atención a aquellas características que describan la variación de la banda de frecuencias del hablante ya que estas son las más susceptibles al aumento de la rigidez del aparato sonoro. No obstante, se hace necesario tener en cuenta las características que permitan tomar una mirada más amplia respecto del contexto fonético del hablante como la sonoridad, fluctuaciones, MFCC, LPCC entre otras, para asegurar un panorama completo de la condición del hablante. En síntesis, una adecuada caracterización de la ansiedad-estado requiere como mínimo las variables de tono, centroides, asimetría y pendiente espectral acompañadas de las variables contextuales sonoridad integrada y MFCC, no olvidando que contar con otras características facilita la identificación de la ansiedad-estado

Similarmente, las tablas 4.2 y 4.3 presentan los resultados para la determinación de la existencia de diferencias estadísticamente significativas intra sujetos. En otras palabras, determina si existen diferencias estadísticamente significativas entre los valores de las variables

no lingüísticas de la señal de voz entre los momentos evaluativos LB y MPI. A partir de los resultados de dichas tablas, se puede afirmar que:

1. No se observan diferencias estadísticamente significativas entre las variables de los registros de LB y MPI de palabras para los sujetos de grupo experimental
2. Por el contrario, sí se observa una diferencia significativa para las variables de los sujetos del grupo control.

Este resultado es congruente con el resultado asociado a la biorealimentación antes mencionado. Por otra parte, sí se observan diferencias estadísticamente significativas para las comparaciones LB y MPI intra-sujeto de TN tanto para el grupo experimental como para el grupo control.

De esta forma, los resultados de las pruebas estadísticas apuntan a que la estrategia de biorealimentación tiene un efecto estadísticamente relevante y, por lo tanto, es posible afirmar que la biorealimentación provista a los participantes del grupo experimental durante el proceso de medición de palabras post intervención, contribuyó satisfactoriamente a que los participantes retornaran a niveles de activación fisiológica estadísticamente equivalentes a aquellos registrados para la línea base. Adicionalmente, es posible afirmar que esta recuperación de los estados basales no se debe simplemente a un proceso espontáneo asociado únicamente al transcurso del tiempo ya que de ser el caso no se habría observado diferencias estadísticamente significativas en la comparación intra sujetos de las etapas de lectura de texto neutro de línea base y medición post intervención de ansiedad-estado. El resultado de esta prueba estadística concuerda con las variaciones porcentuales anteriormente presentadas y descritas.

Estos resultados, son quizás más apreciables visualmente a partir de las gráficas de caja y bigote entre grupos e intra sujetos. Para la ejemplificación visualmente de los datos de la tabla 4.3 se optó por emplear la variable *tono* (f_0) ya que, como se mencionó anteriormente, es una de las características no lingüísticas de la señal de voz más susceptibles a la ansiedad-estado según lo reportado por la literatura [36]. Las figuras 4.7 a 4.9 correspondientes a las

Cuadro 4.2: Resultados pruebas estadísticas intra sujetos (LB vs MPI) - Etapa TN

Intra Sujetos (LB vs MPI) TN		G1			G2				
		1	2	3	1	2	3	4	5
MFCC	Hipótesis	H1							
	P-Valor	0.054	0.048*	0.875	0.010*	0.072	0.035*	0.021*	0.045*
Centroide Espectral	Hipótesis	H1							
	P-Valor	0.961	0.035*	0.025*	0.039*	0.047*	0.111	0.316	0.053*
Entropía Espectral	Hipótesis	H1	H0	H1	H1	H1	H1	H1	H0
	P-Valor	0.046*	0.035*	0.015*	0.049*	0.038*	0.058	0.062	0.003*
Curtosis Espectral	Hipótesis	H1	H0	H0	H1	H1	H1	H1	H1
	P-Valor	0.01*	0.023*	0.041*	0.061	0.013*	0.048*	0.058	0.044*
Asimetría Espectral	Hipótesis	H1	H0	H1	H1	H1	H0	H1	H1
	P-Valor	0.984	0.012*	0.879*	0.023*	0.062	0.036*	0.05*	0.023*
Pendiente Espectral	Hipótesis	H1							
	P-Valor	0.045*	0.042*	0.784	0.090	0.018*	0.067	0.042*	0.021*
Dispersión Espectral	Hipótesis	H1	H0	H1	H0	H1	H1	H1	H1
	P-Valor	0.028*	0.034*	0.926	0.04*	0.50	0.01*	0.61	0.029*
Tono	Hipótesis	H0	H1						
	P-Valor	0.031*	0.015*	0.034*	0.033*	0.014*	0.009*	0.010*	0.009*
Fluctuaciones	Hipótesis	H1	H0	H1	H1	H1	H1	H1	H1
	P-Valor	0.841	0.452	0.765	0.036*	0.018*	0.016*	0.013*	0.047*
Duración del Segmento	Hipótesis	H1							
	P-Valor	0.554*	0.670	0.725	0.083*	0.016*	0.0629	0.08*	0.230
Sonoridad Integrada	Hipótesis	H1	H0	H1	H1	H1	H1	H0	H1
	P-Valor	0.954	0.032*	0.768	0.067	0.014*	0.014*	0.01*	0.055*

H0: No Existe una diferencia estadísticamente significativa.

H1: Existe una diferencia estadísticamente significativa

Cuadro 4.3: Resultados pruebas estadísticas intra sujetos (LB vs MPI) - Etapa P

Intra Sujetos (LB vs MIP) Palabras		G1			G2				
		1	2	3	1	2	3	4	5
MFCC	Hipótesis	H1	H1	H1	H0	H0	H0	H1	H0
	P-Valor	0.026*	0.019*	0.017*	0.050*	0.044*	0.048*	0.053*	0.044*
Centroide Espectral	Hipótesis	H1	H1	H1	H0	H0	H0	H0	H0
	P-Valor	0.022*	0.036*	0.054*	0.048*	0.039*	0.037*	0.025*	0.033*
Entropía Espectral	Hipótesis	H1	H1	H1	H0	H0	H0	H0	H1
	P-Valor	0.024*	0.055*	0.043*	0.043*	0.012*	0.011*	0.056	0.100
Curtosis Espectral	Hipótesis	H1	H0	H1	H0	H0	H0	H0	H1
	P-Valor	0.702	0.024*	0.076	0.031*	0.042*	0.039*	0.037*	0.078
Asimetría Espectral	Hipótesis	H1	H1	H1	H1	H0	H0	H0	H0
	P-Valor	0.400	0.024*	0.047*	0.038*	0.054*	0.047*	0.048*	0.011*
Pendiente Espectral	Hipótesis	H1	H1	H1	H0	H0	H0	H0	H0
	P-Valor	0.559	0.621	0.551	0.015*	0.012*	0.028*	0.037*	0.021*
Dispersión Espectral	Hipótesis	H1	H1	H1	H0	H0	H0	H1	H1
	P-Valor	0.945	0.036*	0.067	0.048*	0.015*	0.018*	0.010*	0.095
Tono	Hipótesis	H1	H1	H1	H0	H0	H1	H0	H0
	P-Valor	0.029*	0.044*	0.354	0.015*	0.047*	0.057	0.019*	0.038*
LPCC	Hipótesis	H1	H1	H1	H0	H0	H0	H1	H0
	P-Valor	0.033*	0.073	0.042*	0.072	0.027*	0.048*	0.105	0.023*
Formantes	Hipótesis	H1	H1	H1	H0	H0	H0	H0	H0
	P-Valor	0.027*	0.043*	0.954	0.013*	0.034*	0.055*	0.054*	0.011*
Sonoridad Integrada	Hipótesis	H1	H1	H1	H0	H0	H0	H0	H0
	P-Valor	0.1	0.072	0.046*	0.012*	0.011*	0.014*	0.085	0.049*
Duración del segmento	Hipótesis	H1	H1	H1	H0	H0	H0	H1	H1
	P-Valor	0.75	0.045*	0.049*	0.051*	0.031*	0.041*	0.140	0.015*

H0: No Existe una diferencia estadísticamente significativa.

H1: Existe una diferencia estadísticamente significativa

mediciones de TN (LB y MPI) y Palabras (LB) muestran como la media de las distribuciones no difiere significativamente, mientras que para la medición de Palabras (MPI) sí, como se puede observar en la figura 4.10.

Comparación entre grupos control (G1) y experimental (G2) para la medición texto neutro

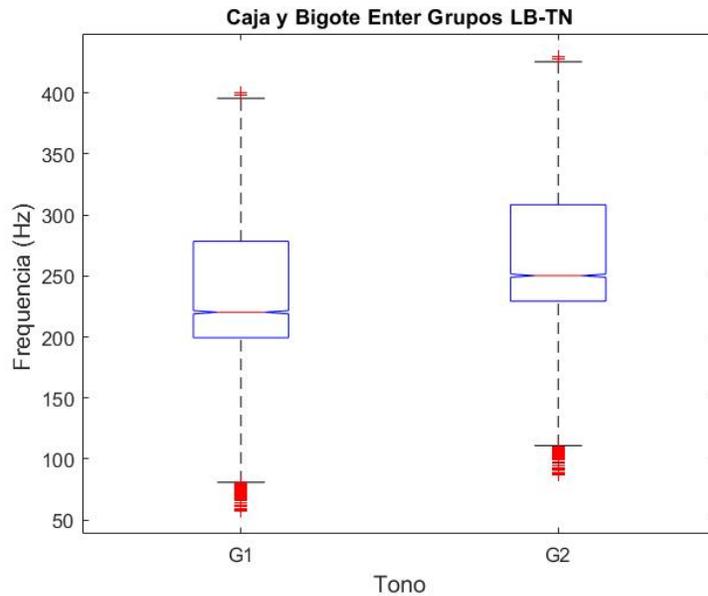


Figura 4.7: Comparación entre grupos del tono para registro durante la etapa de línea base (LB-TN)

Comparación entre grupos control (G1) y experimental (G2) para la medición de palabras (P)

Para ilustrar la comparación del comportamiento intra sujeto anteriormente descrito se eligió el sujeto 2 tanto para el grupo control como para el grupo experimental. En particular se optó por estos participantes para esta ilustración, ya que no presentan valores atípicos que podrían llegar a sugerir un sesgo desmedido en favor de las afirmaciones anteriormente expuestas.

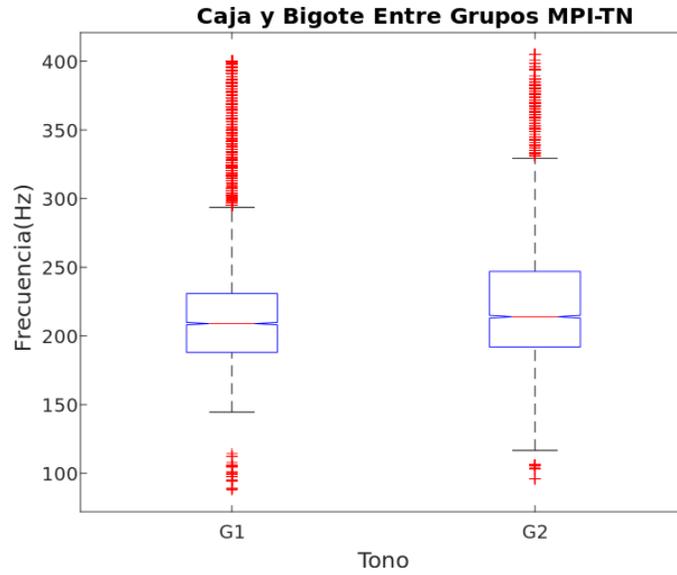


Figura 4.8: Comparación entre grupos del tono para registro durante la etapa de medición post intervención (MPI-TN)

Comparación intra sujeto Sbj 2-G1 entre etapas LB y MPI

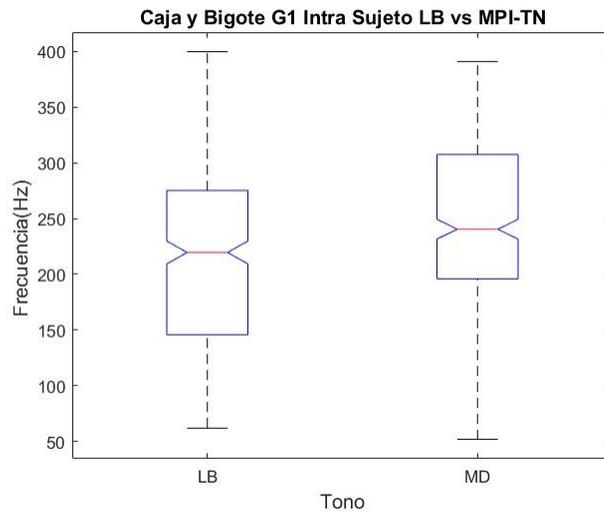


Figura 4.11: Comparación intra sujeto G1 del tono para registro LB y MPI de TN

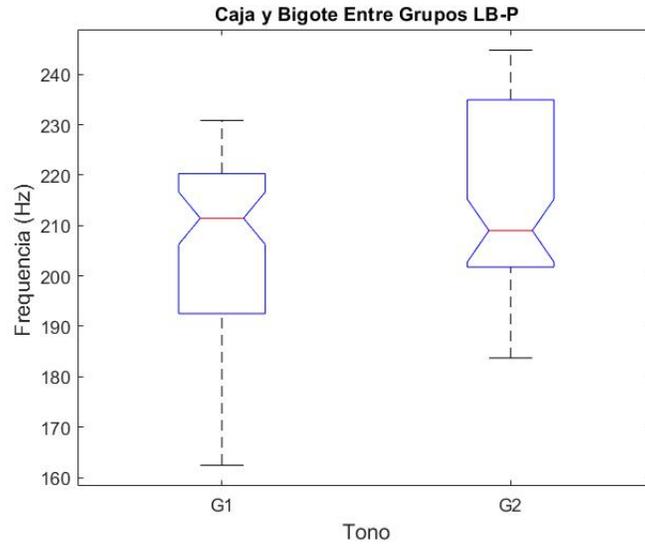


Figura 4.9: Comparación entre grupos del tono para registro durante la etapa de línea base (LB-P)

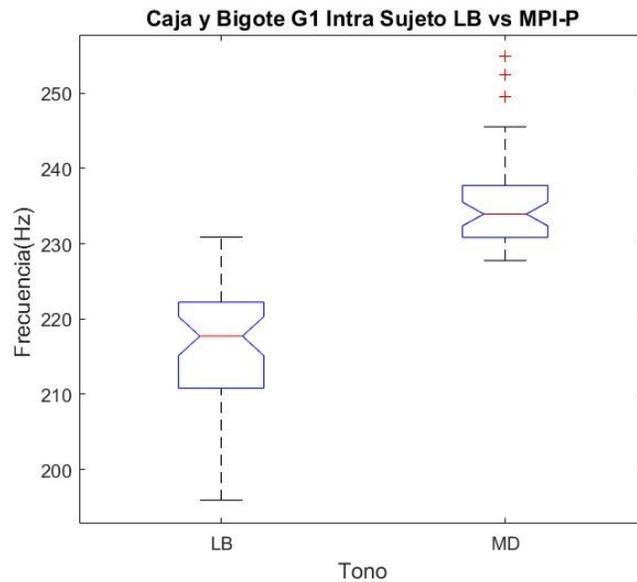


Figura 4.12: Comparación intra sujeto G1 del tono para registro LB y MPI de Palabras

Comparación intra sujeto Sbj 2-G2 entre etapas LB y MPI

Como se puede observar en la gráfica de la figura 4.12, la variación del tono para la etapa de medición de las palabras es claramente diferenciable de la etapa de línea base. Este comportamiento concuerda con la evolución de la señal de voz reportada por la literatura cuando el hablante se encuentra bajo condiciones de ansiedad-estado [36].

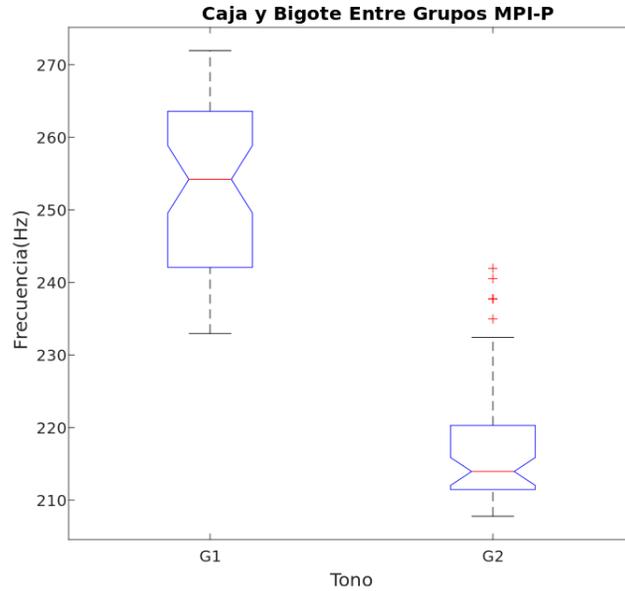


Figura 4.10: Comparación entre grupos del tono para registro durante la etapa de medición post intervención (MPI-P)

4.1.4. Resultados de las pruebas psicométricas

SAM

La prueba SAM contempla tres dimensiones: valencia, dominancia y activación, las cuales se evalúan mediante una puntuación en el rango de 1 a 9 puntos dónde 5 puntos hacen referencia a un estado neutro para esa dimensión en particular. Para la dimensión de la valencia, una puntuación de 1 significa que el sujeto reporta estar percibiendo una emoción muy negativa (deseando alejarse del estímulo) y una puntuación de 9 indica una emoción percibida como muy positiva (deseando acercarse al estímulo). Por otra parte, para la dimensión de activación, 1 punto significa estar lo más pasivo y 9 puntos lo más activo posible. Finalmente, para la dimensión de la dominancia, 1 significa que el participante se siente dominado por la emoción y 9 que se siente dominante sobre la emoción.

El comportamiento esperado de los datos obtenidos a partir de esta prueba psicométrica, suponiendo que la actividad de inducción de estrés cumpliera su objetivo sería el siguiente: Para la dimensión de la valencia se espera ver una disminución en la puntuación respecto

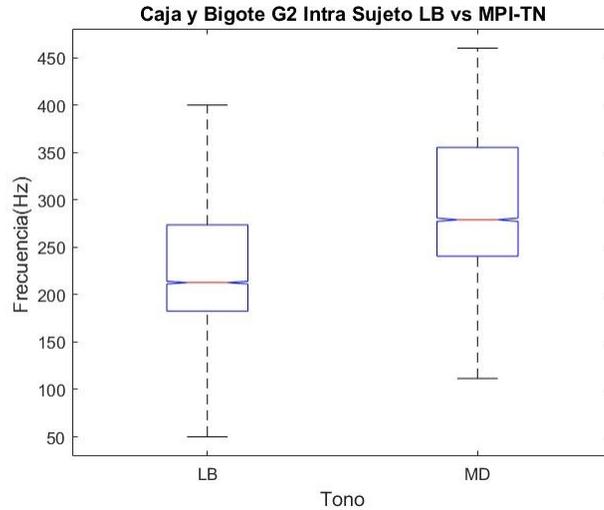


Figura 4.13: Comparación intra sujeto G2 del tono para registro LB y MPI de TN

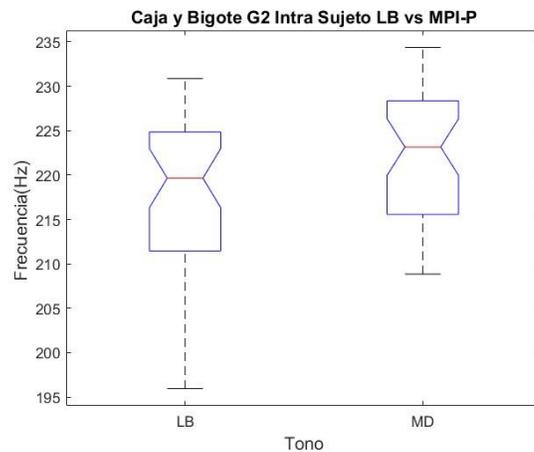


Figura 4.14: Comparación intra sujeto G2 del tono para registro LB y MPI de Palabras

a los niveles de línea base lo cual indica que los participantes perciben la situación como desagradable. Para la dimensión de la dominancia se espera que la puntuación de la prueba disminuya respecto a la puntuación de la línea base, lo cual indica que los participantes perciben tener menos dominio sobre la situación. Por el contrario, para la dimensión de la activación se espera que la puntuación de la prueba aumente respecto a la puntuación de la línea base, lo cual indica que los participantes perciben sentirse más activados o alerta ante las exigencias de la situación [48]. A partir, de lo anterior resulta fácil deducir que la recuperación del estado basal al haber superado la condición de ansiedad-estado se manifiesta

de forma inversa a los cambios antes descritos. Este proceso de recuperación se puede dar de forma activa (mediada por una intervención) o pasiva (proceso homeostático asociado al paso del tiempo) [38].

En las figuras 4.15 a 4.17 se presentan los resultados obtenidos a partir de la prueba psicométrica SAM para las tres mediciones línea base de (LB), post actividad (PA) y medición post intervención (MPI).

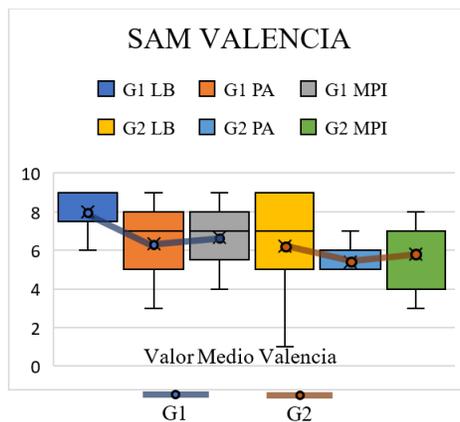


Figura 4.15: Resultados test psicométrico SAM dimensión Valencia

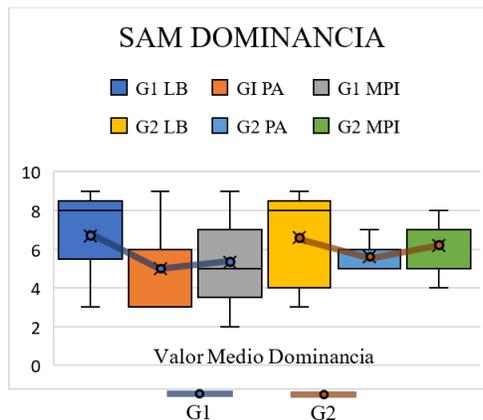


Figura 4.16: Resultados test psicométrico SAM dimensión Dominancia

La figura 4.16 presenta una representación gráfica de los resultados de la prueba psicométrica SAM para sus tres dimensiones. Se puede apreciar cómo las tres dimensiones tuvieron un comportamiento consistente para los 2 grupos experimentales. En general, para las tres dimensiones tanto el grupo control (G1) como el experimental (G2) tuvieron cambios aprecia-

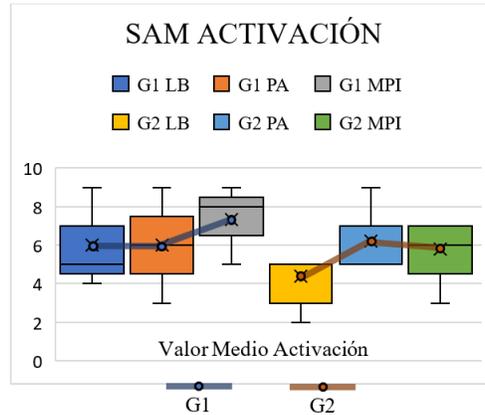


Figura 4.17: Resultados test psicométrico SAM dimensión Activación

bles para los tres momentos evaluativos: línea base (LB), post inducción de ansiedad-estado (PA) y medición post intervención (MPI). Similarmente, se puede apreciar un cambio entre las puntuaciones de LB y PA tanto para G1 como para G2 en las dimensiones de valencia y dominancia. Este resultado sugiere que la actividad de inducción de ansiedad-estado fue efectiva en modificar el estado psicológico percibido por los participantes.

Para las dimensiones de valencia y dominancia es posible apreciar cómo tanto G1 como G2 tuvieron comportamientos similares para las etapas LB y PA, pero se observa una clara diferencia al comparar los puntajes de la etapa evaluativa MPI. Para el G1 se observa una ligera recuperación a estados basales, mientras que para G2 la tendencia a la recuperación es más notable. En el caso de la dimensión de activación, la puntuación de LB del G1 fue equivalente a la puntuación de G2 PA y se mantuvo constante inclusive durante la etapa PA. A pesar que este comportamiento no era de esperarse, puede ser atribuido al mayor grado de variabilidad presente en el G1 en razón al reducido número de participantes de este grupo. Sin embargo, como se puede apreciar en esa misma gráfica, sí hay una diferencia clara entre los puntajes MPI entre los grupos. Mientras que para G2, se aprecia una recuperación hacia los puntajes de LB, para G1 se observa un aumento en el nivel de activación emocional inclusive después de haber transcurrido un periodo de tiempo desde la terminación de la actividad de inducción de la ansiedad-estado. Al contrastar este resultado con el comportamiento observado durante las tres etapas, se observa que sí hay una recuperación a estados basales

activa para el G2. Se sabe que esta recuperación es activa dado que para el mismo período de tiempo la recuperación del G1 es mucho menor o empeora.

A pesar que existe una diferencia apreciable entre los grupos la magnitud de la misma no es la esperada. Este comportamiento inesperado puede deberse a los altos niveles de ansiedad causados por la coyuntura por la que está atravesando el mundo en este momento y por lo mismo los niveles de ansiedad no varían con facilidad. Sin embargo, es importante destacar que los resultados obtenidos a partir del procesamiento de las señales sugieren que la actividad estresante si tuvo el efecto deseado y que se puede atribuir la recuperación de los niveles basales del grupo experimental como un efecto de la biorealimentación.

IDARE

Por otra parte, se empleó el test psicométrico IDARE, el cual consta de 20 afirmaciones que se contestan al asignar una calificación en una escala tipo Likert, de 1 a 4 según el grado de conformidad que se tenga con dicha afirmación. La puntuación final tiene un mínimo de 20 puntos y un máximo de 80 puntos.

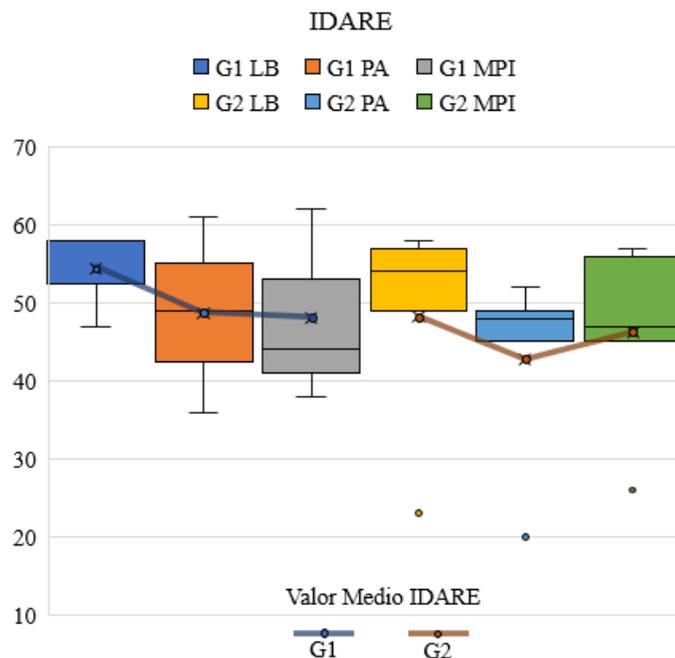


Figura 4.18: Gráfica de caja y bigote para la prueba psicométrica IDARE

La figura 4.18 presenta los resultados obtenidos para la prueba psicométrica IDARE durante las tres etapas evaluativas: línea base (LB), post actividad de inducción de ansiedad-estado (PA) y medición post intervención mediante biorealimentación (MIP) tanto para el grupo control (G1) como para el grupo experimental (G2). A primera vista se observa un comportamiento inusual para el conjunto de datos de G2, ya que la media del conjunto de datos (demarcada por el símbolo X) se encuentra por debajo del gráfico de caja, este comportamiento se debe a la presencia de valores extremos en el conjunto de datos. En una exploración más granular de los datos se identificó que el sujeto 1 del grupo G2 reportó puntuaciones extremadamente bajas durante los tres momentos evaluativos como se puede apreciar en la figura 4.18. Estas puntuaciones pueden deberse a que el sujeto no entendió el debido procedimiento para el diligenciamiento de la prueba psicométrica IDARE o que el sujeto se encontraba en una condición emocional extrema. Dado el hallazgo que las puntuaciones del test SAM y los valores de las características no lingüísticas de la señal de voz de este participante no son coherentes con estas puntuaciones extremas, se opta por la explicación asociada a la falta de comprensión del participante, con lo cual se decide retirar sus puntuaciones del análisis de los resultados de la prueba IDARE. No obstante, este resultado se presta para ilustrar como las pruebas psicométricas autoevaluativas son susceptibles a las subjetividades de los participantes y por lo tanto lo valioso de contar con herramientas complementarias de evaluación objetivas como la desarrollada a lo largo de esta investigación.

La figura 4.19 presenta los resultados obtenidos para la prueba psicométrica IDARE durante las tres etapas evaluativas corregida mediante la exclusión de los valores extremos.

Según los datos obtenidos para la etapa evaluativa LB, se obtuvo una media de 54.3 y 54.5 para el G1 y G2 respectivamente. Esto sugiere que los niveles de ansiedad-estado para los dos grupos no son significativamente diferentes y por lo tanto el comportamiento posterior no se relaciona con factores externos a la prueba realizada. Adicionalmente se observa que la variabilidad de los puntajes para esta etapa evaluativa es comparable entre los grupos, mientras que difiere notablemente para el resto de las etapas.

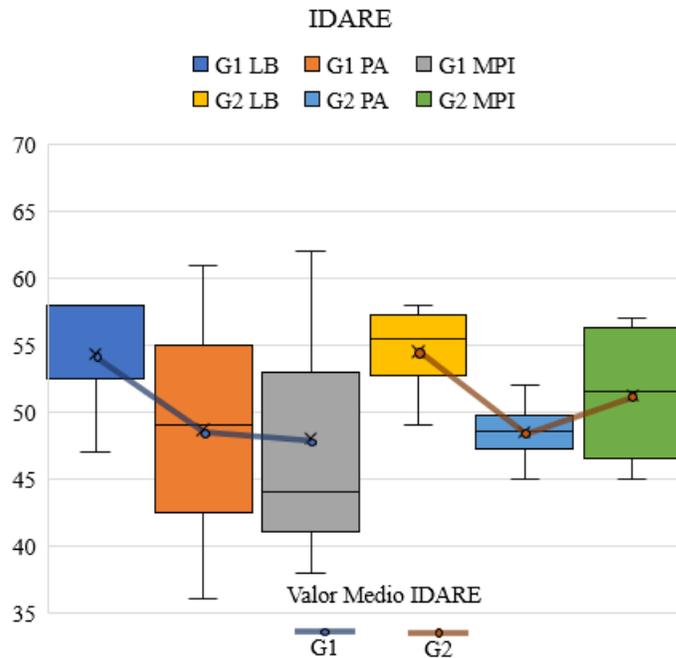


Figura 4.19: Gráfica de caja y bigote para la prueba psicométrica IDARE corregida

En cuanto a los resultados de la etapa PA, se observa una disminución del valor medio a 48.6 y 48.5 para G1 y G2 respectivamente. Este resultado confirma nuevamente que la actividad de inducción de ansiedad-estado cumple con su objetivo y que se observa una respuesta coherente entre las pruebas psicométricas y los resultados de la interpretación de las características extraídas de la señal de voz. Sin embargo, la gran diferencia en la variabilidad de los conjuntos de datos reafirma las apreciaciones anteriormente mencionadas en torno a la sensibilidad de este tipo de pruebas autoevaluativas a la subjetividad de los participantes. Mientras que el grupo G1 presentó unas puntuaciones máximas de 61 y una mínima de 36, el grupo G2 fue considerablemente más consistente con puntuaciones máxima de 58 y mínima de 49. Por consiguiente, según la revisión de literatura, se puede catalogar que en promedio los participantes de ambos grupos experimentaron niveles moderados de ansiedad [59].

Finalmente, los datos obtenidos para la etapa evaluativa MPI, se obtuvo una media de 48 y 51.3 para el G1 y G2 respectivamente. Se puede apreciar tanto visual como cuantitativamente que el grupo G2 tiende a recuperar los niveles basales de la etapa LB mientras que el G1

permanece bajo un estado de ansiedad-estado e incluso su puntuación empeora ligeramente (48.6 a 48.0). Este resultado es de particular relevancia ya que ilustra como la intervención por medio de la biorealimentación contribuye a la recuperación de los estados basales en los participantes del grupo experimental.

4.2. Aportes más Relevantes de la Presente Investigación

4.2.1. Diseño e implementación de sistema de biorealimentación para el manejo de la ansiedad-estado

Mediante la integración del diseño de software y el procesamiento de la señal de voz fue posible generar una herramienta para biorealimentación de la ansiedad-estado. A partir de la caracterización de la señal de voz en el dominio del tiempo y frecuencia fue posible identificar diferencias estadísticamente significativas entre el nivel de ansiedad-estado previo y posterior a la biorealimentación. Por lo tanto, se puede afirmar que la biorealimentación se constituye como una estrategia viable para el manejo de la ansiedad-estado.

4.2.2. Validación de la metodología experimental mediante resultados de pruebas psicométricas

Los hallazgos anteriormente mencionados fueron contrastados con la aplicación de dos tipos de pruebas psicométricas autoevaluativas (SAM e IDARE). En cuanto a la prueba SAM, se pudo observar cómo las dimensiones de valencia, activación y dominancia fueron sensibles a las diferentes etapas experimentales exhibiendo el comportamiento esperado para un proceso de inducción y recuperación de ansiedad-estado para el grupo experimental. Este comportamiento coincide con los resultados de otros estudios [60–63]. Similarmente, la evaluación IDARE, presentó una susceptibilidad a la metodología experimental análoga a la de la prueba psicométrica SAM. Por lo cual fue posible, corroborar que el diseño experimental logró inducir la ansiedad-estado, así como proveer una recuperación a niveles basales para los sujetos del grupo experimental mediante el uso de la biorealimentación. Los valores ob-

servados para esta prueba concuerdan con aquellos reportados por estudios que aplican este mismo instrumento para medir el estrés en entornos tan diversos como en la sala de operaciones [64] como en el monitoreo de niveles de ansiedad y habilidades de afrontamiento entre reclutas [65]. Sin embargo, vale la pena resaltar que en la revisión de la literatura relevante no se encontraron estudios recientes para los cuales se integrara el análisis de bioseñales y el uso de pruebas psicométrías para su validación. En general, se encontraron estudios que abordaban los aspectos técnicos de la medición y clasificación de la ansiedad y empleando métodos de validación basados en "teorías psicológicas existentes" pero no en pruebas psicométricas establecidas [66–70].

4.2.3. Identificación de variables no lingüísticas de la señal de voz relevantes para la identificación de la ansiedad-estado

En síntesis, el análisis estadístico de las características en tiempo y frecuencia de la señal de voz contrastado con los resultados obtenidos a partir de las pruebas psicométricas SAM e IDARE ponen de manifiesto que es posible identificar la presencia de ansiedad-estado de manera menos intrusiva que los sistemas de medición comerciales actuales, sin incurrir en la subjetividad de las herramientas tradicionales de autoevaluación. Sin embargo, teniendo en cuenta el tamaño de muestra empleado en este estudio, es prudente mencionar que la reproducción de estos resultados en con una muestra de mayor tamaño daría lugar a un mayor grado de impacto de los mismo. De este modo, para una adecuada identificación de la ansiedad-estado se debe prestar particular atención a aquellas características no lingüísticas de la señal de voz que describan la variación de la banda de frecuencias del hablante sin dejar de lado aquellas que aporten información del contexto fonético del mismo. Según como se puede apreciar a partir de las figuras 4.5 y 4.6 y de las tablas 4.1, 4.2 y 4.3 las características de tono, centroides, asimetría y pendiente espectral acompañadas de la sonoridad integrada y MFCC resultan de particular interés.

Capítulo 5

Conclusión y Trabajos Futuros

5.1. Conclusión

El objetivo de este trabajo fue el estudiar el efecto de un sistema de biorealimentación en el manejo del estrés situacional, en una población estudiantil del área de la salud. Para este fin se diseñó, implementó y aplicó de forma remota un protocolo experimental mediante un sistema de biorealimentación para el adecuado manejo de la ansiedad-estado. En una comparación con 2 pruebas psicométricas comúnmente empleadas para la identificación de la ansiedad-estado se observó algún grado de afinidad entre los resultados obtenidos. Se comprobó que la variación de las características extraídas, concuerda con lo reportado en la literatura científica [25, 36, 44, 71–73]. Así mismo, se observó que, la variable de la frecuencia fundamental o tono es particularmente sensible a la ansiedad-estado, hecho que se menciona en gran parte de la literatura científica relevante [36]. Sin embargo, dada la complejidad de la identificación de esta condición se hace necesario tener en cuenta el comportamiento de otras características de la señal de voz a fin de robustecer su identificación.

Esta investigación tuvo éxito en plantear un protocolo experimental útil para inducir la ansiedad-estado y diferenciar el habla bajo ansiedad-estado, promoviendo una reducción de la misma mediante la técnica de la biorealimentación. En términos de la variación porcentual se halló que para las características evaluada el grupo experimental obtuvo un aumento de 14.8 % entre la medición de línea base y la medición post intervención comparada con el 73 % de aumento para el grupo control.

5.2. Trabajos Futuros

Los resultados de este estudio dan paso a futuras investigaciones dirigidas a una validación del modelo aumentando el tamaño de la muestra con una población más grande y más diversa en términos de áreas del conocimiento y etapas del proceso educativo. En adición, es de interés identificando y evaluando diferentes vías sensoriales que podría demostrar ser más eficaz en el manejo de la ansiedad-estado con técnicas de biorealimentación.

Glosario

f_0 Frecuencia fundamental o Tono.

AAPB Asociación de Psicofisiología Aplicada y Biofeedback.

Ansiedad estado Estado emocional inmediato, caracterizado por un sentimiento de miedo, aprensión y tensión, acompañado o asociado con la activación del sistema nervioso autónomo.

Ansiedad rasgo la predisposición de una persona como rasgo de su personalidad a percibir ciertas situaciones como amenazadoras o no.

BCIA Alianza Internacional de Certificación de Biofeedback.

G1 Grupo Control.

G2 Grupo Experimental.

HHA Eje Hipotálamo-Hipófisis-Suprarrenal.

IDARE Inventario De la Ansiedad Rasgo-Estado.

ISNR Sociedad Internacional de Neurofeedback e Investigación.

LB Momento evaluativo Línea Base.

Lombard tendencia involuntaria a incrementar el esfuerzo vocal cuando se habla en un lugar ruidoso para mejorar la audibilidad de la voz.

LPC Coeficientes predictivos lineales.

MFCC Coeficientes cepstrales de frecuencia de Mel.

MPI Momento evaluativo Post Intervención.

P Registro de Señal de Voz para la Lectura del Conjunto de Palabras Individuales.

SAM Self Assesment Manikin.

SNA Sistema Nervioso Autónomo.

TN Registro de Señal de Voz para la Lectura del Texto Neutro.

Variables Congruentes variable no lingüísticas de la señal de voz para las cuales todos los sujetos del un grupo experimental determinado presenta un mismo tipo de variación bien sea positiva o negativa.

Bibliografía

- [1] A. Seidler, M. Thinschmidt, S. Deckert, F. Then, J. Hegewald, K. Nieuwenhuijsen, and S. G. Riedel-Heller, “The role of psychosocial working conditions on burnout and its core component emotional exhaustion – a systematic review,” *Journal of Occupational Medicine and Toxicology*, vol. 9, p. 10, mar 2014.
- [2] D. A. J. Salvagioni, F. N. Melanda, A. E. Mesas, A. D. González, F. L. Gabani, and S. M. de Andrade, “Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies.,” *PloS one*, 2017.
- [3] C. Maslach and S. E. Jackson, “The measurement of experienced burnout,” *Journal of Organizational Behavior*, vol. 2, pp. 99–113, apr 1981.
- [4] A. M. Richardsen and R. J. Burke, “Occupational stress and job satisfaction among physicians: sex differences.,” *Social science & medicine (1982)*, vol. 33, no. 10, pp. 1179–87, 1991.
- [5] D. S. Tawfik, J. B. Sexton, P. Kan, P. J. Sharek, C. C. Nisbet, J. Rigdon, H. C. Lee, and J. Profit, “Burnout in the neonatal intensive care unit and its relation to healthcare-associated infections,” *Journal of Perinatology*, vol. 37, pp. 315–320, mar 2017.
- [6] A. Griffiths, S. Leka, and T. Cox, “La organización del trabajo y el estrés: estrategias sistemáticas de solución de problemas para empleadores, personal directivo y representantes sindicales,” *WHO*, 2004.
- [7] S. Folkman and R. S. Lazarus, *Stress, appraisal, and coping*. New York: Springer Publishing Company, 1984.
- [8] R. S. W. Weinberg and D. Gould, *Foundations of sport and exercise psychology: Study guide*. Human Kinetics, 1999.
- [9] C. D. Spielberger, R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs, “Manual for the state-trait anxiety inventory (palo alto, ca, consulting psychologists press),” *Inc*, 1983.

- [10] G. Grossi, A. Perski, W. Osika, and I. Savic, "Stress-related exhaustion disorder—clinical manifestation of burnout? A review of assessment methods, sleep impairments, cognitive disturbances, and neurobiological and physiological changes in clinical burnout," *Scandinavian Journal of Psychology*, vol. 56, no. 6, pp. 626–636, 2015.
- [11] K. Azam, A. Khan, and M. T. Alam, "Causes and adverse impact of physician burnout: a systematic review," *J Coll Physicians Surg Pak*, vol. 27, no. 8, pp. 495–501, 2017.
- [12] D. S. Tawfik, J. Profit, T. I. Morgenthaler, D. V. Satele, C. A. Sinsky, L. N. Dyrbye, M. A. Tutty, C. P. West, and T. D. Shanafelt, "Physician burnout, well-being, and work unit safety grades in relationship to reported medical errors," in *Mayo Clinic Proceedings*, vol. 93, pp. 1571–1580, Elsevier, 2018.
- [13] L. B. Mills and E. S. Huebner, "A prospective study of personality characteristics, occupational stressors, and burnout among school psychology practitioners," *Journal of school psychology*, vol. 36, no. 1, pp. 103–120, 1998.
- [14] L. H. Hall, J. Johnson, I. Watt, A. Tsipa, and D. B. O'Connor, "Healthcare staff wellbeing, burnout, and patient safety: a systematic review," *PloS one*, vol. 11, no. 7, 2016.
- [15] G. S. de Oliveira Jr, R. Chang, P. C. Fitzgerald, M. D. Almeida, L. S. Castro-Alves, S. Ahmad, and R. J. McCarthy, "The prevalence of burnout and depression and their association with adherence to safety and practice standards: a survey of United States anesthesiology trainees," *Anesthesia & Analgesia*, vol. 117, no. 1, pp. 182–193, 2013.
- [16] G. Mayzell and B. Flareau, "Burnout: A healthcare crisis for us all," *The Resilient Healthcare Organization: How to Reduce Physician and Healthcare Worker Burnout*, 2020.
- [17] R. Salud, "Al año, 180.000 personas que son hospitalizadas mueren a causa de errores médicos," *El espectador*, 2016.
- [18] O. L. Paredes and P. A. Sanabria-Ferrand, "Prevalencia del síndrome de burnout en residentes de especialidades médico quirúrgicas, su relación con el bienestar psicológico y con variables sociodemográficas y laborales," *Revista Med*, vol. 16, no. 1, pp. 25–32, 2008.
- [19] J. McConville, R. McAleer, and A. Hahne, "Mindfulness training for health profession students—the effect of mindfulness training on psychological well-being, learning and clinical performance of health professional students: a systematic review of randomized and non-randomized controlled trials," *Explore*, vol. 13, no. 1, pp. 26–45, 2017.

- [20] S. Mahmoudzadeh, P. Mohammadkhani, B. Dolatshahi, and S. Moradi, “Prediction of psychological well-being based on dispositional mindfulness and cognitive emotion regulation strategies in students,” *Practice In Clinical Psychology*, 2015.
- [21] K. Olson, K. J. Kemper, and J. D. Mahan, “What factors promote resilience and protect against burnout in first-year pediatric and medicine-pediatric residents?,” *Journal of evidence-based complementary & alternative medicine*, vol. 20, no. 3, pp. 192–198, 2015.
- [22] M. C. Beach, D. Roter, P. T. Korhuis, R. M. Epstein, V. Sharp, N. Ratanawongsa, J. Cohn, S. Eggly, A. Sankar, R. D. Moore, and Others, “A multicenter study of physician mindfulness and health care quality,” *The Annals of Family Medicine*, vol. 11, no. 5, pp. 421–428, 2013.
- [23] P. Ratanasiripong, K. Sverduk, D. Hayashino, and J. Prince, “Setting up the next generation biofeedback program for stress and anxiety management for college students: A simple and cost-effective approach,” *College Student Journal*, vol. 44, no. 1, pp. 97–100, 2010.
- [24] I. Z. Khazan, *The clinical handbook of biofeedback: A step-by-step guide for training and practice with mindfulness*. John Wiley & Sons, 2013.
- [25] M. B. Danhof-Pont, T. van Veen, and F. G. Zitman, “Biomarkers in burnout: a systematic review,” *Journal of psychosomatic research*, vol. 70, pp. 505–24, jun 2011.
- [26] L. Kennedy and S. H. Parker, “Biofeedback as a stress management tool: a systematic review,” *Cognition, Technology & Work*, vol. 21, no. 2, pp. 161–190, 2019.
- [27] S. Melamed, A. Shirom, S. Toker, S. Berliner, and I. Shapira, “Burnout and risk of cardiovascular disease: Evidence, possible causal paths, and promising research directions,” *Psychological Bulletin*, vol. 132, no. 3, pp. 327–353, 2006.
- [28] T. Chandola, A. Britton, E. Brunner, H. Hemingway, M. Malik, M. Kumari, E. Badrick, M. Kivimaki, and M. Marmot, “Work stress and coronary heart disease: what are the mechanisms?,” *European heart journal*, vol. 29, no. 5, pp. 640–648, 2008.
- [29] E. Demerouti, P. M. Le Blanc, A. B. Bakker, W. B. Schaufeli, and J. Hox, “Present but sick: a three-wave study on job demands, presenteeism and burnout,” *Career Development International*, vol. 14, pp. 50–68, feb 2009.
- [30] M. Borritz, R. Rugulies, K. B. Christensen, E. Villadsen, and T. S. Kristensen, “Burnout as a predictor of self-reported sickness absence among human service workers: prospective findings from three year follow up of the PUMA study,” *Occupational and environmental medicine*, vol. 63, pp. 98–106, feb 2006.

- [31] M. Borritz, K. B. Christensen, U. Bültmann, R. Rugulies, T. Lund, I. Andersen, E. Villadsen, F. Dirichsen, and T. S. Kristensen, "Impact of burnout and psychosocial work characteristics on future long-term sickness absence. Prospective results of the Danish PUMA Study among human service workers.," *Journal of occupational and environmental medicine*, vol. 52, pp. 964–70, oct 2010.
- [32] L. Hallsten, M. Voss, S. Stark, and M. Josephson, "Job burnout and job wornout as risk factors for long-term sickness absence," *Work*, vol. 38, pp. 181–192, jan 2011.
- [33] C. A. M. Roelen, M. F. A. van Hoffen, J. W. Groothoff, J. de Bruin, W. B. Schaufeli, and W. van Rhenen, "Can the Maslach Burnout Inventory and Utrecht Work Engagement Scale be used to screen for risk of long-term sickness absence?," *International Archives of Occupational and Environmental Health*, vol. 88, pp. 467–475, may 2015.
- [34] S. Toppinen-Tanner, A. Ojajärvi, A. Väänäänen, R. Kalimo, and P. Jäppinen, "Burnout as a Predictor of Medically Certified Sick-Leave Absences and Their Diagnosed Causes," *Behavioral Medicine*, vol. 31, pp. 18–32, apr 2005.
- [35] G. Aronsson and K. Gustafsson, "Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research.," *Journal of occupational and environmental medicine*, vol. 47, pp. 958–66, sep 2005.
- [36] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.
- [37] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Computer methods and programs in biomedicine*, vol. 108, no. 3, pp. 1287–1301, 2012.
- [38] I. R. Murray, C. Baber, and A. South, "Towards a definition and working model of stress and its effects on speech," *Speech Communication*, vol. 20, no. 1-2, pp. 3–12, 1996.
- [39] E. J. Peacock and P. T. Wong, "The stress appraisal measure (sam): A multidimensional approach to cognitive appraisal," *Stress medicine*, vol. 6, no. 3, pp. 227–236, 1990.
- [40] I. Lefter, L. J. M. Rothkrantz, D. A. Van Leeuwen, and P. Wiggers, "Automatic stress detection in emergency(telephone) calls," *International Journal of Intelligent Defence Support Systems*, vol. 4, no. 2, pp. 148–168, 2011.
- [41] G. Zhou, J. H. L. Hansen, and J. F. Kaiser, "Nonlinear feature based classification of speech under stress," *IEEE Transactions on speech and audio processing*, vol. 9, no. 3, pp. 201–216, 2001.

- [42] M. Garnier and N. Henrich, “Speaking in noise: How does the Lombard effect improve acoustic contrasts between speech and ambient noise?,” *Computer Speech & Language*, vol. 28, no. 2, pp. 580–597, 2014.
- [43] J. R. Deller, J. G. Proakis, and J. H. L. Hansen, “Discrete-time processing,” in *Discrete-time processing of speech signals*, Institute of Electrical and Electronics Engineers, 2000.
- [44] P. R. Cook, “SPASM, a real-time vocal tract physical model controller; and singer, the companion software synthesis system,” *Computer Music Journal*, vol. 17, no. 1, pp. 30–44, 1993.
- [45] D. L. Frank, L. Khorshid, J. F. Kiffer, C. S. Moravec, and M. G. McKee, “Biofeedback in medicine: who, when, why and how?,” *Mental health in family medicine*, vol. 7, no. 2, p. 85, 2010.
- [46] M. S. Schwartz and F. Andrasik, *Biofeedback: A practitioner’s guide*. Guilford Publications, 2017.
- [47] C. Yucha and D. Montgomery, *Evidence-based practice in biofeedback and neurofeedback*. AAPB Wheat Ridge, CO, 2008.
- [48] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.
- [49] N. B. Lasa and A. I. V. Iraeta, *Diseños de investigación experimental en psicología: modelos y análisis de datos mediante el SPSS 10.0*. Pearson Educación, 2002.
- [50] J. Bustamante, “College enrollment statistics [2020]: Total + by demographic,” *Educationdata: college enrollment statistics*, 2019.
- [51] D. A. N. de Estadística, “Saber para decidir boletín nacional 2018,” *Boletan Nacional*, 2018.
- [52] P. Hill, *Audio and Speech Processing with MATLAB*. CRC Press, 2018.
- [53] J. H. Hansen and S. Patil, “Speech under stress: Analysis, modeling and recognition,” in *Speaker classification I*, pp. 108–137, Springer, 2007.
- [54] H. A. Murthy, F. Beaufays, L. P. Heck, and M. Weintraub, “Robust text-independent speaker identification over telephone channels,” *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 5, pp. 554–568, 1999.
- [55] M. Sahidullah and G. Saha, “Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition,” *Speech communication*, vol. 54, no. 4, pp. 543–565, 2012.
- [56] L. B. Jackson and L. B. Jackson, *Solutions manual, Digital filters and signal processing, second edition*. Kluwer, 1989.

- [57] H. W. Lilliefors, "On the kolmogorov-smirnov test for the exponential distribution with mean unknown," *Journal of the American Statistical Association*, vol. 64, no. 325, pp. 387–389, 1969.
- [58] D. Groppe, "mult_comp_perm_t2(data1,data2,n_perm,tail,alpha_level,mu,t_stat,reports,seed_state)," mar 2018.
- [59] J. S. Villarreal Sánchez, L. V. Ortíz Martínez, F. Vigoth, K. Viviana, and Others, "Prevalencia de ansiedad en población universitaria de la ciudad de Bucaramanga, según el cuestionario de Ansiedad Estado-Rasgo (STAI), 2019," *Universidad Cooperativa de Colombia*, 2019.
- [60] A. Vartanov, V. Ivanov, and I. Vartanova, "Facial expressions and subjective assessments of emotions," *Cognitive Systems Research*, vol. 59, pp. 319–328, 2020.
- [61] T.-M. Bynion and M. T. Feldner, "Self-assessment manikin," *Encyclopedia of personality and individual differences*, pp. 1–3, 2017.
- [62] B. Geethanjali, K. Adalarasu, A. Hemaprabha, S. P. Kumar, and R. Rajasekeran, "Emotion analysis using sam (self-assessment manikin) scale," *Biomedical Research*, 2017.
- [63] A. Darzi, H. Azami, and R. Khosrowabadi, "Brain functional connectivity changes in long-term mental stress," *Journal of Neurodevelopmental Cognition*, 2019.
- [64] S. Arora, T. Tierney, N. Sevdalis, R. Aggarwal, D. Nestel, M. Woloshynowych, A. Darzi, and R. Kneebone, "The imperial stress assessment tool (isat): a feasible, reliable and valid approach to measuring stress in the operating room," *World journal of surgery*, vol. 34, no. 8, pp. 1756–1763, 2010.
- [65] E. P. Clemons, "Monitoring anxiety levels and coping skills among military recruits," *Military medicine*, vol. 161, no. 1, pp. 18–21, 1996.
- [66] W. Liao, W. Zhang, Z. Zhu, and Q. Ji, "A real-time human stress monitoring system using dynamic bayesian network," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pp. 70–70, IEEE, 2005.
- [67] A. Cheema and M. Singh, "An application of phonocardiography signals for psychological stress detection using non-linear entropy based features in empirical mode decomposition domain," *Applied Soft Computing*, vol. 77, pp. 24–33, 2019.
- [68] A. Asif, M. Majid, and S. M. Anwar, "Human stress classification using eeg signals in response to music tracks," *Computers in biology and medicine*, vol. 107, pp. 182–196, 2019.

- [69] A. Anusha, J. Jose, S. Preejith, J. Jayaraj, and S. Mohanasankar, “Physiological signal based work stress detection using unobtrusive sensors,” *Biomedical Physics & Engineering Express*, vol. 4, no. 6, p. 065001, 2018.
- [70] G. Acerbi, E. Rovini, S. Betti, A. Tirri, J. F. Rónai, A. Sirianni, J. Agrimi, L. Eusebi, and F. Cavallo, “A wearable system for stress detection through physiological data analysis,” in *Italian Forum of Ambient Assisted Living*, pp. 31–50, Springer, 2016.
- [71] L. Ciabattoni, F. Ferracuti, S. Longhi, L. Pepa, L. Romeo, and F. Verdini, “Real-time mental stress detection based on smartwatch,” in *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 110–111, IEEE, 2017.
- [72] S. Clarke, L. G. Jaimes, and M. A. Labrador, “mstress: A mobile recommender system for just-in-time interventions for stress,” in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–5, IEEE, 2017.
- [73] K. W. Godin, T. Hasan, and J. H. L. Hansen, “Glottal waveform analysis of physical task stress speech,” in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.