



UNIVERSIDAD  
AUTÓNOMA  
DE QUERÉTARO



FACULTAD  
DE INGENIERÍA

Universidad Autónoma de Querétaro

Facultad de Ingeniería

Maestría en Ciencias en Ingeniería Matemática

**Análisis de ondículas para señales de EEG en el habla imaginada**

Tesis

Que como parte de los requisitos para obtener el Grado de

**Maestro en Ciencias en Ingeniería Matemática**

Presenta

Carlos Agosto Rodríguez de la Gala Ureña

Dirigido por:

Dr. Samuel Estala Arias

Dr. Samuel Estala Arias

Presidente

Dr. Victor Antonio Aguilar Arteaga

Secretario

M.C. Iván González García

Vocal

M.C. Luisa Ramírez Granados

Suplente

Dra. María Florencia Assaneo

Suplente

Centro Universitario, Querétaro, Qro.

Fecha de aprobación por el Consejo Universitario (septiembre 2023)

México



Dirección General de Bibliotecas y Servicios Digitales  
de Información



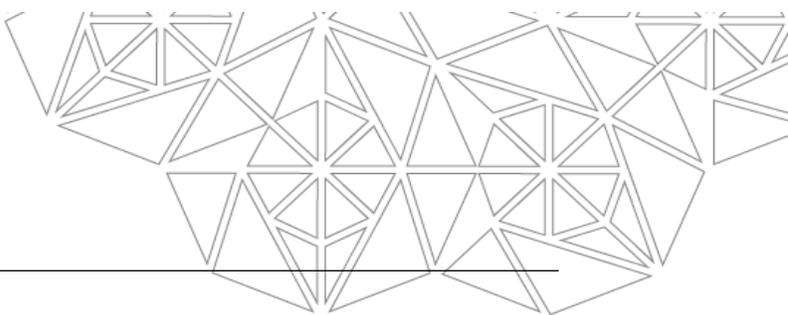
Análisis de ondículas para señales de EEG en el habla  
imaginada

**por**

Carlos Agosto Rodríguez de la Gala Ureña

se distribuye bajo una [Licencia Creative Commons  
Atribución-NoComercial-SinDerivadas 4.0  
Internacional](#).

**Clave RI:** IGMAN-309255



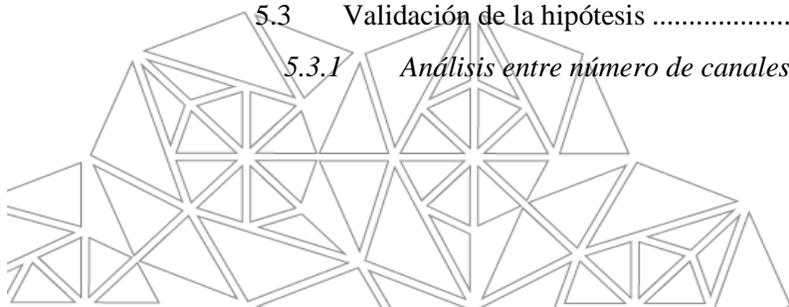
---

# Tabla de Contenidos

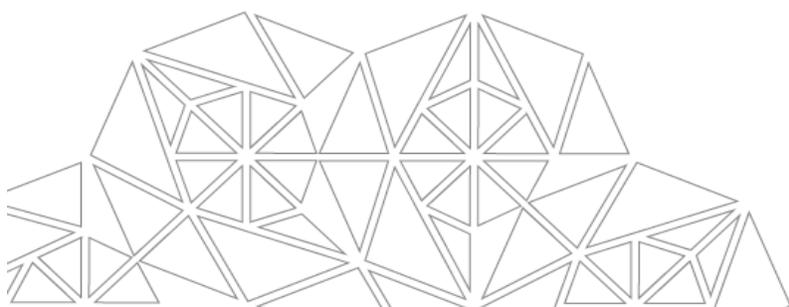
---

Lista de Figuras.....	IV
Lista de Tablas.....	VI
Resumen.....	1
Abstract.....	3
Capítulo 1 .....	5
1 INTRODUCCIÓN .....	5
1.1 Antecedentes .....	11
1.2 Justificación.....	16
1.3 Definición del problema.....	17
1.4 Hipótesis.....	18
1.5 Objetivos .....	18
1.5.1 <i>Objetivo general</i> .....	18
1.5.2 <i>Objetivos específicos</i> .....	18
1.6 Alcances y limitaciones.....	19
1.7 Estructura del documento.....	19
Capítulo 2.....	21
2 FUNDAMENTO TEÓRICO.....	21
2.1 Anatomía del cerebro .....	21
2.1.1 <i>Modelo Geschwind-Wernicke</i> .....	22
2.1.2 <i>Corteza visual</i> .....	23
2.2 Electroencefalograma (EEG) .....	24
2.2.1 <i>Ritmos cerebrales</i> .....	25
2.2.2 <i>Áreas de Brodmann</i> .....	26
2.2.3 <i>Montajes de electrodos</i> .....	28
2.2.4 <i>Artefactos</i> .....	30
2.3 Interfaces Cerebro-Computadora (BCI's).....	31
2.3.1 <i>Componentes de un sistema BCI</i> .....	32

2.3.2	<i>Señales electrofisiológicas</i> .....	33
2.4	Transformadas Wavelet.....	35
2.4.1	<i>Transformada wavelet discreta (DWT)</i> .....	37
2.4.2	<i>Caracterización con energías wavelets</i> .....	39
2.5	Clasificación.....	41
2.5.1	<i>Máquina de Vectores de Soporte (SVM)</i> .....	42
2.5.2	<i>Random Forest (RF)</i> .....	43
2.5.3	<i>Medición de la evaluación de los clasificadores</i> .....	43
2.5.4	<i>Validación cruzada con K-pliegues</i> .....	45
Capítulo 3	.....	46
3	ESTADO DEL ARTE.....	46
3.1	Habla imaginada.....	46
3.1.1	<i>Enfoque basado en palabras</i> .....	47
3.2	Potenciales evocados visuales .....	50
Capítulo 4	.....	54
4	METODOLOGÍA PROPUESTA .....	54
4.1	Adquisición de las señales EEG.....	55
4.2	Preprocesamiento .....	56
4.3	Caracterización de las señales .....	57
4.3.1	<i>Normalización con la energía relativa wavelet (RWE)</i> .....	58
4.3.2	<i>Selección de canales del EEG</i> .....	59
4.4	Selección de características .....	60
4.4.1	<i>Eliminación de valores RWE</i> .....	60
4.4.2	<i>Eliminación de outliers</i> .....	61
4.5	Clasificación de las neuroseñales.....	61
Capítulo 5	.....	63
5	EXPERIMENTACIÓN Y RESULTADOS .....	63
5.1	Diseño experimental.....	63
5.2	Experimentación con 8 individuos .....	65
5.2.1	<i>32 canales de EEG</i> .....	66
5.2.2	<i>16 canales de EEG</i> .....	70
5.3	Validación de la hipótesis .....	74
5.3.1	<i>Análisis entre número de canales</i> .....	78



5.3.2	<i>Análisis entre tipo de ventana</i> .....	81
5.3.3	<i>Análisis entre ondículas madre</i> .....	83
Capítulo 6	.....	87
6	CONCLUSIONES .....	87
6.1	Discusión.....	88
6.2	Trabajo futuro.....	88
Anexos	.....	90
<b>Anexo 1</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, IS, 32 ch .....	90
<b>Anexo 2</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, +VEP, 32 ch.....	90
<b>Anexo 3</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con db7, IS, 32 ch .....	91
<b>Anexo 4</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con db7, +VEP, 32 ch .....	91
<b>Anexo 5</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, IS, 16 ch .....	92
<b>Anexo 6</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, +VEP, 16 ch.....	92
<b>Anexo 7</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con db7, IS, 16 ch .....	93
<b>Anexo 8</b>	Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con db7, +VEP, 16 ch .....	93
Referencias	.....	94

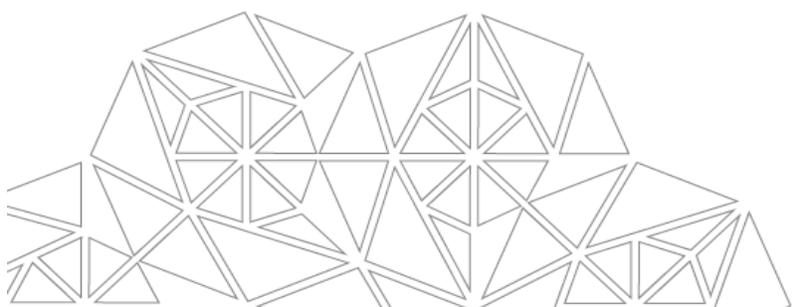


## Lista de Figuras

<b>Figura 1.1</b> Modelo de Norman Geschwind.....	6
<b>Figura 1.2</b> Simulación interna de una conversación .....	7
<b>Figura 2.1</b> Secciones del cerebro.....	22
<b>Figura 2.2</b> Modelo Geschwind-Wernicke .....	23
<b>Figura 2.3</b> Corteza visual. Áreas 17, 18 y 19 de Brodmann.....	23
<b>Figura 2.4</b> Patrones en la actividad de los ritmos del EEG .....	26
<b>Figura 2.5</b> Áreas de Brodmann.....	27
<b>Figura 2.6</b> Tipos de montajes para medición EEG.....	29
<b>Figura 2.7</b> Posición de electrodos de acuerdo al Sistema Internacional 10/20 .....	30
<b>Figura 2.8</b> Diagrama de bloques de un sistema BCI.....	33
<b>Figura 2.9</b> P300.....	34
<b>Figura 2.10</b> Potenciales evocados visuales .....	35
<b>Figura 2.11</b> Selección del hiperplano óptimo en SVM .....	42
<b>Figura 4.1</b> Metodología aplicada en la investigación .....	54
<b>Figura 4.2</b> Sistema internacional 10/20 para 32 canales .....	55
<b>Figura 4.3</b> Mapas topográficos de componentes ICA obtenidos con MNE .....	57
<b>Figura 4.4</b> Estructura de la matriz $A_{20,160 \times 6}$ .....	59
<b>Figura 4.5</b> 16 canales asociados a las áreas cerebrales de Broca, Wernicke y corteza visual.....	60
<b>Figura 5.1</b> Diagrama del diseño experimental del trial .....	63
<b>Figura 5.2</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre bior 3.7 para IS.....	66
<b>Figura 5.3</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre bior 3.7 para +VEP. 67	67
<b>Figura 5.4</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre db7 para IS.....	68
<b>Figura 5.5</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre db7 para +VEP.....	68
<b>Figura 5.6</b> Exactitud promedio obtenida de la clasificación en 8 sujetos con 32 canales para cada tipo de ventana (IS y +VEP) y su respectiva ondícula madre .....	69
<b>Figura 5.7</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre bior 3.7 para IS.....	70
<b>Figura 5.8</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre bior 3.7 para +VEP. 71	71

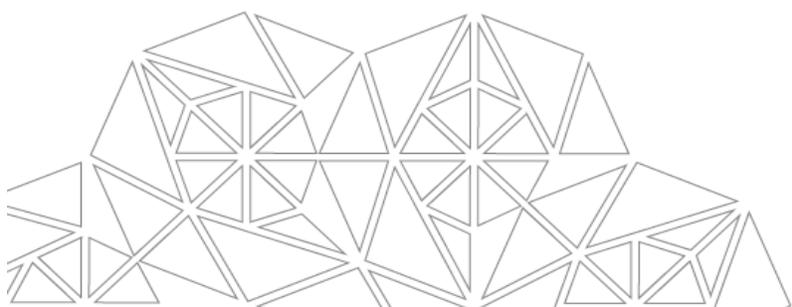


<b>Figura 5.9</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre db7 para IS.....	72
<b>Figura 5.10</b> Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre db7 para +VEP .....	72
<b>Figura 5.11</b> Exactitud promedio obtenida de la clasificación en 8 sujetos con 16 canales para cada tipo de ventana (IS y +VEP) y su respectiva ondícula madre .....	73
<b>Figura 5.12</b> Prueba Shapiro-Wilk para la población (bior 3.7, IS, 16 ch).....	75
<b>Figura 5.13</b> Prueba Shapiro-Wilk para la población (bior 3.7, +VEP, 16 ch) .....	75
<b>Figura 5.14</b> Prueba Shapiro-Wilk para la población (db7, IS, 16 ch).....	76
<b>Figura 5.15</b> Prueba Shapiro-Wilk para la población (db7, +VEP, 16 ch).....	76
<b>Figura 5.16</b> Prueba Shapiro-Wilk para la población (bior 3.7, IS, 32 ch).....	76
<b>Figura 5.17</b> Prueba Shapiro-Wilk para la población (bior 3.7, +VEP, 32 ch) .....	77
<b>Figura 5.18</b> Prueba Shapiro-Wilk para la población (db7, IS, 32 ch).....	77
<b>Figura 5.19</b> Prueba Shapiro-Wilk para la población (db7, +VEP, 32 ch).....	77
<b>Figura 5.20</b> Prueba t student para (bior 3.7, IS, 16 ch) y (bior 3.7, IS, 32 ch).....	79
<b>Figura 5.21</b> Prueba t student para (bior 3.7, +VEP, 16 ch) y (bior 3.7, +VEP, 32 ch) .....	79
<b>Figura 5.22</b> Prueba t student para (db7, IS, 16 ch) y (db7, IS, 32 ch).....	80
<b>Figura 5.23</b> Prueba Mann-Whitney para (db7, IS, 16 ch) y (db7, IS, 32 ch).....	80
<b>Figura 5.24</b> Prueba t student para (bior 3.7, IS, 16 ch) y (bior 3.7, +VEP, 16 ch).....	81
<b>Figura 5.25</b> Prueba t student para (bior 3.7, IS, 32 ch) y (bior 3.7, +VEP, 32 ch).....	82
<b>Figura 5.26</b> Prueba t student para (db7, IS, 16 ch) y (db7, +VEP, 16 ch) .....	82
<b>Figura 5.27</b> Prueba Mann-Whitney para (db7, IS, 32 ch) y (db7, +VEP, 32 ch).....	83
<b>Figura 5.28</b> Prueba t student para (bior 3.7, IS, 16 ch) y (db7, IS, 16 ch).....	84
<b>Figura 5.29</b> Prueba t student para (bior 3.7, +VEP, 16 ch) y (db7, +VEP, 16 ch).....	84
<b>Figura 5.30</b> Prueba t student para (bior 3.7, IS, 32 ch) y (db7, IS, 32 ch).....	85
<b>Figura 5.31</b> Prueba Mann-Whitney para (bior 3.7, +VEP, 32 ch) y (db7, +VEP, 32 ch).....	85



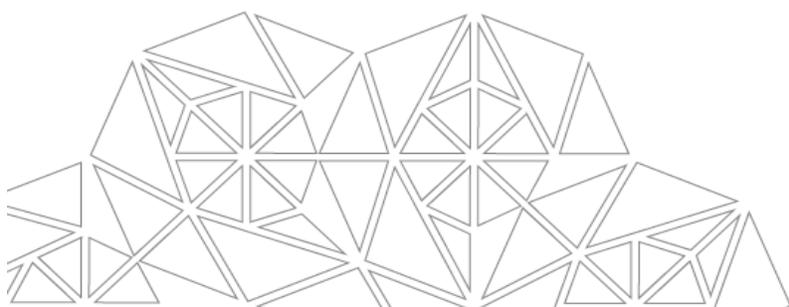
## Lista de Tablas

<b>Tabla 1.1</b> Resumen de las tecnologías disponibles de sensores para el monitoreo de la actividad cerebral ....	8
<b>Tabla 2.1</b> Artefactos fisiológicos y no fisiológicos.....	31
<b>Tabla 2.2</b> Matriz de confusión para la clase A .....	44
<b>Tabla 3.1</b> Estado del arte del habla imaginada con el enfoque basado en palabras.....	50
<b>Tabla 4.1</b> Niveles de descomposición de la DWT .....	58
<b>Tabla 4.2</b> Clases para la clasificación.....	62
<b>Tabla 5.1</b> Exactitud promedio obtenida de la clasificación en 8 sujetos con 32 canales .....	69
<b>Tabla 5.2</b> Exactitud promedio obtenida de la clasificación en 8 sujetos con 16 canales .....	73
<b>Tabla 5.3</b> Pares de poblaciones de datos a comparar, considerando las variables ondícula madre ( $w_1, w_2$ ), tipo de ventana (IS, +VEP) y número de canales (16 ch, 32 ch).....	74



## Dedicatoria

*A ti, querido lector, que por alguna razón encontraste este documento (o quizá él te encontró) donde el autor pretende ayudarte con tu investigación mediante su granito de arena.*



## Agradecimientos

*Al pueblo de México, que a través del Consejo Nacional de Ciencia y Tecnología (CONACyT) ayudó a la investigación del proyecto de esta tesis con la beca del autor.*

*Hay tres conjuntos de personas quienes pretendo considerar en mi gratitud,  $A$ ,  $B$ ,  $C$  tal que*

$$A \cap B = \emptyset,$$

$$C: \{x | x \in A \cup B \cup C\}$$

*$A$ : {Mi familia, especialmente a mis amados padres Lourdes María y Carlos Humberto}*

*Gracias por su esfuerzo, motivación y dedicación en mi formación, logrando que mis recuerdos con ustedes hayan sido felices.*

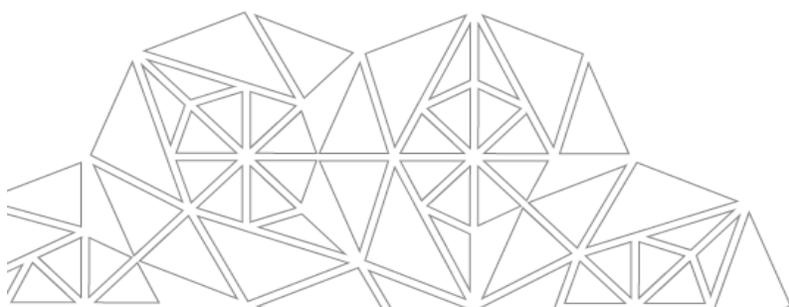
*$B$ : {Dr. Samuel, Dr. Victor, Dra. Florencia, Dr. Roberto, Ricardo}*

*Mi enorme gratitud por sus consejos, cooperación y ayuda, que permitieron perfeccionar este proyecto hasta llegar a la conclusión del mismo.*

*$C$ : {Aquellas personas que, de alguna manera, directa e indirecta, contribuyeron en esta etapa de mi vida}*

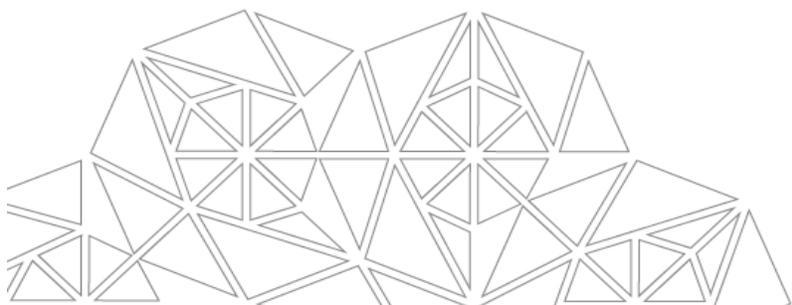
*Hay otro conjunto, que no especifiqué previamente al no considerar sus elementos como humanos, pero que sí quiero mencionar como  $E$ .*

*Gracias a la Entropía de este universo en el que habito ( $E$ ), volviendo posible la existencia de la información causal para llegar al momento actual y permitirme lograr un propósito más durante el lapso de mi vida.*



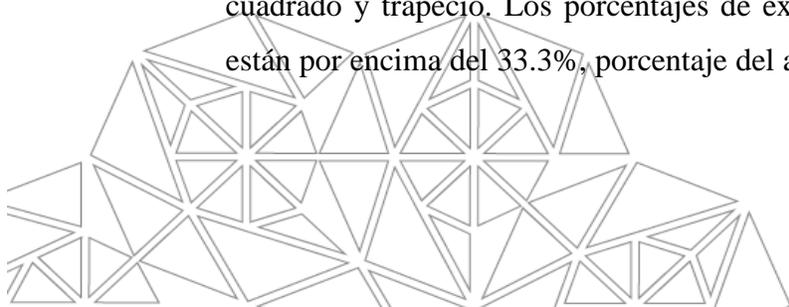
*“Es mi orgullo haber nacido  
En la novia de los mares, vieja tierra colonial  
Se adormece entre colinas  
Y entre verdes naranjales, y se arrulla con el mar.  
Tus bellos atardeceres  
Y tus noches enlunadas no se pueden comparar  
Pero lo que más me gusta  
Es mirar de mar adentro  
Es mirar de mar adentro las torres de Catedral...”*

Extracción de la canción “Las Torres de Catedral” del compositor campechano José Narvárez Márquez “Pepe Narvárez”.



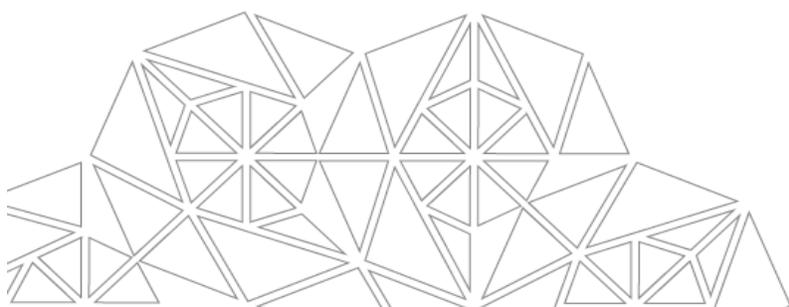
## Resumen

En esta investigación, se tiene como propósito el desarrollo de un método de preprocesamiento, procesamiento y clasificación de neuroseñales EEG para el reconocimiento de palabras de un reducido léxico, usando como señal electrofisiológica el habla imaginada. De manera específica, se pretende conseguir una clasificación (de neuroseñales) que, para propósitos posteriores, podría permitir el control mental de algún dispositivo BCI que pudiese requerir la identificación de ciertas figuras. Si bien es cierto que, el habla imaginada no requiere estímulos externos visuales como P300, VEP o SSVEP, ¿ayudaría agregar la información de las neuroseñales obtenidas de estos estímulos, generados por imágenes de figuras geométricas, a la información de las neuroseñales del habla imaginada de ciertas palabras para mejorar la exactitud de su clasificación? Con dicho fin, el vocabulario está compuesto de las palabras del lenguaje español: “círculo”, “cuadrado” y “trapecio”. En el comienzo del proceso, las señales EEG se obtuvieron mediante un protocolo que se basa en marcadores mediante “triggers”, para identificar a priori en cuál sección de la señal adquirida el individuo imagina la pronunciación (habla imaginada) de la palabra que se indica, así como la sección donde se presentó el estímulo visual de la figura. Estos marcadores generan los límites inferiores de las ventanas (épocas). Para lograr el propósito, las señales de los canales más cercanos a las áreas cerebrales de Broca, Wernicke y corteza visual son de interés. Dichas señales se filtraron empleando un filtro IIR (Infinite Impulse Response) pasa-bandas en el rango de 0 a 32 Hz. Adicionalmente, se caracterizaron las neuroseñales mediante la transformada wavelet discreta (DWT) con distintas wavelets madre (biorthogonal 3.7 y daubechies 7) y luego se obtuvieron valores de energía relativa wavelet, para después hacer una selección de características. Las características seleccionadas se emplearon para el entrenamiento y evaluación de dos clasificadores usando validación cruzada: Máquina de Vectores de Soporte (SVM) y Random Forest (RF). Al final, se evaluaron los clasificadores con la medida de exactitud (accuracy) y se compararon los resultados obtenidos, para determinar cuál tuvo la mejor. El clasificador RF tuvo mejores resultados en la exactitud que SVM para el vocabulario conformado de las palabras círculo, cuadrado y trapecio. Los porcentajes de exactitud conseguidos en promedio por ventana, están por encima del 33.3%, porcentaje del azar para 3 clases (véase las Figuras 5.6 y 5.11),



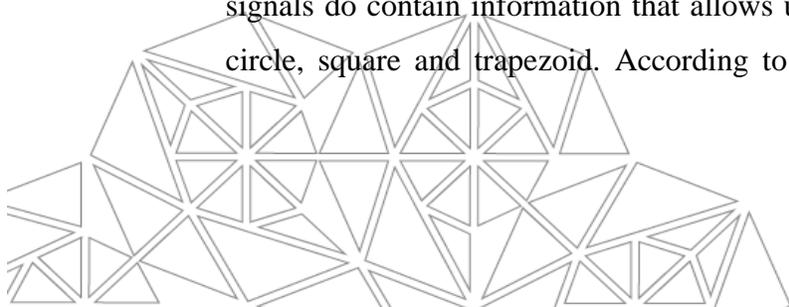
lo que permite inferir que, las señales EEG sí contienen información que permita distinguir el habla imaginada de las palabras círculo, cuadrado y trapecio. De acuerdo a los tres análisis, al comparar las poblaciones de datos que se tuvieron mediante inferencia estadística, considerando las variables *ondícula madre*, *tipo de ventana* y *número de canales*, no hay diferencia significativa en la exactitud (accuracy) de la clasificación con la caracterización de las ondículas madre biorthogonal 3.7 (bior 3.7) y daubechies 7 (db7), entre tipos de ventana IS y +VEP, así como la información obtenida con distinto número de canales EEG (16 y 32).

***Palabras clave:*** *Habla imaginada, VEP, ondículas, neuroseñales, EEG.*



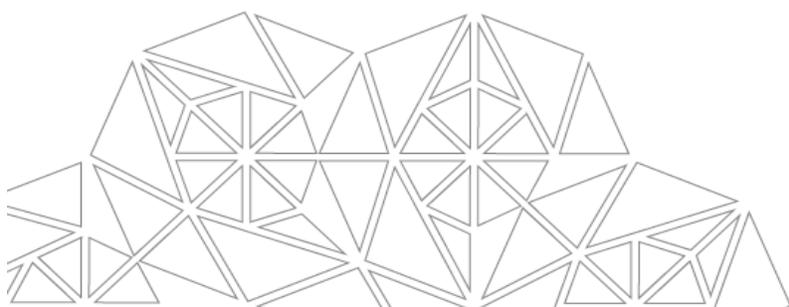
## Abstract

The purpose of this research is to develop a method of preprocessing, processing and classification of EEG neurosignals for the recognition of words from a reduced vocabulary, using imagined speech as an electrophysiological signal. Specifically, the aim is to achieve a classification (of neurosignals) that, for subsequent purposes, could allow the mental control of a BCI device that could require the identification of certain figures. While it is true that imagined speech does not require external visual stimuli such as P300, VEP or SSVEP, *would it help to add the neurosignal information obtained from these stimuli, generated by images of geometric figures, to the neurosignal information of imagined speech of certain words to improve the accuracy of its classification?* For this purpose, the vocabulary is made up of the words of the spanish language: “círculo”, “cuadrado” and “trapecio” (circle, square, trapezoid; respectively). At the beginning of the process, the EEG signals were obtained using a protocol that is based on markers using “triggers”, to identify a priori in which section of the acquired signal the individual imagines the pronunciation (imagined speech) of the indicated word, as well as the section where the visual stimulus of the figure was presented. These markers generate the lower limits of the windows (epochs). To achieve the purpose, signals from the channels closest to the brain areas of Broca, Wernicke and visual cortex are of interest. These signals were filtered using a band-pass IIR (Infinite Impulse Response) filter in the range of 0 to 32 Hz. Additionally, the neurosignals were characterized using the discrete wavelet transform (DWT) with different mother wavelets (biorthogonal 3.7 and daubechies 7) and then relative wavelet energy values were obtained, to then perform feature selection. The selected features were used for the training and evaluation of two classifiers using cross-validation: Support Vector Machine (SVM) and Random Forest (RF). In the end, the classifiers were evaluated with the accuracy measure and the results obtained were compared to determine which one was the best. The RF classifier had better results in accuracy than SVM for the vocabulary made up of the words circle, square and trapezoid. The accuracy percentages achieved on average per window are above 33.3%, a percentage of chance for 3 classes (see Figures 5.6 and 5.11), which allows us to infer that the EEG signals do contain information that allows us to distinguish imagined speech of the words circle, square and trapezoid. According to the three analyses, when comparing the data



populations that were obtained through statistical inference, considering the variables *mother wavelet*, *type of window* and *number of channels*, there is no significant difference in the accuracy of the classification with the characterization of the biorthogonal 3.7 (bior 3.7) and daubechies 7 (db7) mother wavelets, between IS and +VEP window types, as well as the information obtained with different numbers of EEG channels (16 and 32).

***Keywords:*** *Imagined speech, VEP, wavelets, neurosignals, EEG.*

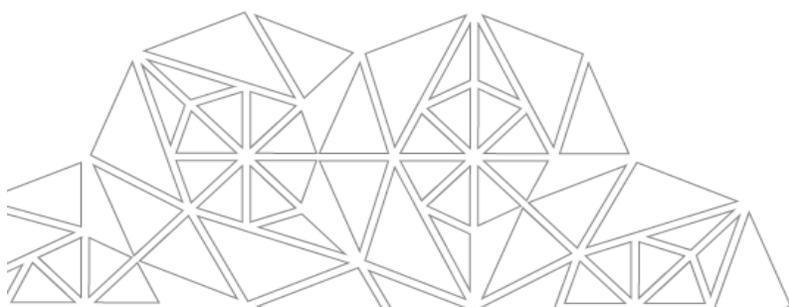


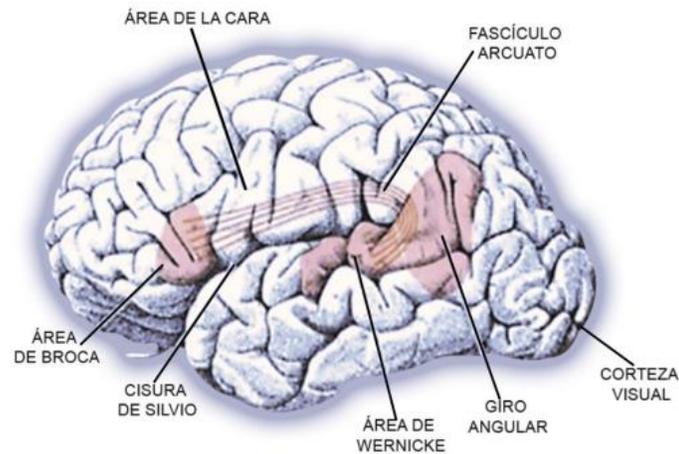
# Capítulo 1

## INTRODUCCIÓN

Si nos preguntamos, ¿cuál es el propósito de comunicarnos? Una posible respuesta es transmitir información, pero para lograr ese proceso, se requiere identificar los elementos que lo permiten. Al menos, se necesita un emisor y un receptor, es decir, el que envía y recibe el mensaje con la información. Dicho mensaje está codificado y para lograr que la información sea correctamente transmitida, el que lo recibe, debe saber decodificarla. También se necesita un medio por el cual la información viajará, el canal donde idealmente el mensaje no se alterará hasta su destino final, pero se tiene el riesgo que esto pueda no ocurrir. Aquello que se encuentre en el mensaje y no forme parte de la información original se le puede llamar ruido.

Por otro lado, el ser humano por su biología, al crecer es potencialmente apto para transmitir información de diversas maneras. La más eficiente por la facilidad y rapidez para comunicarse con otras personas es el lenguaje oral. Se ha descubierto que, para lograr la comunicación oral, hay regiones específicas en el cerebro encargadas del procesamiento del lenguaje hablado. El sistema léxico-semántico y el sistema gramatical, relacionados con el *área de Wernicke* (segmento posterior de la primera circunvolución del lóbulo temporal) y el *área de Broca* (pars opercularis y pars triangularis) respectivamente, ubicados normalmente en la región perisilviana (alrededor de la cisura de Silvio) del hemisferio izquierdo cerebral (Ardila, et al., 2016), (Aguirre-Cruz, et al., 2022) participan esencialmente en el reconocimiento, producción y la comprensión del lenguaje. Ambas áreas están conectadas mediante un haz de fibras nerviosas el cual recibe el nombre de *fascículo arqueado* que pertenece al fascículo longitudinal superior (Aguirre-Cruz, et al., 2022) (véase en **Figura 1.1**).



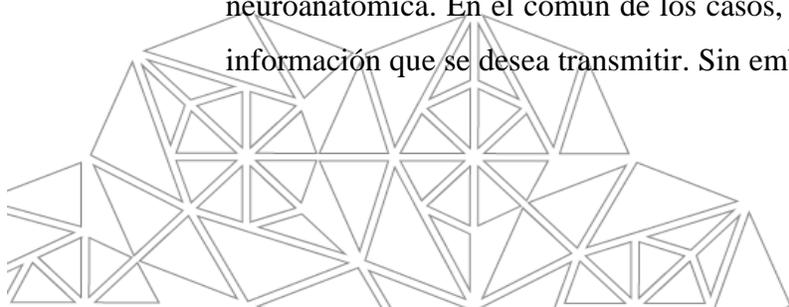


**Figura 1.1** Modelo de Norman Geschwind. Sugiere la conexión de las áreas motoras y sensitivas del lenguaje mediante el fascículo arqueado (o arcuato) (Adaptación de (Aguirre-Cruz, et al., 2022)).

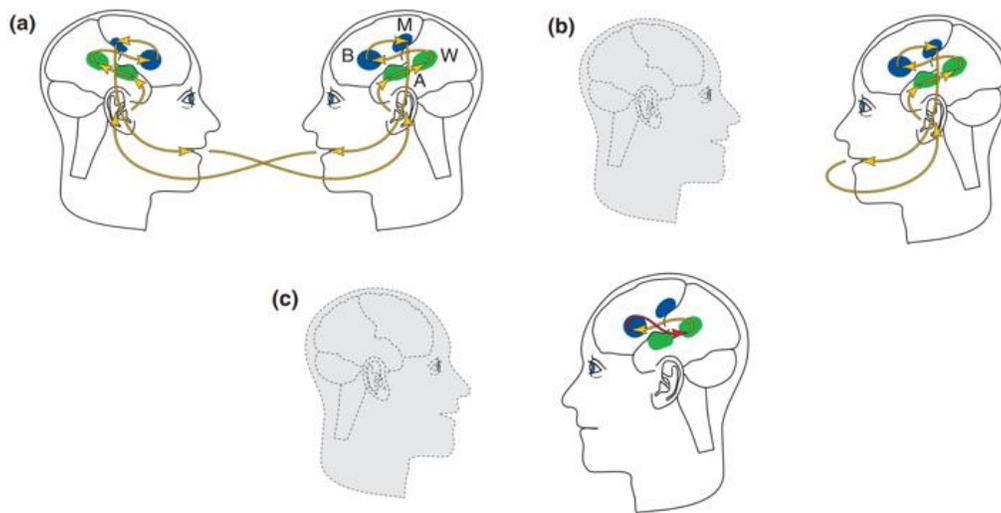
Entre los intentos por mostrar un modelo que explique cómo funciona el lenguaje en el cerebro, tanto en una conversación interna como con alguien más, se encuentra el mostrado por Germund Hesslow en (Hesslow, 2002), donde se muestra el circuito por el que pasa la señal verbal, identificando cuatro regiones cerebrales: la corteza auditiva primaria (A), el área de Wernicke (W), el área de Broca (B) y la corteza motora primaria (M) (véase en **Figura 1.2**). Cuando el habla es interna o imaginada, se sugiere que la señal es generada en la corteza auditiva o área de Wernicke para después interaccionar con el área de Broca. En este caso, no se requiere de la corteza motora primaria. Tanto en un soliloquio como en una conversación con alguien más, esencialmente se utilizan los mismos mecanismos, la señal verbal entra a la corteza auditiva primaria y después al área de Wernicke, provocando que se genere una respuesta tanto en el área de Broca como en la corteza motora primaria.

Sin embargo, es importante mencionar que, la dominancia exclusiva del hemisferio izquierdo ha sido cuestionada. Por ejemplo, en (Yazbek , et al., 2021) se hace referencia de la variación del dominio hemisférico del lenguaje tanto en sujetos zurdos como diestros.

Hasta ahora, se tiene la noción del procesamiento del lenguaje desde la perspectiva neuroanatómica. En el común de los casos, se tiene una señal acústica emitida que tiene la información que se desea transmitir. Sin embargo, se puede identificar estas cuatro posibles



situaciones: la señal acústica es entendible y puede o no haber ruido en el ambiente o la señal acústica no es entendible y puede o no haber ruido en el ambiente. En los primeros dos casos, se entiende que la información puede llegar sin problemas de decodificación al receptor. En los últimos dos casos, tenemos el problema de una señal no entendible. Ejemplos de estas situaciones son, el caso de recibir una llamada en un lugar exterior como en una calle o en una feria, o cuando el emisor tiene problemas de transmitir la señal acústica por algún impedimento, ya sea por alguna discapacidad u otro factor. Para estas circunstancias, es conveniente sistemas que posibiliten comunicación hablada sin la presencia de una señal acústica inteligible.



**Figura 1.2** Simulación interna de una conversación (Adaptación de (Hesslow, 2002)). a) Conversación entre dos personas. La señal verbal entra a la corteza auditiva primaria (A) y después al área de Wernicke (W), provocando que se genere una respuesta tanto en el área de Broca (B) como en la corteza motora primaria (M). b) Tanto en un soliloquio como en una conversación con alguien más, esencialmente se utilizan los mismos mecanismos. c) Cuando el habla es interna o imaginada, se sugiere que la señal es generada (flecha roja) en la corteza auditiva o área de Wernicke para después interactuar con el área de Broca.

De acuerdo con (Denby, et al., 2010), una *interface de habla silenciosa* (SSI, por sus siglas en inglés *Silent Speech Interface*) es un sistema que permite la comunicación hablada cuando no hay una señal acústica audible. Estos sistemas generan una representación digital del habla sintetizable directamente e interpretada como datos o enrutada a una red de comunicaciones, mediante la adquisición de datos con sensores en los elementos de la

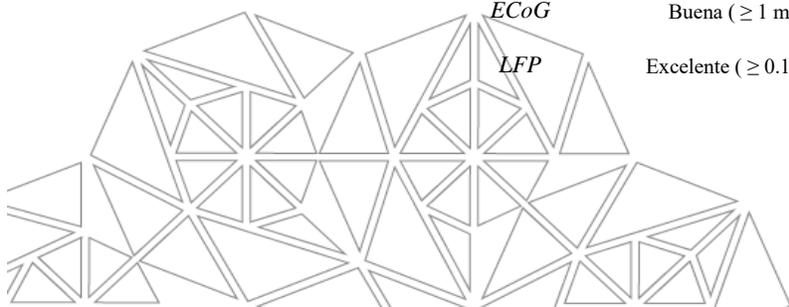
producción del habla como articuladores (lengua, labios, mandíbula, laringe, entre otros), vías neuronales o el cerebro en sí mismo. En esta investigación, se tiene especial interés en aquellos que involucran al monitoreo de las señales cerebrales, pues hay la posibilidad de desenvolverse en ambientes donde la señal acústica no es entendible, ya sea por ruido o por otros factores como el caso donde el emisor tiene dificultades para transmitir una señal audible.

Para la observación de la actividad cerebral, se tiene disponible varias opciones o alternativas. En (Gonzalez-Lopez, et al., 2020) se presenta un breve resumen (véase en **Tabla 1.1**) donde se exponen las tecnologías disponibles de sensores para el monitoreo de la actividad cerebral, como las imágenes de resonancia magnética funcional (fMRI), la espectroscopía del infrarrojo cercano (fNIRS), magnetoencefalografía (MEG), electroencefalografía (EEG), electrocorticografía (ECoG) y potencial de campo local (LFP).

Si tenemos la opción de decidir cuál tecnología convendría en una investigación, dependerá en gran medida del propósito con que se realiza, pues se debe considerar que hay ventajas y desventajas en varios aspectos. Por ejemplo, ECoG y EEG tienen una buena resolución temporal, sin embargo, ECoG es una tecnología invasiva, es decir, requiere de cirugía, además de acceso algo difícil para la investigación por los riesgos que implica y su costo. En cambio, EEG es no invasiva, relativamente sencilla y económica, siendo de esta manera una alternativa para obtener información de la actividad cerebral.

**Tabla 1.1** Resumen de las tecnologías disponibles de sensores para el monitoreo de la actividad cerebral (Adaptación de (Gonzalez-Lopez, et al., 2020)).

Enfoque	Técnica de detección	Resolución temporal	Resolución espacial	Intrusión	Portabilidad
Hemodinámico	<i>fMRI</i>	Baja ( $\geq 1$ s)	Buena (1.5 – 2 mm)	No intrusivo	No portable
	<i>fNIRS</i>	Baja ( $\geq 1$ s)	Media (1 – 2 cm)	No intrusivo	Moderada
Electrodinámico	<i>MEG</i>	Buena ( $\geq 1$ ms)	Buena (1 – 2 mm)	No intrusivo	No portable
	<i>EEG</i>	Buena ( $\geq 1$ ms)	Baja ( $\approx 10$ cm <sup>2</sup> )	No intrusivo	Moderada
	<i>ECoG</i>	Buena ( $\geq 1$ ms)	Excelente (0.5 – 1 mm)	Intrusivo	Alta
	<i>LFP</i>	Excelente ( $\geq 0.1$ ms)	Excelente (5 – 100 $\mu$ m)	Muy intrusivo	Alta



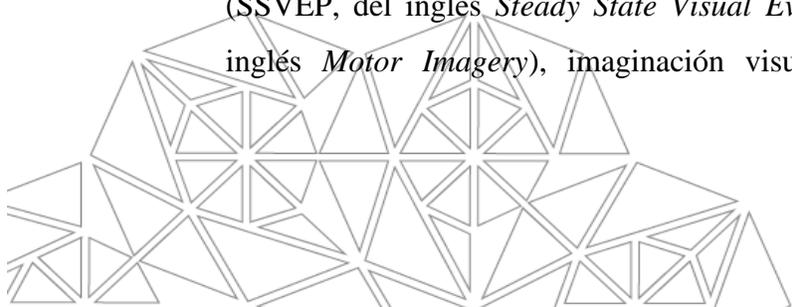
Las señales EEG son mediciones de la actividad eléctrica cerebral, por medio de electrodos no invasivos, que se colocan sobre la superficie del cuero cabelludo. El número de electrodos o canales del EEG es variable, puede depender de las características del dispositivo o de su uso. En (Malete, et al., 2019) se utilizó el dispositivo Emotivo Eloc+ de 14 canales para una interface cerebro-computadora (BCI), para el control de un juego en 3D, en (Montoya-Martínez, et al., 2021) se empleó hasta 64 electrodos, donde se propuso una estrategia de selección de canales que permite una evaluación cuantitativa rápida, para determinar el efecto del número y ubicación de los electrodos de EEG en la medición del seguimiento neuronal del habla, mientras que en (Fiedler, et al., 2022) se probó la electroencefalografía de alta densidad (HD-EEG) de 256 electrodos, comparándola con una convencional a base de gel, obteniendo una reducción del tiempo de preparación del 69%.

Cuando se pretende utilizar las señales obtenidas del EEG para el control de otros dispositivos o su interacción con ellos, se ingresa al área de investigación de las *interfaces cerebro-computadora* (BCI's, por las siglas en inglés *Brain-Computer Interfaces*).

De acuerdo con (Martínez Peón, 2016), la definición tomada de *Brain-computer interfaces for communication and control*, Wolpaw et al;2002 para BCI es:

*“Una **interface cerebro-computadora** basada en EEG es un sistema que permite un canal alternativo de comunicación y control al cerebro con el exterior, sin utilizar las salidas normales de los nervios periféricos y los músculos.”*

Si nos preguntamos el cómo un usuario de la BCI puede producir las señales de mensajes o comandos para la comunicación o control al exterior, vienen a relucir los paradigmas utilizados en los sistemas BCI, también conocidos como *señales electrofisiológicas*, entre las que se pueden destacar los potenciales P300, los potenciales corticales lentos (SCPs, del inglés *Slow Cortical Potentials*), potenciales evocados visuales (VEPs, del inglés *Visual Evoked Potentials*), potencial evocado visual de estado estacionario (SSVEP, del inglés *Steady State Visual Evoked Potential*), imaginación motriz (MI, del inglés *Motor Imagery*), imaginación visual (VI, del inglés *Visual Imagery*), ritmos



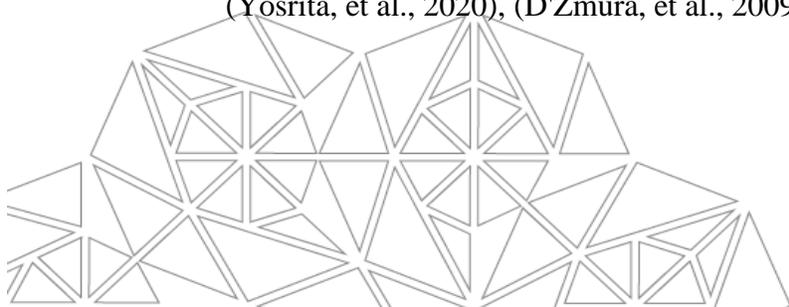
sensoriales motrices mu y beta (SMRs, del inglés *Sensorimotor Rhythms*), entre otros (Wolpaw, et al., 2002), (Jeunet, et al., 2019), (Martínez Peón, 2016), (González Castañeda, 2015), (Bermúdez Cicchino, 2013).

¿Tendrán algún punto de encuentro los sistemas SSI con los BCI? Anteriormente, se mencionó que los sistemas SSI también generan una representación digital del habla sintetizable directamente y la adquisición de datos es mediante sensores en los elementos de la producción del habla, los cuales incluyen al cerebro y, por ende, el monitoreo de las señales de actividad cerebral. Este principio también lo realiza un sistema BCI, con el propósito de crear un canal alternativo de comunicación. Aquí surge una intersección en el campo de investigación de ambos sistemas, donde se encuentra la señal electrofisiológica *habla imaginada*.

El habla imaginada, también conocido como habla no pronunciada, habla interna o habla silente (*Imagined, Unspoken, Inner or Silent Speech*) hace referencia a pronunciar internamente o de manera imaginada palabras, pero sin la emisión de sonidos o el uso de articuladores (Moctezuma Pascual, 2017), (González Castañeda, 2015).

La ventaja que tiene el habla imaginada respecto a otras señales electrofisiológicas, es su producción consciente por el usuario, lo que genera como consecuencia cortos periodos de entrenamiento (Moctezuma, et al., 2019). Por otro lado, paradigmas como SCP, P300, SMRs, SSVEP operan a velocidades lentas comparadas con el habla conversacional o casi conversacional (tasas de producción de una palabra o menos), debido a razones como requerir “mapeo” o traducción en el dominio del habla (Brumberg, et al., 2010).

Debido a lo anterior, en la presente investigación se tiene interés en el habla imaginada. Algunos trabajos previos en la exploración del uso de esta señal electrofisiológica se encuentran (Moctezuma Pascual, 2017), (Moctezuma, et al., 2017), (González Castañeda, 2015), (Torres García, et al., 2013), (Moctezuma, et al., 2019), (Yosrita, et al., 2021), (Yosrita, et al., 2020), (D'Zmura, et al., 2009), (Proix, et al., 2022), (Lee, et al., 2022).



## 1.1 Antecedentes

Hay una diversidad de investigaciones respecto al procesamiento de señales en EEG enfocado a Interfaces Cerebro-Computadora (BCI) con el fin de darle un uso práctico. Se pretende que la explicación expuesta de los antecedentes a este proyecto sea entendible, en consecuencia, se desarrollará en dos secciones: *preprocesamiento* y *procesamiento* de neuroseñales de EEG.

### *Preprocesamiento.*

El preprocesamiento de señales no estacionarias, como lo son las neuroseñales en (EEG), ha ocasionado la necesidad de la búsqueda y desarrollo de diversas técnicas que permitan extraer características tanto del ruido en ellas como de la señal original.

El dispositivo EEG graba involuntariamente actividades eléctricas que nacen de otras regiones que no son el cerebro. Dichas actividades eléctricas se denominan artefactos. En la literatura consultada, se puede encontrar un espectro bastante amplio de técnicas de preprocesamiento para la eliminación de ruido presente en las neuroseñales EEG. De acuerdo al tipo de herramientas matemáticas que se empleen, se sugiere la siguiente agrupación: procesos estadísticos, separación ciega de fuentes (BSS), transformada wavelet (WT), descomposición de modo empírico (EMD), filtros, inteligencia artificial, entre otras.

#### A. Procesos estadísticos

En (Romero , et al., 2008), se mencionan métodos basados en regresión, donde se supone que las neuroseñales EEG son una superposición lineal e invariantes en el tiempo de fuentes distintas y se empleó *análisis de regresión múltiple* para eliminar actividad cerebral registrada en los canales de Electrooculografía (EOG), aunque también se encontró información reciente contrastante, puesto que en (Bisht , et al., 2020) se dice que se puede procesar fuentes de artefactos limitadas (no factible para señales altamente no estacionarias como EEG).



## B. Separación ciega de fuentes (BSS)

La técnica supone que el conjunto de  $m$  canales en registro se compone de una mezcla de  $n$  componentes fuente estadísticamente independientes, donde  $n \leq m$  (Romero , et al., 2008), (Bisht , et al., 2020). Ejemplos de BSS se encuentran el *Análisis de componentes independientes (ICA)*, *Análisis de componentes principales (PCA)*, *Análisis de componentes morfológicos (ACM)*, *Análisis de correlación canónica (CCA)*, entre otros. En (Torres García, 2016) se utilizó la técnica ICA en sus esquemas de eliminación de artefactos, así como en el procesamiento aplicado en (Guadarrama Rendón, 2018), en ambos casos se estudió la fuente electrofisiológica “unspoken speech”.

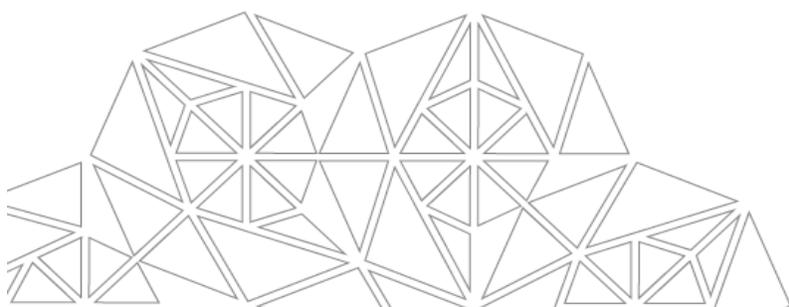
## C. Transformada wavelet (WT)

El análisis tiene su fundamento en la descomposición de una señal en versiones dilatadas y trasladadas de funciones básicas conocidas como wavelet madre (Bisht , et al., 2020), (Islam, et al., 2016). Se utilizó la *Transformada wavelet estacionaria (SWT)* en (Mejía, et al., 2016) para el reconocimiento de emociones en señales EEG multicanal, logrando altos índices de acierto con el clasificador que propusieron.

## D. Descomposición de modo empírico (EMD)

EMD se basa por completo en los datos (Bisht , et al., 2020) y esencialmente se identifica de manera empírica los modos oscilatorios intrínsecos por sus escalas de tiempo para en consecuencia la descomposición de los datos (Elías Díaz, 2020).

En (Bisht , et al., 2020) se hace mención de la eficiencia de EMD para la eliminación de la interferencia de potencia y el ruido de la electromiografía (EMG), artefactos fisiológicos relacionados a la actividad eléctrica de los músculos.



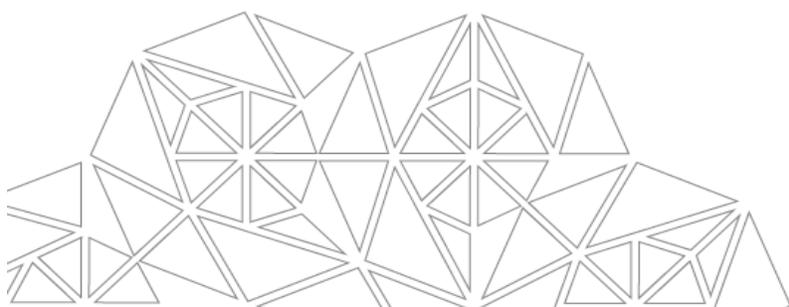
## E. Filtros

Se ha presentado una variedad de filtros en la literatura consultada, podemos mencionar el *Filtro basado en Cauchy*. Se utiliza en chip's de detección de convulsiones. Este hardware filtro implementado se aplica bajo la limitación de las dimensiones de los chips de las interfaces cerebro-computadora (BCI's) de detección médica. La distribución de Cauchy puede caracterizar artefactos como EOG y EMG (Elsayed, et al., 2017).

En (Bansal & Mahajan, 2019) se menciona que tanto el filtro de *Respuesta de Impulso Finito* (FIR) como el de *Respuesta de Impulso Infinito* (IIR) son considerados filtros digitales eficientes para el preprocesamiento de señales EEG y son seleccionados de acuerdo a las especificaciones requeridas en las frecuencias de la respuesta, mientras que (Torres García, et al., 2011) utilizó el filtro FIR como filtro pasa-bandas y como fuente electrofisiológica unspoken speech en el rango de 4 a 25 Hz. En (Sreeshakthy, et al., 2016) se empleó el *Filtrado Laplaciano de Superficie* para eliminar el ruido de artefactos.

## F. Inteligencia artificial (Clasificadores)

La inteligencia artificial ha mostrado ser una herramienta muy poderosa para el procesamiento de datos, en el cual también se incluyen las señales. En (Ranganathan , 2021) se menciona el uso de Máquinas de Vectores de Soporte (SVM) con el filtro Butterworth de segundo orden en el preprocesamiento para la clasificación de artefactos EMG, dando como resultado una exactitud del 99.23%. La *Red Neuronal Cuántica Recurrente* (RQNN) presenta una nueva arquitectura de procesamiento de información neuronal cuya inspiración se fundamenta en la mecánica cuántica, la cual incorpora la ecuación de onda de Schrödinger. Se han realizado investigaciones respecto a este modelo como un mecanismo de filtrado en el preprocesamiento de neuroseñales EEG (Gandhi , et al., 2013). El clasificador *Naive Bayes* se ha propuesto como posible método para distinguir entre los componentes relacionados con los artefactos EMG y el cerebro (Xu, et al., 2019).



## G. Otras

La diversidad de técnicas es amplia, por ende, es comprensible que existan algunas que no entren en las categorías anteriormente mencionadas, sin quitarles su particular relevancia dentro de la literatura revisada. Con el propósito de mejorar la relación señal a ruido de la señal de EEG, se puede utilizar el método de *Referencia promedio común* (CAR) puesto que se pretende eliminar lo que es común en todas las lecturas simultáneas de los electrodos. En (Torres García, et al., 2013), (Moctezuma, et al., 2017), (González Castañeda, 2015) se empleó este método. En (Aler , et al., 2012) se combinó un filtro espacial  $S$  con una *Transformada rápida de Fourier*

$$(FFT): FFT(M * S) = FFT(M) * S \quad \text{ec. 1.1}$$

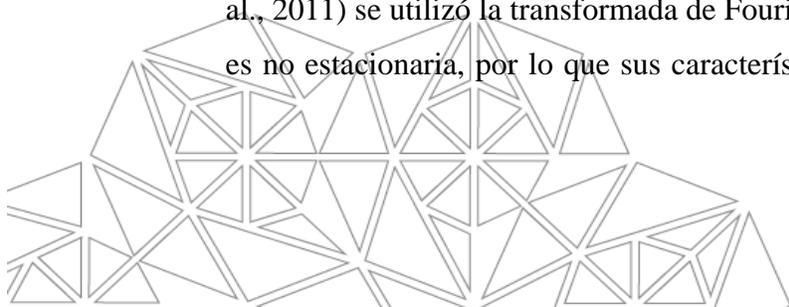
siendo  $M$  la matriz que contiene las series temporales de todos los electrodos del EEG.

En (Yuan, et al., 2017) se produjeron espectrogramas de EEG empleando la *Transformada de Fourier de tiempo corto* (STFT).

### *Procesamiento.*

El procesamiento digital de señales conlleva una metodología conocida y muy utilizada que normalmente consta de las etapas de la caracterización de las señales, selección de características y la clasificación, ésta última nos orienta a qué tanto podemos mejorar para conseguir los resultados deseados. Sin embargo, esto no implica que los resultados se deban exclusivamente al clasificador o a la técnica de inteligencia artificial elegida, pues depende en gran medida de lo obtenido en las etapas previas.

Se han realizado diversas investigaciones respecto al tema del procesamiento y extracción de características de señales EEG. En trabajos mencionados en (Torres García, et al., 2011) se utilizó la transformada de Fourier, aunque se debe considerar que la señal EEG es no estacionaria, por lo que sus características a extraer deben contener información del

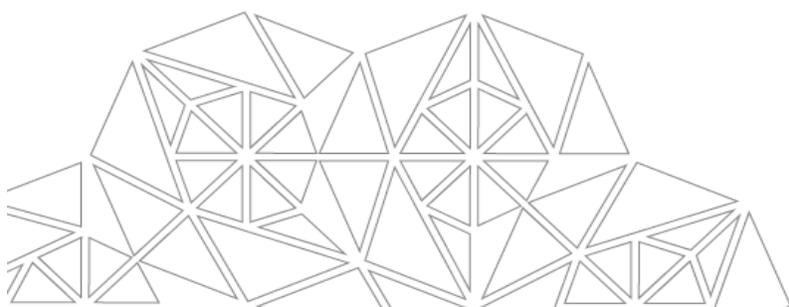


tiempo. Aunque en algunos trabajos se asumió que los modelos para reconocer características del habla común servirían también para las extraídas del EEG, las señales EEG y del habla tienen características diferentes, se presenta como ejemplo las frecuencias y los canales para captarlas. Mientras que en las señales EEG las frecuencias llegan hasta 60 Hz y se pueden captar incluso hasta en 128 canales, las del habla humana están en el rango de 125 a 4000 Hz y se captan sólo en un canal. En (Moctezuma, et al., 2017) se hace referencia de trabajos relacionados, la cual en uno se empleó para extraer características Double-Tree Complex Wavelet Transform (DTCWT) y para la etapa de clasificación offline se utilizó Janus recognition toolkit (Jrtk). Su tasa de reconocimiento alcanzada fue de 45.50%. En otros se aplicaron para la extracción de características un modelo de regresión lineal (Método de Burg), patrones espaciales comunes (CSP) y en la clasificación análisis de discriminación lineal y máquinas de soporte vectorial con una función de base radial.

En (Torres García, et al., 2011) se utilizó la DWT con 8 niveles de descomposición y una Daubechies de segundo orden (db2) como wavelet madre para extraer características y 4 clasificadores: Perceptrón Multicapa, Máquinas de Soporte Vectorial, Naive Bayes (NB) y Random Forest (RF). Tanto en (Torres García, et al., 2011) como en (Torres García, et al., 2013), RF fue el de mejores resultados, aunque no se muestran cambios significativos entre los clasificadores usados.

En (Moctezuma, et al., 2017) se extrajeron dos tipos de características. En un tipo, para calcularse se usó la DWT bior 2.2 con distintos niveles de descomposición y a cada uno se le calculó varios tipos de energía con el propósito de compararlos. El otro tipo, sus características fueron basadas en información estadística sobre la señal cruda. Para la clasificación utilizaron RF.

Lo anterior, nos brinda una noción del camino a seguir para lograr el propósito de esta investigación.



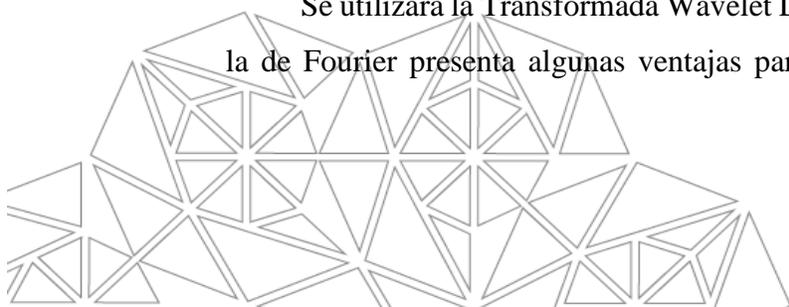
## 1.2 Justificación

Existen diversas causas para el interés en el procesamiento de bioseñales, específicamente las neuroseñales. En México, de acuerdo con el Censo de Población y Vivienda 2020, el número de personas con algún tipo de discapacidad supera los 6 millones, representando aproximadamente el 5% de la población total del país. Por tipo de discapacidad se tiene que las relacionadas con caminar, subir o bajar usando piernas es del 48%, con mover o usar los brazos o manos del 19% y el 15% con hablar o comunicarse (INEGI, 2020). También hay discapacidades motrices severas, algunas son la parálisis cerebral, las lesiones de médula espinal o cerebral, la esclerosis lateral amiotrófica (ELA), distrofias musculares, entre otras (Torres García, et al., 2011).

Las Interfaces Cerebro-Computadora (BCI) son una opción para aquellas personas que presentan las discapacidades antes mencionadas, puesto que se pretende que sean un nuevo canal de comunicación para el cerebro. Existe una variedad en los tipos de dispositivos BCI, aunque pueden generalmente clasificarse en dos grupos: invasivo y no invasivo. Las técnicas invasivas (requieren cirugía) como la electrocorticografía (ECoG), potencial de campo local (LFP) y la actividad de una unidad única (SUA) realizan mediciones para el registro de la actividad cerebral, mientras que, el electroencefalograma (EEG) o la magnetoencefalografía (MEG) son técnicas no invasivas. Es por ello, que para la adquisición de datos se empleará el EEG, al ser no invasivo y necesita un equipo relativamente sencillo y económico (Torres García, et al., 2011), (Martínez Peón, 2016), (Brumberg, et al., 2010).

En la etapa de preprocesamiento se empleará la técnica BSS con la herramienta Análisis de Componentes Independientes (ICA) para la eliminación de artefactos, como movimiento ocular, latidos del corazón, actividad muscular, entre otros, debido a su uso común en el preprocesamiento de señales EEG por su efectividad en la descomposición de fuentes (Klug & Gramann, 2021).

Se utilizará la Transformada Wavelet Discreta (DWT), puesto que en comparación con la de Fourier presenta algunas ventajas para el tipo de señal EEG, como las técnicas de



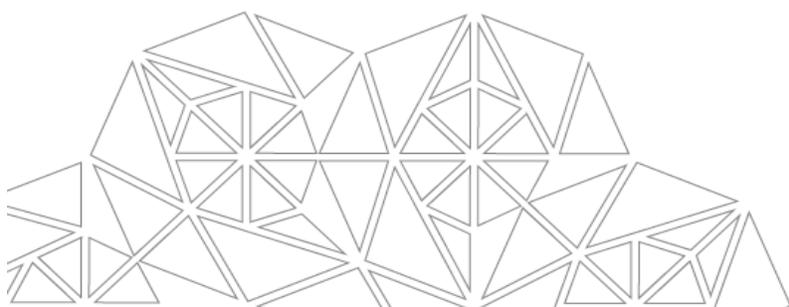
filtrado de ruido basadas en wavelets, ya que proporcionan mejores resultados, mayor rapidez desde el punto de vista computacional y en varias ocasiones brinda un mejor ajuste a los datos con menos coeficientes.

Por otro lado, se propone utilizar los clasificadores Random Forest (RF) y Máquina de Vectores de Soporte (SVM) ya que en (Torres García, et al., 2011), (Torres García, et al., 2013), (Moctezuma, et al., 2017), se mostró que presentan mejores resultados respecto a otros clasificadores como el Perceptrón Multicapa.

### 1.3 Definición del problema

Esta investigación tiene la intención de contribuir en la mejora de los dispositivos BCI, reduciendo su tiempo de procesamiento y optimizando en lo posible la clasificación de las neuroseñales de la fuente electrofisiológica habla imaginada, de un vocabulario reducido del idioma español.

Se tiene como propósito, lograr interpretar las señales EEG asociadas a la dicción imaginada de las palabras del idioma español “círculo”, “cuadrado” y “trapecio”. Se identifica a priori en cuál sección de la señal EEG el individuo imagina la pronunciación de las palabras que se le indica, así como la presentación del estímulo visual de la figura. Se tendrán dos tipos de ventana, una con sólo la información de la señal captada del habla imaginada y otra que incluya tanto la información del habla imaginada como la del estímulo visual. El problema es analizado con el criterio de clasificación, empleando clasificadores de distinta naturaleza e identificar si existe diferencia significativa en la exactitud al agregar la información del estímulo visual de las figuras geométricas en la ventana.



## 1.4 Hipótesis

Utilizando la DWT con wavelets madre seleccionadas a priori para la caracterización de las señales EEG registradas con el equipo Neuroscan Neuvo, como fuente electrofisiológica el habla imaginada, de un vocabulario de 3 palabras del idioma español y con estímulos visuales de figuras geométricas, existirá diferencia significativa en la exactitud de la clasificación al incluir la información del potencial evocado visual a la de la fuente electrofisiológica.

## 1.5 Objetivos

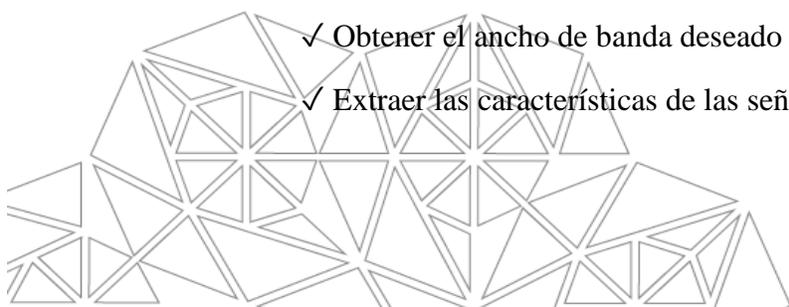
En esta sección, se presenta el objetivo general del proyecto de investigación del presente trabajo, así como los objetivos específicos que permitirán consecuentemente lograrlo.

### 1.5.1 *Objetivo general*

Interpretar palabras imaginadas no pronunciadas en señales EEG mediante el procesamiento digital, empleando un número específico de canales que proporcionen información suficiente para la caracterización de las neuroseñales utilizando la DWT, incluyendo la información del potencial evocado visual a las ventanas de la fuente electrofisiológica “habla imaginada” para medir y comparar la exactitud de la clasificación de las palabras.

### 1.5.2 *Objetivos específicos*

- ✓ Elaboración y programación del diseño experimental.
- ✓ Adquirir las señales del equipo EEG mediante un protocolo adecuado al experimento.
- ✓ Implementar un método para el filtrado de artefactos en las señales EEG.
- ✓ Obtener el ancho de banda deseado para analizar mediante un filtro pasa-banda.
- ✓ Extraer las características de las señales EEG.



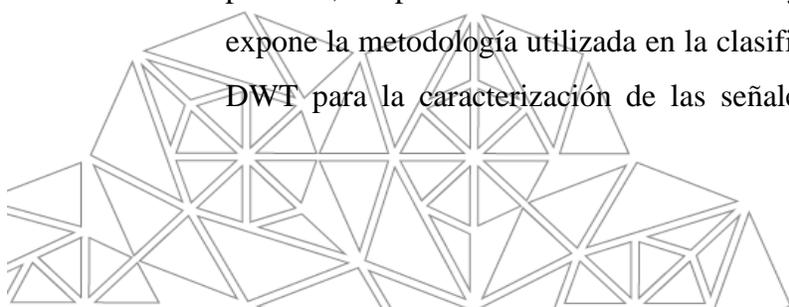
- ✓ Ajustar la dimensionalidad del/los vector(es) de características de las señales EEG.
- ✓ Emplear métodos de clasificación para las señales EEG.
- ✓ Evaluar la exactitud de los clasificadores.
- ✓ Validar la hipótesis.

## 1.6 Alcances y limitaciones

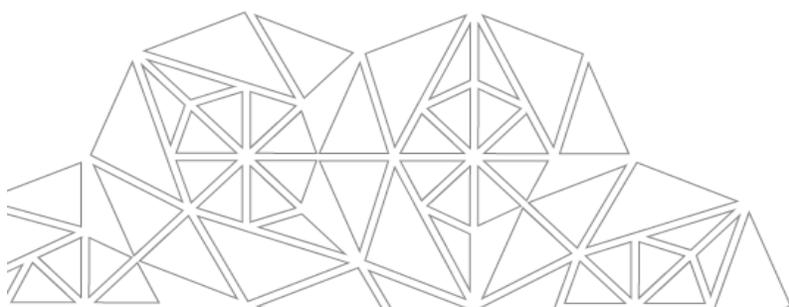
En este proyecto de investigación se experimentará con el habla imaginada, utilizando una base de datos obtenida en un laboratorio de EEG, de 8 sujetos sanos (5 hombres y 3 mujeres), en un rango de edad  $30 \pm 10$  años, cuya lengua madre es el español. Al ser una nueva base de datos con palabras no empleadas en otros experimentos (al menos no se encontró en la literatura consultada), se tiene la limitante de la comparación con otros resultados previos. Sin embargo, se abre la posibilidad de continuar con esta investigación para trabajos posteriores y expandir el vocabulario actual. Se llegará hasta la etapa de clasificación con la evaluación de la exactitud de las palabras del habla imaginada (con/sin VEP), usando las características obtenidas con la técnica DWT, por lo que, no se explorará con otras que también se emplearon en trabajos previos y se consideran viables para conseguir resultados prometedores. Además, en esta investigación no se realizarán todas las etapas necesarias para la construcción completa de un sistema BCI. Se utilizará información disponible de 16 y 32 canales con la misma configuración para todos los sujetos.

## 1.7 Estructura del documento

Para una mejor comprensión del contenido subsecuente a este capítulo, a continuación, se dará una breve descripción de su estructura. En el capítulo 2 se presenta el fundamento teórico, donde se encuentra la explicación de los conceptos que se consideran necesarios para entender los temas que se tratan en esta tesis. En el capítulo 3 se exponen los trabajos consultados donde se realizaron investigaciones del habla imaginada y el enfoque basado en palabras, los potenciales evocados visuales y la selección de canales de EEG. El capítulo 4 expone la metodología utilizada en la clasificación de palabras del habla imaginada usando DWT para la caracterización de las señales. En el capítulo 5 se expone los resultados



obtenidos de la experimentación, así como la validación de la hipótesis. El capítulo 6 tiene las conclusiones de la investigación, así como propuestas de ideas para trabajo futuro. Por último, en el documento se presenta una sección de Anexos, la cual incluye tablas de la exactitud obtenida con las distintas wavelets madre, en cada tipo de ventana y con cada clasificador, por sujeto, en las configuraciones de 16 y 32 canales.



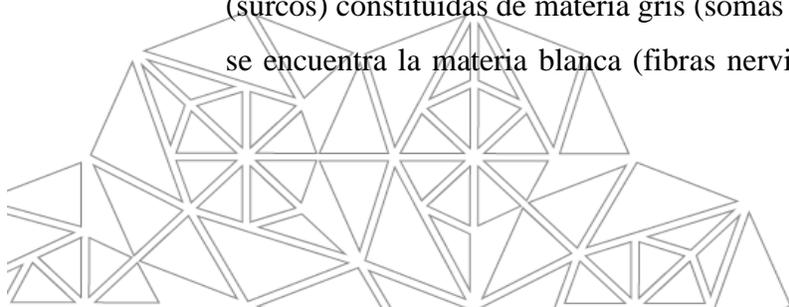
## Capítulo 2

### FUNDAMENTO TEÓRICO

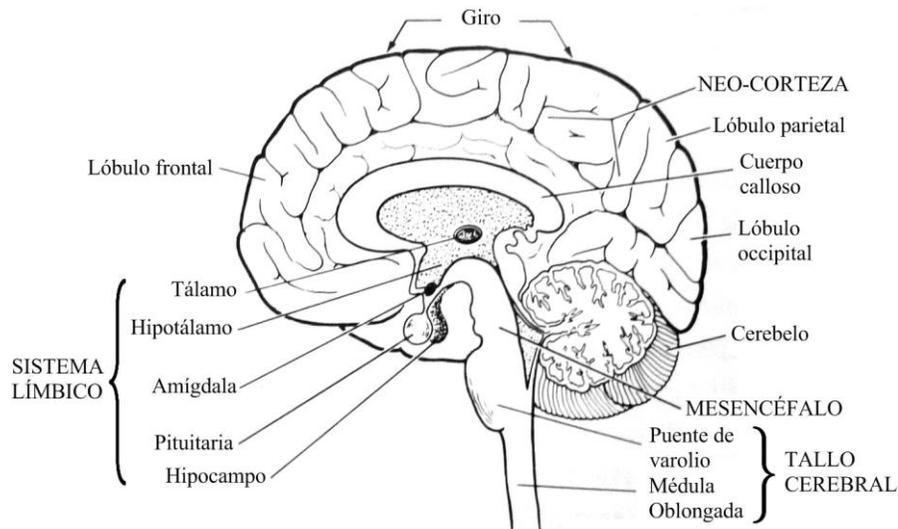
Para fines de esta investigación, se debe considerar la información requerida para el entendimiento de la metodología a emplear. Se puede identificar las siguientes áreas: *la anatomía del cerebro, el electroencefalograma (EEG)* para la obtención de las neuroseñales que son el objeto de estudio, *las interfaces cerebro-computadora (BCI's), wavelets* (ondículas) y *la transformada wavelet*, siendo esta última la herramienta matemática que se utilizará como técnica para la caracterización de las señales y su procesamiento digital. Para la etapa de clasificación se deben entender la naturaleza de los clasificadores que se emplearán, siendo para este caso la *máquina de soporte vectorial y random forest*.

#### 2.1 Anatomía del cerebro

Veamos al cerebro como parte de un sistema más grande. En (Maton, et al., 1993) se menciona que el cerebro forma parte del sistema nervioso. El sistema nervioso se divide en dos: *sistema nervioso central* (encéfalo y la médula espinal) y el *sistema nervioso periférico* (nervios motores y sensitivos que surgen del sistema nervioso central). El encéfalo humano se puede dividir a grandes rasgos en tres partes: cerebro, cerebelo y tronco cerebral, donde el cerebro es la de mayor peso y volumen. Las partes principales que consta el encéfalo son: *prosencefalo*, conformado por el telencefalo (corteza cerebral, hipocampo, amígdala cerebral, cuerpo estriado, bulbo olfatorio, ganglios basales) y el diencefalo (epitálamo, tálamo, subtálamo, hipotálamo), el *mesencefalo* y el *rombencefalo*, conformado por el metencefalo (cerebelo, puente de Varolio) y el mielencefalo (bulbo raquídeo) (Jarvis, et al., 2005), (Kandel, 2006), (Puelles López, et al., 2008). El cerebro humano está dividido en dos hemisferios, el derecho e izquierdo, separados por la cisura interhemisférica y se comunican por medio del cuerpo caloso. La superficie se denomina corteza cerebral y está formada por plegamientos denominados circunvoluciones (giros) y dividido en regiones por las fisuras (surcos) constituidas de materia gris (somas neuronales, neuroglías). Subyacente a la misma se encuentra la materia blanca (fibras nerviosas mielinizadas) (Testut & Latariet, s.f.). La



neurona es la unidad básica del cerebro. De acuerdo a algunos estudios, un cerebro humano medio dispone de alrededor de 86, 000, 000, 000 (ochenta y seis mil millones) de neuronas (Herculano Houzel & Neurosci, 2009).



**Figura 2.1** Secciones del cerebro

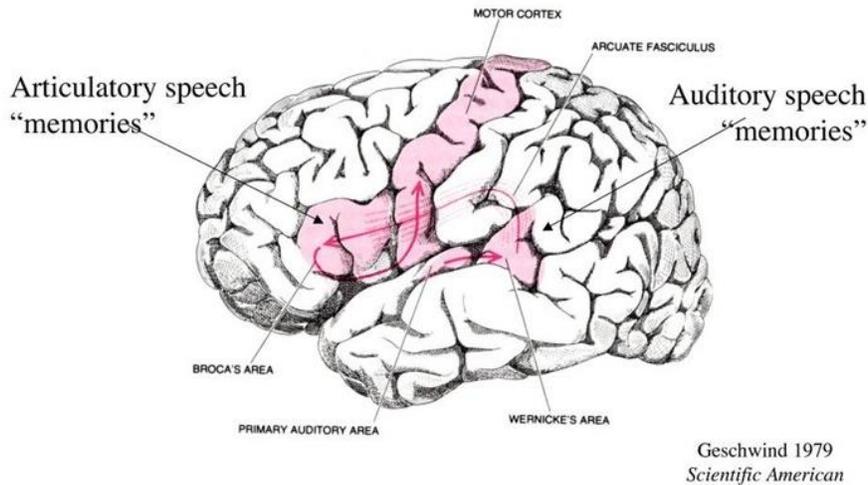
(Sagan, 1977)

### 2.1.1 Modelo Geschwind-Wernicke

De acuerdo con (Vallabhaneni, et al., 2005), se tiene identificado ciertas funciones en cada hemisferio cerebral. El izquierdo se enfoca en actividades que impliquen lógica y el lenguaje. El hemisferio derecho tiene mayor actividad en la imaginación de movimientos y las relaciones espaciales.

Hay dos áreas identificadas, consideradas vitales para la comunicación humana: el área de Wernicke y el área de Broca, localizadas en el hemisferio dominante (izquierdo en el 97% de las personas) y son consideradas las más importantes en cuanto a procesamiento de lenguaje (Purves, et al., 2004). El desarrollo del área de Wernicke permite lograr altos niveles de comprensión y procesar la mayoría de las funciones intelectuales del cerebro. Además, es el encargado de la decodificación de lo oído y de preparar posibles respuestas. También es importante para la comprensión de palabras. Mientras que, el área de Broca, también conocida como el área motora de las palabras y se conecta con el área de Wernicke mediante

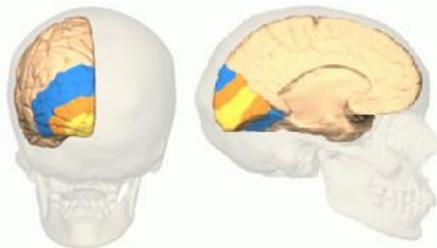
el fascículo longitudinal superior, tiene la función de permitir la realización de los patrones motores para la expresión de las palabras, articulando el lenguaje hablado y también el escrito. El modelo Geschwind-Wernicke (véase en **Figura 2.2**) pretende dar explicación al funcionamiento del lenguaje en el cerebro.



**Figura 2.2** Modelo Geschwind-Wernicke  
(Rohde, 2006)

### 2.1.2 Corteza visual

Otra región de relevancia es la corteza de asociación visual, identificada en el área 19 de Brodmann. Su función es integrar información visual y compararla con experiencias previas, de forma que su lesión impide reconocer objetos en el campo visual contralateral (García de Sola, 2014). Otras áreas relacionadas son la corteza visual primaria y secundaria (áreas 17 y 18 de Brodmann respectivamente).



**Figura 2.3** Corteza visual. Áreas 17, 18 y 19 de Brodmann (colores amarillo, naranja y azul respectivamente)  
(Science (DBCLS), s.f.)

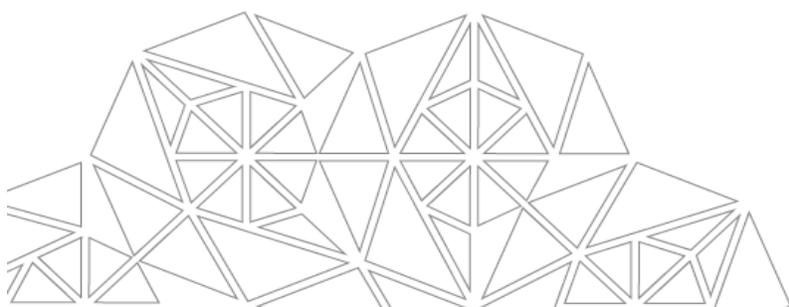
## 2.2 Electroencefalograma (EEG)

El registro electroencefalográfico (EEG) es una señal que representa el estado funcional del cerebro al detectar con alto grado de sensibilidad cambios eléctricos debidos a la oxigenación y/o perfusión cerebral, entre otras causas (Gutiérrez , 2001). Dicha señal se puede obtener mediante electrodos colocados en la superficie del cuero cabelludo. Por otro lado, en (Torres García, et al., 2011) se hace mención que las señales EEG son no estacionarias, o sea, presentan una frecuencia variable en el tiempo. Se ha medido las amplitudes de las señales en EEG y los valores típicos son entre 10 y 200  $\mu$ V.

Para medir la actividad eléctrica del cerebro, existe una clasificación de técnicas diferenciadas por el nivel de profundidad del lugar de colocación de los electrodos para la adquisición de la señal (Bermúdez Cicchino, 2013).

- ✓ *Electroencefalografía profunda*: con una intervención quirúrgica, se implantan microelectrodos en el interior del cerebro.
- ✓ *Electrocorticograma (ECoG)*: se emplean electrodos corticales sobre la corteza cerebral. También requiere intervención quirúrgica para la colocación de los electrodos.
- ✓ *Electroencefalograma*: técnica no invasiva que utiliza electrodos superficiales que se colocan sobre el cuero cabelludo.

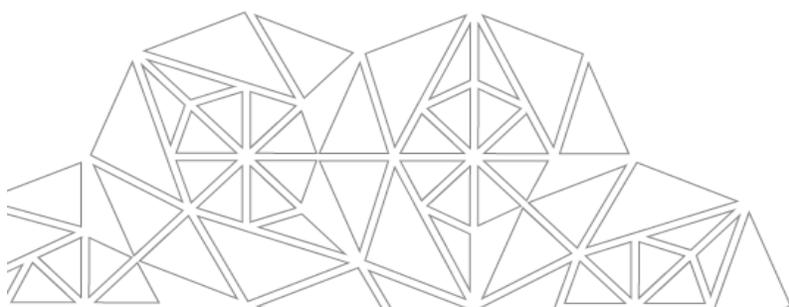
Por otro lado, se consigue una mejor relación señal-ruido y mejor resolución espacial, conforme a un mayor acercamiento al cerebro para la toma de registros, debido al reflejo de la actividad de un menor número de neuronas.

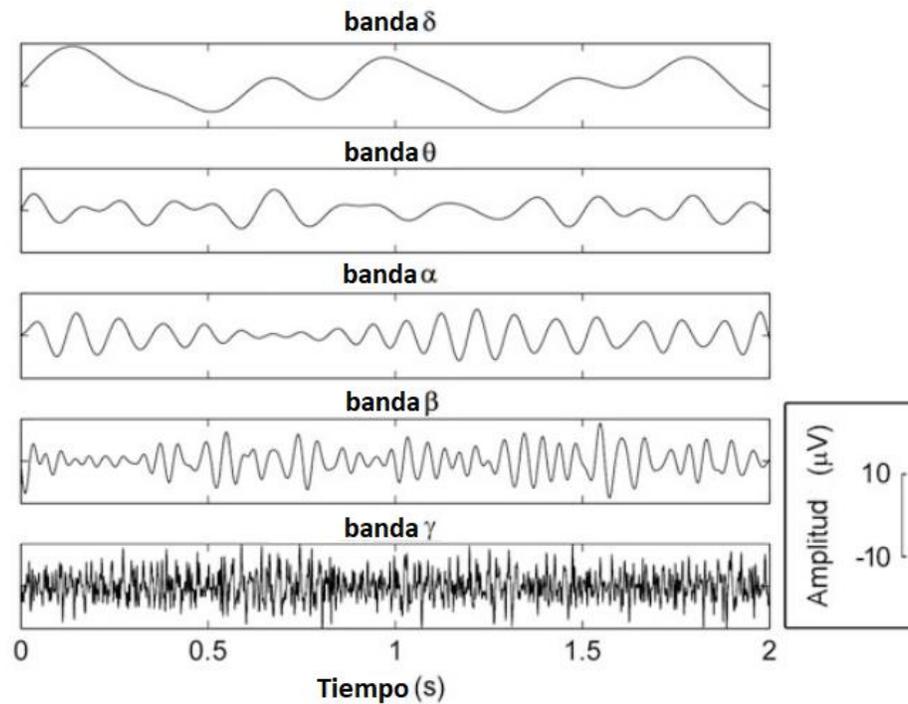


### 2.2.1 Ritmos cerebrales

Los ritmos cerebrales se definen como ondas regulares a lo largo del tiempo y que pueden obtenerse con la medición de un EEG y se han dividido en bandas denominadas delta, theta, alfa, mu, beta y gamma (Martínez Peón, 2016), (Bermúdez Cicchino, 2013).

- ✓ *Banda Delta ( $\delta$ )*, rango de frecuencia de 0.5 a 4 Hz. Se presenta durante el sueño profundo.
- ✓ *Banda Theta ( $\theta$ )*, rango de frecuencia de 4 a 8 Hz. Actividades relacionadas: la creatividad, meditación profunda y cambios emocionales.
- ✓ *Banda Alfa ( $\alpha$ )*, rango de frecuencia de 8 a 13 Hz. (Imaginación visual, concentración mental o atención).
- ✓ *Banda Mu ( $\mu$ )*, rango de frecuencia de 8 a 13 Hz. Se encuentra en el mismo rango de frecuencia que la banda  $\alpha$ , pero los cambios se detectan en la corteza motriz. La banda  $\mu$  cambia cuando el sujeto realiza imaginación motriz.
- ✓ *Banda Beta ( $\beta$ )*, rango de frecuencia de 13 a 30 Hz. (Cálculo mental y durante percepción háptica).
- ✓ *Banda Gamma ( $\gamma$ )*, rango de frecuencia de 25 a 100 Hz. Se presenta durante la sincronización de eventos relacionados (ERS).

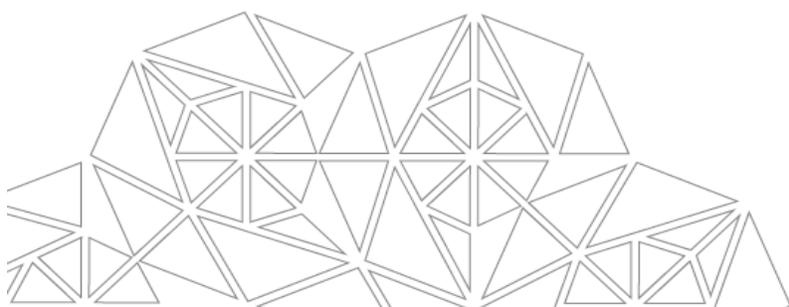


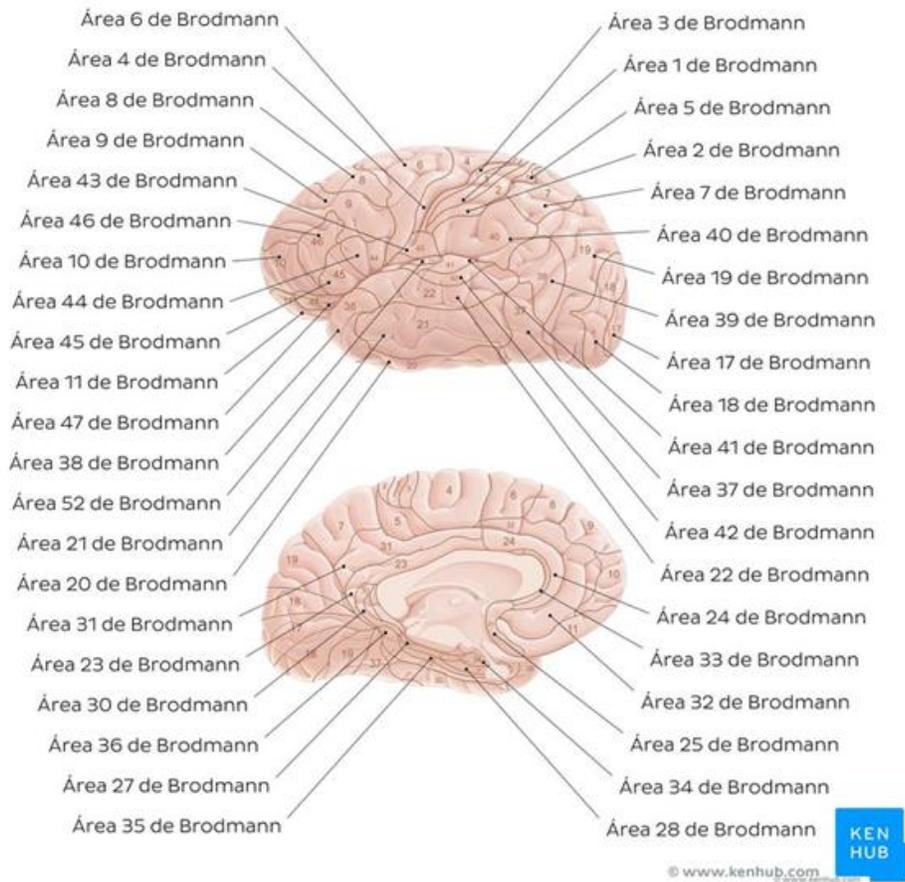


**Figura 2.4** Patrones en la actividad de los ritmos del EEG. La actividad rítmica del EEG a menudo se divide en bandas por frecuencia. Empíricamente, la actividad EEG espontánea se divide en cinco bandas de frecuencia: delta ( $\delta$ ,  $< 4$  Hz), theta ( $\theta$ ,  $4 - 8$  Hz), alfa ( $\alpha$ ,  $8 - 13$  Hz), beta ( $\beta$ ,  $13 - 30$  Hz) y gamma ( $\gamma$ ,  $> 30$  Hz) (Adaptación de (Hu & Zhang, 2019)).

### 2.2.2 Áreas de Brodmann

A comienzos de los años 1900, el anatomista alemán Korbinian Brodmann definió y enumeró 52 áreas de la corteza cerebral con base en su citoarquitectura (estructura histológica y organización celular). Estas 52 regiones se conocen actualmente como las áreas de Brodmann (Navarro & Torres, 2022).



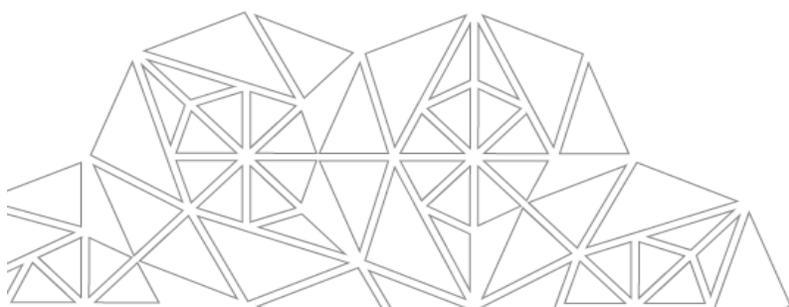


**Figura 2.5** Áreas de Brodmann  
(Navarro & Torres, 2022)

Para interés de esta investigación, nos enfocaremos en las áreas de Broca, Wernicke y la corteza visual.

#### Áreas 22, 39, 40 (área de Wernicke)

El giro que comprende esta área puede ser de mayor o menor tamaño dependiendo del individuo y es la responsable de la comprensión del lenguaje. Esta área nos permite hilar palabras separadas para completar oraciones que tengan sentido (Navarro & Torres, 2022).



### Áreas 44, 45 (área de Broca)

Esta región está asociada con la práctica del habla (programación motora del habla) (Navarro & Torres, 2022). Esto incluye:

- ✓ Reunir los elementos que vinculan al lenguaje.
- ✓ Seleccionar información entre fuentes que compiten entre sí.
- ✓ Secuenciación motora/elementos expresivos.
- ✓ Mecanismos de control cognitivos para el procesamiento sintáctico de oraciones.
- ✓ Construcción de oraciones complejas y patrones del habla.

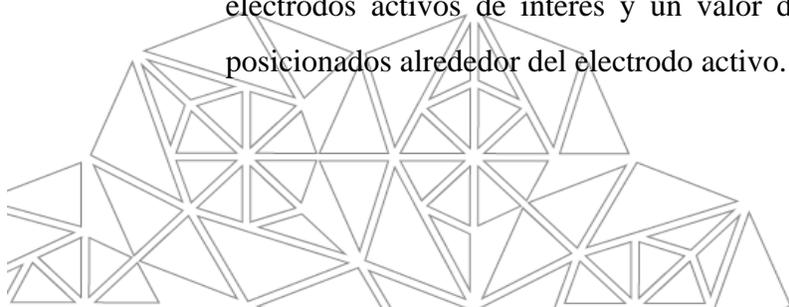
### Áreas 17, 18 y 19 (Corteza visual)

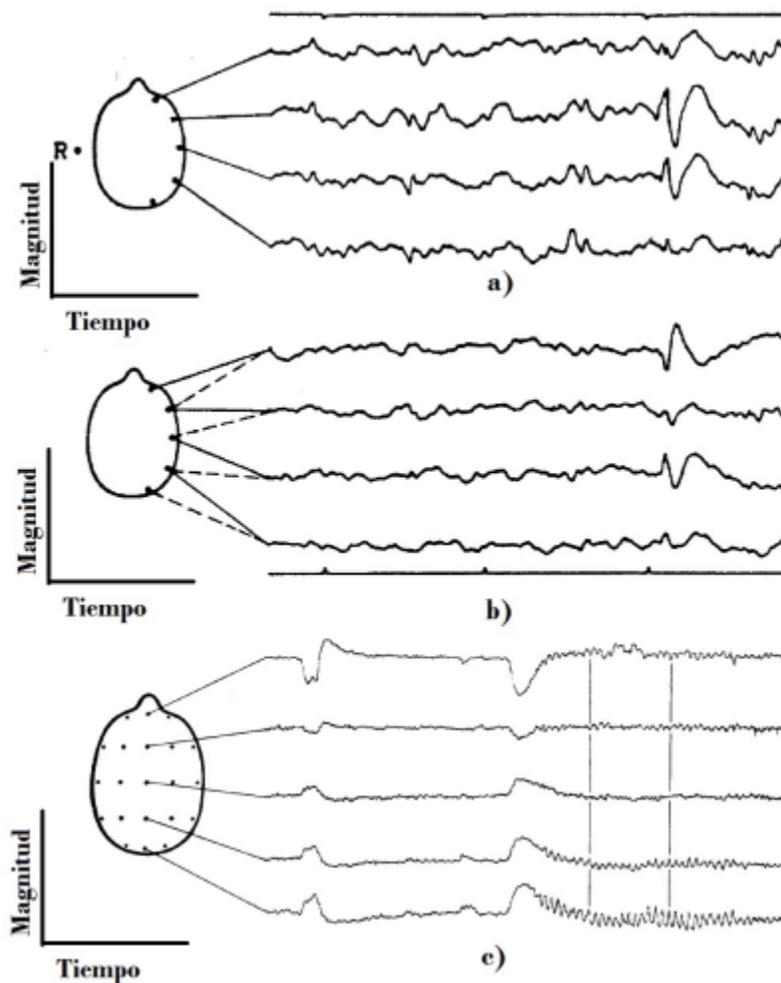
De acuerdo con (Navarro & Torres, 2022), la corteza visual se localiza en el lóbulo occipital, en la parte posterior del encéfalo, y contiene un mapa bien definido de la información espacial necesaria para la visión. Nos permite integrar información visual y compararla con experiencias previas, de forma que su lesión impide reconocer objetos en el campo visual contralateral (García de Sola, 2014).

#### 2.2.3 Montajes de electrodos

Cuando se va a realizar el montaje de los electrodos en el cuero cabelludo, según la configuración, se pueden identificar tres topologías, (Martínez Peón, 2016), (González Castañeda, 2015), (Bermúdez Cicchino, 2013): *monopolar (unipolar) o de referencia*, *bipolar* y *laplaciana o de referencia promedio*.

- ✓ *Monopolar o referencial*: Hay registro de la diferencia de potencial entre un electrodo de referencia (preferentemente ubicado en una zona con la menor actividad posible) y los electrodos ubicados en la zona de interés.
- ✓ *Bipolar*: Se adquieren las diferencias de potencial entre los puntos de las distintas zonas de interés donde se posicionarán los electrodos.
- ✓ *Laplaciana*: Parecido al monopolar, se obtienen las diferencias de potencial entre los electrodos activos de interés y un valor de referencia promedio de distintos electrodos posicionados alrededor del electrodo activo.





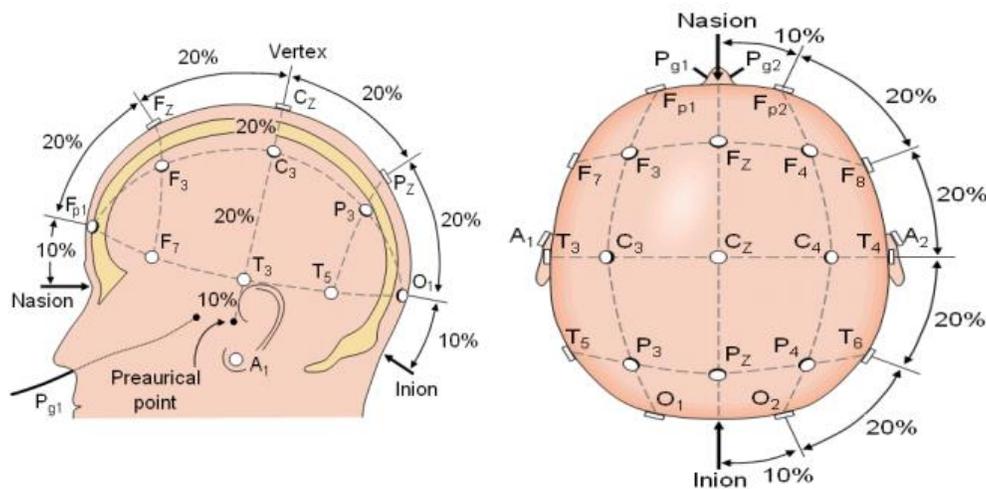
**Figura 2.6** Tipos de montajes para medición EEG: a) Montaje unipolar, b) Montaje bipolar, c) Montaje de referencia promedio. (Martínez Peón, 2016)

En (Martínez Peón, 2016) se menciona que, los dispositivos de adquisición de señales EEG empleados en los sistemas BCI requieren un montaje unipolar, la cual se toma un canal como referencia.

Hay distintos sistemas que posibilitan estandarizar la colocación de los electrodos no invasivos, es decir, aquellos que se colocan en la superficie del cuero cabelludo. Entre ellos, se encuentra el *sistema internacional 10/20* (véase en **Figura 2.7**), siendo el sistema 10/10 su extensión. Los números 10/20 son referencia al porcentaje (10 y 20) de la distancia en la que se deben colocar los electrodos en el lado izquierdo y derecho del cráneo, la sección

frontal y occipital. Se emplea una letra para identificar el lóbulo y un número para la identificación del hemisferio (UNAM, 2019). El electrodo F corresponde al lóbulo frontal, T al temporal, C al central, P al parietal y O al occipital. El hemisferio izquierdo adquiere los números impares y el hemisferio derecho los pares. Para la ubicación de la parte central del cráneo, se utiliza como referencia la Z (zero) para la mitad del cráneo. Existen cuatro puntos esenciales para la colocación de electrodos:

- ✓ 1.- Nasion (punto entre la frente y la nariz).
- ✓ 2.- Inión (punto más bajo del cráneo). Sección posterior de la cabeza, habitualmente se identifica por un tope prominente.
- ✓ 3 y 4.- Puntos preauriculares (se colocan anterior al oído).



**Figura 2.7** Posición de electrodos de acuerdo al Sistema Internacional 10/20.

(UNAM, 2019)

#### 2.2.4 Artefactos

Algo relevante a considerar, es conocer cuáles señales son las que se van a estudiar, puesto que lo adquirido que no sea de nuestro interés puede considerarse como ruido.

El EEG graba involuntariamente actividades eléctricas que nacen de otras regiones que no son el cerebro. Dichas actividades eléctricas se denominan *artefactos*. En (Elsayed, et al., 2017), (Muñoz Sánchez, 2014), (Hu & Zhang, 2019), se menciona que se pueden clasificar

en dos tipos dependiendo de su origen: los *fisiológicos* o *biológicos* (producidos por el propio paciente como movimiento de los ojos, latidos del corazón, muscular, entre otros) y los *extrafisiológicos* (*no fisiológicos*) o *técnicos* (interferencias captadas por el EEG como ondas electromagnéticas de distintas frecuencias y amplitudes del entorno como red eléctrica, ruido del amplificador operacional, electrodos, entre otros), (véase en la **Tabla 2.1**).

**Tabla 2.1** Artefactos fisiológicos y no fisiológicos  
(Adaptación de (Hu & Zhang, 2019))

<i>Artefactos fisiológicos</i>	<i>Artefactos no fisiológicos</i>
<u>Electrooculografía (EOG)</u> <i>Relacionados a los movimientos de los ojos y parpadeos.</i>	Interferencia de la línea eléctrica (50 Hz en Europa, 60 Hz en Estados Unidos).
<u>Electromiografía (EMG)</u> <i>Relacionados a la actividad eléctrica de los músculos.</i>	Vinculados a la mala colocación de los electrodos en el cuero cabelludo.
<u>Electrocardiograma (ECG)</u> <i>Relacionados a los latidos del corazón.</i>	Mal funcionamiento de cualquier parte dentro de la grabación del sistema EEG.
<u>Transpiración y movimientos del cuero cabelludo.</u>	Artefactos digitales como el cableado suelto, fallos en las conexiones de las placas de los circuitos.

Cabe recalcar, que los artefactos EMG suelen estar en frecuencias superiores a los 25 Hz (Brigham & Kumar, 2010).

### 2.3 Interfaces Cerebro-Computadora (BCI's)

Anteriormente, se mencionó que una interfaz cerebro-computadora basada en EEG es un sistema que permite un canal alternativo de comunicación y control al cerebro con el exterior, sin utilizar las salidas normales de los nervios periféricos y los músculos. En (Ramírez Diez, 2017) se define como un sistema de recolección de señales cerebrales con distintos mecanismos de neuroimagen, cuyo propósito es utilizarse mediante un computador, generando interacción de un sujeto con el medio, sin llegar a necesitar sus nervios periféricos o sus músculos. Para la observación de las señales cerebrales hay alternativas, como el



electroencefalograma (EEG), las imágenes de resonancia magnética funcional (fMRI), el electrocorticograma (ECoG), la actividad de una neurona (SUA), entre otras.

Estos sistemas deben ser capaces de transmitir las intenciones de las personas a un dispositivo externo como un computador, sintetizadores de habla, tecnologías asistidas o neuroprótesis, entre otros, siendo de interés a personas con reducción en la capacidad de la movilidad, teniendo la posibilidad de mejorar su calidad de vida y, también, reducir los costos de un cuidado intensivo externo.

Las BCI's pueden reconocer ciertos patrones en las señales de EEG, mediante 3 etapas: preprocesamiento de la señal de EEG, extracción de características y la etapa de clasificación (Ilyas, et al., 2015).

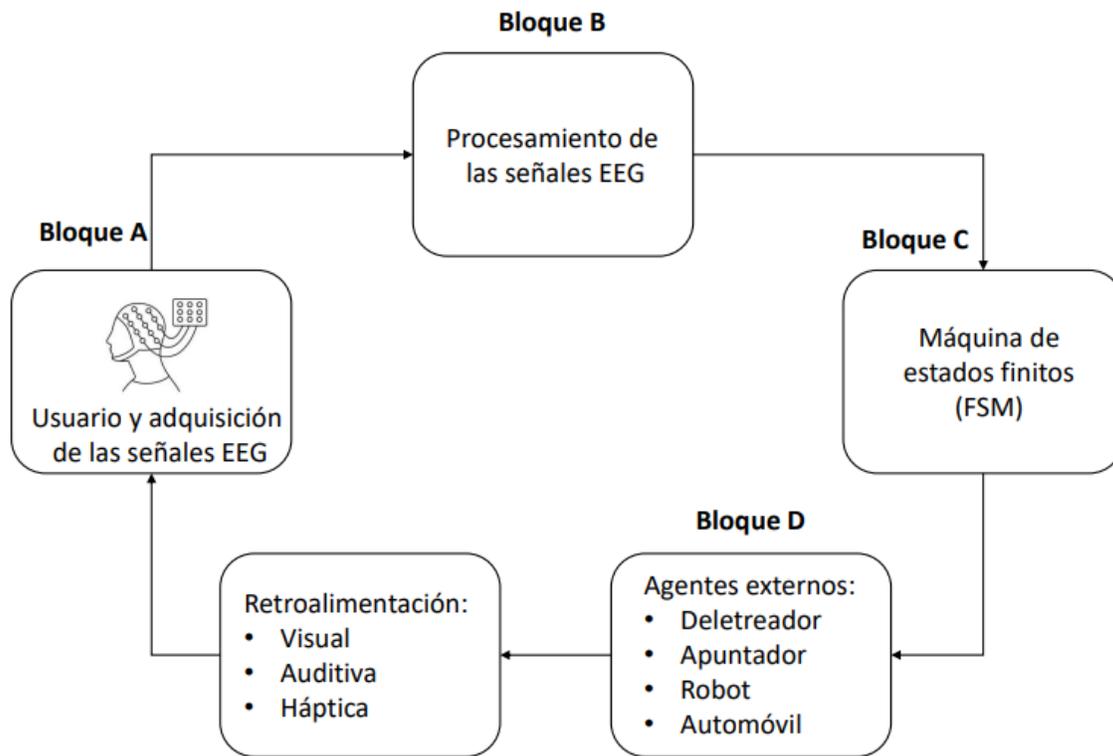
De acuerdo con (Moctezuma Pascual, 2017), podemos clasificar a las BCI's en dos categorías: invasivas y no invasivas. Las técnicas invasivas realizan mediciones para el registro de la actividad cerebral en forma intracraneal o intracortical mediante ECoG, potencial del campo local (LFP) o SUA, mientras que las técnicas no invasivas (no requieren cirugía) a través del EEG o magnetoencefalografía (MEG).

### 2.3.1 Componentes de un sistema BCI

Generalmente, podemos mencionar cuatro componentes o bloques principales que conforman un sistema BCI: *el usuario y adquisición de las señales EEG*, mismo que se le entrena, pidiéndole que realice una tarea, que puede ser cognitiva motriz (imaginarse movimiento o fuerza de algún miembro (imaginación motriz (MI, del inglés *motor imagery*))) o cognitiva no motriz (imaginación visual (VI, del inglés *visual imagery*), cálculo mental, habla imaginada, entre otros), *procesamiento de las señales EEG*, conformado por las etapas de preprocesamiento, caracterización y clasificación, *la máquina de estados finitos* (FSM, del inglés *Finite State Machine*) que mapea los comandos provenientes del bloque de procesamiento en comandos inteligibles para el agente externo y *el agente externo* como apuntadores de pantalla o sistemas dinámicos (robots, exoesqueletos, automóviles, entre



otros). Para cerrar el bucle del sistema, hay retroalimentación ya sea de forma visual, háptica o auditivamente, entre otras, condicionando la respuesta del usuario en el bloque B (véase en la **Figura 2.8**).



**Figura 2.8** Diagrama de bloques de un sistema BCI (Adaptación de (Martínez Peón, 2016)). Bloque A: Las señales EEG del usuario son medidas. Bloque B: Se limpian, caracterizan y clasifican las señales EEG para la inferencia de comandos. Bloque C: la FSM transforma los comandos a tareas factibles para el agente externo que serán por un controlador del bloque D. El bucle del sistema se cierra, retroalimentando al sujeto (visual, auditiva y/o háptica).

### 2.3.2 Señales electrofisiológicas

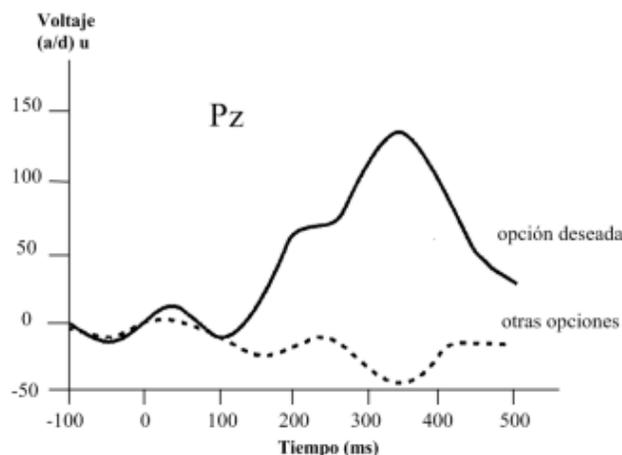
Los paradigmas utilizados en los sistemas BCI, también conocidos como señales electrofisiológicas, son las señales eléctricas generadas por el usuario que permiten el control de dicho sistema y producir las señales de mensajes o comandos para la comunicación o control al exterior. Es importante entender que, las vías de salida naturales del cerebro no son las que pretende manejar un sistema BCI (músculos, nervios periféricos). Más bien, uno de los principales objetivos es manipular las señales eléctricas emitidas por las neuronas

cerebrales. Anteriormente, se mencionó la descripción de la fuente electrofisiológica del habla imaginada, sin embargo, hay diferentes tipos de potenciales de EEG que también pueden utilizarse como señales o fuentes electrofisiológicas.

De acuerdo con (Bermúdez Cicchino, 2013), una manera de clasificar a los potenciales de EEG es por su origen en *potenciales evocados* si son producidos como consecuencia de estímulos externos y *potenciales espontáneos* al ser, como su nombre lo indica, de origen naturalmente espontáneo (ritmos cerebrales).

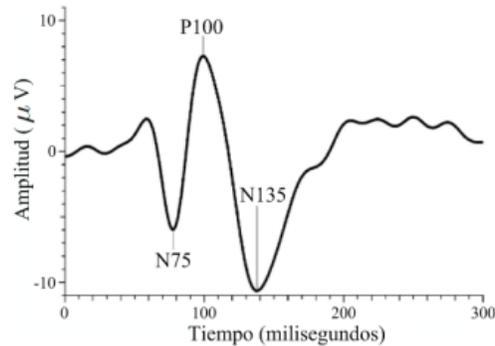
Los potenciales evocados se encuentran en un subconjunto de los *potenciales relacionados a eventos* (ERP, del inglés *Event Related Potentials*), los *ERP exógenos*, aquellos cuando los eventos que generan señales de respuesta de EEG son externos al sujeto, debido a que los potenciales evocados tienen su origen de una respuesta a estímulos físicos. Dentro de los ERP exógenos se encuentran los potenciales P300, potenciales evocados visuales (VEPs, del inglés *Visual Evoked Potentials*), potencial evocado visual de estado estacionario (SSVEP, del inglés *Steady State Visual Evoked Potential*), entre otros.

P300: Son picos positivos en la EEG a causa de estímulos visuales, auditivos o sensoriomotores no frecuentes. El P300 aparece en la EEG a los 300 ms después de un estímulo extraño o atípico entre estímulos normales y frecuentes.



**Figura 2.9** P300 (Torres García, 2011)

VEP: Normalmente aparecen en la región occipital de la corteza cerebral (Lee, et al., 2019), (Torres García, 2011) se identifican con tres picos, con una sucesiva variación de la polaridad (negativa-positiva-negativa), con latencias de los picos de 70 ms, 100 ms y 135 ms respectivamente (Ojeda, 2005).



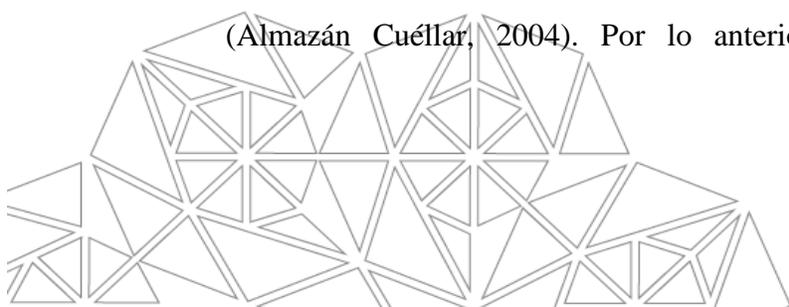
**Figura 2.10** Potenciales evocados visuales (Adaptación de (Torres García, 2011)). Latencias promedio (75, 100 y 135 ms) después del estímulo, acompañadas de las letras N y P (picos negativos y positivos respectivamente).

SSVEP: Señal de respuesta eléctrica oscilatoria continua, ocasionada por una estimulación visual a una frecuencia específica (Martínez Peón, 2016), (Torres García, 2011).

Los ERP endógenos se presentan como respuesta a un evento interno del sujeto. Entre ellos se encuentran imaginación motriz (MI, del inglés *Motor Imagery*), imaginación visual (VI, del inglés *Visual Imagery*), ritmos sensoriales motrices mu y beta (SMRs, del inglés *Sensorimotor Rhythms*), imaginación kinestésica motriz (KMI, del inglés *Kinesthetic Motor Imagery*), entre otros (Martínez Peón, 2016), (González Castañeda, 2015), (Bermúdez Cicchino, 2013).

## 2.4 Transformadas Wavelet

Las wavelets son una herramienta muy poderosa que permite análisis simultáneo en el tiempo y frecuencia, gracias a la flexibilidad de sus fundamentos matemáticos, siendo útil para el análisis de fenómenos transitorios, no estacionarios, o variantes en el tiempo (Almazán Cuéllar, 2004). Por lo anterior, son una herramienta “ad hoc” para el



procesamiento de señales EEG, ya que están clasificadas como no estacionarias, dado que presentan una frecuencia variable en el tiempo (Torres García, 2016).

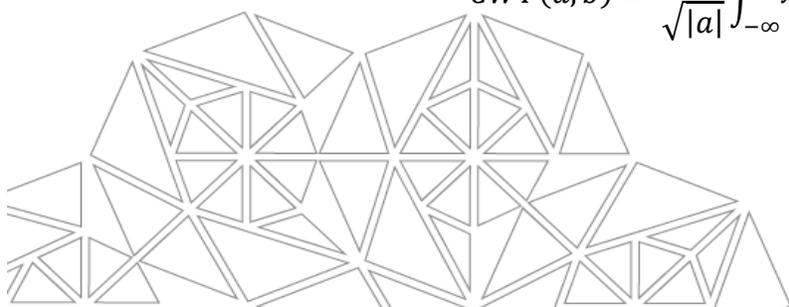
Se define a una *onda* como una función oscilatoria en el tiempo o espacio. Una *wavelet*, también conocida como “*ondícula*” u “*ondeleta*” es una onda pequeña la cual tiene su energía concentrada en el tiempo o espacio (Almazán Cuéllar, 2004), es de duración limitada en el tiempo y tiene un valor medio de cero (Barcala Riveira, et al., 2003).

El análisis de una transformada wavelet se basa en la descomposición de una señal o función en versiones dilatadas y trasladadas de funciones básicas conocidas como *ondículas madre* (Islam, et al., 2016), (Bisht , et al., 2020). Con la transformada wavelet, la señal (en este caso, las señales EEG) se compara con versiones desplazadas y distinta escala de las funciones de la ondícula madre. Al comparar la ondícula madre y la señal, se pueden obtener los coeficientes de correlación entre esta última y las diferentes versiones de la ondícula madre, y con estos se puede reconstruir la señal original mediante la *Transformada wavelet inversa* (Santos, et al., 2004). Comparando con la transformada de Fourier con ventana, la transformada wavelet tiene una mayor capacidad para hacer un acercamiento en fenómenos de alta frecuencia con una muy corta vida (Melara Estrada, 2015).

Hay diversos tipos de transformada wavelet, todas operando bajo el mismo principio. Entre los tipos comúnmente utilizados están la *Transformada wavelet continua* y la *Transformada wavelet discreta*.

La transformada wavelet continua (CWT, del inglés *Continuous Wavelet Transform*) Es muy útil para el análisis de datos multi-escala. Es una transformada integral que se puede utilizar para explorar y obtener características de señales no estacionarias (Domingues , et al., 2016). La *CWT* de una señal  $x(t)$  está dada por:

$$CWT(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad \text{ec. 2.1}$$



donde  $\psi$  es la función wavelet madre,  $b$  realiza traslaciones de la función  $\psi$  a lo largo de  $x(t)$  y  $a = \frac{f_0}{f}$  genera distintas escalas (dilataciones o contracciones) de la función  $\psi$ , siendo  $f_0$  la frecuencia central de la wavelet. Además,  $\psi^*$  representa el complejo conjugado de  $\psi$  (Almazán Cuéllar, 2004), y el factor de normalización  $\frac{1}{\sqrt{|a|}}$  asegura que la energía sea la misma para todos los valores de  $a$  (Torres García, 2011).

Un detalle a considerar, es que la *CWT* presenta la desventaja de ser redundante, y se debe al sobre-muestreo de la señal original  $x(t)$ , o sea, se producen más coeficientes de los necesarios para especificar la señal. Dicha redundancia genera ineficiencia computacional, mostrando la transformada wavelet discreta (DWT, del inglés *Discrete Wavelet Transform*) una ventaja en este aspecto.

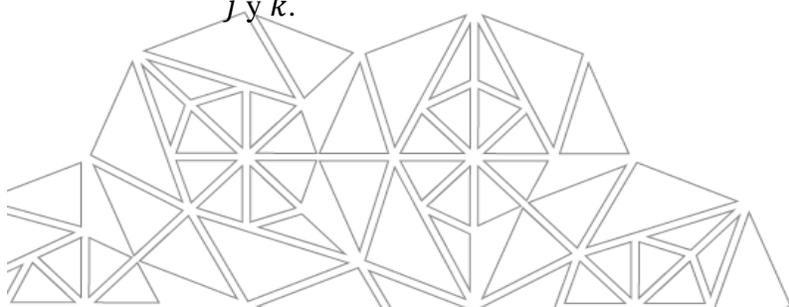
#### 2.4.1 Transformada wavelet discreta (DWT)

Si se eligen las escalas y posiciones en la forma  $2^x$ , el análisis muestra una evidente mayor rapidez e igual eficiencia. Al análisis wavelet que utiliza estas escalas diádicas se le conoce como *transformada wavelet discreta* (DWT) (Barcala Riveira, et al., 2003) . Las frecuencias se particionan o dividen en octavas consecutivas o bandas de frecuencia, con el propósito de conseguir la máxima eficiencia computacional. Es por ello, que se utilizan potencias enteras de base 2 (Almazán Cuéllar, 2004) .

En ese caso, la DWT es algunas veces llamada transformada wavelet diádica (Torres García, 2016), (Torres García, 2011), (Wu, et al., 2009):

$$DWT(j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t - 2^j k}{2^j} \right) dt \quad \text{ec. 2.2}$$

donde  $a = 2^j$ ,  $b = 2^j k$  y  $DWT(j, k)$  es una muestra de  $CWT(a, b)$  en los puntos discretos  $j$  y  $k$ .



Otra expresión para la ec. 2.2 que también define la *DWT* es (Torres García, et al., 2013):

$$DWT(j, k) = \sum_j \sum_k f(x) \psi_{j,k}(x), \quad \text{ec. 2.3}$$

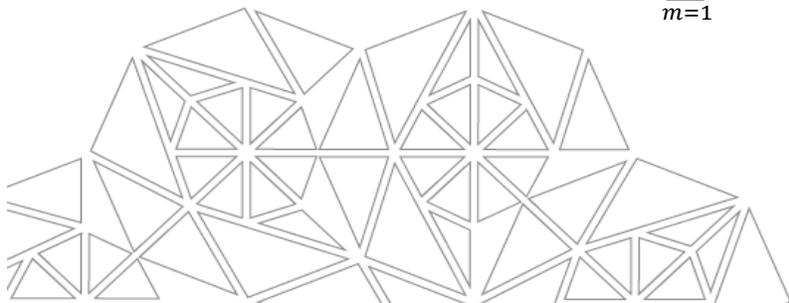
donde  $\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k)$  y  $j, k$  números enteros que indican el ancho y la posición de la wavelet.

$\psi_{j,k}(n)$  es el conjunto de funciones a las que se puede hacer referencia como la familia de wavelets provenientes de  $\psi(n)$  (wavelet madre), una función del tiempo que es una onda pequeña cuya energía finita se concentra y decae rápidamente (González González, 2010). Mediante el uso de una familia de ondículas ortonormales, las bases del espacio wavelet tienen una correspondencia con estas funciones después de las operaciones de escalado y desplazamiento.

La DWT está determinada por dos funciones, la función wavelet madre  $\psi$  y la función escala  $\phi$  (también conocida como wavelet padre) (Almazán Cuéllar, 2004), (Rufiner Di Persia, 1996). Los coeficientes wavelet o de detalle  $h_j(k)$  y los coeficientes de escala o aproximación  $a_j(k)$  están asociados con la función wavelet madre y la función escala respectivamente. Las ecuaciones de descomposición 2.4 y 2.5 se pueden escribir como convoluciones discretas con submuestras porque las secuencias se muestrean sólo en los números enteros pares (Torres García, et al., 2013), (Medina Sánchez, 2017):

$$a_{j+1}(k) = \sum_{m=1}^N l(m - 2k) a_j(m), \quad \text{ec. 2.4}$$

$$h_{j+1}(k) = \sum_{m=1}^N g(m - 2k) a_j(m), \quad \text{ec. 2.5}$$



donde  $l$  es el filtro pasa-bajas (filtro de escala) y  $g$  es el filtro pasa-altas (filtro wavelet) y  $N$  es la longitud de la secuencia de entrada  $a_j, a_0 = f(x)$ .

#### 2.4.2 Caracterización con energías wavelets

Una manera de caracterizar las señales EEG, es normalizando la información disponible en cada ventana o época que se obtiene, gracias al número de coeficientes wavelet en cada nivel de descomposición. Hay alternativas que permiten calcular la distribución de varios tipos de energía de una señal, entre las que se encuentran la *relativa*, *instantánea*, *jerárquica* y *teager* (Moctezuma, et al., 2017) .

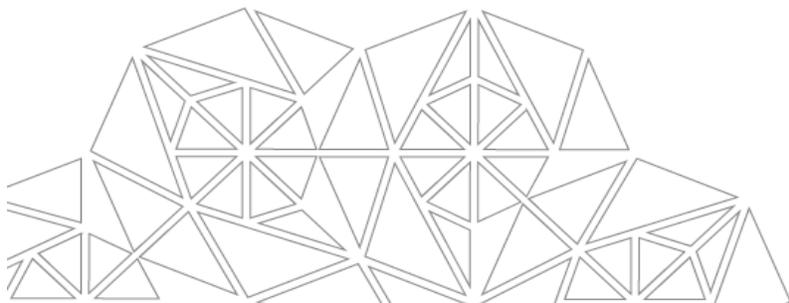
Sea  $o_m(n)$  el coeficiente ondícula en el tiempo  $n$  y banda de frecuencia  $m$ . Si  $T$  es la longitud de la ventana (época) de análisis, entonces,  $o_m(n)$  tiene  $T_m = T/2^m$  muestras en el nivel de descomposición  $m$ :

Energía Relativa: se puede considerar como una densidad tiempo-escala su distribución de energía y su información permite caracterizar en distintas bandas de frecuencia la distribución de energía de la señal.

$$RWE_m = \frac{E_m}{E_{total}} \quad \text{ec. 2.6}$$

donde  $E_m$  representa la energía en el  $m$ -ésimo nivel de descomposición y  $E_{total}$  la energía total de los coeficientes wavelet de una señal.  $E_m$  se define como:

$$E_m = \begin{cases} \sum_j |h_{m,j}|^2, & \text{si } m \leq T; \\ \sum_j |a_j|^2, & \text{caso contrario.} \end{cases} \quad \text{ec. 2.7}$$



En el caso de  $E_{total}$  se calcula de la siguiente manera:

$$E_{total} = \sum_{m=1}^{T+1} E_m \quad \text{ec. 2.8}$$

Se puede deducir que,  $\sum_m RWE_m = 1$ , logrando caracterizar la distribución de energía de la señal en distintos rangos (bandas) de frecuencia. Los coeficientes de energía se calculan para cada canal (electrodo) por ventana y se concatenan para generar un vector de características que represente a la señal EEG.

Energía Instantánea: refleja la amplitud de la señal.

$$f_m = \log_{10} \left( \frac{1}{T_m} \sum_{n=1}^{T_m} (o_m(n))^2 \right) \quad \text{ec. 2.9}$$

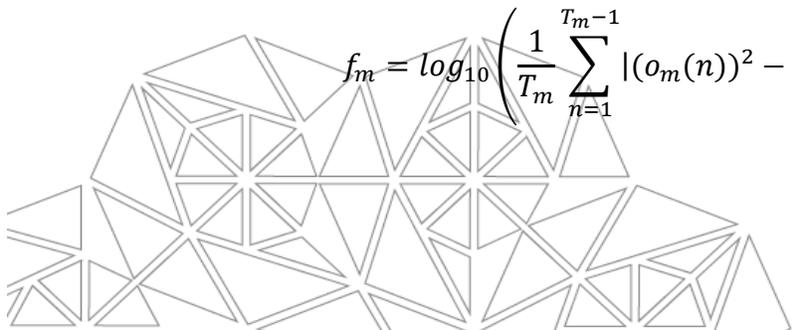
Energía Jerárquica: brinda una resolución jerárquica en tiempo y le da más importancia al centro de la ventana de análisis.

$$f_m = \log_{10} \left( \frac{1}{T_m} \sum_{n=(T_m-T_M)/2}^{(T_m+T_M)/2} (o_m(n))^2 \right) \quad \text{ec. 2.10}$$

donde  $T_M$  es el número de muestras en el último nivel de descomposición.

Energía Teager: refleja las variaciones tanto en amplitud como en frecuencia de la señal. Es un parámetro robusto para el reconocimiento de habla, ya que atenúa el ruido auditivo.

$$f_m = \log_{10} \left( \frac{1}{T_m} \sum_{n=1}^{T_m-1} |(o_m(n))^2 - o_m(n-1) * o_m(n+1)| \right) \quad \text{ec. 2.11}$$



## 2.5 Clasificación

Para entender con claridad el concepto de clasificación, debemos partir de dos conjuntos de datos a los que llamaremos *conjunto de entrenamiento* y *conjunto de prueba*. La capacidad de predecir o tomar una decisión depende del conjunto de entrenamiento, es decir, de la información histórica disponible, del cual surge un modelo que será evaluado con un conjunto distinto de información, el cual llamamos conjunto de prueba y de este modo, identificar si el modelo puede generalizarse con nuevos casos. Se podría decir que es una modalidad de aprendizaje supervisado (Orenes Casanova, 2022)

Llamémosle  $B$  a la base de datos de la cual se obtienen los conjuntos de entrenamiento y de prueba y podemos definir como:

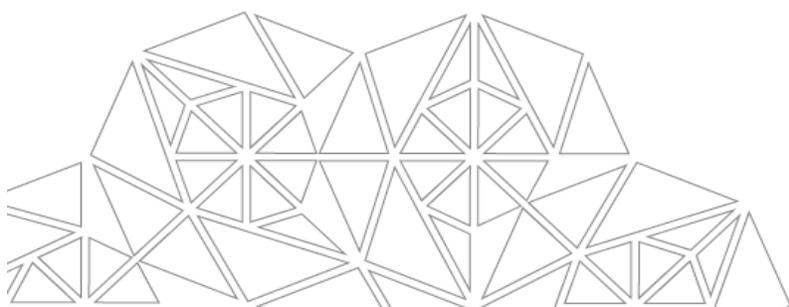
$$B = \{\langle u_1, v_1 \rangle, \langle u_2, v_2 \rangle, \dots, \langle u_m, v_m \rangle\} = \langle U, V \rangle \quad \text{ec. 2.12}$$

donde los valores  $u_i \in U$  de manera típica son vectores multidimensionales  $u_i = \{w_1, w_2, \dots, w_n\}$  y los  $w_k$  pueden ser valores reales o discretos denominados *atributos* o *características*. Se tiene como propósito inferir una función o relación  $R$  donde

$$R : U \rightarrow V \quad \text{ec. 2.13}$$

donde los valores de  $V$  se encuentran en un conjunto finito de clases  $C = \{C_1, \dots, C_k\}$  que caracterizan los datos dados (Torres García, et al., 2013).

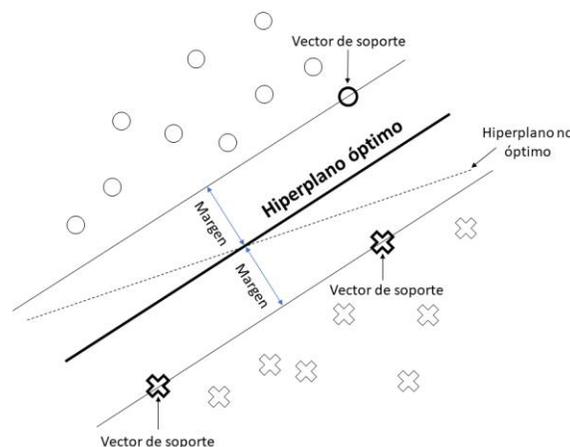
En este trabajo se utilizaron dos clasificadores de distinta naturaleza, la Máquina de Soporte Vectorial (SVM) y Bosques Aleatorios (RF).



### 2.5.1 Máquina de Vectores de Soporte (SVM)

La naturaleza de este clasificador consiste esencialmente en utilizar un hiperplano que separe los elementos de una clase y otra. La intención es encontrar aquel hiperplano que tenga el margen máximo de separación, el cual depende y es sensible a los vectores de soporte (James , et al., 2021).

En ocasiones, emplear fronteras de decisión lineal para la clasificación no genera los mejores resultados, incluso la clasificación resulta imposible o muy poco favorable. Para ello, se aumenta la complejidad del clasificador, modificando la función del núcleo (*kernel*), creando límites no lineales de decisión, asignando implícitamente (mapeo) los datos a otro espacio, por lo general, de dimensión mayor.



**Figura 2.11** Selección del hiperplano óptimo en SVM

Para la clasificación multiclase, SVM emplea la estrategia OVA (*One Versus All*), que consiste en etiquetar la n-ésima clase como positiva y al resto como negativas, repitiendo el proceso iterando la n-ésima clase positiva con todas las clases.

En (Torres García, et al., 2013), (Auria & Moro, 2008) se mencionan algunas ventajas de la máquina de vectores de soporte (SVM, del inglés *Support Vector Machine*) como la insensibilidad al sobre-entrenamiento, requerir que pocos hiper-parámetros se definan

manualmente y proporciona una buena generalización, es decir, las SVM's pueden ser robustas, aunque la muestra de entrenamiento tenga sesgo (outliers).

### 2.5.2 *Random Forest (RF)*

El clasificador de bosques aleatorios (RF, del inglés *Random Forest*) son un conjunto de árboles de decisión, los cuales se forman con una muestra aleatoria del conjunto de entrenamiento con reemplazo. Además, cada árbol tendrá una muestra aleatoria de  $m$  predictores del conjunto completo de  $p$  predictores, logrando reducir la correlación entre los árboles de decisión. Cada uno arrojará un voto para una entrada  $u$  brindada y el voto mayoritario, es decir, la variable categórica (clase) con mayor frecuencia, generará la clase pronosticada.

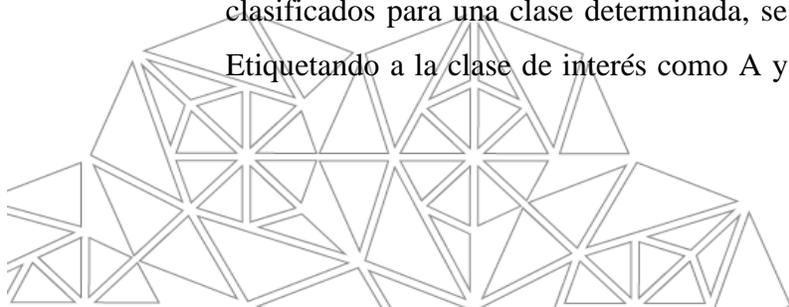
La búsqueda consiste, teniendo en cuenta únicamente los  $m$  predictores seleccionados, la mejor división de los datos de entrenamiento (Moctezuma Pascual, 2017). Para conseguir dicho propósito, se debe implementar una función objetivo que, comúnmente suele ser la entropía o el índice de Gini.

Algunas de las ventajas del RF son la reducción del riesgo de sobreajuste y que facilita la determinación de la relevancia de las características o predictores (James , et al., 2021), (IBM, s.f.), (Amat Rodrigo, 2020).

### 2.5.3 *Medición de la evaluación de los clasificadores*

Al momento de medir el desempeño de un clasificador, el criterio para evaluar es un factor muy relevante. Si bien es cierto que, la exactitud (*accuracy*) es uno de los más utilizados, éste aporta poca o insuficiente información.

Para entender mejor el número de elementos que han sido correcta e incorrectamente clasificados para una clase determinada, se utiliza la matriz de confusión para dos clases. Etiquetando a la clase de interés como A y a la otra como B, se tiene que, los verdaderos



positivos  $VP$  son los elementos etiquetados y clasificados como A, los falsos negativos  $FN$  son los elementos cuya etiqueta es A y son clasificados como B, los falsos positivos  $FP$  son los elementos cuya etiqueta es B y son clasificados como A, y los verdaderos negativos  $VN$  son los elementos etiquetados y clasificados como B.

**Tabla 2.2** Matriz de confusión para la clase A

Clase	Clasificado como	
	A	B
A	VP	FN
B	FP	VN

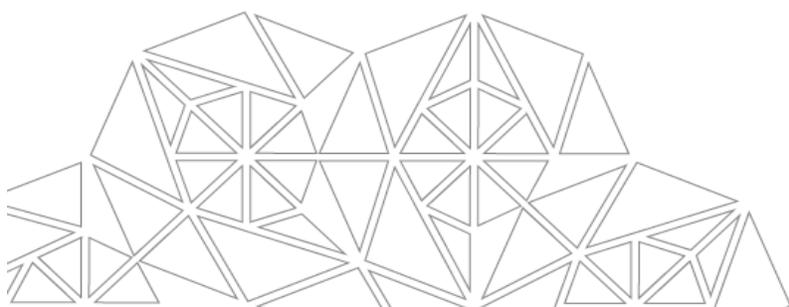
Algunos de los criterios para evaluar el desempeño de la clasificación, surgen de la matriz de confusión anterior, como la exactitud, precisión, sensibilidad o recuerdo y la medida F, siendo esta última una apropiada manera de validar la clasificación con clases desbalanceadas (Moctezuma Pascual, 2017).

*Exactitud*: fracción de la suma de las observaciones clasificadas correctamente tanto positivas como negativas, entre el número total de observaciones.

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN} \quad \text{ec. 2.14}$$

*Precisión*: fracción de observaciones clasificadas correctamente como positivas, entre todas las observaciones clasificadas como positivas.

$$precisión = \frac{VP}{VP + FP} \quad \text{ec. 2.15}$$



*Sensibilidad o recuerdo*: fracción de observaciones clasificadas correctamente como positivas, entre todas las observaciones positivas.

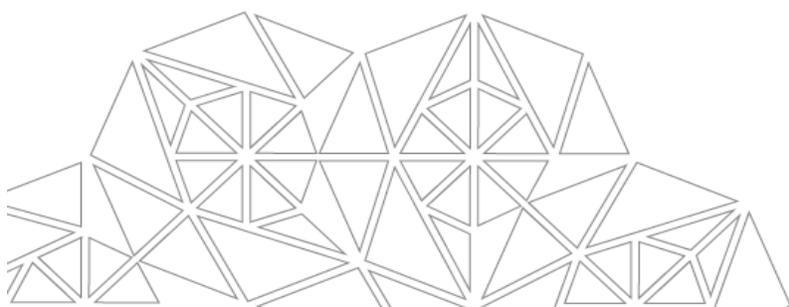
$$recuerdo = \frac{VP}{VP + FN} \quad \text{ec. 2.16}$$

*Medida F*: media armónica de las medidas de precisión y recuerdo. Es una manera de combinar ambas medidas, debido a que, al tratar de mejorar una medida, con frecuencia se deteriora la otra.

$$Medida F = \frac{2 \cdot precisión \cdot recuerdo}{precisión + recuerdo} = \frac{2 \cdot VP}{FP + FN + 2 \cdot VP} \quad \text{ec. 2.17}$$

#### 2.5.4 Validación cruzada con *K*-pliegues

La validación cruzada con *K*-pliegues (*K-fold cross validation*), consiste en dividir de manera aleatoria en *K* submuestras la muestra original, donde una se retiene para utilizarla como prueba y validación del modelo y las *K*-1 restantes se emplean como datos de entrenamiento. El proceso de validación cruzada es repetido *K* veces, debido a que, cada una de las *K* submuestras es utilizada una vez como datos de validación. Después, se pueden promediar o combinar los *K* resultados de los pliegues.



## Capítulo 3

### ESTADO DEL ARTE

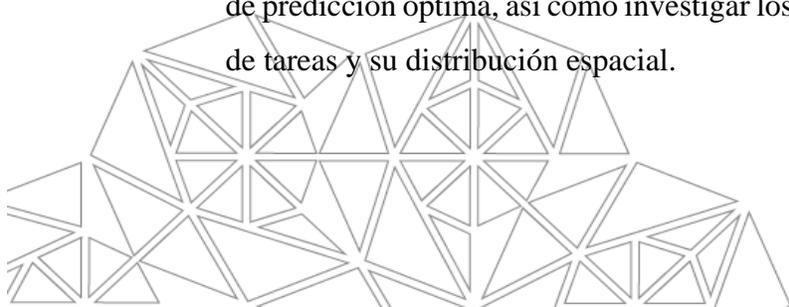
Se puede entender al estado del arte, como la exposición documental de las investigaciones que se encuentran en el horizonte del conocimiento actual, referente al tema de trabajo o investigación a realizar. Es de gran importancia consultar previamente el estado del arte para encaminarse hacia el planteamiento de un problema que no se ha resuelto, aportando un nuevo conocimiento.

Los trabajos de la literatura consultada se presentarán en dos secciones. En la primera sección, se tratará sobre los resultados conseguidos en la clasificación del habla imaginada utilizando señales EEG, mientras que, en la segunda sección, se mencionará de lo obtenido en trabajos con respecto a potenciales evocados visuales.

#### 3.1 Habla imaginada

Anteriormente se mencionó que, el habla imaginada, también conocido como habla no pronunciada, habla interna o habla silente, refiere a pronunciar internamente o de manera imaginada palabras, pero sin la emisión de sonidos o el uso de articuladores. Sin embargo, se han realizado investigaciones bajo diferentes enfoques. En (Moctezuma Pascual, 2017), (Torres García, 2016) se hace referencia de trabajos basados en el enfoque de sílabas, fonemas y vocales, además del enfoque de palabras, incluso el análisis de la distinción de actividad e inactividad cerebral lingüística (Moctezuma, et al., 2017).

Otro tema a considerar, es el número y selección óptima de canales de EEG con el fin de obtener ventajas y/o beneficios. Entre las más relevantes se mencionan en (Torres García, 2016), siendo éstas la reducción de la redundancia de información a tratar, consiguiendo un subconjunto de canales igual de efectivo como el conjunto completo, conseguir la exactitud de predicción óptima, así como investigar los canales de mayor importancia para la distinción de tareas y su distribución espacial.



Para los fines de la presente investigación, la literatura consultada se concentra en las aportaciones de trabajos recientes bajo el enfoque basado en palabras.

### 3.1.1 Enfoque basado en palabras

En (Lee, et al., 2019), siete sujetos realizaron tanto habla imaginada como imaginación visual (VI) de 12 palabras/frases ('ambulance', 'clock', 'hello', 'help me', 'light', 'pain', 'stop', 'thank you', 'toilet', 'TV', 'water', 'yes'). Se mostró que, la presencia de los dos paradigmas, tienen regiones cerebrales distintas entre sí. La exactitud máxima de clasificación de trece clases, incluida la clase de descanso fue del 34.2 % para habla imaginada y del 26.7 % para imaginación visual. Para la extracción de características se utilizó patrón común espacial (CSP, del inglés *Common Spatial Pattern*). Se realizó clasificación multiclase con el clasificador RF, con la estrategia uno contra el resto (OVR, del inglés *One Versus Rest*) puesto que CSP es un método de extracción de características para clasificación binaria.

En (Cooney, et al., 2019), dos diferentes redes neuronales convolucionales (CNN, del inglés *Convolutional Neural Networks*) fueron entrenados con 15 pares de 6 palabras del habla imaginada del idioma español ("arriba", "abajo", "adelante", "atrás", "derecha", "izquierda"), y su rendimiento en comparación con un clasificador de análisis discriminante lineal (LDA, del inglés *Linear Discriminant Analysis*) entrenado, usando un banco de filtros de características de patrones espaciales comunes (FBCSP, del inglés *Filterbank Common Spatial Patterns*). Los clasificadores fueron entrenados usando validación cruzada anidada para permitir optimización de hiperparámetros. Los resultados obtenidos mostraron que las CNN superaron a las FBCSP con una exactitud promedio (accuracy) de 62.37% y 60.88% vs. 57.80%. Además, en este trabajo utilizaron 6 canales del EEG en 15 sujetos.

Por otro lado, en (Cooney, et al., 2020) se entrenó a tres redes neuronales convolucionales distintas, utilizando un enfoque de validación cruzada anidada para la optimización de hiperparámetros. Cada una de las CNN evaluadas fue diseñada



específicamente para la decodificación de EEG. Un conjunto de datos de EEG de voz imaginaria que consta de palabras y vocales facilitaron el entrenamiento en ambos conjuntos de forma independiente. Los resultados de CNN se compararon con tres métodos de aprendizaje automático de referencia: SVM, RF y LDA. Se probaron métodos intra e intersujetos de optimización de hiperparámetros y se analizaron estadísticamente. Las exactitudes obtenidas por las CNN fueron significativamente mayores que los métodos de referencia cuando se entrenaron en ambos conjuntos de datos (palabras: 24.97 % , vocales: 30.00 %) El conjunto de datos se obtuvo de 15 sujetos, siendo el español su idioma nativo, en una edad promedio de 25 años, que realizaban habla imaginada en respuesta a dos grupos (vocales: /a/, /e/, /i/, /o/, /u/; palabras: “arriba”, “abajo”, “derecha”, “izquierda”, “adelante”, “atrás”), utilizando 6 canales EEG (F3, F4, C3, C4, P3, P4).

En (Bakhshali, et al., 2020), se evaluaron matrices de densidad espectral de correntropía (CSD, del inglés *Correntropy Spectral Density*) para señales de EEG, obtenidas de diferentes canales (16 canales: T7, C5, C3, C1, CP5, FC1, FC3, FC5, FC6, FT7, F1, F3, F5, F7, FT8, TP7) seleccionados mediante análisis estadístico, y las distancias entre estas matrices se consideraron medidas para el reconocimiento de voz imaginario mediante distancia Riemanniana (DR). Se utilizó la base de datos “Kara One” que incluye señales de EEG de ocho sujetos durante el discurso imaginario de cuatro palabras en inglés (“gnaw”, “knew”, “pot”, “pat”). El porcentaje de exactitud promedio en la clasificación, utilizando k-vecinos más cercanos (K-NN, del inglés *K-Nearest Neighbors*) fue de 90.25%.

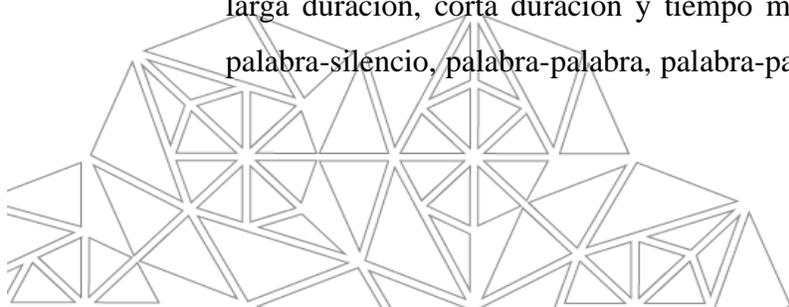
A diferencia de los trabajos anteriores, en (Saha, et al., 2019) la predicción de 11 clases para la clasificación emplea información combinada disponible de 6 categorías fonológicas (por ejemplo: nasal, bilabial), obtenida de una arquitectura novedosa con CNN de 6 capas, para extraer las características espaciales de una matriz de covarianza de los electrodos, CNN temporal (TCNN) de 6 capas para las características temporales, y un codificador automático profundo no supervisado (DAE, del inglés *Deep AutoEncoder*) de 7 capas entrenado con las características heterogéneas producidas por la combinación de la información de CNN y TCNN, consiguiendo un método jerárquico CNN-TCNN-DAE como paso intermedio para clasificar fonemas y palabras, encontrando así la señal discriminativa responsable de la



síntesis natural del habla en lugar de utilizar el formato de data cruda multicanal de alta dimensión de EEG. El conjunto de datos fue obtenido de “Kara One” como en (Bakhshali, et al., 2020), compuesto de base de datos multimodal para el estado del habla articulada, imaginada y basada en estímulos correspondientes a 7 fonético/silábico (/iy/, /piy/, /tiy/, /diy/, /uw/, /m/, /n/), así como 4 palabras (“gnaw”, “knew”, “pot”, “pat”) de 14 sujetos, empleando 64 canales EEG. Consiguieron una exactitud promedio de 83.42% en la clasificación binaria para las 6 categorías fonológicas, y una de 53.36% para las 11 clases, utilizando las características fonológicas, comparado con 28.08% utilizando la data cruda de EEG.

Explorando con los paradigmas del habla imaginada e imaginación visual para BCI's, (Lee, et al., 2020) trabajó con 22 sujetos, 12 palabras/frases (‘ambulance’, ‘clock’, ‘hello’, ‘help me’, ‘light’, ‘pain’, ‘stop’, ‘thank you’, ‘toilet’, ‘TV’, ‘water’, ‘yes’). La clasificación de trece clases, incluida la clase de descanso de los dos paradigmas, fue realizada usando características de patrón común espacial (CSP), usando tres clasificadores básicos: una contracción análisis discriminante lineal regularizado (RLDA shrink, del inglés *shrinkage of Regularized Linear Discriminant Analysis*), RF y SVM. Para SVM se utilizó el kernel *rbf* (función de base radial). Determinaron la región de frecuencia más poderosa que contiene información significativa para cada paradigma, dividiendo los grupos de bandas (BG) en BG1 (delta, 0,5-4 Hz), BG2 (theta, 4-8 Hz), BG3 (mu, 8-12 Hz), BG4 (alfa, 8-13 Hz), BG5 (beta, 13-30 Hz) y BG6 (gamma, 30-125 Hz), consiguiendo una mejor clasificación con BG6 y con el clasificador SVM. La exactitud fue de 39.73% y 40.14% para el habla imaginada y la imaginación visual, respectivamente.

(Asghari Bejestani, et al., 2022) empleó habla imaginada con 6 palabras persas {(bΛΛ), (pΛyin), (chæp), (rΛst), (bæle), (kheir)} equivalentes a las palabras (“arriba”, “abajo”, “izquierda”, “derecha”, “sí”, “no”), respectivamente. Utilizó 19 canales EEG con 5 sujetos. El vector de características fue la amplitud de cada frecuencia en canales seleccionados. Dichas características se obtuvieron mediante la transformada rápida de Fourier (FFT, del inglés *Fast Fourier Transform*). Para la clasificación se utilizó SVM, definiendo tres modos: larga duración, corta duración y tiempo mixto. Las exactitudes promedio para los casos palabra-silencio, palabra-palabra, palabra-palabra-silencio y 7 clases (6 palabras + silencio)



en los tres modos fueron (58%, 60%, 40%, 32%), (96%, 75%, 79%, 55%) y (97%, 97%, 92%, 62%) respectivamente.

**Tabla 3.1** Estado del arte del habla imaginada con el enfoque basado en palabras

<i>Trabajo</i>	<i>Datos</i>	<i>Idioma</i>	<i>Tipo de clasificación</i>	<i>Extracción de características</i>	<i>Exactitud en la clasificación</i>
(Lee, et al., 2019)	12 palabras-frases/7 sujetos/18 canales	Inglés	Multiclase (13)	CSP	RF 34.2%
(Cooney, et al., 2019)	6 palabras/15 sujetos/6 canales	Español	Binaria	FBCSP	LDA (FBCSP) 57.80% CNN 62.37%
(Saha, et al., 2019)	4 palabras-7 fonemas/14 sujetos/64 canales	Inglés	Multiclase (11)	CNN-TCNN-DAE (Características fonológicas)	CNN-TCNN-DAE 53.36%
(Cooney, et al., 2020)	6 palabras/15 sujetos/6 canales	Español	Multiclase (6)	FBCSP	CNN 24.97%
(Bakhshali, et al., 2020)	4 palabras/8 sujetos/16 canales	Inglés	Binaria	CSD DR	K-NN 90.25%
(Lee, et al., 2020)	12 palabras-frases /22 sujetos/64 canales	Inglés	Multiclase (13)	CSP	SVM 39.73%
(Asghari Bejestani, et al., 2022)	6 palabras/5 sujetos/19 canales	Persa	Binaria/Multiclase (7)	FFT	SVM (58%, 60%, 40%, 32%), (96%, 75%, 79%, 55%), (97%, 97%, 92%, 62%)

### 3.2 Potenciales evocados visuales

En (Sayilgan, et al., 2021) se exploró la wavelet madre que mejor se adapte para representar señales SSVEP con fines de clasificación en BCI, con tres características comunes basadas en wavelet (varianza, energía y entropía) para cinco bandas de frecuencias de EEG (delta, theta, alfa, beta y gamma), con el propósito de determinar tres comandos de usuario diferentes. El estudio se repitió para seis funciones de wavelet madre diferentes de uso común (haar, daubechies, symlet, coiflet, biorthogonal y biorthogonal inversa). El conjunto de datos constó de mediciones EEG de las respuestas desencadenadas de las señales



SSVEP de 4 individuos sanos mientras miraban el objetivo parpadeante con tres frecuencias diferentes (6, 8.2 y 10 Hz). La mejor discriminación se obtuvo con una exactitud del 100% y un promedio del 75.85%. La wavelet Haar tuvo el mejor desempeño en la representación de señales SSVEP entre otras todas las wavelets madre adoptadas. Por otro lado, (Sayilgan, et al., 2021) reporta que las tres características de energía, varianza y entropía deben usarse juntas, ya que ninguna de estas características tuvo un rendimiento superior del clasificador por sí sola.

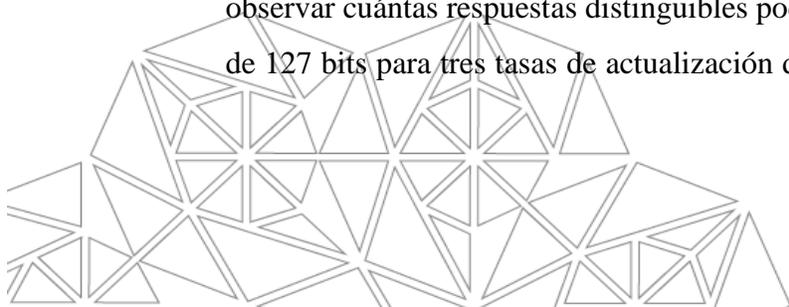
Los SSVEP's se han empleado de manera amplia en sistemas BCI. Los estudios tradicionales utilizaban principalmente SSVEP de baja/media frecuencia como señales de control del sistema. Sin embargo, los parpadeos visuales con frecuencias bajas/medias son incómodos e incluso causan fatiga visual y convulsiones de epilepsia. (Yue, et al., 2020) menciona que, desarrollaron un paradigma BCI para mejorar la relación señal-ruido de SSVEP de alta frecuencia, denominado potencial evocado visual asimétricamente en estado estable (SSaVEP, del inglés *Steady-State asymmetrically Visual Evoked Potential*). Diez caracteres fueron codificados por diez pares de parpadeos asimétricos, cuyas duraciones sólo duraron un segundo y las frecuencias oscilaron entre 31 y 40 Hz con un paso de 1 Hz. Se utilizó la coincidencia de patrones canónicos discriminativos (DCPM, del inglés *Discriminative Canonical Pattern Matching*) para decodificar las señales SSaVEP de alta frecuencia. Cuatro sujetos participaron en el experimento fuera de línea. Como resultado, la exactitud alcanzó un promedio del 87.5% con un pico del 97.1%. La velocidad de transferencia de información en línea simulada alcanzó 87.2 bits/min en promedio y 111.2 bits/min como máximo. Con base a sus resultados, refirieron que, el paradigma SSaVEP de alta frecuencia es un enfoque prometedor para aliviar el malestar causado por los estímulos visuales.

(Zhang, et al., 2019) menciona que, una cuestión clave en BCI es la detección de estados de control intencional (CI) y estados de control no intencionales (NCI) de forma asincrónica. Además, para los sistemas BCI de potencial evocado visual en estado estacionario (SSVEP), existen múltiples estados (sub-estados) dentro del estado CI. Los métodos de reconocimiento existentes se basan en una técnica de umbral, que es difícil de



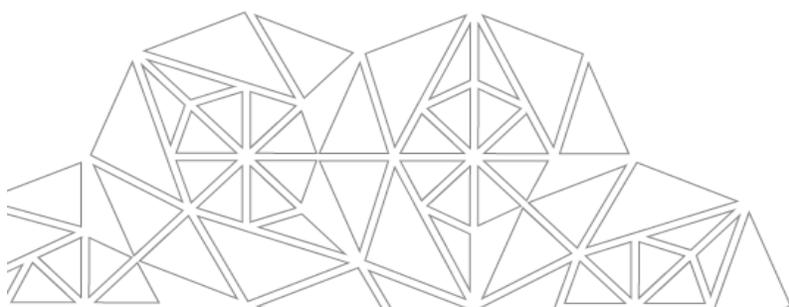
lograr con una alta exactitud, es decir, una alta tasa de verdaderos positivos y una baja tasa de falsos positivos simultáneamente. Para ello, propusieron una novedosa red neuronal convolucional (CNN) para detectar estados CI y NCI en un sistema SSVEP-BCI por primera vez, específicamente, eligieron como paradigma experimental el de potenciales evocados visuales de movimiento en estado estacionario (SSMVEP, del inglés *Steady-State Motion Visual Evoked Potentials*), la cual ha demostrado que induce menos malestar visual. Propusieron dos métodos de procesamiento para la detección de estados CI y NCI. El primero fue usando CNN como clasificador multiclase para discriminar entre todos los estados al mismo tiempo, en el estado CI y NCI, utilizando como datos de entrada para la red, la información transformada en su representación en el dominio de la frecuencia (3 Hz-33 Hz) de las ventanas (épocas) con FFT (FFT-CNN). El segundo fue usar CNN para discriminar entre los estados CI y NCI, y después un análisis de correlación canónica (CCA, del inglés *Canonical Correlation Analysis*) para realizar tareas de clasificación dentro del CI (FFT-CNN-CCA). Ambos métodos lograron un aumento significativo en la exactitud para participantes sanos que consiguieron bajo rendimiento cuando se utilizaron algoritmos tradicionales como el umbral CCA. Además, el método FFT-CNN-CCA logró un mejor rendimiento que el método FFT-CNN de acuerdo con los datos de los pacientes con accidente cerebrovascular.

En (Başaklar, et al., 2019) se investigaron los efectos de las altas tasas de presentación de estímulos sobre las características de las respuestas del potencial evocado visual modulado por código (c-VEP, del inglés *code-modulated Visual Evoked Potential*) junto con el cambio en el rendimiento dependiendo de la tasa de presentación de estímulos mediante la utilización de un BCI deletreador basado en c-VEP. Veinte sujetos participaron en tres experimentos diferentes con tasas de actualización de 60 Hz (E1), 120 Hz (E2) y 240 Hz (E3), la cual se utilizó una secuencia  $m$  de 127 bits de longitud. Para observar el cambio en el contenido de frecuencia de las respuestas de c-VEP, se evaluaron las estimaciones de la densidad espectral de potencia (PSD, del inglés *Power Spectral Density*) de las respuestas de c-VEP. Se aplicó el análisis de componentes principales (PCA, del inglés *Principal Component Analysis*) para observar cuántas respuestas distinguibles podrían evocarse con una secuencia  $m$  de longitud de 127 bits para tres tasas de actualización diferentes. Entre los resultados principales de la



investigación, los valores promedio de la tasa de transferencia de información y exactitud fueron de  $85.87 \text{ bits min}^{-1}$  y 92% para E1,  $94.21 \text{ bits min}^{-1}$  y 97% para E2,  $78.65 \text{ bits min}^{-1}$  y 87% para E3, respectivamente. Además, PCA mostró que, el número de respuestas bien distinguibles disminuye con el aumento de la tasa de actualización.

Para mejorar el rendimiento de BCI, la clave es mejorar la exactitud de la clasificación de las señales de EEG, lo que siempre ha sido el foco de la investigación y el desarrollo. En (Gao, et al., 2020), se propuso un método novedoso que integra una red compleja y un sistema de aprendizaje amplio (BLS, del inglés *Broad Learning System*) para la investigación de una BCI basada en VEP. Se llevaron a cabo experimentos cerebrales sistemáticos basados en VEP para obtener señales de EEG, incluidos SSVEP y SSMVEP. Luego, se emplea el gráfico de visibilidad penetrable limitada (LPVG, del inglés *Limited Penetrable Visibility Graph*) y su secuencia de grados para implementar la extracción de características preliminar. Todas estas características, finalmente se introducen en un BLS para estudiar y clasificar las señales SSVEP y SSMVEP, respectivamente. Los resultados de la clasificación mostraron que su BLS basado en LPVG puede clasificar eficazmente señales de EEG basadas en VEP, con una exactitud de clasificación promedio del 96.22 % para SSVEP y del 74.54 % para SSMVEP. Estos resultados son significativamente mejores que otros métodos de comparación, así como los métodos tradicionales basados en CCA.

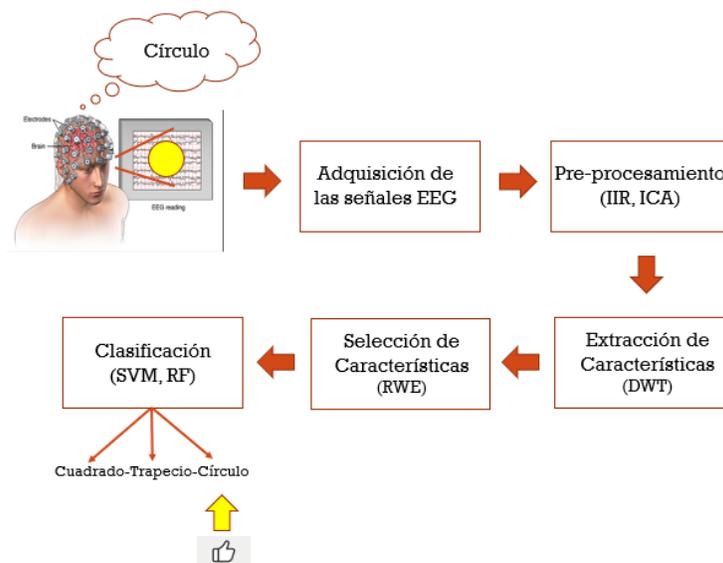


## Capítulo 4

### METODOLOGÍA PROPUESTA

En el capítulo 1, se mencionó que el problema a tratar es analizado con el criterio de clasificación. Dado a lo anterior, se pretende encontrar un método que permita distinguir entre las 3 palabras del vocabulario propuesto cuál fue pronunciada mentalmente por el individuo, mientras se graban las señales EEG cuando es pensada. Como se ilustra en la **Figura 4.1**, se utiliza la señal eléctrica cerebral producida por el sujeto al imaginar la pronunciación de las palabras propuestas. Al ser de naturaleza no estacionaria las señales de EEG que se pretenden clasificar para uso posterior en una BCI, es necesario tener ejemplos. Por ello, es indispensable una etapa para su adquisición, seguida de otras etapas con propósitos específicos que permitan cumplir el objetivo planteado, mismas que se utilizan comúnmente en BCI's.

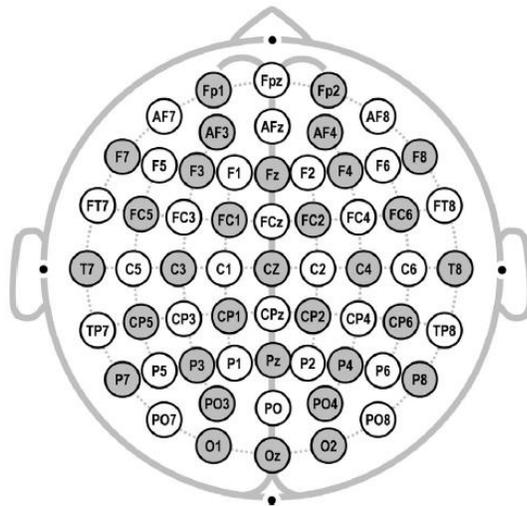
En consecuencia a lo anteriormente explicado, se aplicarán las siguientes etapas: adquisición de las señales, preprocesamiento, caracterización de las neuroseñales, selección de características y clasificación.



**Figura 4.1** Metodología aplicada en la investigación.

#### 4.1 Adquisición de las señales EEG

Se empleó la configuración 32 electrodos de EEG de acuerdo al sistema internacional 10-20 (Fp-1/2, AF-3/4, F-7/3/z/4/8, FC-5/1/2/6, T-7/8, C-3/z/4, CP-5/1/2/6, P-7/3/z/4/8, PO-3/4, O-1/z/2) (Ali, et al., 2016) con una topología *monopolar o de referencia* (Bermúdez Cicchino, 2013), (González Castañeda, 2015). Los electrodos correspondientes a la referencia (Ref) y la tierra (Gnd) se posicionaron en los mastoideos izquierdo y derecho respectivamente, para un total de 34 electrodos utilizados. Además, se verificó que la impedancia de cada electrodo no fuese mayor a 50 k $\Omega$ .



**Figura 4.2** Sistema internacional 10/20 para 32 canales.

(Ali, et al., 2016)

Es conocida la susceptibilidad del EEG al ruido, debido a ello, las condiciones del entorno donde se presentará el experimento deben garantizar en lo posible la eliminación de artefactos. En la cámara de Faraday se mantiene homogénea la luminosidad y la temperatura, elimina en gran medida sonidos provenientes del exterior, así como algunos campos electromagnéticos no deseados provenientes de dispositivos electrónicos, entre otros artefactos.

En (Torres García, et al., 2011) se hace referencia a las frecuencias de las señales EEG las cuales llegan hasta 60 Hz. De acuerdo al teorema de muestreo (*Teorema de Nyquist*), la

exacta reconstrucción de una señal continua de manera matemática es posible si la frecuencia de muestreo  $f_s$  es superior al doble de la frecuencia más alta de la señal analógica  $F_{\text{máx}}$ , es decir,  $f_s \geq 2F_{\text{máx}}$ . A  $2F_{\text{máx}}$  se le conoce como la velocidad de Nyquist y a  $F_{\text{máx}}$  como frecuencia de Nyquist (Oppenheim, et al., 1998). Debido a lo anterior, es importante que la frecuencia de muestreo del equipo EEG a utilizar para la adquisición de las neuroseñales sea superior a  $2(60 \text{ Hz})$ , o sea, 120 Hz.

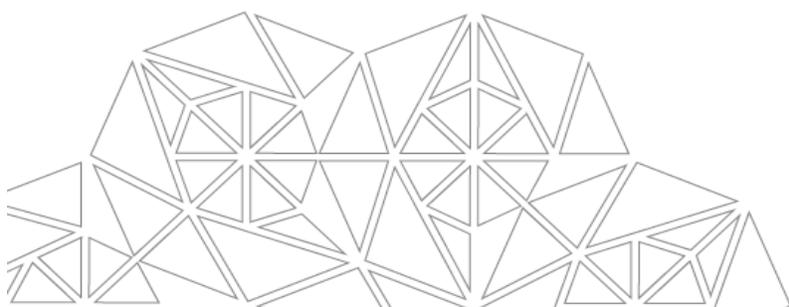
## 4.2 Preprocesamiento

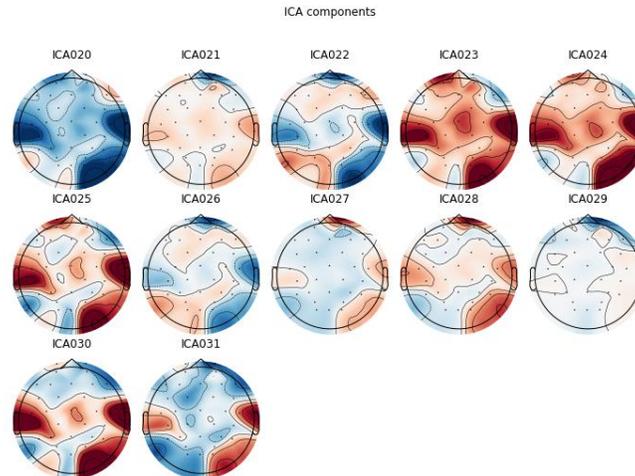
Se puede decir que el preprocesamiento es la etapa donde se elimina lo que no deseamos de las señales, como artefactos o ruido en general, para quedarnos con la información que nos interesa, en este caso, las neuroseñales. Podemos distinguir tres pasos: *cropping*, *aplicación del filtro pasa-banda* y *eliminación de artefactos*.

i. *Cropping*. - Se recortan las partes que no incluye información de nuestro interés, antes del primer trigger y después del último trigger en cada una de las grabaciones de las neuroseñales adquiridas.

ii. *Filtro pasa-banda*. - En este paso, se elige y utiliza el filtro que permitirá pasar sólo el rango de banda de frecuencias de interés, eliminando aquellas que no se encuentren en dicho rango. Se aplicó el filtro *IIR (Infinite Impulse Response)* como filtro pasa-banda del rango de nuestro interés (0 a 32 Hz).

iii. *Eliminación de artefactos*. – Se aplicó la herramienta ICA en la señal obtenida en cada bloque para la eliminación de artefactos, como latidos del corazón, parpadeos y fuentes no reconocidas como posibles neuroseñales (componentes ICA considerados como ruido), con un paquete de Python de código abierto llamado MNE (*Minimum-Norm Estimation*) (versión 1.3.0). Aunque anteriormente se aplicó un filtro pasa-bandas, también se empleó como complemento un filtro notch, para eliminar la frecuencia de 60 Hz proveniente del ruido ocasionado por la línea de corriente alterna (Blanco-Díaz & Ruiz-Olaya, 2019).





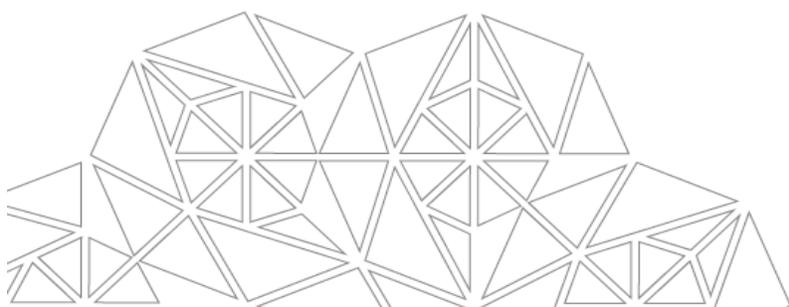
**Figura 4.3** Mapas topográficos de componentes ICA obtenidos con MNE. El número de ICA's coincide con el número de canales EEG utilizados.

#### 4.3 Caracterización de las señales

Se asume que, después de la etapa de preprocesamiento, se han conseguido las señales de nuestro interés, las neuroseñales. El siguiente paso es caracterizarlas. Para esta etapa, se utilizó el software Matlab R2023a.

Cada ventana de 2.4 s donde se incluye la información del VEP y el habla imaginada debe tener la cantidad de 1201 datos, debido a la frecuencia de muestreo más un dato que genera el trigger. En el caso de la ventana de 1.7 s donde solamente se incluye la información del habla imaginada hay 851 debido a lo explicado anteriormente.

En este trabajo, se empleó como wavelet madre la Daubechies de séptimo orden (db7) y la Biorthogonal 3.7 (bior 3.7) con 5 niveles de descomposición para obtener los coeficientes wavelets en cada ventana. La interpretación de la caracterización en cada nivel de descomposición está con base a los rangos de frecuencia.



**Tabla 4.1** Niveles de descomposición de la DWT

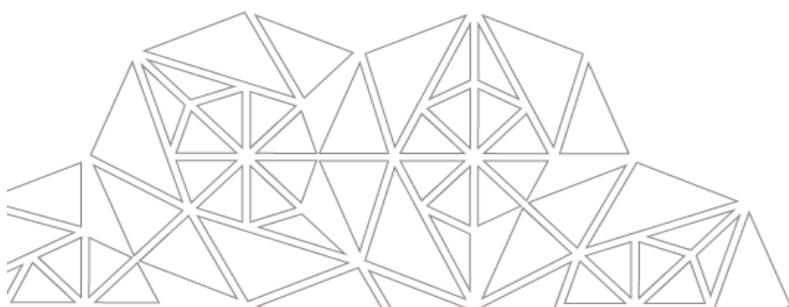
NIVEL	Bandas de frecuencias (Hz)
A5	0-2
D5	2-4
D4	4-8
D3	8-16
D2	16-32
D1	32-64

Con lo anterior, conseguimos un vector con el mismo número de coeficientes wavelets en cada ventana de cada canal.

Nota: Anteriormente, se mencionó que se aplicó el filtro *IIR* como filtro pasa-banda del rango de nuestro interés (0 a 32 Hz). Podría decirse que no surgirían coeficientes en el rango 32-64 Hz, sin embargo, se tomó en consideración por la posibilidad de que el filtro dejase algún remanente.

#### 4.3.1 Normalización con la energía relativa wavelet (*RWE*)

Para normalizar a los coeficientes wavelet se empleó la *energía relativa wavelet* (*RWE*). Se representó cada ventana por canal (tanto para las que tienen el VEP como las que sólo tienen el habla imaginada) con 6 valores de energía relativa wavelet ( $RWE_1 - RWE_6$ ) asociados con los coeficientes wavelet ( $D1 - D5$  y  $A5$ ) respectivamente. Por cada sujeto se obtienen 630 ventanas (210 por figura geométrica: círculo, cuadrado, trapecio), tanto para el habla imaginada, como de aquellas que tienen incluido el VEP, al tener 32 canales, da origen a 20,160 muestras por cada conjunto de datos, pudiéndolo entender como una matriz  $A$  de dimensión  $A_{20,160 \times 6}$ .

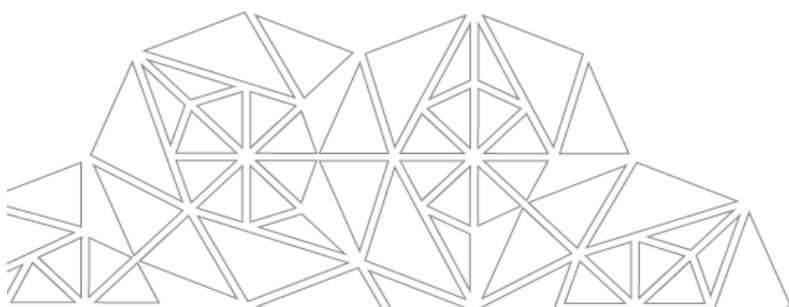


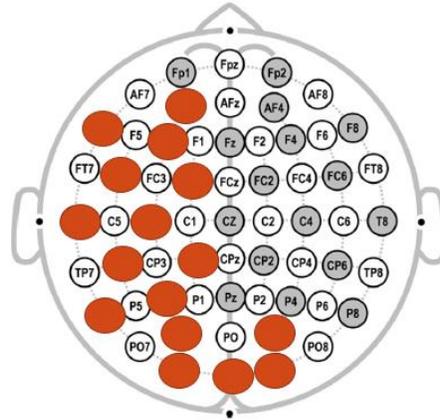
$$\begin{array}{c}
 \text{Matriz } A \\
 \left. \begin{array}{l}
 \text{Figura}_1 \\
 \vdots \\
 \text{Figura}_2 \\
 \vdots \\
 \text{Figura}_3
 \end{array} \right\}
 \begin{array}{l}
 \left. \begin{array}{l}
 \text{ventana}_1 \\
 \vdots \\
 \text{ventana}_{210} \\
 \vdots \\
 \text{ventana}_{211} \\
 \vdots \\
 \text{ventana}_{420} \\
 \vdots \\
 \text{ventana}_{421} \\
 \vdots \\
 \text{ventana}_{630}
 \end{array} \right\}
 \begin{array}{l}
 \left\{ \begin{array}{l}
 RWE_{1,1} \quad RWE_{1,2} \quad RWE_{1,3} \quad RWE_{1,4} \quad RWE_{1,5} \quad RWE_{1,6} \\
 RWE_{2,1} \quad RWE_{2,2} \quad RWE_{2,3} \quad RWE_{2,4} \quad RWE_{2,5} \quad RWE_{2,6} \\
 \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1}
 \end{array} \right. \\
 \vdots \\
 \left\{ \begin{array}{l}
 RWE_{1,1} \quad RWE_{1,2} \quad RWE_{1,3} \quad RWE_{1,4} \quad RWE_{1,5} \quad RWE_{1,6} \\
 RWE_{2,1} \quad RWE_{2,2} \quad RWE_{2,3} \quad RWE_{2,4} \quad RWE_{2,5} \quad RWE_{2,6} \\
 \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1}
 \end{array} \right. \\
 \vdots \\
 \left\{ \begin{array}{l}
 RWE_{1,1} \quad RWE_{1,2} \quad RWE_{1,3} \quad RWE_{1,4} \quad RWE_{1,5} \quad RWE_{1,6} \\
 RWE_{2,1} \quad RWE_{2,2} \quad RWE_{2,3} \quad RWE_{2,4} \quad RWE_{2,5} \quad RWE_{2,6} \\
 \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1}
 \end{array} \right. \\
 \vdots \\
 \left\{ \begin{array}{l}
 RWE_{1,1} \quad RWE_{1,2} \quad RWE_{1,3} \quad RWE_{1,4} \quad RWE_{1,5} \quad RWE_{1,6} \\
 RWE_{2,1} \quad RWE_{2,2} \quad RWE_{2,3} \quad RWE_{2,4} \quad RWE_{2,5} \quad RWE_{2,6} \\
 \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1}
 \end{array} \right. \\
 \vdots \\
 \left\{ \begin{array}{l}
 RWE_{1,1} \quad RWE_{1,2} \quad RWE_{1,3} \quad RWE_{1,4} \quad RWE_{1,5} \quad RWE_{1,6} \\
 RWE_{2,1} \quad RWE_{2,2} \quad RWE_{2,3} \quad RWE_{2,4} \quad RWE_{2,5} \quad RWE_{2,6} \\
 \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1} \quad RWE_{32,1}
 \end{array} \right.
 \end{array}
 \end{array}
 \end{array}$$

**Figura 4.4** Estructura de la matriz  $A_{20,160 \times 6}$

### 4.3.2 Selección de canales del EEG

De los 32 canales, se seleccionaron 16, aquellos canales más cercanos a las áreas cerebrales de Broca, Wernicke y corteza visual ( $AF3$ ,  $F7$ ,  $F3$ ,  $FC5$ ,  $FC1$ ,  $T7$ ,  $C3$ ,  $P7$ ,  $P3$ ,  $CP5$ ,  $CP1$ ,  $PO3$ ,  $PO4$ ,  $O1$ ,  $Oz$ ,  $O2$ ).





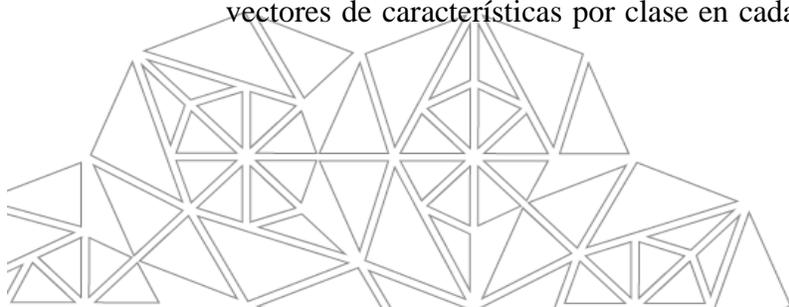
**Figura 4.5** 16 canales (círculos rojos) asociados a las áreas cerebrales de Broca, Wernicke y corteza visual.

#### 4.4 Selección de características

En (Torres García, et al., 2013) se sugirió eliminar el valor de la energía wavelet obtenido de los coeficientes en el nivel de detalle (D1) debido a que están relacionados con frecuencias superiores a 25 Hz, más vinculadas con actividad electromiográfica (EMG).

##### 4.4.1 Eliminación de valores RWE

Al estar asociado  $RWE_1$  con el coeficiente  $D1$ , se sigue también con su eliminación. Por lo tanto, el subconjunto seleccionado de características está compuesto de los valores de energía relativa wavelet  $RWE_2 - RWE_6$  asociados con los coeficientes wavelet de detalle ( $D2 - D5$ ) y de aproximación  $A5$  respectivamente. Como consecuencia de lo anterior, se reduce la dimensión de los vectores de características y a su vez, el impacto del *problema de la dimensionalidad* (Guerrero, 2016), con lo que cada una de las ventanas de cada uno de los canales queda representada con los valores  $RWE_2$  a  $RWE_6$ , es decir, 5 valores de energía relativa wavelet, concatenados en el orden  $AF3, Fp1, Fp2, AF4, F7, F3, Fz, F4, F8, FC5, FC1, FC2, T7, C3, Cz, C4, FC6, P7, P3, Pz, P4, P8, CP5, CP1, CP2, PO3, PO4, CP6, O1, Oz, O2, T8$ , obteniendo un vector de características de 160 elementos por época del mismo intervalo de tiempo. Cada vector viene etiquetado con su respectiva clase, teniendo 210 vectores de características por clase en cada sujeto. Por ejemplo, la ecuación ec. 4.1 es un



vector de características para la clase Cir-IS (ventana del habla imaginada de la palabra círculo).

$$(RWE_{1_{AF3}} \ RWE_{2_{AF3}} \ \dots \ RWE_{5_{AF3}} \ RWE_{6_{AF3}} \ \dots \ RWE_{155_{T8}} \ RWE_{156_{T8}} \ \dots \ RWE_{159_{T8}} \ RWE_{160_{T8}} | \text{Cir} - \text{IS}) \quad \text{ec. 4.1}$$

Análogamente, se obtuvo el vector de características para 16 canales, es decir, 5 valores de energía relativa wavelet, concatenados en el orden  $AF3$ ,  $F7$ ,  $F3$ ,  $FC5$ ,  $FC1$ ,  $T7$ ,  $C3$ ,  $P7$ ,  $P3$ ,  $CP5$ ,  $CP1$ ,  $PO3$ ,  $PO4$ ,  $O1$ ,  $Oz$ ,  $O2$ , obteniendo un vector de características de 80 elementos por época del mismo intervalo de tiempo (véase ec. 4.2).

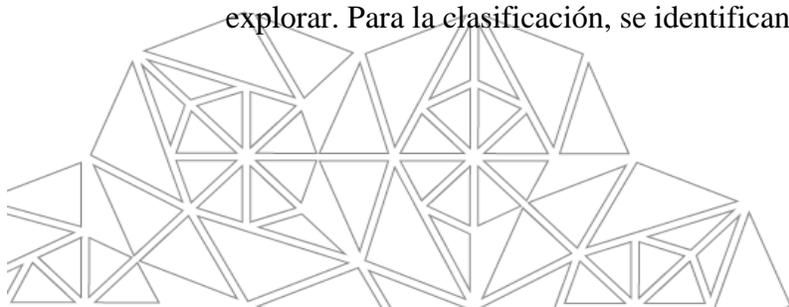
$$(RWE_{1_{AF3}} \ RWE_{2_{AF3}} \ \dots \ RWE_{5_{AF3}} \ RWE_{6_{AF3}} \ \dots \ RWE_{75_{Oz}} \ RWE_{76_{Oz}} \ \dots \ RWE_{79_{Oz}} \ RWE_{80_{Oz}} | \text{Cir} - \text{IS}) \quad \text{ec. 4.2}$$

#### 4.4.2 Eliminación de outliers

La eliminación de valores atípicos se realiza en cada conjunto de vectores de los valores  $RWE$  correspondientes a una clase por cada sujeto. Es decir, aquellos vectores de 160 elementos que tengan un comportamiento atípico (por ejemplo, de la clase Cir-IS; véase **Tabla 4.2**), se eliminan antes, para limpiar el conjunto de datos que ingresará a la clasificación.

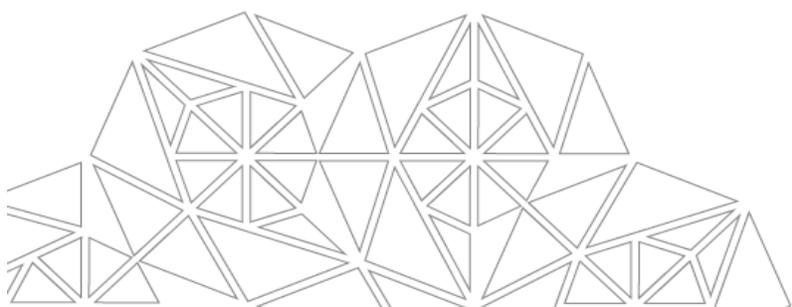
#### 4.5 Clasificación de las neuroseñales

Se utilizaron los clasificadores RF y SVM. Se aplicó validación cruzada de 10 pliegues, así como el proceso de *tuning* para la búsqueda de la mejor combinación de hiperparámetros en cada clasificador. En el caso de RF, se exploraron los hiperparámetros *criterion* (criterio para decidir una división en los árboles de decisión): “Gini”, “entropy”, *max\_samples* (tamaño de la submuestra del conjunto de datos para entrenar cada árbol): “0.5”, “0.7” y *n\_estimators* (número de árboles): “50”, “100”, “200”, obteniendo 12 combinaciones posibles para explorar. Para el clasificador SVM se exploraron los hiperparámetros *kernel*: “rbf”, “sigmoid”, *gamma*: “10”, “100”, consiguiendo 4 combinaciones posibles para explorar. Para la clasificación, se identifican las siguientes clases:



**Tabla 4.2** Clases para la clasificación

<i>Clase</i>	<i>Descripción</i>
1	Círculo/Habla Imaginada (Cir- <b>IS</b> )
2	Cuadrado/Habla Imaginada ( <b>Squ-IS</b> )
3	Trapezio/Habla Imaginada ( <b>Tra-IS</b> )
4	Círculo/Habla Imaginada + VEP (Cir + <b>VEP</b> )
5	Cuadrado/Habla Imaginada + VEP ( <b>Squ +VEP</b> )
6	Trapezio/Habla Imaginada + VEP ( <b>Tra +VEP</b> )



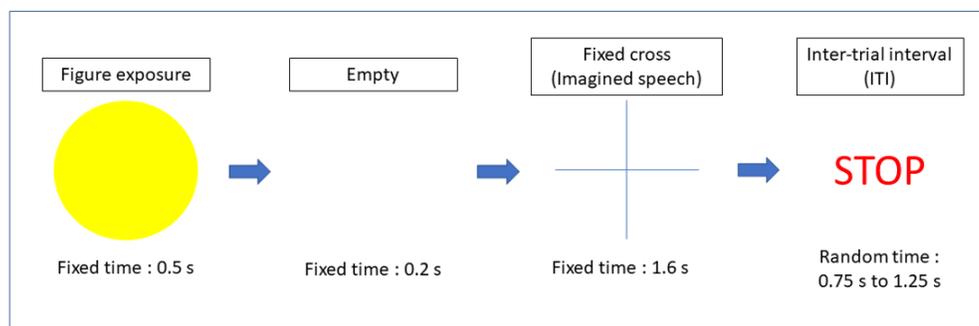
## Capítulo 5

### EXPERIMENTACIÓN Y RESULTADOS

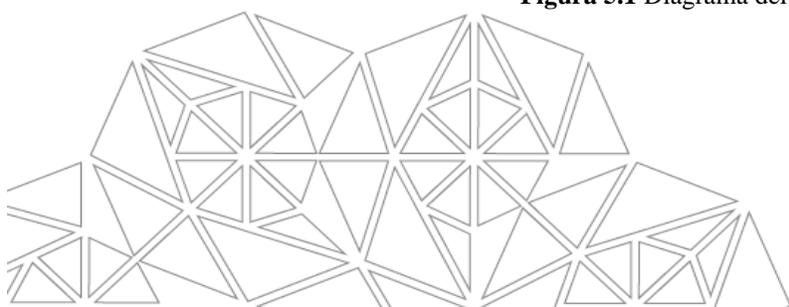
En este capítulo se muestran los resultados obtenidos de la experimentación. Se evaluaron las clasificaciones con la exactitud y se compararon las evaluaciones de los distintos clasificadores utilizados para determinar si hay diferencias significativas entre las ventanas del habla imaginada (IS) y de aquellas que tienen incluida la información del estímulo visual (+VEP), considerando factores como la wavelet madre y el número de canales.

#### 5.1 Diseño experimental

Para distinguir las neuroseñales generadas por los estímulos de las figuras y el habla imaginada, es importante que se presenten en ventanas distintas en cada trial. El tiempo de la presentación de la figura geométrica es fijo, con una duración de 0.5 segundos, el tiempo de la ventana del habla imaginada es de 1.6 segundos mientras aparece una cruz fija en medio de la pantalla, después se mostrará la palabra **STOP** con un tiempo aleatorio en un rango de 0.75 a 1.25 segundos para indicar el final del trial. Entre la presentación del estímulo visual y la realización del habla imaginada habrá un espacio de tiempo fijo de 0.2 segundos (como se muestra en la **Figura 5.1**). El tiempo medio de la duración del trial es de 3.3 segundos.



**Figura 5.1** Diagrama del diseño experimental del trial.

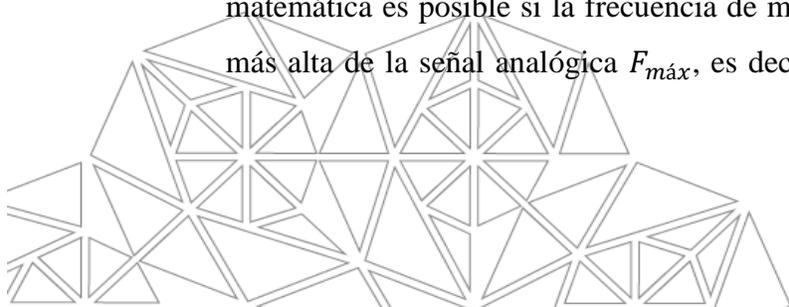


El experimento para cada sujeto se dividió en tres bloques. Por bloque se presentó 70 veces cada figura, obteniendo 210 trials en un tiempo medio de 11.55 minutos, para un total de 34.55 minutos por experimento. En cada trial se consiguen dos ventanas, una que contiene el habla imaginada de 1.7 s y otra cuyo límite inferior es el inicio del estímulo visual y límite superior el final del habla imaginada de 2.4 s (ambas ventanas tienen incluido 100 ms que permite referenciar y filtrar ruido previo en cada época original de longitud 1.6 s y 2.3 s respectivamente), para un total de 1260 ventanas por sujeto, es decir, 630 del habla imaginada y 630 que incluya el VEP. En cada trial aparecerá una de las tres figuras de manera aleatoria, así como la distancia entre trial y trial (0.75 a 1.25 s), esto con el propósito de no generar patrones o secuencias predecibles, para que el estímulo sea auténtico en cada sujeto.

Se consideró el mismo tamaño y color (amarillo) para las figuras, así como en sus respectivos nombres en el idioma español que contengan el mismo número de sílabas, para que solamente la forma de cada figura sea la característica que pudiera distinguirse (y no otra, como el tamaño, color o complejidad del nombre). Se fijó una distancia entre el sujeto y la pantalla donde aparecían los estímulos para homogeneizar las condiciones experimentales. El software que se utilizó en el diseño experimental fue *Psychopy* versión 2022.2.4.

El equipo de EEG utilizado es de la marca Neuroscan Neuvo, que permite hasta 64 canales y que contiene un amplificador de las señales adquiridas. También se empleó el dispositivo StimTraker de la compañía Cedrus, para garantizar que los pulsos sean enviados en el instante en que se desea que aparezcan los marcadores que delimitarán las ventanas donde aparecerá la información tanto del VEP como del habla imaginada.

Con el software CURRY 8.0.6 se pudo visualizar las señales que fueron recepcionadas por el amplificador, así como grabar las señales del EEG de todas las sesiones con los sujetos. En él, se configuró el *sample rate* a 500Hz y como las frecuencias de las señales EEG llegan hasta 60 Hz (Torres García, et al., 2011) se cumple el *teorema de muestreo* (Teorema de Nyquist), que menciona que la exacta reconstrucción de una señal continua de manera matemática es posible si la frecuencia de muestreo  $f_s$  es superior al doble de la frecuencia más alta de la señal analógica  $F_{m\acute{a}x}$ , es decir,  $f_s \geq 2F_{m\acute{a}x}$ . A  $2F_{m\acute{a}x}$  se le conoce como la



velocidad de Nyquist y a  $F_{m\acute{a}x}$  como frecuencia de Nyquist (Oppenheim, et al., 1998). Debido a lo anterior, es importante que la frecuencia de muestreo del equipo EEG a utilizar para la adquisición de las neuroseñales sea superior a  $2(60 \text{ Hz})$ , o sea, 120 Hz.

## 5.2 Experimentación con 8 individuos

Previo a la exposición de los tres bloques, se explicó a detalle a los sujetos lo que debían hacer durante el experimento, indicando que después de que apareciera la figura, en el tiempo de presentación de la cruz fija, debían imaginar la pronunciación del nombre de la misma. Además, realizaron un breve entrenamiento como práctica para la velocidad de la dicción de los nombres de las figuras y ajustarse al tiempo establecido. Durante los bloques, sostenían un batelenguas con la boca para evitar el artefacto de los micromovimientos de la mandíbula (EMG).

Se procuró que la distancia entre el voluntario y el monitor donde se presenta el diseño experimental fuera constante. La información adquirida fue de 8 voluntarios sanos (3 mujeres, 5 hombres) en un rango de edad de  $30 \pm 10$  años, cuya lengua madre es el español.

A los voluntarios se les sugirió de manera previa al día que les correspondía pasar al laboratorio para la adquisición de las señales, que consideraran los siguientes aspectos para no afectar los resultados:

- ✓ No estar desvelados.
- ✓ No estar bajo los efectos de estimulantes o estupefacientes.
- ✓ No asistir con químicos o sustancias en el cabello como gel, shampoo, acondicionador o productos parecidos, puesto que pueden afectar la conductividad de los electrodos.

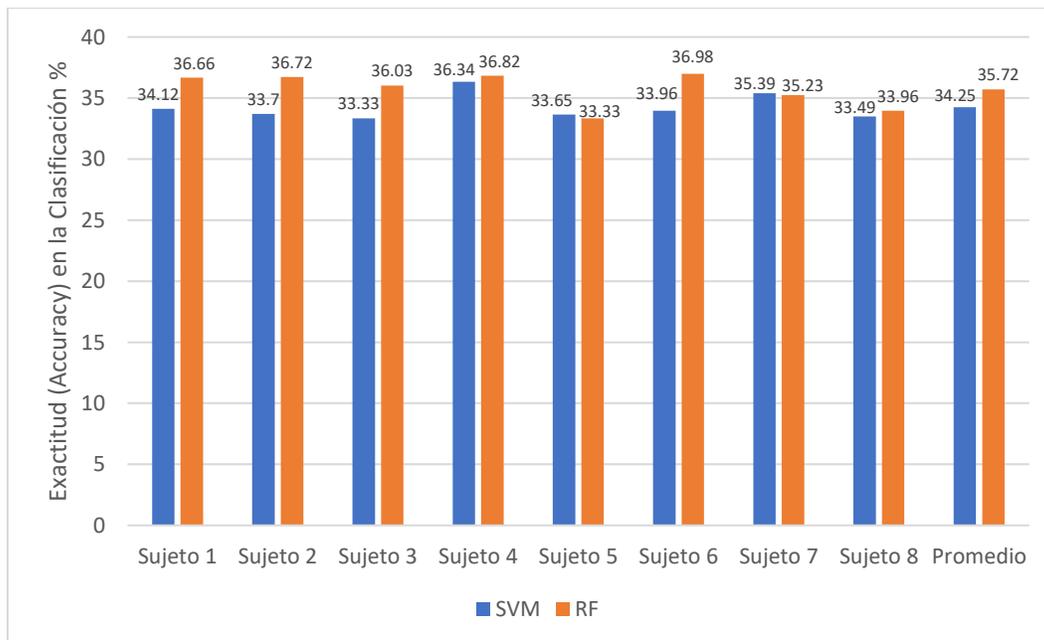
El experimento se realizó en una cámara que actúa como una jaula de Faraday del laboratorio de electroencefalografía del Instituto de Neurobiología de la UNAM campus Juriquilla, en la ciudad de Querétaro, México.



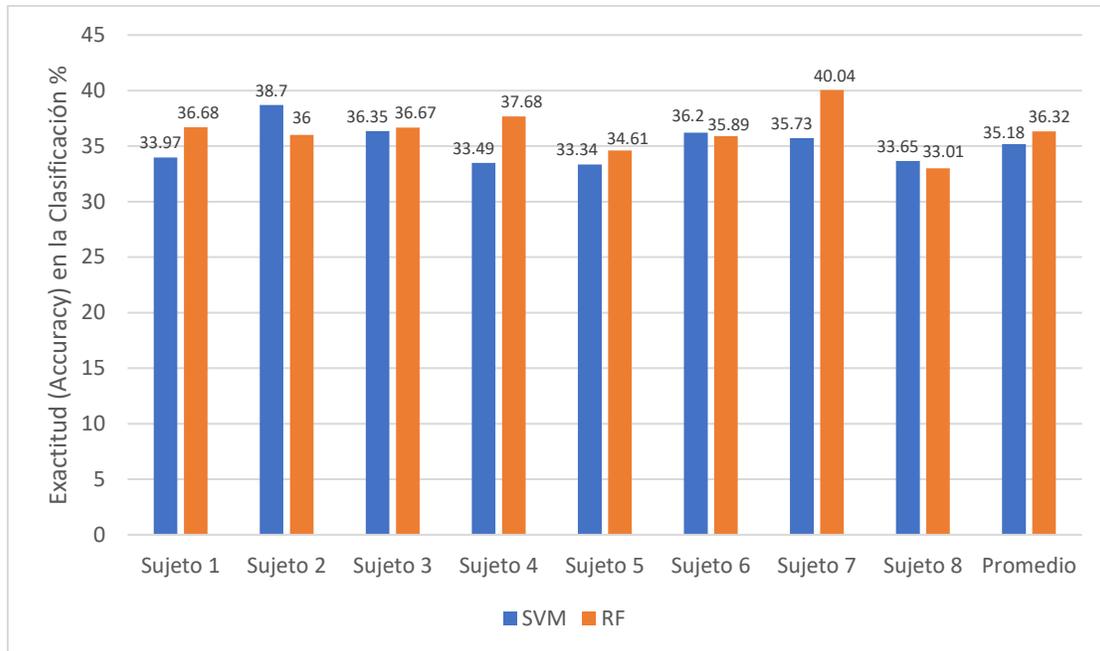
### 5.2.1 32 canales de EEG

La exactitud promedio conseguida por los dos clasificadores con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre bior 3.7 para **IS** y **+VEP** se presenta en las Figuras 5.2 y 5.3 respectivamente, donde **IS** se relaciona con las ventanas que tienen la información del habla imaginada y (**+VEP**) con aquellas que tienen la información tanto del habla imaginada como del estímulo visual. La mejor combinación de hiperparámetros por sujeto con 32 canales y bior 3.7 en ambos clasificadores se pueden consultar en los Anexos 1 y 2.

Se obtuvo una exactitud promedio con **IS** de 34.25% y 35.72% para SVM y RF respectivamente, en el caso de **+VEP** se consiguió 35.18 % y 36.32% para SVM y RF respectivamente. Es claro notar que, el clasificador RF tuvo mejores resultados con respecto a SVM en ambos tipos de ventanas y que, hay una diferencia entre la exactitud de **IS** y **+VEP**. Ahora, ¿es significativa dicha diferencia? Se tratará de responder esa pregunta en la sección de la validación de la hipótesis.



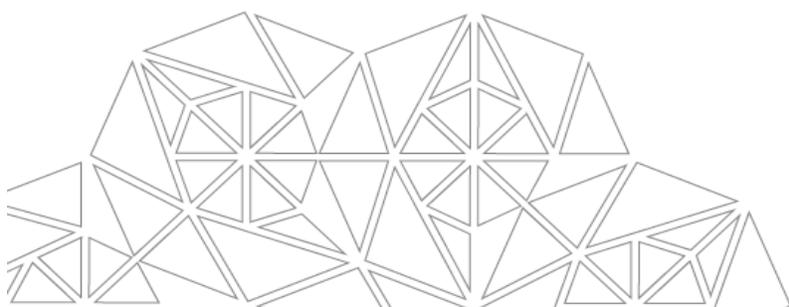
**Figura 5.2** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre bior 3.7 para **IS**.

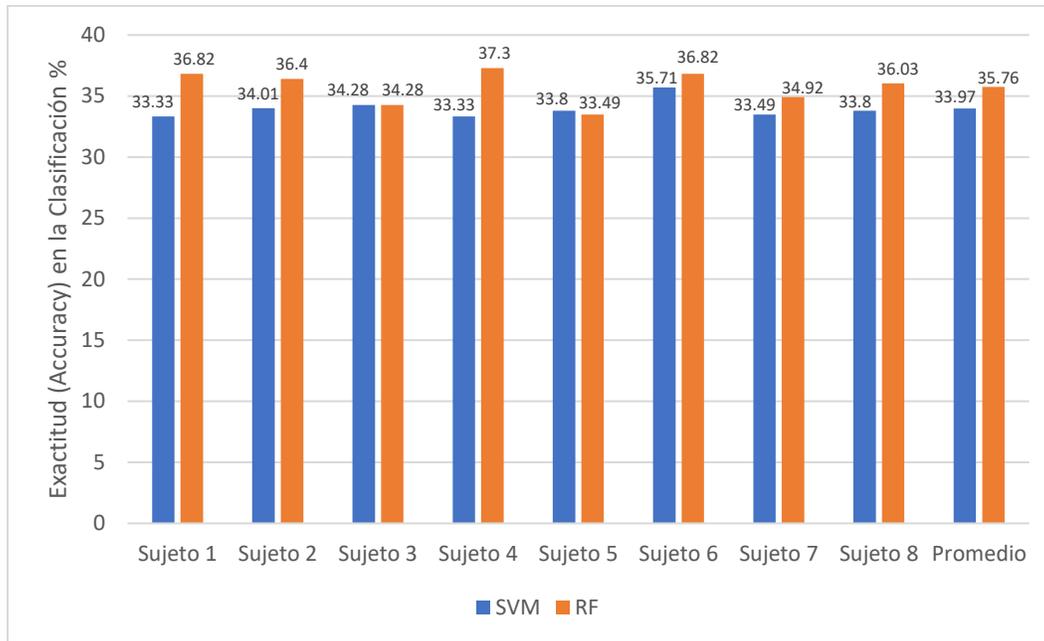


**Figura 5.3** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre bior 3.7 para +VEP.

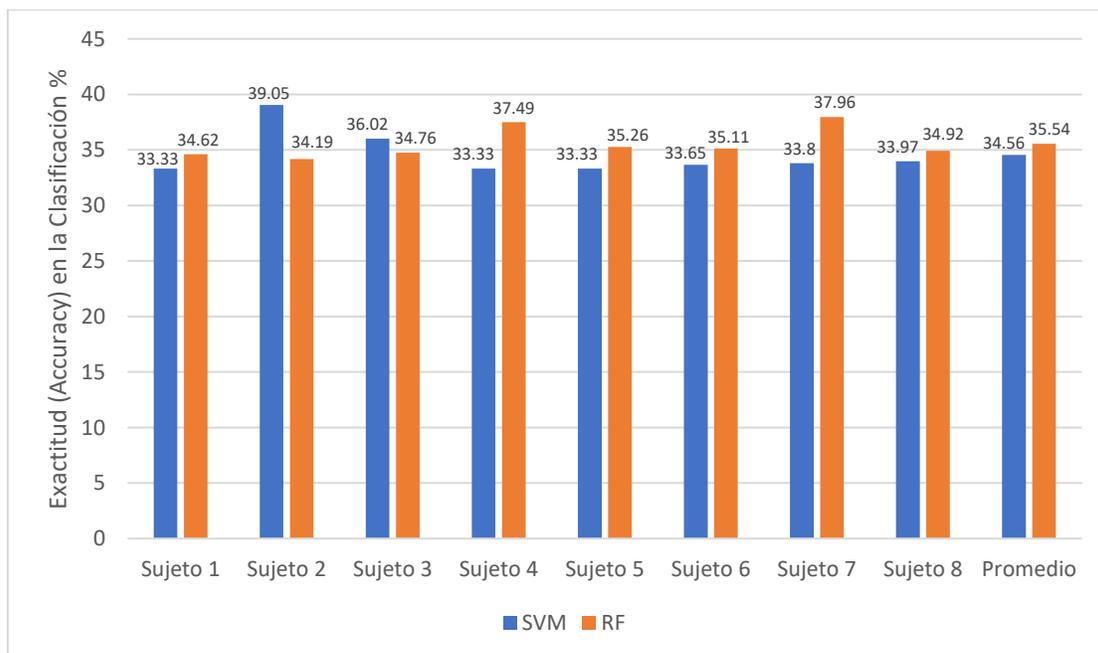
La exactitud promedio conseguida por los dos clasificadores con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre db7 para **IS** y **+VEP** se presenta en las Figuras 5.4 y 5.5 respectivamente. La mejor combinación de hiperparámetros por sujeto con 32 canales y db7 en ambos clasificadores se pueden consultar en los Anexos 3 y 4.

La exactitud promedio con **IS** fue de 33.97% y 35.76% para SVM y RF respectivamente, en el caso de **+VEP** se consiguió 34.56 % y 35.54% para SVM y RF respectivamente. Nuevamente, el clasificador RF tuvo mejores resultados con respecto a SVM en ambos tipos de ventanas y hay una diferencia entre la exactitud de **IS** y **+VEP**. ¿Es significativa dicha diferencia? ¿Es significativa la diferencia entre la exactitud obtenida en ambas ondículas? Se tratará de responder estas preguntas en la sección de la validación de la hipótesis.





**Figura 5.4** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre db7 para IS.



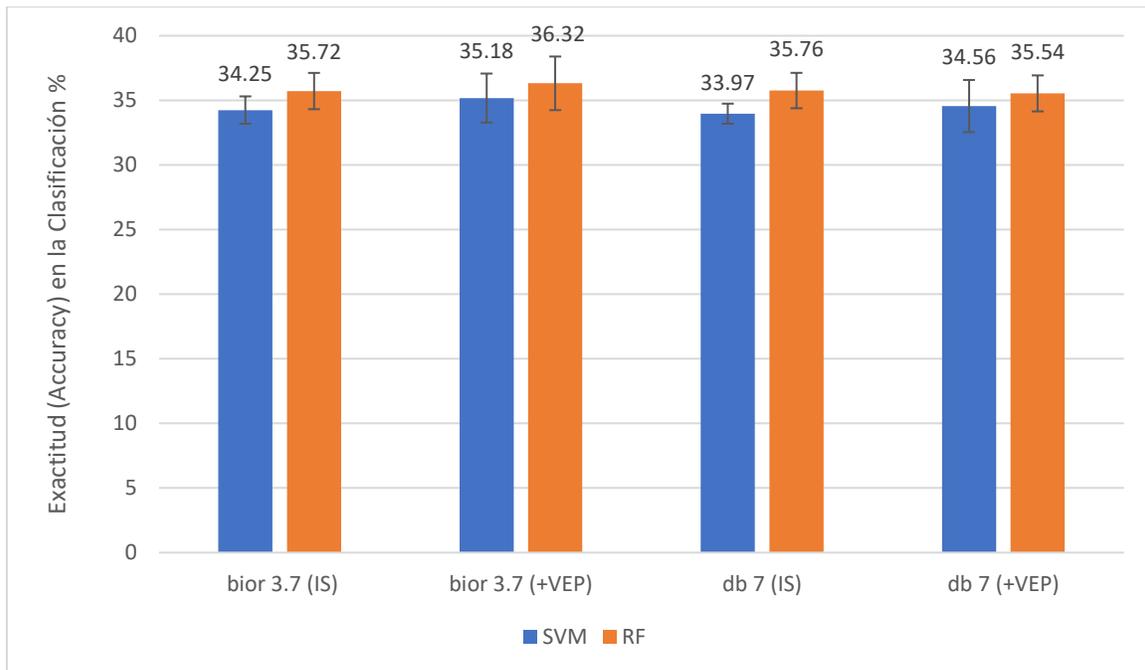
**Figura 5.5** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 32 canales de interés, usando como ondícula madre db7 para +VEP.

El resumen de la exactitud promedio obtenida de la clasificación en 8 sujetos con 32 canales en cada clasificador con ambas ondículas madre, se muestra en la **Tabla 5.1** y en la

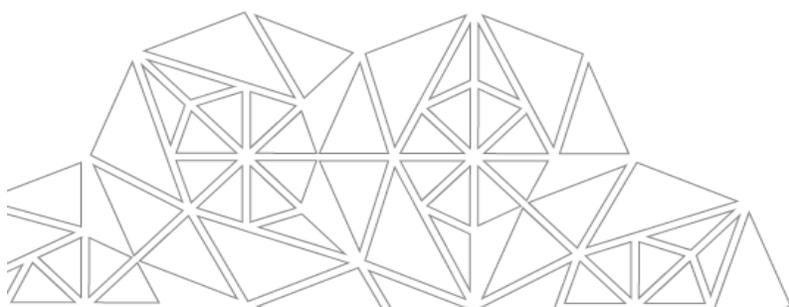
**Figura 5.6.** La mejor combinación de hiperparámetros se obtuvo mediante el proceso de tuning (ajuste de hiperparámetros) y que, a su vez, fue la de mayor frecuencia de aparición entre todos los sujetos.

**Tabla 5.1** Exactitud promedio obtenida de la clasificación en 8 sujetos con 32 canales

Ondícula Madre	Época	Clasificador	Exactitud Promedio	Desviación Estándar	Mejor combinación de hiperparámetros
bior 3.7	IS	SVM	34.25	$\pm 1.059$	gamma:10; kernel: rbf
		RF	<b>35.72</b>	$\pm 1.405$	gini; 0.7; 100
	+VEP	SVM	35.18	$\pm 1.895$	gamma:10; kernel: rbf
		RF	<b>36.32</b>	$\pm 2.073$	entropy; 0.5; 50
db 7	IS	SVM	33.97	$\pm 0.777$	gamma:10; kernel: rbf
		RF	<b>35.76</b>	$\pm 1.370$	entropy; 0.5; 50
	+VEP	SVM	34.56	$\pm 2.020$	gamma:10; kernel: rbf
		RF	<b>35.54</b>	$\pm 1.393$	entropy; 0.5; 50



**Figura 5.6** Exactitud promedio obtenida de la clasificación en 8 sujetos con 32 canales para cada tipo de ventana (IS y +VEP) y su respectiva ondícula madre.

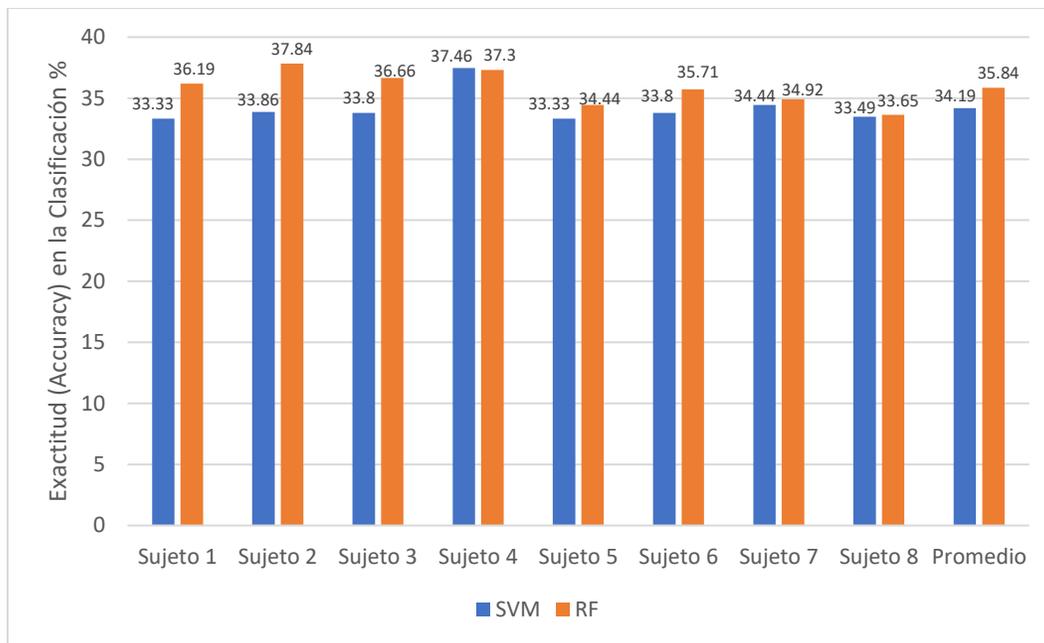


Es observable que, el clasificador con mejor resultado es el de bosques aleatorios (RF) en ambas ondículas y ventanas, para el caso de 32 canales.

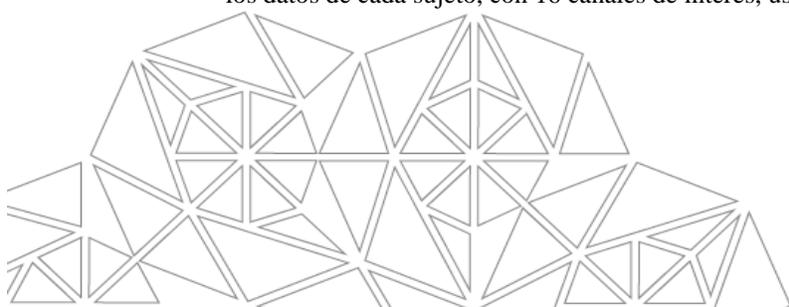
### 5.2.2 16 canales de EEG

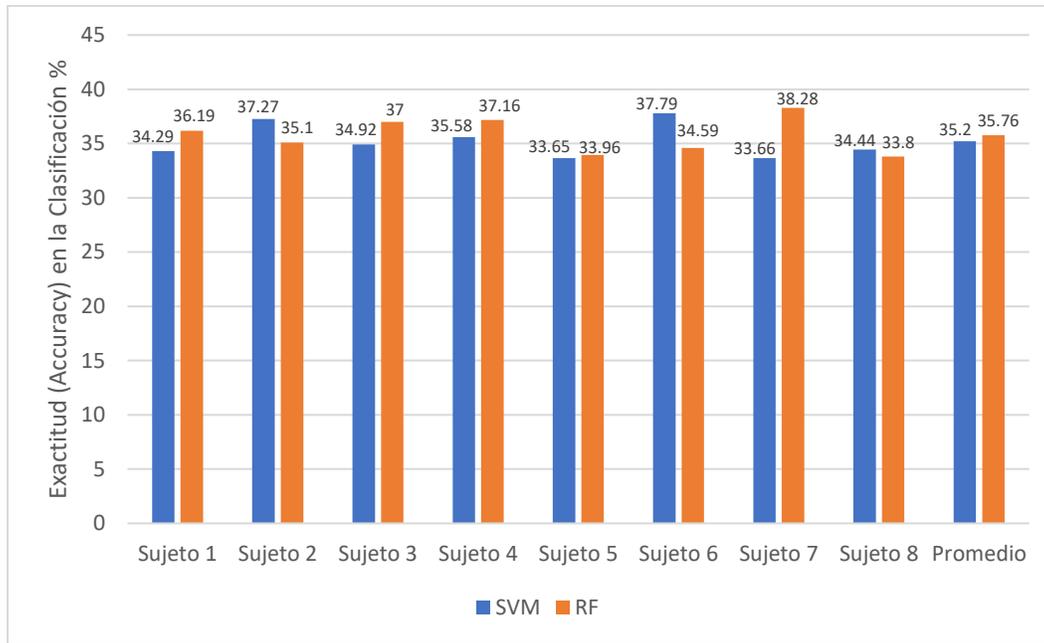
Análogamente con 32 canales, se realizó el mismo análisis para 16 canales. La mejor combinación de hiperparámetros por sujeto con 16 canales y bior 3.7 en ambos clasificadores se pueden consultar en los Anexos 5 y 6.

Se obtuvo una exactitud promedio con IS de 34.19% y 35.84% para SVM y RF respectivamente, en el caso de +VEP se consiguió 35.2 % y 35.76% para SVM y RF respectivamente (véase las Figuras 5.7 y 5.8). Como en los casos anteriores, el clasificador RF tuvo mejores resultados con respecto a SVM en ambos tipos de ventanas y hay una diferencia entre la exactitud de IS y +VEP. ¿Es significativa dicha diferencia? Se tratará de responder esa pregunta en la sección de la validación de la hipótesis.



**Figura 5.7** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre bior 3.7 para IS.

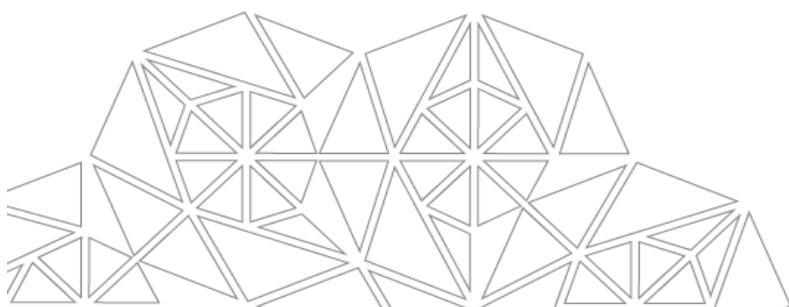


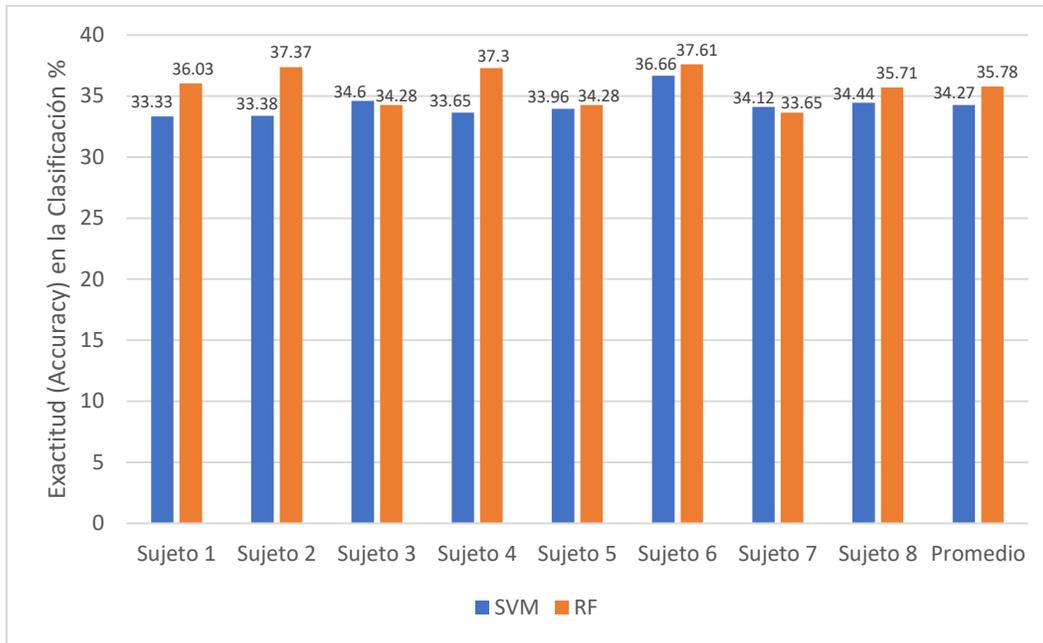


**Figura 5.8** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre bior 3.7 para +VEP.

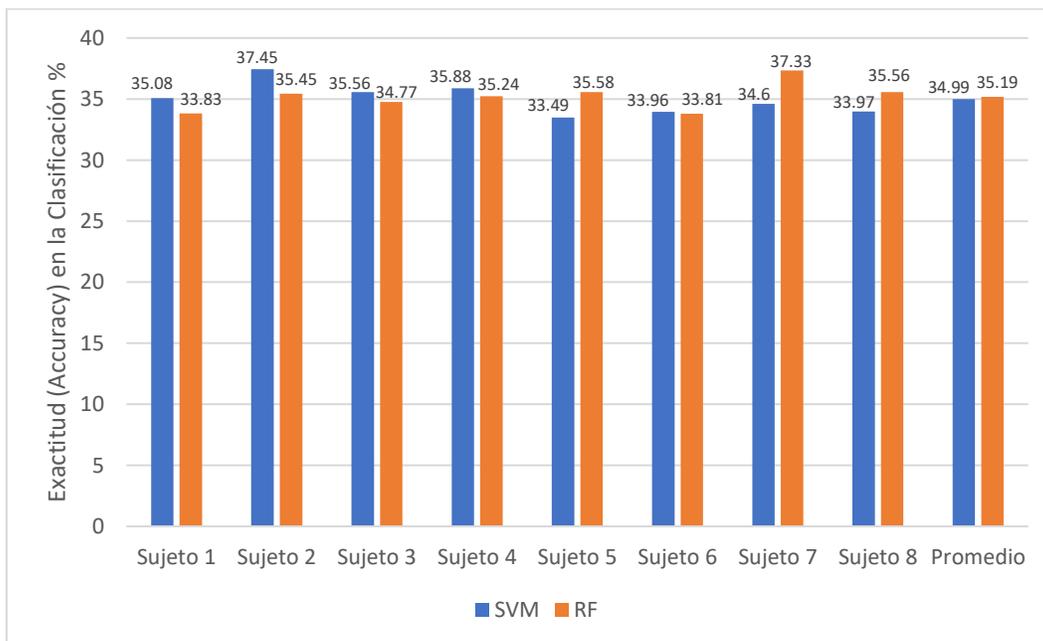
La exactitud promedio conseguida por los dos clasificadores con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre db7 para IS y +VEP se presenta en las Figuras 5.9 y 5.10 respectivamente. La mejor combinación de hiperparámetros por sujeto con 16 canales y db7 en ambos clasificadores se pueden consultar en los Anexos 7 y 8.

La exactitud promedio con IS fue de 34.27% y 35.78% para SVM y RF respectivamente, en el caso de +VEP se consiguió 34.99 % y 35.19% para SVM y RF respectivamente. El clasificador RF tuvo mejores resultados con respecto a SVM en ambos tipos de ventanas, además de haber una diferencia entre la exactitud de IS y +VEP. ¿Es significativa dicha diferencia? ¿Es significativa la diferencia entre la exactitud obtenida en ambas ondículas? Se tratará de responder estas preguntas en la sección de la validación de la hipótesis.





**Figura 5.9** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre db7 para IS.



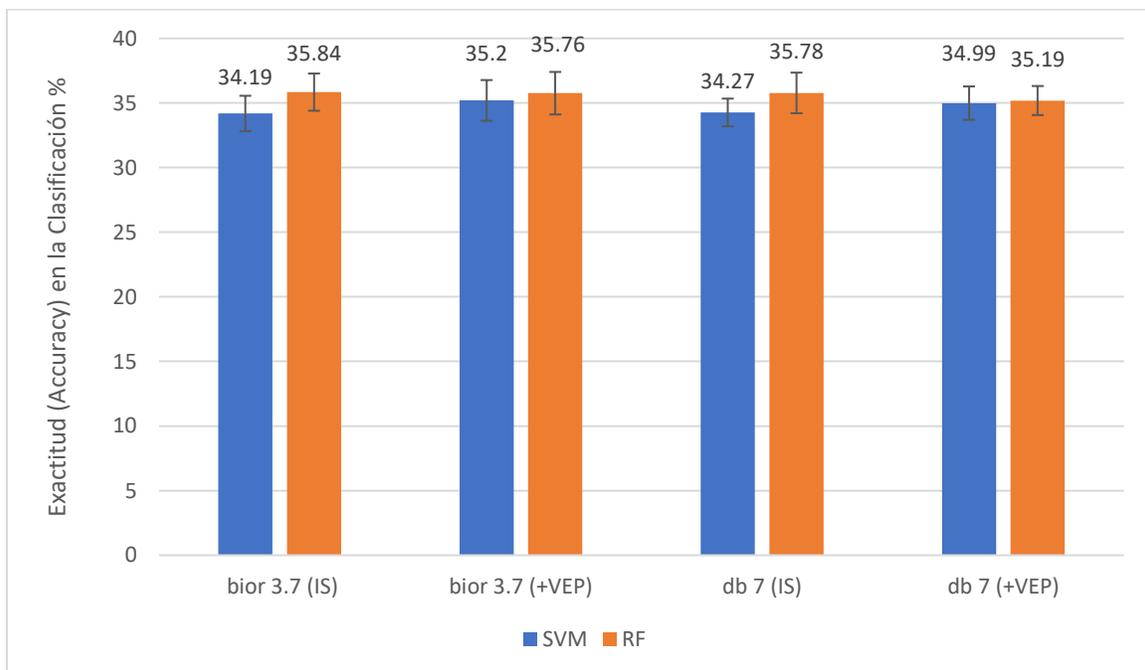
**Figura 5.10** Exactitud promedio conseguida en cada clasificador con la validación cruzada de 10 particiones con los datos de cada sujeto, con 16 canales de interés, usando como ondícula madre db7 para +VEP.

El resumen de la exactitud promedio obtenida de la clasificación en 8 sujetos con 16 canales en cada clasificador con ambas ondículas madre, se muestra en la **Tabla 5.2** y en la

**Figura 5.11.** La mejor combinación de hiperparámetros se obtuvo mediante el proceso de tuning (ajuste de hiperparámetros) y que, a su vez, fue la de mayor frecuencia de aparición entre todos los sujetos.

**Tabla 5.2** Exactitud promedio obtenida de la clasificación en 8 sujetos con 16 canales. N/A implica que todas las combinaciones con los hiperparámetros del clasificador tuvieron la misma frecuencia de aparición entre todos los sujetos.

Ondícula Madre	Época	Clasificador	Exactitud Promedio	Desviación Estándar	Mejor combinación de hiperparámetros
bior 3.7	IS	SVM	34.19	$\pm 1.369$	gamma:100; kernel: rbf
		RF	<b>35.84</b>	$\pm 1.441$	N/A
	+VEP	SVM	35.2	$\pm 1.576$	gamma:10; kernel: rbf
		RF	<b>35.76</b>	$\pm 1.643$	entropy; 0.7; 100
db 7	IS	SVM	34.27	$\pm 1.070$	gamma:10; kernel: rbf
		RF	<b>35.78</b>	$\pm 1.571$	entropy; 0.5; 50
	+VEP	SVM	34.99	$\pm 1.291$	gamma:10; kernel: rbf
		RF	<b>35.19</b>	$\pm 1.125$	gini; 0.7; 100



**Figura 5.11** Exactitud promedio obtenida de la clasificación en 8 sujetos con 16 canales para cada tipo de ventana (IS y +VEP) y su respectiva ondícula madre.

Como en el caso de 32 canales, para 16 canales, el clasificador con mejor resultado es el de bosques aleatorios (RF) en ambas ondículas y ventanas.

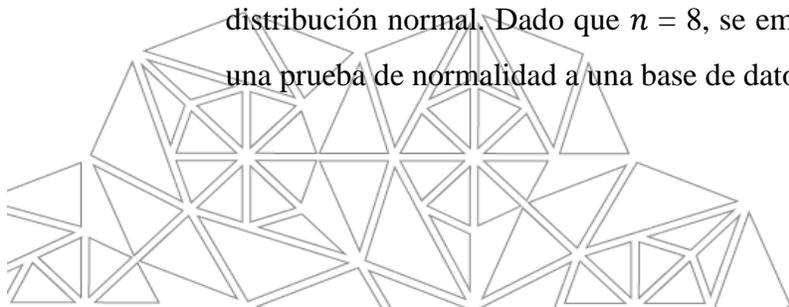
### 5.3 Validación de la hipótesis

Para validar la hipótesis de la presente investigación, es necesario comparar las poblaciones de datos que se tienen mediante inferencia estadística, considerando las variables (ondícula madre, tipo de ventana y número de canales). Se ordenaron los pares de 8 poblaciones a comparar, de acuerdo al análisis (véase en **Tabla 5.3**). Se utilizarán los valores obtenidos del RF para los análisis posteriores, dado que tuvo un mejor rendimiento en la exactitud en todas las clasificaciones.

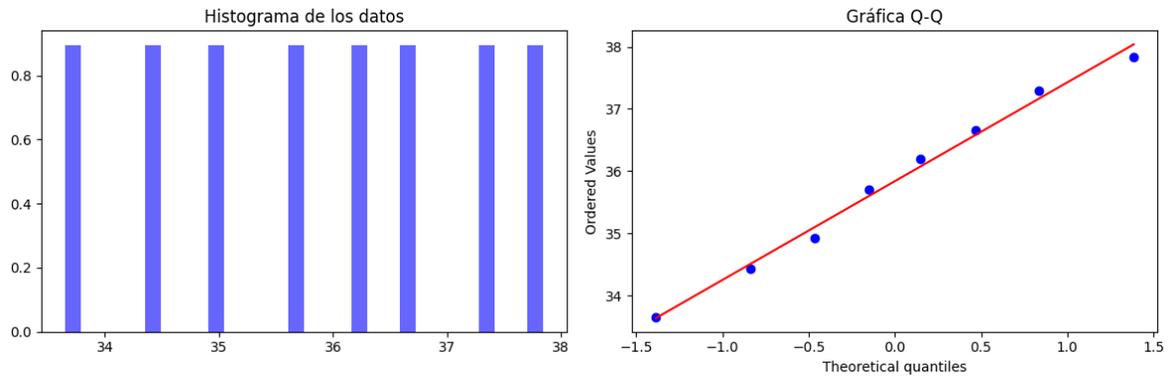
**Tabla 5.3** Pares de poblaciones de datos a comparar, considerando las variables *ondícula madre* ( $w_1, w_2$ ), *tipo de ventana* (IS, +VEP) y *número de canales* (16 ch, 32 ch). ( $w_1$ : bior 3.7,  $w_2$ : db7).

Análisis entre número de canales	$(w_1, \text{IS}, 16 \text{ ch})$ vs $(w_1, \text{IS}, 32 \text{ ch})$ $(w_1, +\text{VEP}, 16 \text{ ch})$ vs $(w_1, +\text{VEP}, 32 \text{ ch})$ $(w_2, \text{IS}, 16 \text{ ch})$ vs $(w_2, \text{IS}, 32 \text{ ch})$ $(w_2, +\text{VEP}, 16 \text{ ch})$ vs $(w_2, +\text{VEP}, 32 \text{ ch})$
Análisis entre tipo de ventana	$(w_1, \text{IS}, 16 \text{ ch})$ vs $(w_1, +\text{VEP}, 16 \text{ ch})$ $(w_1, \text{IS}, 32 \text{ ch})$ vs $(w_1, +\text{VEP}, 32 \text{ ch})$ $(w_2, \text{IS}, 16 \text{ ch})$ vs $(w_2, +\text{VEP}, 16 \text{ ch})$ $(w_2, \text{IS}, 32 \text{ ch})$ vs $(w_2, +\text{VEP}, 32 \text{ ch})$
Análisis entre ondículas madre	$(w_1, \text{IS}, 16 \text{ ch})$ vs $(w_2, \text{IS}, 16 \text{ ch})$ $(w_1, +\text{VEP}, 16 \text{ ch})$ vs $(w_2, +\text{VEP}, 16 \text{ ch})$ $(w_1, \text{IS}, 32 \text{ ch})$ vs $(w_2, \text{IS}, 32 \text{ ch})$ $(w_1, +\text{VEP}, 32 \text{ ch})$ vs $(w_2, +\text{VEP}, 32 \text{ ch})$

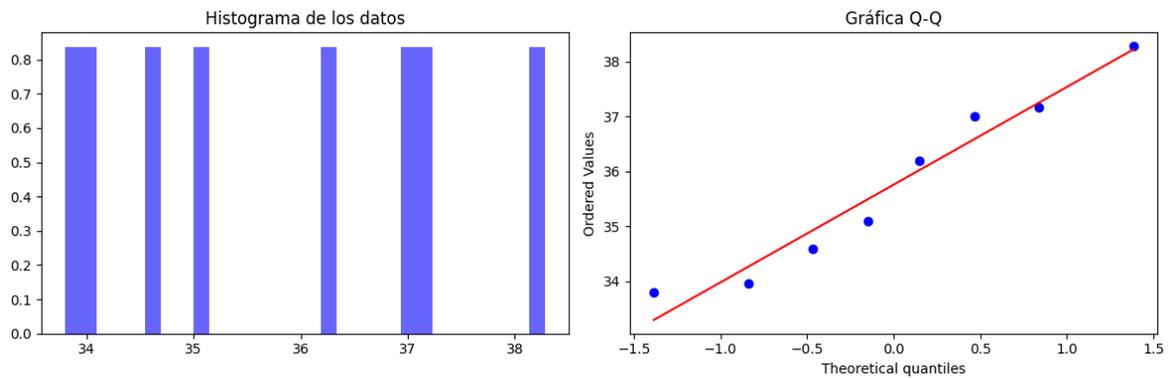
Antes de realizar algún test estadístico para inferir en algún parámetro, es importante verificar si los supuestos para aplicar dicho test se están cumpliendo, y no llegar a conclusiones falaces. Se tiene un tamaño muestral  $n$  de 8 (número de sujetos) y se pretende hacer una prueba de normalidad, es decir, si las 8 poblaciones tienen un comportamiento de distribución normal. Dado que  $n = 8$ , se empleará el test Shapiro-Wilk, el cual es útil para una prueba de normalidad a una base de datos con  $n < 50$ .



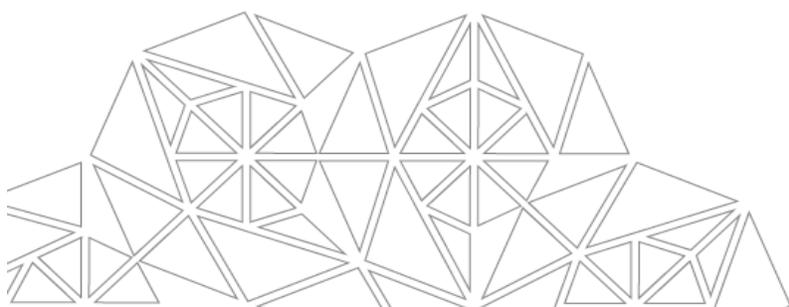
Tenemos entonces, como  $H_0$ : La distribución es normal y  $H_1$ : La distribución no es normal.

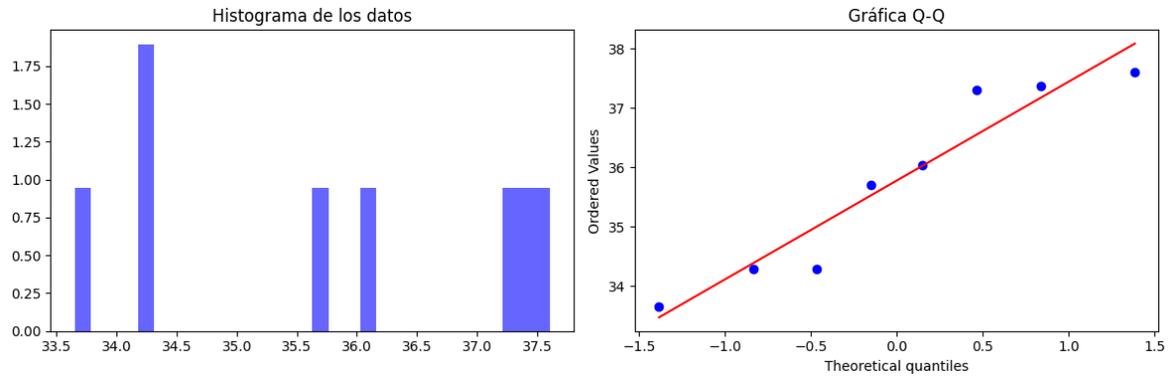


**Figura 5.12** Prueba Shapiro-Wilk para la población (bior 3.7, IS, 16 ch)  $W$ : 0.977,  $p$ : 0.952 ( $H_0$ : La distribución es normal).

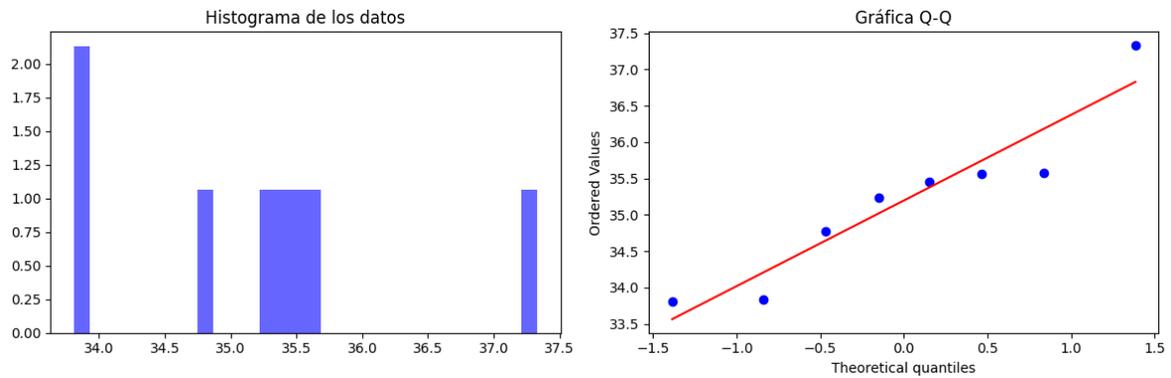


**Figura 5.13** Prueba Shapiro-Wilk para la población (bior 3.7, +VEP, 16 ch)  $W$ : 0.935,  $p$ : 0.565 ( $H_0$ : La distribución es normal).

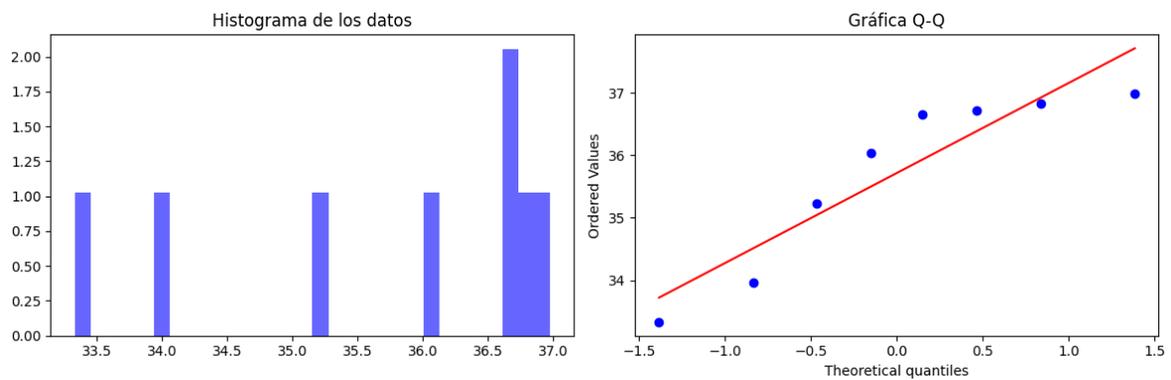




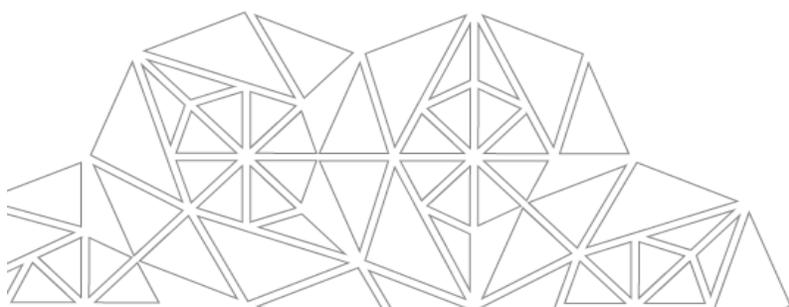
**Figura 5.14** Prueba Shapiro-Wilk para la población (db7, IS, 16 ch)  $W: 0.888$ ,  $p: 0.228$  ( $H_0$ : La distribución es normal).

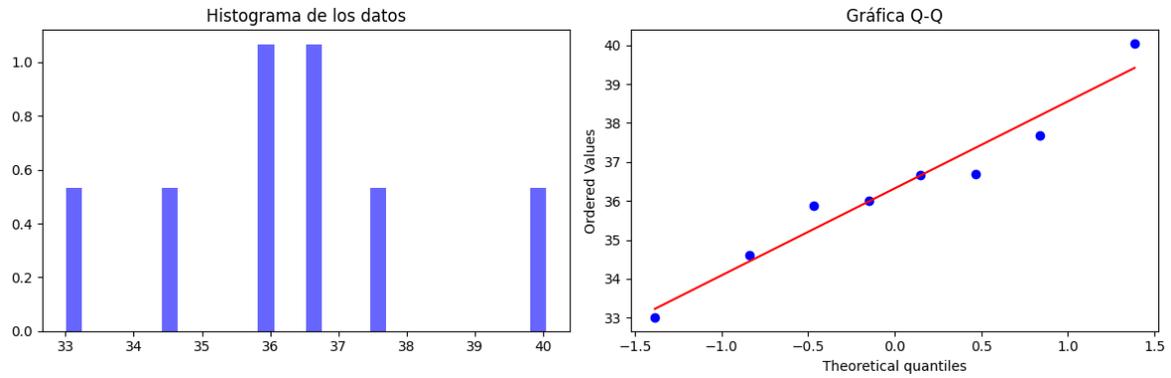


**Figura 5.15** Prueba Shapiro-Wilk para la población (db7, +VEP, 16 ch)  $W: 0.905$ ,  $p: 0.322$  ( $H_0$ : La distribución es normal).

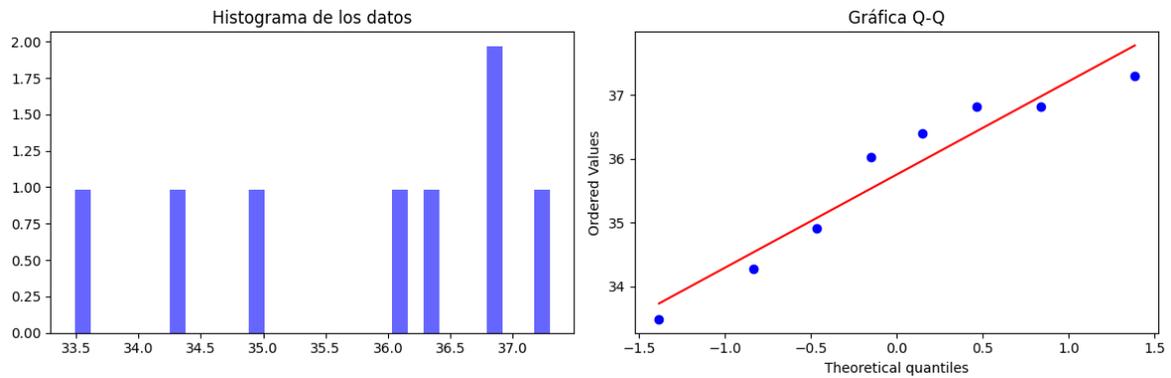


**Figura 5.16** Prueba Shapiro-Wilk para la población (bior 3.7, IS, 32 ch)  $W: 0.840$ ,  $p: 0.076$  ( $H_0$ : La distribución es normal).

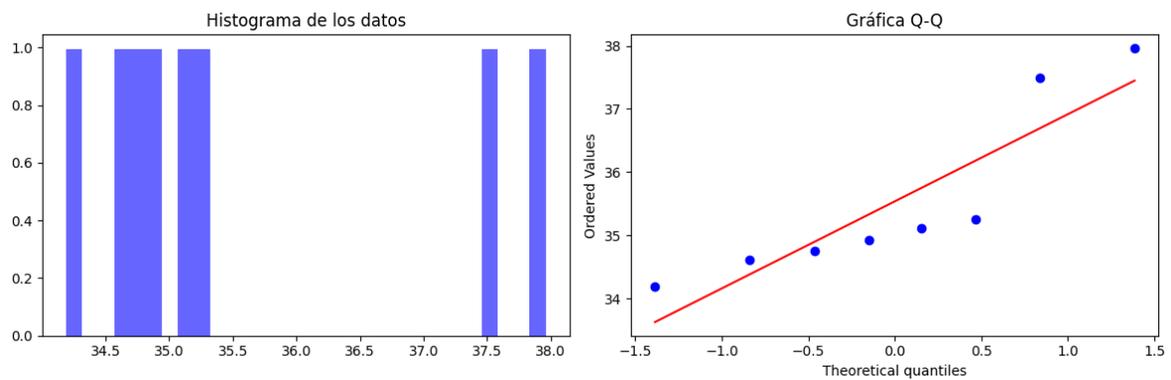




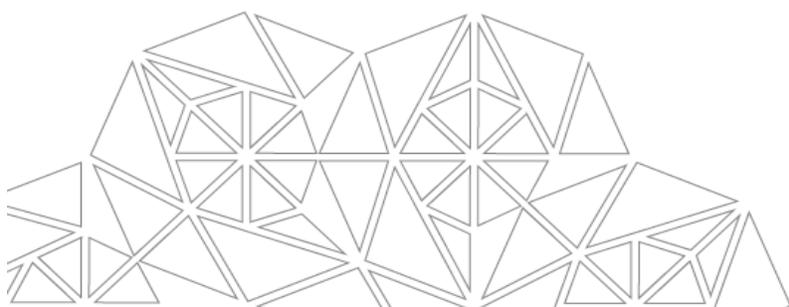
**Figura 5.17** Prueba Shapiro-Wilk para la población (bior 3.7, +VEP, 32 ch)  $W: 0.968$ ,  $p: 0.888$  ( $H_0$ : La distribución es normal).



**Figura 5.18** Prueba Shapiro-Wilk para la población (db7, IS, 32 ch)  $W: 0.911$ ,  $p: 0.365$  ( $H_0$ : La distribución es normal).



**Figura 5.19** Prueba Shapiro-Wilk para la población (db7, +VEP, 32 ch)  $W: 0.795$ ,  $p: 0.025$  ( $H_1$ : La distribución no es normal).



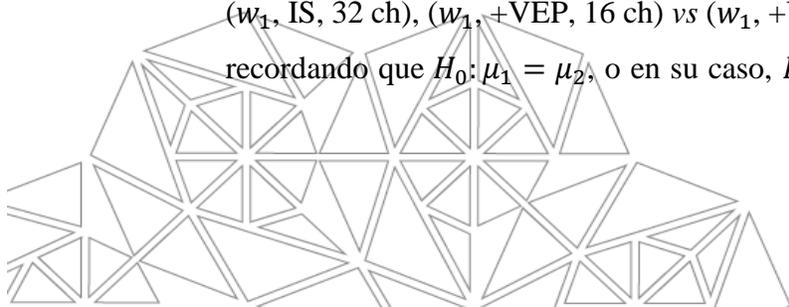
De acuerdo con las pruebas de Shapiro-Wilk realizadas a todas las poblaciones de datos, encontramos que la población (db7, +VEP, 32 ch), es decir, los resultados de la clasificación con RF de los 8 sujetos cuyas neuroseñales de 32 canales EEG se caracterizaron con la ondícula madre db7 y cuyas ventanas incluyen información del VEP, no presentan un comportamiento normal. Esta información es relevante, puesto que, para comparar esta población de datos con otra, se requerirá un test que permita no cumplir los supuestos de normalidad.

Estudiar la relación entre medidas de tendencia central de poblaciones de datos, es una manera de inferir si las diferencias entre estas poblaciones son significativas o no, en nuestro caso la exactitud promedio obtenida en la clasificación.

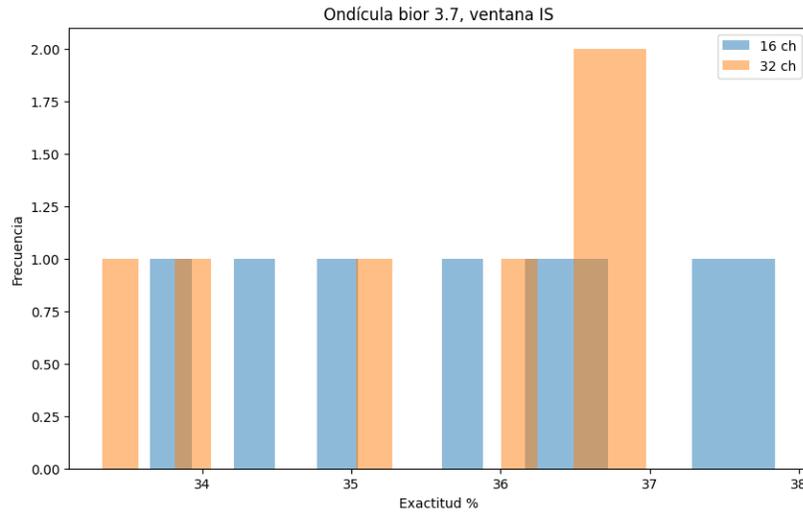
Ahora bien, si las poblaciones cumplen los supuestos de normalidad, se les pueden aplicar test que permitan inferir si sus medias son estadísticamente iguales o no, es decir, se plantea una hipótesis nula  $H_0: \mu_1 = \mu_2$ , o en su caso, una hipótesis alternativa  $H_1: \mu_1 \neq \mu_2$ . Cuando  $n < 30$ , tenemos la prueba *t student*. Pero en el caso de no cumplir los supuestos de normalidad, situación que ocurre cuando la distribución verdaderamente no tiene un comportamiento normal, o bien, la muestra es muy pequeña, siendo no representativa de la población, se debe optar por alternativas no paramétricas para las pruebas de comparación de medias, también con el supuesto de si los datos son independientes, la cual, una alternativa para la *t student* es la de *Mann-Whitney*, teniendo como hipótesis nula  $H_0: Me_1 = Me_2$ , o en su caso, una hipótesis alternativa  $H_1: Me_1 \neq Me_2$ , donde  $Me_1$  y  $Me_2$  son las medianas de los conjuntos de datos 1 y 2 respectivamente.

### 5.3.1 Análisis entre número de canales

En la **Tabla 5.3** se muestran los pares de poblaciones a comparar para determinar si la diferencia en la exactitud es significativa o no, dependiendo de la información del número de canales EEG. Se realizó la prueba *t student* para los casos de comparación ( $w_1$ , IS, 16 ch) vs ( $w_1$ , IS, 32 ch), ( $w_1$ , +VEP, 16 ch) vs ( $w_1$ , +VEP, 32 ch) y ( $w_2$ , IS, 16 ch) vs ( $w_2$ , IS, 32 ch), recordando que  $H_0: \mu_1 = \mu_2$ , o en su caso,  $H_1: \mu_1 \neq \mu_2$ . Para el caso ( $w_2$ , +VEP, 16 ch) vs

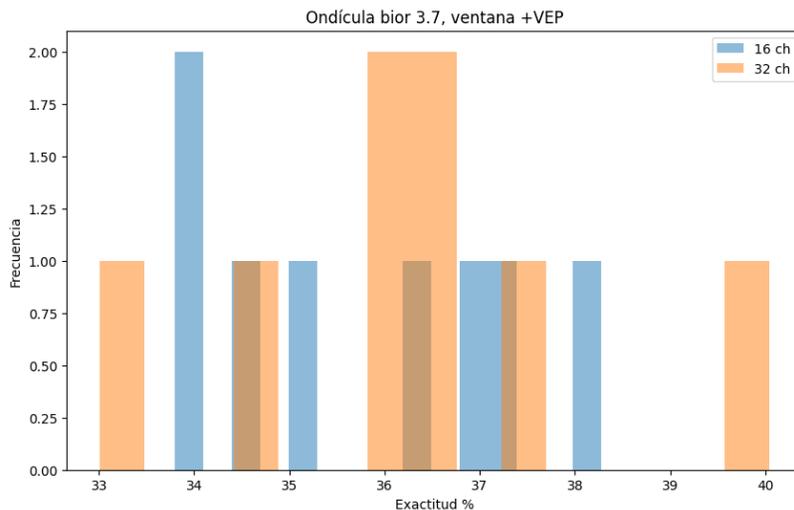


( $w_2$ , +VEP, 32 ch), se observó que la población ( $w_2$ , +VEP, 32 ch) no cumplía los supuestos de normalidad, de acuerdo a la prueba Shapiro-Wilk (véase la **Figura 5.19**), por lo que para este par de poblaciones se aplicó la prueba *Mann-Whitney*, donde  $H_0: Me_1 = Me_2$ , o en su caso,  $H_1: Me_1 \neq Me_2$ .



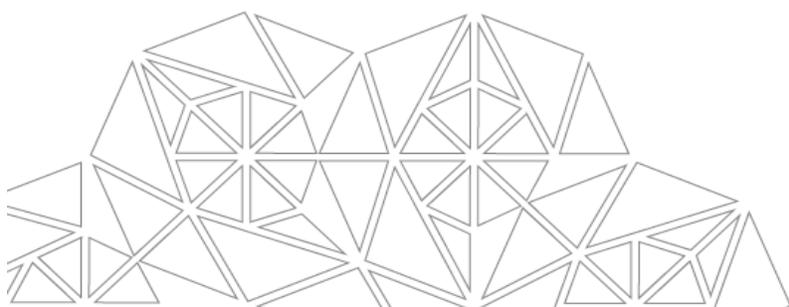
**Figura 5.20** Prueba *t student* para (bior 3.7, IS, 16 ch) y (bior 3.7, IS, 32 ch)  $t: 0.172$ ,  $p: 0.865$

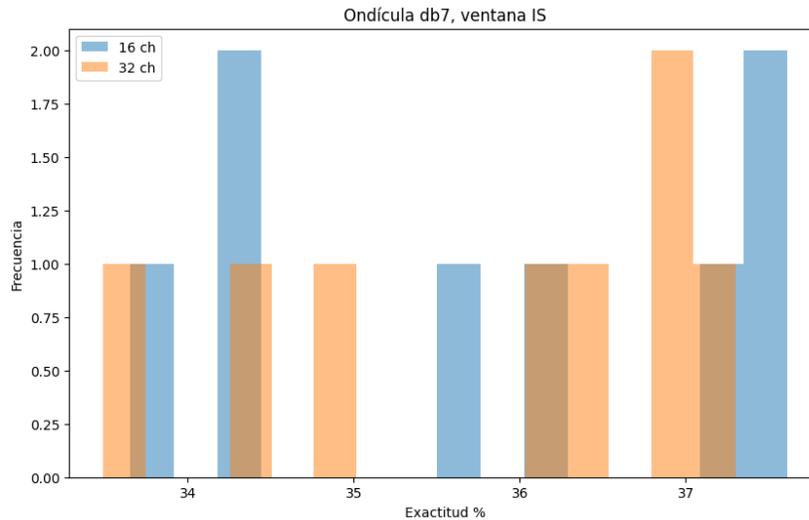
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.



**Figura 5.21** Prueba *t student* para (bior 3.7, +VEP, 16 ch) y (bior 3.7, +VEP, 32 ch)  $t: -0.601$ ,  $p: 0.557$

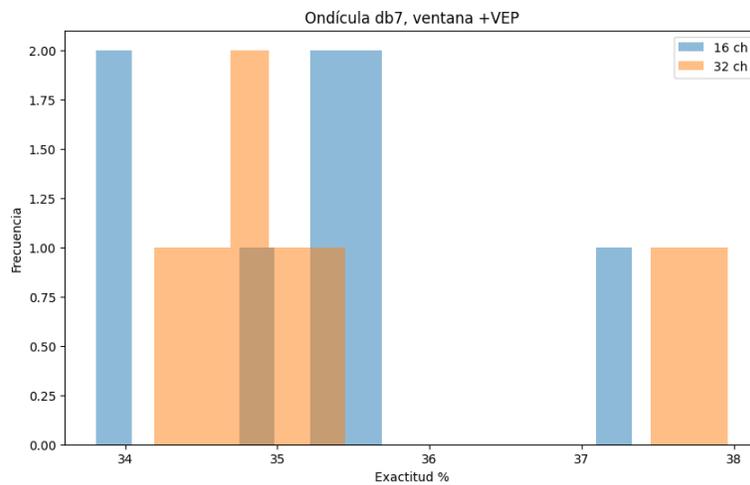
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.





**Figura 5.22** Prueba *t student* para (db7, IS, 16 ch) y (db7, IS, 32 ch)  $t: 0.028$ ,  $p: 0.977$

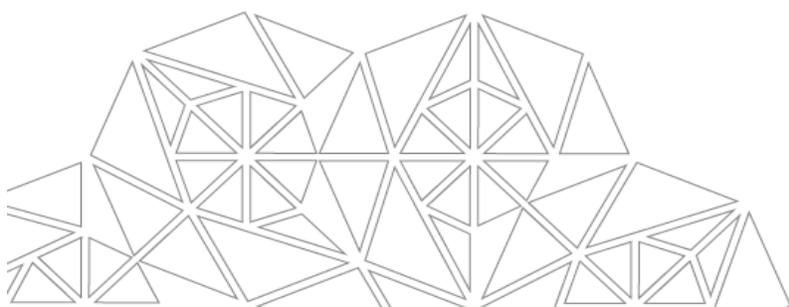
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.



**Figura 5.23** Prueba *Mann-Whitney* para (db7, IS, 16 ch) y (db7, IS, 32 ch)  $U: 32$ ,  $p: 1$

( $H_0: Me_1 = Me_2$ ) Las muestras provienen de poblaciones con medianas iguales.

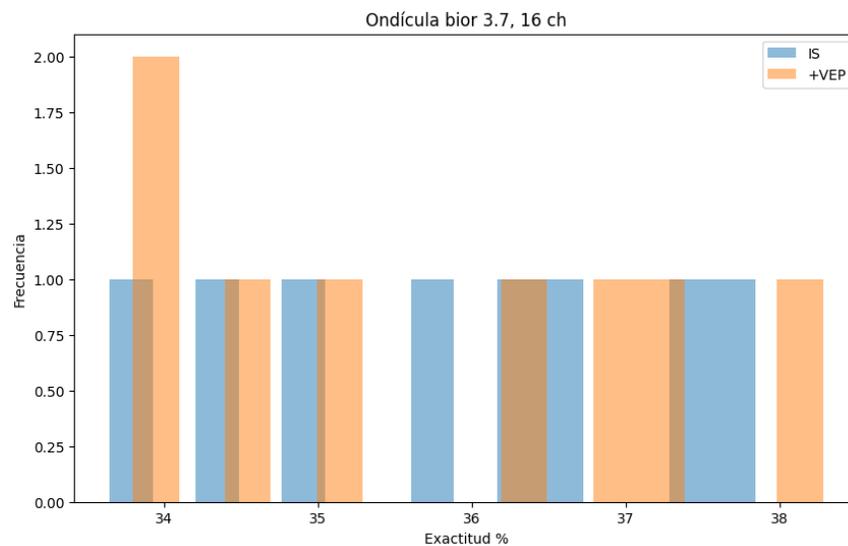
Las pruebas aplicadas en todos los pares de poblaciones, no rechazan la hipótesis nula, ya sea  $H_0: \mu_1 = \mu_2$  o bien,  $H_0: Me_1 = Me_2$ , por lo que, con base a la evidencia estadística, no hay diferencia significativa en la exactitud (accuracy) de la clasificación con la información de 16 y 32 canales.



### 5.3.2 Análisis entre tipo de ventana

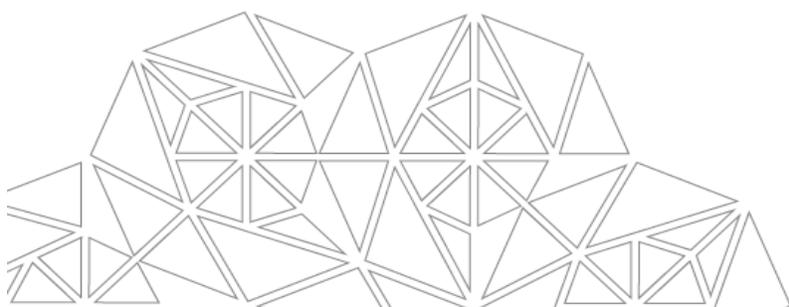
Otra manera de analizar los resultados obtenidos es mediante el análisis entre tipo de ventana, siendo IS las ventanas que tienen la información del habla imaginada y +VEP aquellas que tienen la información tanto del habla imaginada como del estímulo visual. Este análisis es crucial para la hipótesis de esta investigación.

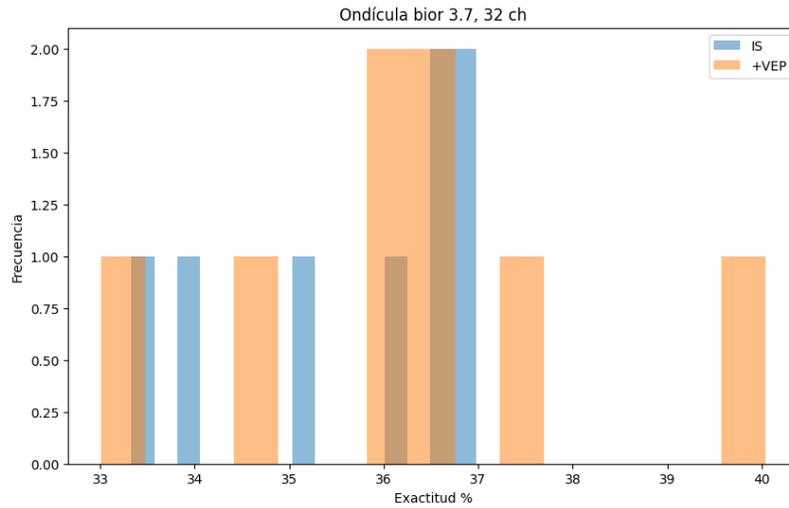
Se realizó la prueba *t student* para los casos de comparación ( $w_1$ , IS, 16 ch) vs ( $w_1$ , +VEP, 16 ch), ( $w_1$ , IS, 32 ch) vs ( $w_1$ , +VEP, 32 ch) y ( $w_2$ , IS, 16 ch) vs ( $w_2$ , +VEP, 16 ch), recordando que  $H_0: \mu_1 = \mu_2$ , o en su caso,  $H_1: \mu_1 \neq \mu_2$ . Para el caso ( $w_2$ , IS, 32 ch) vs ( $w_2$ , +VEP, 32 ch), se observó que la población ( $w_2$ , +VEP, 32 ch) no cumplía los supuestos de normalidad, de acuerdo a la prueba Shapiro-Wilk (véase la **Figura 5.19**), por lo que para este par de poblaciones se aplicó la prueba *Mann-Whitney*, donde  $H_0: Me_1 = Me_2$ , o en su caso,  $H_1: Me_1 \neq Me_2$ .



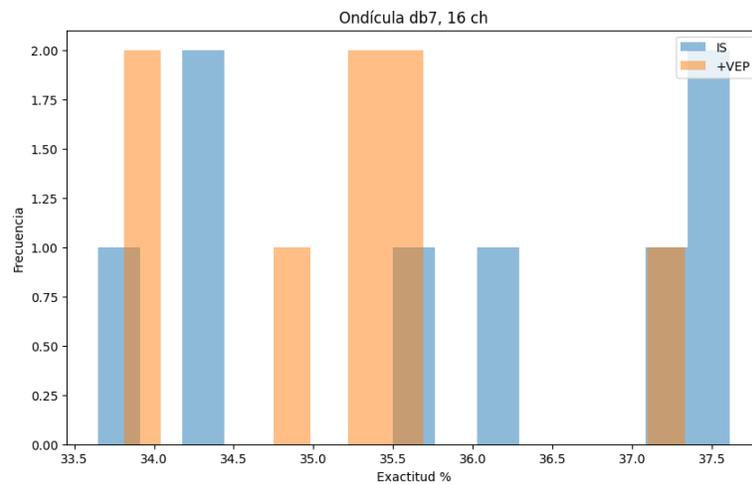
**Figura 5.24** Prueba *t student* para (bior 3.7, IS, 16 ch) y (bior 3.7, +VEP, 16 ch)  $t: 0.101$ ,  $p: 0.920$

( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.

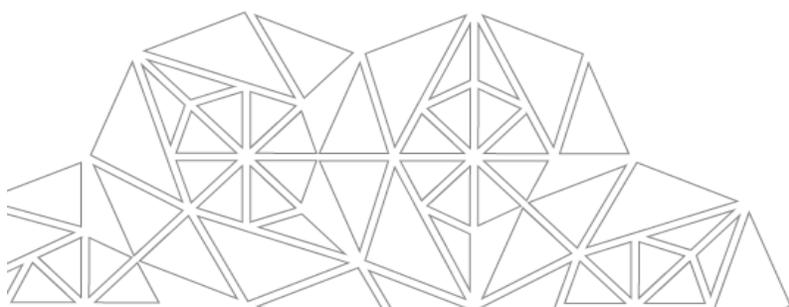


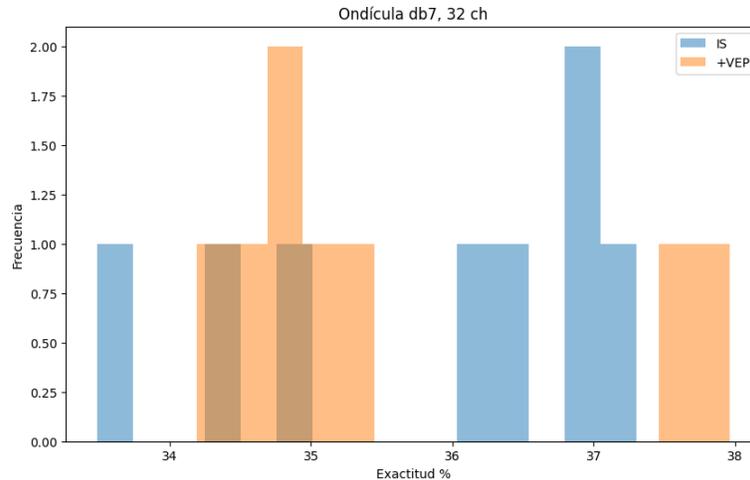


**Figura 5.25** Prueba *t student* para (bior 3.7, IS, 32 ch) y (bior 3.7, +VEP, 32 ch)  $t: -0.684$ ,  $p: 0.504$   
 $(H_0: \mu_1 = \mu_2)$  Las muestras provienen de poblaciones con medias iguales.



**Figura 5.26** Prueba *t student* para (db7, IS, 16 ch) y (db7, +VEP, 16 ch)  $t: 0.852$ ,  $p: 0.408$   
 $(H_0: \mu_1 = \mu_2)$  Las muestras provienen de poblaciones con medias iguales.





**Figura 5.27** Prueba *Mann-Whitney* para (db7, IS, 32 ch) y (db7, +VEP, 32 ch)  $U: 34.5$ ,  $p: 0.833$

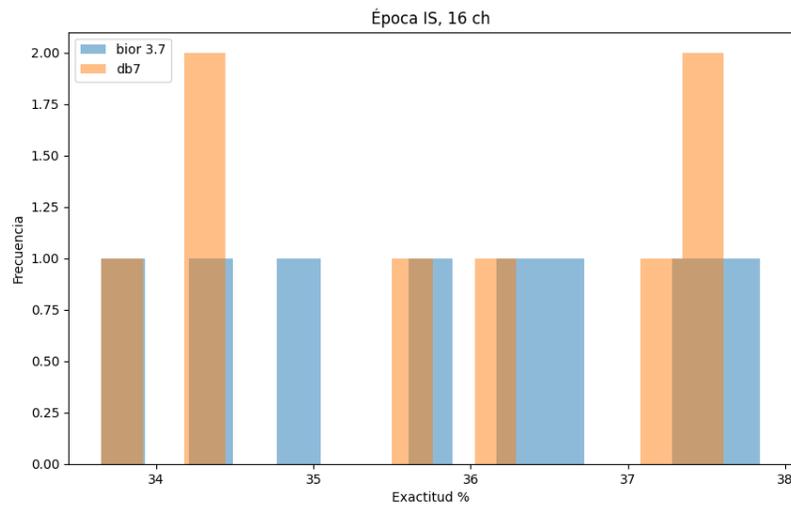
( $H_0: Me_1 = Me_2$ ) Las muestras provienen de poblaciones con medianas iguales.

Las pruebas aplicadas en todos los pares de poblaciones, no rechazan la hipótesis nula, ya sea  $H_0: \mu_1 = \mu_2$  o bien,  $H_0: Me_1 = Me_2$ , por lo que, con base a la evidencia estadística, no hay diferencia significativa en la exactitud (accuracy) de la clasificación con la información de las ventanas (épocas) IS y +VEP.

### 5.3.3 Análisis entre ondículas madre

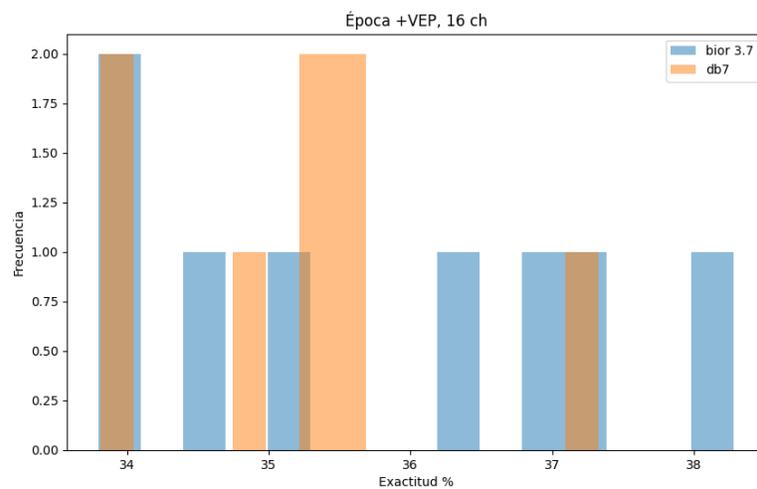
Otra variable para el análisis de los resultados, es la ondícula madre que se utilizó para caracterizar las neuroseñales. Observar si alguna aportó mejor información para la eficiencia de la clasificación es relevante para esta investigación.

Se realizó la prueba *t student* para los casos de comparación ( $w_1$ , IS, 16 ch) vs ( $w_2$ , IS, 16 ch), ( $w_1$ , +VEP, 16 ch) vs ( $w_2$ , +VEP, 16 ch) y ( $w_1$ , IS, 32 ch) vs ( $w_2$ , IS, 32 ch), recordando que  $H_0: \mu_1 = \mu_2$ , o en su caso,  $H_1: \mu_1 \neq \mu_2$ . Para el caso ( $w_1$ , +VEP, 32 ch) vs ( $w_2$ , +VEP, 32 ch), se observó que la población ( $w_2$ , +VEP, 32 ch) no cumplía los supuestos de normalidad, de acuerdo a la prueba Shapiro-Wilk (véase la **Figura 5.19**), por lo que para este par de poblaciones se aplicó la prueba *Mann-Whitney*, donde  $H_0: Me_1 = Me_2$ , o en su caso,  $H_1: Me_1 \neq Me_2$ .



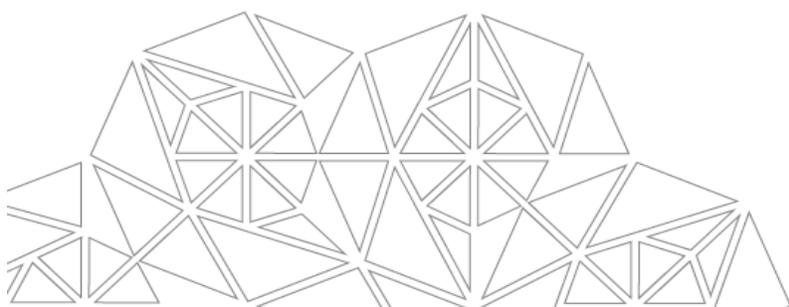
**Figura 5.28** Prueba *t student* para (bior 3.7, IS, 16 ch) y (db7, IS, 16 ch)  $t: 0.079$ ,  $p: 0.937$

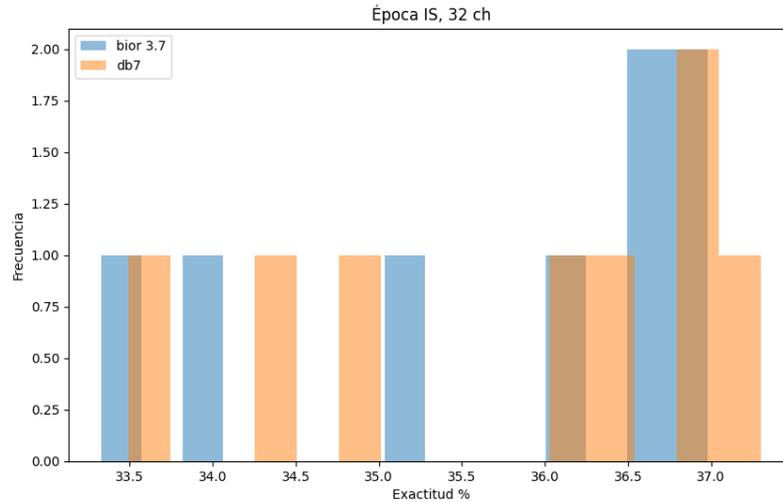
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.



**Figura 5.29** Prueba *t student* para (bior 3.7, +VEP, 16 ch) y (db7, +VEP, 16 ch)  $t: 0.800$ ,  $p: 0.436$

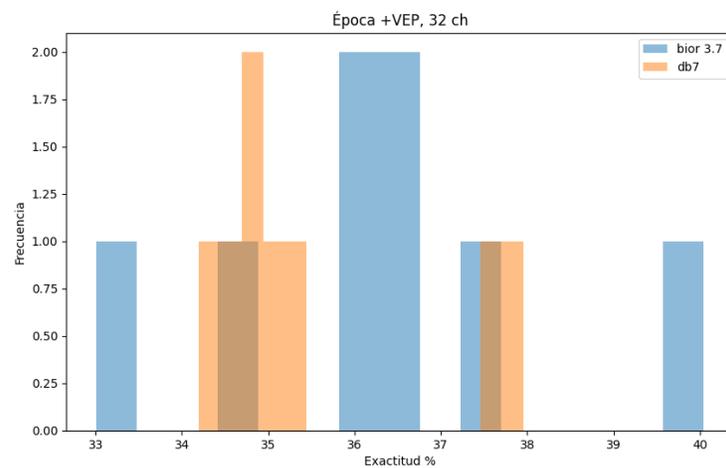
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.





**Figura 5.30** Prueba *t student* para (bior 3.7, IS, 32 ch) y (db7, IS, 32 ch)  $t: -0.059$ ,  $p: 0.953$

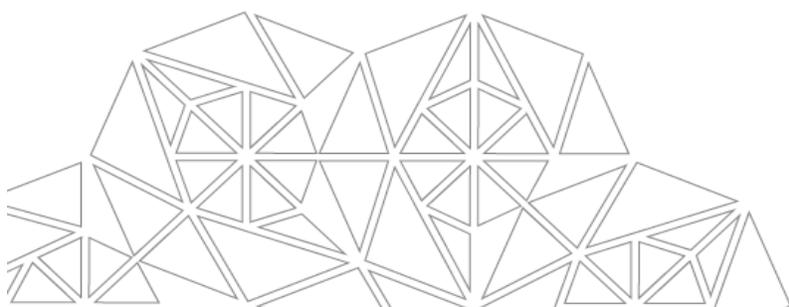
( $H_0: \mu_1 = \mu_2$ ) Las muestras provienen de poblaciones con medias iguales.



**Figura 5.31** Prueba *Mann-Whitney* para (bior 3.7, +VEP, 32 ch) y (db7, +VEP, 32 ch)  $U: 40.0$ ,  $p: 0.441$

( $H_0: Me_1 = Me_2$ ) Las muestras provienen de poblaciones con medianas iguales.

Las pruebas aplicadas en todos los pares de poblaciones, no rechazan la hipótesis nula, ya sea  $H_0: \mu_1 = \mu_2$  o bien,  $H_0: Me_1 = Me_2$ , por lo que, con base a la evidencia estadística, no hay diferencia significativa en la exactitud (accuracy) de la clasificación con la caracterización de las ondículas madre biorthogonal 3.7 (bior 3.7) y daubechies 7 (db7).



Recopilando los resultados de los distintos análisis y con referencia a la hipótesis planteada para esta investigación, se llega a lo siguiente:

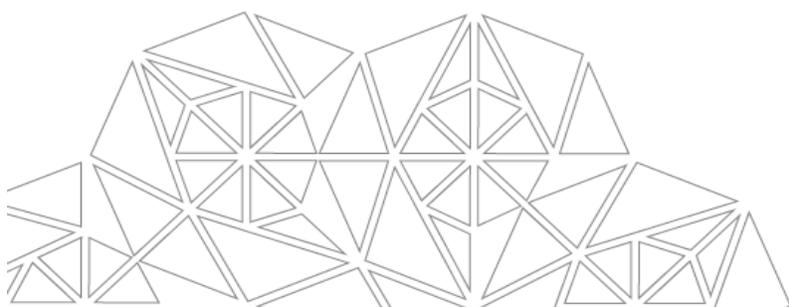
**Hipótesis:** Utilizando la DWT con las wavelets (ondículas) madre biorthogonal 3.7 (bior 3.7) y daubechies 7 (db7) para la caracterización de las señales EEG registradas con el equipo Neuroscan Neuvo, como fuente electrofisiológica el habla imaginada, de un vocabulario de 3 palabras del idioma español (círculo, cuadrado, trapecio) y con estímulos visuales de figuras geométricas, existirá diferencia significativa en la exactitud de la clasificación al incluir la información del potencial evocado visual a la de la fuente electrofisiológica.

En síntesis,

$H_0$ : *Hay diferencia significativa entre IS y +VEP*

$H_1$ : *No hay diferencia significativa entre IS y +VEP*

Conclusión, se rechaza  $H_0$  y se acepta  $H_1$ , es decir, **no hay diferencia significativa en la exactitud de la clasificación entre las ventanas IS y +VEP.**



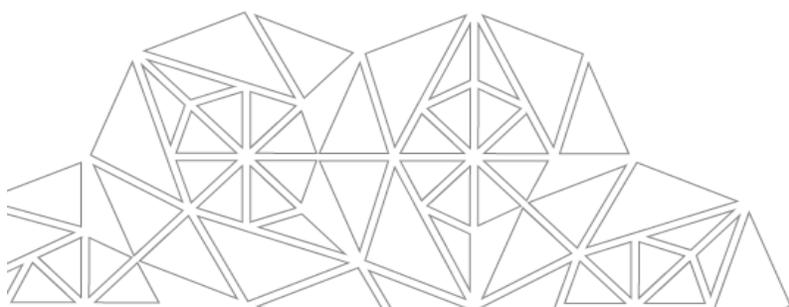
## Capítulo 6

### CONCLUSIONES

Si bien es cierto que, los resultados obtenidos en esta investigación distan por mucho para un control eficiente en los sistemas BCI basados en habla imaginada, también es conveniente hacer notar que la exploración con distintas herramientas disponibles nos da un bosquejo de la realidad respecto al comportamiento de las neuroseñales, incluso la oportunidad de desarrollar nuevos métodos. El clasificador RF tuvo mejores resultados en la exactitud que SVM para el vocabulario conformado de las palabras círculo, cuadrado y trapecio. Los porcentajes de exactitud conseguidos en promedio por ventana, están por encima del 33.3%, porcentaje del azar para 3 clases (véase las Figuras **5.6** y **5.11**), lo que permite inferir que, las señales EEG sí contienen información que permita distinguir el habla imaginada de las palabras círculo, cuadrado y trapecio.

De acuerdo a los tres análisis, al comparar las poblaciones de datos que se tuvieron mediante inferencia estadística, considerando las variables ondícula madre, tipo de ventana y número de canales, no hay diferencia significativa en la exactitud (accuracy) de la clasificación con la caracterización de las ondículas madre biorthogonal 3.7 (bior 3.7) y daubechies 7 (db7), entre tipos de ventana IS y +VEP, así como la información obtenida con distinto número de canales EEG (16 y 32).

También es importante mencionar que, este trabajo de investigación contribuyó en la obtención de un corpus de datos mediante la grabación de señales EEG, relacionadas al habla imaginada y a potenciales evocados visuales de 8 sujetos. Dicho corpus está basado en un vocabulario de 3 palabras del idioma español, siguiendo el diseño experimental anteriormente descrito. Para lograr lo anterior, se dedicaron aproximadamente 6 meses (octubre 2022 – abril 2023) continuando con las etapas de la metodología propuesta (desde la adquisición de las señales hasta la clasificación).



## 6.1 Discusión

El hecho de que, la diferencia en la exactitud de la clasificación respecto al número de canales EEG utilizados no sea significativa, nos permite inferir que no es necesaria la información de más canales para este método, más bien, sugiere enfocarse en aquellos que se encuentren próximos a las áreas cerebrales relacionadas al lenguaje (áreas de Broca y Wernicke) para el habla imaginada y la corteza visual para VEP's, reduciendo el impacto de la maldición de la dimensionalidad y aumentando la eficiencia del procesamiento.

Los porcentajes de exactitud están apenas por encima del porcentaje del azar para tres clases, lo cual abre la posibilidad de mejorar en varios aspectos de todas las etapas de la metodología implementada. También, es cierto que no se consideraron los resultados de otros criterios para evaluar el desempeño de la clasificación.

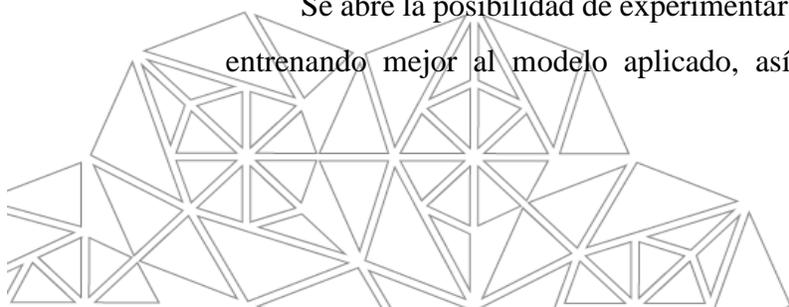
Por otro lado, el número de sujetos como muestra de la metodología sugerida es pequeño; aumentar la base de datos ayudaría en mejorar la robustez del modelo implementado y llegar a conclusiones más certeras.

Además, que la información del potencial evocado visual no sea significativa al agregarse a las ventanas del habla imaginada para un mejor resultado en la exactitud de la clasificación, nos permite sugerir otra manera la cual se pueda aprovechar esta información para lograr sistemas BCI's más eficientes.

## 6.2 Trabajo futuro

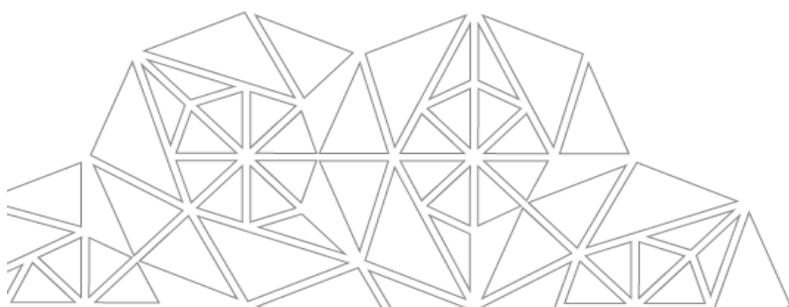
De las conclusiones obtenidas se desprenden implicaciones interesantes, puesto que para el caso de la ondícula madre, sugieren explorar otra familia de wavelets para la etapa de clasificación.

Se abre la posibilidad de experimentar con más sujetos, aumentando la base de datos y entrenando mejor al modelo aplicado, así como explorar con otros parámetros de los



clasificadores en el tuning, posibilitando conseguir mejores combinaciones que aumenten la exactitud, así como utilizar otros criterios de evaluación para medir el desempeño de los clasificadores.

Hay muchas palabras que faltan por explorar del idioma español, usando la fuente electrofisiológica del habla imaginada. Por otro lado, el diseño experimental puede modificarse con la intención de mejorar los resultados, adaptándose a otros paradigmas con propósitos de implementarse en BCI's.



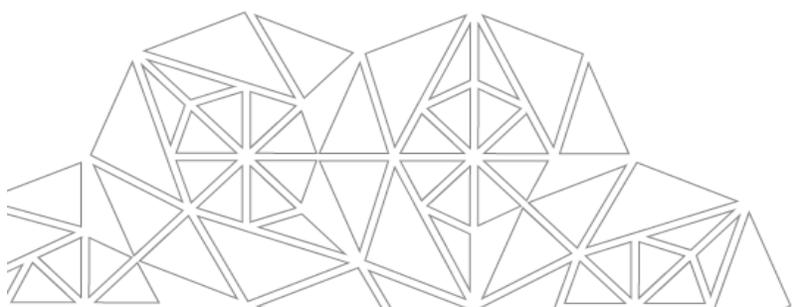
## Anexos

**Anexo 1** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, IS, 32 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=rbf)	34.12	(gini, 0.7, 50)	<b>36.66</b>
S2	(gamma=100, kernel=rbf)	33.70	(gini, 0.7, 100)	<b>36.72</b>
S3	(gamma=10, kernel=sigmoid)	33.33	(gini, 0.7, 50)	<b>36.03</b>
S4	(gamma=10, kernel=rbf)	36.34	(gini, 0.5, 100)	<b>36.82</b>
S5	(gamma=10, kernel=sigmoid)	<b>33.65</b>	(entropy, 0.7, 50)	33.33
S6	(gamma=100, kernel=rbf)	33.96	(entropy, 0.7, 100)	<b>36.98</b>
S7	(gamma=10, kernel=rbf)	<b>35.39</b>	(gini, 0.7, 100)	35.23
S8	(gamma=100, kernel=rbf)	33.49	(entropy, 0.7, 50)	<b>33.96</b>
$\mu \pm \sigma$	34.25 $\pm$ 1.06		<b>35.72 <math>\pm</math> 1.40</b>	

**Anexo 2** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto con bior 3.7, +VEP, 32 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=rbf)	33.97	(gini, 0.7, 100)	<b>36.68</b>
S2	(gamma=10, kernel=rbf)	<b>38.70</b>	(entropy, 0.5, 50)	36
S3	(gamma=10, kernel=rbf)	36.35	(entropy, 0.7, 50)	<b>36.67</b>
S4	(gamma=100, kernel=rbf)	33.49	(entropy, 0.5, 100)	<b>37.68</b>
S5	(gamma=10, kernel=rbf)	33.34	(gini, 0.7, 50)	<b>34.61</b>
S6	(gamma=10, kernel=rbf)	<b>36.2</b>	(entropy, 0.5, 50)	35.89
S7	(gamma=10, kernel=rbf)	35.73	(gini, 0.7, 200)	<b>40.04</b>
S8	(gamma=100, kernel=rbf)	<b>33.65</b>	(gini, 0.5, 100)	33.01
$\mu \pm \sigma$	35.18 $\pm$ 1.89		<b>36.32 <math>\pm</math> 2.07</b>	

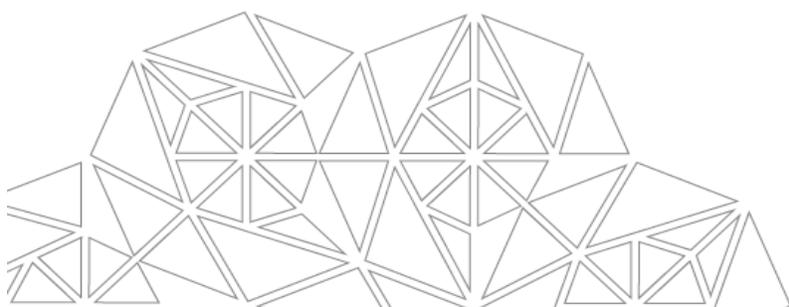


**Anexo 3** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con db7, IS, 32 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=sigmoid)	33.33	(entropy, 0.5, 50)	<b>36.82</b>
S2	(gamma=10, kernel=rbf)	34.01	(entropy, 0.5, 50)	<b>36.4</b>
S3	(gamma=10, kernel=rbf)	34.28	(entropy, 0.5, 50)	34.28
S4	(gamma=10, kernel=sigmoid)	33.33	(91ini, 0.5, 200)	<b>37.3</b>
S5	(gamma=100, kernel=rbf)	<b>33.80</b>	(entropy, 0.5, 100)	33.49
S6	(gamma=10, kernel=rbf)	35.71	(91ini, 0.5, 100)	<b>36.82</b>
S7	(gamma=100, kernel=rbf)	33.49	(91ini, 0.5, 50)	<b>34.92</b>
S8	(gamma=10, kernel=rbf)	33.80	(91ini, 0.7, 200)	<b>36.03</b>
$\mu \pm \sigma$	33.96 $\pm$ 0.77		<b>35.76 <math>\pm</math> 1.37</b>	

**Anexo 4** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con db7, +VEP, 32 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=sigmoid)	33.33	(entropy, 0.5, 50)	<b>34.62</b>
S2	(gamma=10, kernel=rbf)	<b>39.05</b>	(entropy, 0.5, 50)	34.19
S3	(gamma=10, kernel=rbf)	<b>36.02</b>	(entropy, 0.7, 50)	34.76
S4	(gamma=10, kernel=rbf)	33.33	(entropy, 0.7, 50)	<b>37.49</b>
S5	(gamma=10, kernel=sigmoid)	33.33	(gini, 0.7, 200)	<b>35.26</b>
S6	(gamma=100, kernel=rbf)	33.65	(entropy, 0.5, 50)	<b>35.11</b>
S7	(gamma=10, kernel=rbf)	33.80	(gini, 0.7, 50)	<b>37.96</b>
S8	(gamma=100, kernel=rbf)	33.97	(entropy, 0.5, 50)	<b>34.92</b>
$\mu \pm \sigma$	34.56 $\pm$ 2.02		<b>35.53 <math>\pm</math> 1.39</b>	

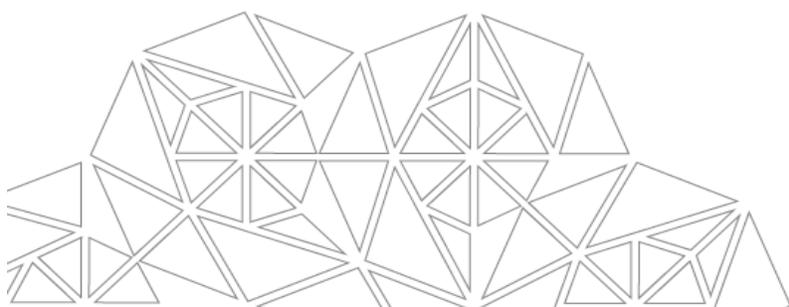


**Anexo 5** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con bior 3.7, IS, 16 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=sigmoid)	33.33	(gini, 0.5, 200)	<b>36.19</b>
S2	(gamma=100, kernel=rbf)	33.86	(gini, 0.5, 100)	<b>37.84</b>
S3	(gamma=100, kernel=rbf)	33.80	(entropy, 0.5, 200)	<b>36.66</b>
S4	(gamma=10, kernel=rbf)	<b>37.46</b>	(gini, 0.7, 50)	37.30
S5	(gamma=100, kernel=sigmoid)	33.33	(gini, 0.7, 100)	<b>34.44</b>
S6	(gamma=10, kernel=sigmoid)	33.80	(entropy, 0.5, 100)	<b>35.71</b>
S7	(gamma=100, kernel=rbf)	34.44	(gini, 0.5, 50)	<b>34.92</b>
S8	(gamma=100, kernel=rbf)	33.49	(entropy, 0.7, 100)	<b>33.65</b>
$\mu \pm \sigma$	34.19 $\pm$ 1.36		<b>35.84 <math>\pm</math> 1.44</b>	

**Anexo 6** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con bior 3.7, +VEP, 16 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=rbf)	34.29	(entropy, 0.5, 50)	<b>36.19</b>
S2	(gamma=10, kernel=sigmoid)	<b>37.27</b>	(entropy, 0.7, 100)	35.10
S3	(gamma=100, kernel=rbf)	34.92	(entropy, 0.7, 100)	<b>37.00</b>
S4	(gamma=10, kernel=rbf)	35.58	(entropy, 0.7, 50)	<b>37.16</b>
S5	(gamma=100, kernel=rbf)	33.65	(gini, 0.7, 50)	<b>33.96</b>
S6	(gamma=10, kernel=rbf)	<b>37.79</b>	(gini, 0.7, 100)	34.59
S7	(gamma=10, kernel=rbf)	33.66	(gini, 0.5, 200)	<b>38.28</b>
S8	(gamma=100, kernel=rbf)	<b>34.44</b>	(gini, 0.5, 100)	33.80
$\mu \pm \sigma$	35.20 $\pm$ 1.57		<b>35.76 <math>\pm</math> 1.64</b>	

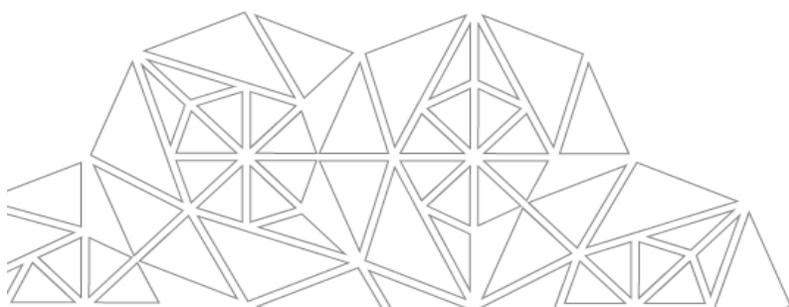


**Anexo 7** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con db7, IS, 16 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=sigmoid)	33.33	(entropy, 0.5, 50)	<b>36.03</b>
S2	(gamma=10, kernel=sigmoid)	33.38	(gini, 0.5, 200)	<b>37.37</b>
S3	(gamma=100, kernel=rbf)	<b>34.60</b>	(gini, 0.5, 200)	34.28
S4	(gamma=10, kernel=rbf)	33.65	(entropy, 0.7, 100)	<b>37.30</b>
S5	(gamma=100, kernel=rbf)	33.96	(gini, 0.5, 100)	<b>34.28</b>
S6	(gamma=10, kernel=rbf)	36.66	(entropy, 0.5, 50)	<b>37.61</b>
S7	(gamma=10, kernel=rbf)	<b>34.12</b>	(entropy, 0.7, 200)	33.65
S8	(gamma=10, kernel=rbf)	34.44	(entropy, 0.5, 100)	<b>35.71</b>
$\mu \pm \sigma$	34.27 $\pm$ 1.07		<b>35.78 <math>\pm</math> 1.57</b>	

**Anexo 8** Mejor combinación de hiperparámetros para el porcentaje de exactitud por sujeto  
con db7, +VEP, 16 ch

Sujeto	SVM		RF	
S1	(gamma=10, kernel=rbf)	<b>35.08</b>	(gini, 0.5, 100)	33.83
S2	(gamma=10, kernel=rbf)	<b>37.45</b>	(gini, 0.7, 100)	35.45
S3	(gamma=100, kernel=rbf)	<b>35.56</b>	(gini, 0.7, 50)	34.77
S4	(gamma=10, kernel=rbf)	<b>35.88</b>	(entropy, 0.7, 50)	35.24
S5	(gamma=10, kernel=rbf)	33.49	(gini, 0.5, 50)	<b>35.58</b>
S6	(gamma=100, kernel=rbf)	<b>33.96</b>	(gini, 0.7, 100)	33.81
S7	(gamma=100, kernel=rbf)	34.6	(gini, 0.7, 200)	<b>37.33</b>
S8	(gamma=100, kernel=rbf)	33.97	(gini, 0.7, 100)	<b>35.56</b>
$\mu \pm \sigma$	34.99 $\pm$ 1.29		<b>35.19 <math>\pm</math> 1.12</b>	



## Referencias

Aguirre-Cruz, Sánchez, L., Landero, A. I. & Rendón, A., 2022. Bases Neuroanatómicas del Lenguaje y Afasia Secundaria a Enfermedad Vascul ar Cerebral.. *Ictus*, 3(3), pp. e03132203013-e03132203013.

Aler, R., Galván, I. M. & Valls, J. M., 2012. Applying evolution strategies to preprocessing EEG signals for brain–computer interfaces. *Information Sciences*, Volumen 215, pp. 53-66.

Ali, M., Al Machot, F., Mosa, A. & Kyamakya, K., 2016. A novel EEG-based emotion recognition approach for e-healthcare applications.. *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, pp. 162-164.

Almazán Cuéllar, S., 2004. Tesis "Algoritmos de Dos Pasos para su uso con Técnicas de Desplazamiento de Fase". León, Guanajuato: s.n.

Amat Rodrigo, J., 2020. *Random Forest con Python*. [En línea]  
Available at: [https://www.cienciadedatos.net/documentos/py08\\_random\\_forest\\_python.html](https://www.cienciadedatos.net/documentos/py08_random_forest_python.html)

Ardila, A., Bernal, B. & Rosselli, M., 2016. Área cerebral del lenguaje: una reconsideración funcional.. *Rev Neurol*, 62(3), pp. 97-106.

Asghari Bejestani, M. R., Khani, M., Nafisi, V. R. & Darakeh, F., 2022. Eeg-based multiword imagined speech classification for Persian words. *BioMed Research International*, Volumen 2022.

Auria, L. & Moro, R. A., 2008. Support Vector Machines (SVM) as a technique for solvency analysis.. Issue 811.

Bakhshali, M. A., Khademi, M., Ebrahimi-Moghadam, A. & Moghimi, S., 2020. EEG signal classification of imagined speech based on Riemannian distance of correntropy spectral density. *Biomedical Signal Processing and Control*, Volumen 59, p. 101899.

Bansal, D. & Mahajan, R., 2019. *Design and Implementation of Efficient Digital Filter for Preprocessing of EEG Signals*. s.l., s.n.

Barcala Riveira, J. M. y otros, 2003. *Aplicación de las Wavelets y los Cuaterniones a la Clasificación de Espectros NIR*. s.l.:CIEMAT.

Başaklar, T., Tuncel, Y. & Ider, Y. Z., 2019. Effects of high stimulus presentation rate on EEG template characteristics and performance of c-VEP based BCIs. *Biomedical Physics & Engineering Express*, 5(3), p. 035023.

Bermúdez Cicchino, A. N., 2013. Tesis Técnicas de procesamiento de EEG para detección de eventos.

Bisht, A., Kaur, C. & Singh, P., 2020. Recent Advances in Artifact Removal Techniques for EEG Signal Processing. *Intelligent Communication, Control and Devices: Proceedings of ICICCD 2018*, pp. 385-392.

Blanco-Díaz, C. F. & Ruiz-Olaya, A. F., 2019. Caracterización de señales de EEG relacionadas a potenciales evocados visuales en estado estacionario.. *Revista Ontare*, Volumen 7, pp. 11-26.



Brigham, K. & Kumar, B., 2010. *Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy*. s.l., IEEE, pp. 1-4.

Brumberg, J. S., Nieto-Castanon, A., Kennedy, P. R. & Guenther, F. H., 2010. Brain-computer interfaces for speech communication. *Speech Communication*, 52(4), pp. 367-379.

Cooney, C., Korik, A., Folli, R. & Coyle, D., 2020. Evaluation of hyperparameter optimization in machine and deep learning methods for decoding imagined speech EEG. *Sensors*, 20(16), p. 4629.

Cooney, C., Korik, A., Raffaella, F. & Coyle, D., 2019. *Classification of imagined spoken word-pairs using convolutional neural networks*. s.l., s.n., pp. 338-343.

Denby, B. y otros, 2010. Silent Speech Interfaces.. *Speech Communication*, 52(4), pp. 270-287.

Domingues, M. O. y otros, 2016. Explorando a transformada wavelet contínua. Exploring the continuous wavelet transform. *Revista Brasileira de Ensino de Física*, 38(3).

D'Zmura, M. y otros, 2009. Toward EEG Sensing of Imagined Speech. *Human-Computer Interaction, Part I, HCII*, pp. 40-48.

Elías Díaz, Á. G., 2020. *Aplicaciones de la Descomposición Empírica Modal al Diseño de Estrategias de Inversión en Mercados Financieros*. Ciudad de México: Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Unidad Zacatenco.

Elsayed, N., Saad, Z. & Bayoumi, M., 2017. Brain Computer Interface: EEG Signal Preprocessing Issues and Solutions. *International Journal of Computer Applications*, 169(3).

Fiedler, P. y otros, 2022. A high-density 256-channel cap for dry electroencephalography.. *Human brain mapping by Wiley Periodicals*, 43(4), pp. 1295-1308.

Gandhi, V. y otros, 2013. Quantum neural network-based EEG filtering for a brain-computer interface.. *IEEE transactions on neural networks and learning systems*, 25(2), pp. 278-288.

Gao, Z. y otros, 2020. Classification of EEG signals on VEP-based BCI systems with broad learning. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(11), pp. 7143-7151.

García de Sola, R., 2014. *Anatomía y función de la corteza cerebral humana. Áreas de Brodman*, s.l.: Blog del Dr. García de Sola.

González Castañeda, E. F., 2015. *Sonificación de EEG para la clasificación de habla imaginada*. Tonantzintla, Puebla: INAOE.

González González, R. A., 2010. Tesis Algoritmo basado en Wavelets aplicado a la detección de incendios forestales. En: *Repositorio Nacional CONACyT*. s.l.:s.n.

Gonzalez-Lopez, J. A. y otros, 2020. Silent speech interfaces for speech restoration: A review.. *IEEE access*, Volumen 8, pp. 177995-178021.

Guadarrama Rendón, Y. S., 2018. *Tesis "Procesamiento de señales EEG para el reconocimiento de voz imaginada"*. México: Instituto Politécnico Nacional - Centro de Investigación en Computación.



Guerrero, J. A., 2016. El problema de la dimensionalidad. *Indice: Revista de Estadística y Sociedad*, Issue 68, pp. 22-24.

Gutiérrez, J., 2001. Análisis de señales en el neuromonitoreo. *Revista Mexicana de Ingeniería Biomédica*, 22(2), pp. 67-77.

Herculano Houzel, S. & Neurosci, H., 2009. The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. *Frontiers in human neuroscience*, p. 31.

Hesslow, G., 2002. Conscious thought as simulation of behaviour and perception. *TRENDS in Cognitive Sciences*, 6(6), pp. 242-247.

Hu, L. & Zhang, Z., 2019. *EEG Signal Processing and Feature Extraction*. s.l.:Springer.

IBM, s.f. *What is random forest?*. [En línea]

Available at: <https://www.ibm.com/topics/random-forest>

Ilyas, M. Z., Saad, P. & Ahmad, M. I., 2015. *A Survey of Analysis and Classification of EEG Signals for Brain-Computer Interfaces*. s.l., s.n.

INEGI, 2020. *Censo de Población y Vivienda*, s.l.: s.n.

Islam, M. K., Rastegarnia, A. & Yang, Z., 2016. Methods for artifact detection and removal from scalp EEG: A review.. *Neurophysiologie Clinique/Clinical Neurophysiology*, 46(4-5), pp. 287-305.

James, G., Witten, D., Hastie, T. & Tibshirani, R., 2021. *An Introduction to Statistical Learning with Applications in R*. Segunda ed. New York: Springer.

Jarvis, E. D., Güntürkün, O. & Bruce, L., 2005. Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2), pp. 151-159.

Jeunet, C. y otros, 2019. Using EEG-based brain computer interface and neurofeedback targeting sensorimotor rhythms to improve motor skills: Theoretical background, applications and prospects. *Neurophysiologie Clinique*, 49(2), pp. 125-136.

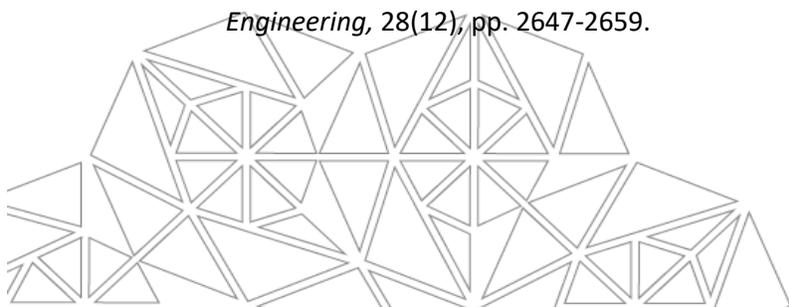
Kandel, E. R., 2006. *Principles of neural science*. Quinta ed. s.l.:Appleton and Lange: McGraw-Hill.

Klug, M. & Gramann, K., 2021. Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *European Journal of Neuroscience*, 54(12), pp. 8406-8420.

Lee, M. H. y otros, 2019. EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy.. *GigaScience*, 8(5).

Lee, S. H., Lee, M., Jeong, J. H. & Lee, S. W., 2019. *Towards an EEG-based intuitive BCI communication system using imagined speech and visual imagery*.. s.l., s.n., pp. 4409-4414.

Lee, S. H., Lee, M. & Lee, S. W., 2020. Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(12), pp. 2647-2659.



Lee, S. H., Lee, Y. E. & Lee, S. W., 2022 . *Toward imagined speech based smart communication system: potential applications on metaverse conditions*. s.l., IEEE, pp. 1-4.

Malete, T. N., Moruti, K., Thapelo, T. S. & Jamisola, R. S., 2019. *EEG-based control of a 3D game using 14-channel Emotiv Epoc+*. s.l., IEEE, pp. 463-468.

Martínez Peón, D. C., 2016. *Tesis Comando Robusto en Interfaz Cerebro-Computadora basada en EEG para Robots Dinámicos*. Apodaca, Nuevo León: s.n.

Maton, A. y otros, 1993. En: *Human Biology and Health*. Englewood Cliffs, New Jersey, USA: Prentice Hall, pp. 132-144.

Medina Sánchez, L. V., 2017. *Trabajo de investigación Análisis de técnicas Wavelet para el desarrollo de compresores de audio*. s.l.:s.n.

Mejía, G., Quintero, O. & Montoya, A. G., 2016. Reconocimiento de Emociones utilizando la Transformada Wavelet Estacionaria en Señales EEG Multicanal. En: Medellín, Colombia: Universidad EAFIT.

Melara Estrada, M. A., 2015. *Introducción a la transformada wavelet y la teoría de análisis de señales*. s.l.:s.n.

Moctezuma Pascual, L. A., 2017. *Distinción de estados de actividad e inactividad lingüística para interfaces cerebro computadora*. Puebla de Zaragoza: Benemérita Universidad Autónoma de Puebla (BUAP).

Moctezuma, L. A., Carrillo, M., Villaseñor-Pineda, L. & Torres-García, A. A., 2017. Hacia la clasificación de actividad e inactividad lingüística a partir de señales de electroencefalogramas (EEG). *Research in Computing Science 140*, Volumen 140, pp. 135-149.

Moctezuma, L. A., Torres-García, A. A., Villaseñor-Pineda, L. & Carrillo, M., 2019. Subjects identification using EEG-recorded imagined speech. *Expert Systems with Applications*, Volumen 118, pp. 201-208.

Montoya-Martínez, J., Vanthornhout, J., Bertrand, A. & Francart, T., 2021. Effect of number and placement of EEG electrodes on measurement of neural tracking of speech.. *Plos one*, 16(2), p. e0246769.

Muñoz Sánchez, F. J., 2014. *Trabajo "Construcción de una nueva Interfaz Cerebro-Computador a partir de un bajo coste (Emotiv Epoc)"*. s.l.:s.n.

Navarro, B. & Torres, A., 2022. *Áreas de Brodmann*, s.l.: s.n.

Ojeda, E., 2005. Potenciales evocados visuales y electroretinograma. En: *Guía Neurológica 7 Neurofisiología*. s.l.:s.n., pp. 127-134.

Oppenheim, A. V., Willsky, A. S. & Hamid Nawab, S., 1998. *Señales y Sistemas*. s.l.:Prentice Hall.

Orenes Casanova, Y., 2022. Tesis "Contribuciones al problema de clasificación en machine learning". En: s.l.:s.n.



Proix, T. y otros, 2022. Imagined speech can be decoded from low-and cross-frequency intracranial EEG features. *Nature communications*, 13(1), p. 48.

Puelles López, L., Martínez Pérez, S. & Martínez de la Torre , M., 2008. *Neuroanatomía*. s.l.:Panamericana.

Purves , D., Augustine, G. J. & Fitzpatrick, D., 2004. *Neuroscience*. s.l.:Sinauer Editors.

Ramírez Diez, F. M., 2017. *Tesis Uso de características espectrales y temporales para clasificación de tareas mentales en señales de electroencefalografía*. s.l.:s.n.

Ranganathan , G., 2021. A study to find facts behind preprocessing on deep learning algorithms.. *Journal of Innovative Image Processing (JIIP)*, 3(1), pp. 66-74.

Rohde, C. G., 2006. *Afasia: Historia, Teoría y Examen Clínico*. s.l.:s.n.

Romero , S., Mañanas, M. Á. & Barbanoj , M., 2008. Estudio Comparativo de Técnicas Automáticas para el filtrado ocular en señales EEG simuladas.

Rufiner Di Persia, H. L., 1996. *Tesis Comparación entre análisis wavelets y Fourier aplicados al reconocimiento automático del habla*. s.l.:s.n.

Sagan, C., 1977. *The dragons of Eden: speculations on the evolution of human intelligence*. New York, USA: Random House.

Saha, P., Abdul-Mageed, M. & Fels, S., 2019. Speak your mind! towards imagined speech recognition with hierarchical deep learning. *arXiv preprint arXiv:1904.05746*.

Santos, M., Farías, G., Fernández Marron, J. L. & Dormido Canto, S., 2004. Determinación de Parámetros de la Transformada Wavelet para la Clasificación de Señales del Diagnóstico Scattering Thomson.

Sayilgan, E., Yüce, Y. K. & İŞLER, Y., 2021. Evaluation of mother wavelets on steady-state visually-evoked potentials for triple-command brain-computer interfaces. *Turkish Journal of Electrical Engineering and Computer Sciences*, 29(5), pp. 2263-2279.

Science (DBCLS), D. C. f. L., s.f. [En línea]

Available at: [De Polygon data were generated by Database Center for Life Science\(DBCLS\)\[2\]. -](#)

[Polygon data are from BodyParts3D\[1\]., CC BY-SA 2.1 jp,](#)

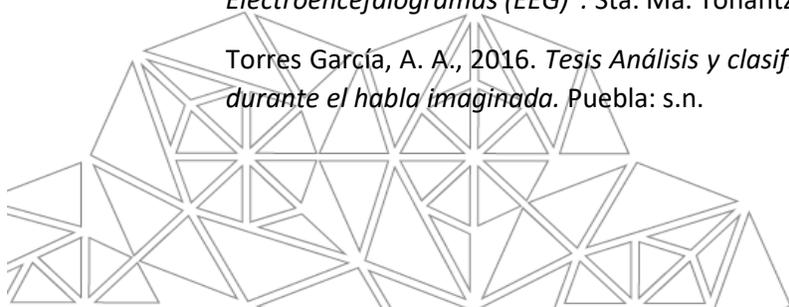
<https://commons.wikimedia.org/w/index.php?curid=43314531>

Sreeshakthy, M., Preethi, J. & Dhilipan, A., 2016. A survey on emotion classification from EEG signal using various techniques and performance analysis.. *International Journal of Information Technology and Computer Science (IJITCS)*, 8(12), pp. 19-26.

Testut, L. & Latariet, A., s.f. *Tratado de anatomía humana*. Barcelona: Salvat Editores.

Torres García, A. A., 2011. *Tesis "Clasificación de palabras no pronunciadas presentes en Electroencefalogramas (EEG)"*. Sta. Ma. Tonantzintla(Puebla): INAOE.

Torres García, A. A., 2016. *Tesis Análisis y clasificación de electroencefalogramas (EEG) registrados durante el habla imaginada*. Puebla: s.n.



Torres García, A. A., Reyes García, C. A. & Villaseñor Pineda, L., 2011. *Hacia la clasificación de habla no pronunciada mediante electroencefalogramas (EEG)*. Ixtapa-Zihuatanejo, Guerrero, México, s.n.

Torres García, A. A., Reyes-García, C. A., Villaseñor-Pineda, L. & Ramírez-Cortés, J. M., 2013. Análisis de Señales Electroencefalográficas para la Clasificación de Habla Imaginada. *Revista Mexicana de Ingeniería Biomédica*, 34(1), pp. 23-39.

UNAM, F. d. M., 2019. *Manual Sistema 10/20 Internacional*. [En línea] Available at: <http://liceaga.facmed.unam.mx/deptos/fisiologia/wp-content/uploads/2019/09/UTI-pr%C3%A1ctica-7-a.-Electroencefalograma.-AnexoManual.pdf>

Vallabhaneni, A., Wang, T. & He, B., 2005. Brain-Computer Interface. *Neural engineering*, pp. 85-121.

Wolpaw, J. R. y otros, 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, Volumen 113, p. 767-791.

Wu, J. D., Hsu, C. C. & Wu, G. Z., 2009. Fault gear identification and classification using discrete wavelet transform and adaptive neuro-fuzzy inference.. *Expert Systems with Applications*, 36(3, Part 2), pp. 6244-6255.

Xu, X., Zhang, Y. & Chen, X., 2019. Removal of Muscle Artifacts From the EEG: A Review and Recommendations. *IEEE Sensors Journal*, 19(14).

Yazbek, S., Hage, S., Mallak, I. & Smayra, T., 2021. Tractography of the arcuate fasciculus in healthy right-handed and left-handed multilingual subjects and its relation to language lateralization on functional MRI.. *Scientific Reports*, 11(1), p. 20936.

Yosrita, E., Heryadi, Y. & Budiharto, W., 2020. Words classifier of imagined speech based on EEG for patients with disabilities.. *ICIC Express Letters*, 14(1), pp. 37-41.

Yosrita, E., Nur Aziza, R., Farah Ningrum, R. & Muhammad, G., 2021. Denoising of eeg signal based on word imagination using ica for artifact and noise removal on unspoken speech.. *Indonesian Journal of Electrical Engineering and Computer Science*, 22(1), pp. 83-88 .

Yuan, Y., Xun, G., Jia, K. & Zhang, A., 2017. *A multi-view deep learning method for epileptic seizure detection using short-time fourier transform*. s.l., s.n.

Yue, L. y otros, 2020. *A brain-computer interface based on high-frequency steady-state asymmetric visual evoked potentials*. s.l., s.n., pp. 3090-3093.

Zhang, X. y otros, 2019. A convolutional neural network for the detection of asynchronous steady state motion visual evoked potential. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(6), pp. 1303-1311.

