

Distinción de estados de actividad e inactividad lingüística para interfaces cerebro computadora



BUAP

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN

TESIS PARA OBTENER EL TÍTULO DE
MAESTRO EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA

Luis Alfredo Moctezuma Pascual

ASESORES

Dra. Maya Carrillo Ruiz

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Dr. Alejandro Antonio Torres García

INSTITUTO NACIONAL DE ASTROFÍSICA, ÓPTICA Y ELECTRÓNICA

CUATRO VECES HEROICA PUEBLA DE ZARAGOZA.

JULIO 2017

S*in duda no hay progreso.*
- *Charles Robert Darwin*

Dedico este trabajo a mi familia, porque siempre me han ayudado a dar el primer paso. Por enseñarme el valor de la vida y por mostrarme, con el ejemplo, que cualquier cosa se puede lograr. A mis hermanos: Alex, Arturo, Gaby, Eduardo y Susana.

Agradecimientos

A mis asesores (Dra. Maya Carrillo y a los Doctores Luis Villaseñor y Alejandro Torres). A ellos, mi admiración y respeto.

Al Consejo Nacional de Ciencia y Tecnología por apoyarme con una beca para realizar mis estudios de Maestría. Al posgrado en ciencias de la computación de la Benemérita Universidad Autónoma de Puebla, por darme todas las facilidades y por prestarme las instalaciones.

Resumen

La capacidad de intercambiar mensajes entre dos o más personas es bien conocida, para ello es necesario que exista un sistema compartido de signos y normas semánticas. Comúnmente se pueden combinar algunas formas de comunicación, por ejemplo, alguien que utilice el lenguaje oral puede comunicarse con alguien que sólo utilice el lenguaje de señas. Sin embargo, un porcentaje significativo de la población mundial, no puede utilizar los modelos tradicionales de comunicación. En el caso de México, se sabe por el Instituto Nacional de Estadística y Geografía que el 6% de la población sufre alguna discapacidad, de las cuales el 18% de ellas se relaciona con discapacidades para hablar o comunicarse, entre ellas la esclerosis lateral amiotrófica, esclerosis múltiple y lesiones de médula espinal o cerebral. Las personas que sufren dichas discapacidades tienen varios riesgos y limitaciones al tratar de convivir día a día con las personas que los rodean, no pueden desempeñar la mayoría de las tareas y no pueden comunicar lo que piensan o sienten.

Una respuesta ampliamente aceptada, es el uso de interfaces que conectan al cerebro humano con una computadora, por sus siglas en inglés, BCI. Las BCI no invasivas basadas en electroencefalogramas (EEG) son las más utilizadas debido a su costo relativamente bajo. Un tipo de BCI, son las conocidas como interfaces de habla silenciosa, cuya finalidad es desarrollar sistemas capaces de permitir la comunicación hablada cuando es imposible emitir una señal acústica audible. Para lograr esto se utiliza un mecanismo neurológico que consiste en imaginar la dicción de las palabras, comúnmente conocido como habla imaginada, para luego identificar de que palabra se trataba con el uso de métodos computacionales. La tarea de identificar las palabras imaginadas ha sido abordada como un problema

de clasificación, y para tratar de resolverlo se han buscado formas de caracterización de las señales y algoritmos que permitan realizar dicha clasificación, pues resulta poco práctico hacer uso de toda la información generada por el cerebro humano.

En este trabajo se evaluaron dos formas de caracterización de las señales de EEG para identificar habla imaginada. Luego las señales de EEG fueron separadas en 2 clases, un conjunto de palabras imaginadas consideradas la clase de *actividad lingüística*; y los estados de reposo o pausa como ejemplos de otra actividad cerebral. Estos últimos denominados por simplicidad de *inactividad lingüística*. Así, caracterizar e identificar cuando se trata de actividad e inactividad lingüística, pues esto es un preproceso indispensable para el desarrollo de una BCI basada en habla imaginada.

La idea de hacer esta clasificación es orientar los trabajos hacia detectar habla imaginada en tiempo real, pues un primer proceso debe encontrar cuando se trata de actividad o inactividad lingüística y posteriormente otro proceso identificar la palabra imaginada. La primera forma de caracterización utilizó el cálculo de la distribución de la energía instantánea, relativa, jerárquica y teager con base en la transformada discreta wavelet (DWT). La segunda caracterización se basó en el cálculo de valores estadísticas directamente sobre la señal. Para la etapa de clasificación se usó *random forest*, *SVM* y *naive Bayes*. Los experimentos se realizaron sobre dos bases de datos de habla imaginada, una con 27 sujetos y otra con 20. Los resultados muestran que al usar los valores estadísticos y el algoritmo de *random forest*, la actividad e inactividad lingüística se distingue mejor que al usar la DWT; sin embargo, cabe señalar que con ambas caracterizaciones se alcanzan altas tasas de exactitud.

Los experimentos expuestos fueron realizados fuera de línea, sin embargo para realizar una BCI en línea es necesario contar con respuestas rápidas y procesos precisos para identificar cuando inicia y termina la actividad e inactividad lingüística. Este trabajo analizó si en segmentos pequeños de la señal de EEG existe información que permita distinguir cuando se trata de actividad e inactividad lingüística, al utilizar diferentes tamaños de ventana, decreciendo su tamaño, la exactitud varía. Con ventanas pequeñas se podrá identificar en un menor tiempo el inicio y fin de la actividad lingüística. Es así que será posible el reconocimiento de habla imaginada en tiempo real.

Abstract

The ability to exchange messages between two or more people is well-known, for this it is necessary that exist a shared system of signs and semantic rules. Commonly you can combine some ways of communication, for example, someone who uses oral language can communicate with someone who only uses sign language. However, a significant percentage of the world's population cannot use traditional communication models. In Mexico case, it is known by the National Institute of Statistics and Geography that 6 % of population suffers a disability, of which 18 % of them are related to disabilities to speak or communicate, including amyotrophic lateral sclerosis, multiple sclerosis, and spinal cord or brain lesions. People who suffer from these disabilities have several risks and limitations when trying to live with people around them, they cannot do most of the tasks and cannot communicate what they think or feel.

One widely accepted answer is the use of interfaces that connect the human brain to a computer (BCI). Non-invasive BCI based on electroencephalograms (EEG) are the most used because of their relatively low cost. A type of BCI are known as silent speech interfaces, whose purpose is to develop systems capable of enabling spoken communication when it is impossible to emit an audible acoustic signal. To achieve this, a neurological mechanism is used which consists in imagining the diction of words, commonly known as imagined speech, and then identifying which word was treated with the use of computational methods. The task of identifying the imagined words has been approached as a classification problem, and to try to solve it we have looked for ways of characterizing the signals and algorithms that allow to make such classification, as it would be impractical and possibly useless to make use of all the information

generated by human brain.

In this work two ways of characterizing EEG signals to identify imagined speech were evaluated. Then EEG signals were separated into 2 classes, a set of imagined words considered the class of *linguistic activity*; And states of rest or pause as examples of other brain activity. The latter for simplicity we call them of *linguistic inactivity*. Thus, characterize and identify when it comes to activity and linguistic inactivity, as this is an indispensable preprocess for the development of a BCI based on imagined speech.

The idea of doing this classification is to guide the work towards detecting imagined speech in real time, because a first process must find when it is a question of activity or linguistic inactivity and then another process to identify the imagined word. The first characterization way used the calculation of the distribution of instantaneous energy, relative, hierarchical and teager based on the discrete wavelet transform (DWT). The second characterization was based on the calculation of statistical values directly from the signal. For the classification step, *random forest*, *SVM*, and *naive Bayes* were used. The experiments were performed on two imagined speech databases, first one with 27 subjects and the other one with 20. The results show that when using the statistical values and *random forest* algorithm, activity and linguistic inactivity are distinguished better than when using the DWT; however, it should be noted that with both characterizations high accuracy rates are reached.

The experiments exposed were performed offline, however to perform an online BCI it is necessary to have fast responses and precise processes to identify when beginning and end the activity and linguistic inactivity. This work analyzed whether in small segments of the EEG signal there is information that allows distinguishing when it comes to activity and linguistic inactivity, when using different window sizes, decreasing its size, accuracy varies. With small windows it will be possible to identify in a shorter time the beginning and end of the linguistic activity. It will be thus possible to recognize imagined speech in real time.

Índice general

Resumen	VII
Abstract	IX
Índice de figuras	1
Índice de tablas	3
1. Introducción	7
1.1. Problemática	8
1.2. Objetivo	9
1.3. Objetivos específicos	9
2. Marco teórico	11
2.1. El cerebro	11
2.1.1. El cerebro y el lenguaje	12
2.2. Electroencefalograma	14
2.3. Interfaces cerebro-computadora	15
2.4. Procesamiento digital de señales	17
2.4.1. Transformada discreta wavelet	17
2.4.2. Distribución de la energía	20
2.5. Valores estadísticos	22
2.6. Clasificación	24
2.6.1. Random Forest	25
2.6.2. Máquina de vectores de soporte	26
2.6.3. Naive Bayes	27

2.6.4. Validación cruzada	28
2.7. Medidas de desempeño para los clasificadores	29
3. Estado del arte	31
4. Método propuesto	37
4.1. Infraestructura	38
4.1.1. Primera base de datos	38
4.1.2. Segunda base de datos	39
4.2. Pre-procesamiento	43
4.3. Extracción de características	43
4.3.1. Transformada discreta wavelet	44
4.3.2. Medidas estadísticas	45
4.4. Clasificación	46
5. Experimentos y resultados	47
5.1. Clasificación de palabras imaginadas	47
5.2. Clasificación actividad e inactividad lingüística	49
5.2.1. Análisis sobre el número de canales y frecuencia de muestreo.	52
5.3. División en segmentos de señal de EEG que aún contienen información lingüística	53
5.4. Identificando periodos de actividad e inactividad lingüística de nuevos sujetos	56
5.4.1. Análisis sobre el número de canales y frecuencia de muestreo en transferencia de clasificador	58
6. Conclusiones y trabajo futuro	61
6.1. Trabajo futuro	63
A. Resultados a detalle de los experimentos	67
Referencias	89

Índice de figuras

2.1. Partes del cerebro.	12
2.2. Áreas corticales, mapa de Brodmann [35].	14
2.3. Modelo funcional de un sistema BCI no invasivo, adaptado de [3].	16
2.4. Funcionamiento de DWT [26]	18
2.5. Funcionamiento de DWT [26] en el nivel de descomposición j	19
2.6. Árbol creado en el nivel de descomposición $j=3$	19
2.7. Wavelet biortogonal 2.2.	20
2.8. Hiperplano óptimo (CO) que maximiza el margen entre los hiperplanos H1 y H2.	27
4.1. Modelo general del método.	37
4.2. Localización de 14 electrodos de acuerdo al sistema internacional 10-20 para EEG.	38
4.3. Diseño del protocolo para adquisición de la señal de EEG para la BD1 usando EMOTIV EPOC.	39
4.4. Localización de 64 electrodos de acuerdo al sistema internacional 10-20 para EEG.	41
4.5. Diseño del protocolo para estimulación y adquisición de la señal de EEG para la BD2 usando Brain Vision.	42
5.1. Promedios de exactitud con los tres clasificadores para los individuos de la BD1, características DWT bior2.2 con 4 niveles de descomposición y estadísticos (EST) para habla imaginada.	48

5.2.	Promedios de exactitud con los tres clasificadores para los individuos de la BD2 en habla imaginada.	49
5.3.	Promedios de exactitud con los tres clasificadores usando la BD1 para distinguir actividad e inactividad lingüística. . . .	51
5.4.	Promedios de exactitud con los tres clasificadores usando la BD2 para distinguir actividad e inactividad lingüística. . . .	51
5.5.	Promedios de exactitud con los tres clasificadores usando la BD1 para distinguir actividad e inactividad lingüística. Considerando 14 canales y una frecuencia de muestreo de 128 Hz.	53
5.6.	Procedimiento para separar las instancias de entrenamiento y validación en la BD1.	55
5.7.	Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD1 y BD2.	55
5.8.	Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD1 al validar sujetos que no estaban en el entrenamiento.	57
5.9.	Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD2 al validar sujetos que no estaban en el entrenamiento.	57
5.10.	Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD2 al validar sujetos que no estaban en el entrenamiento. Comparación del uso de 64 y 14 canales, con frecuencias de 500 y 128 Hz.	59
6.1.	Posible proceso para detectar el inicio y final de la actividad cerebral lingüística.	64

Índice de tablas

2.1. Matriz de confusión para dos clases	29
4.1. índice de palabras imaginadas y almacenadas en la BD1. . .	39
4.2. Descripción de los canales almacenados en la BD1.	40
4.3. índice de palabras imaginadas y almacenadas en la BD2. . .	42
4.4. Comparación de las características de las bases de datos usadas.	43
4.5. Rangos de frecuencias(Hz) para el nivel de descomposición 4 de DWT para la BD1 y los niveles de descomposición 5 y 6 para la BD2.	45
A.1. Exactitud, media y desviación estándar (Std) obtenida con <i>random forest</i> para los individuos de la BD1 para habla imaginada.	68
A.2. Exactitud, media y desviación estándar (Std) obtenida con <i>SVM</i> para los individuos de la BD1 para habla imaginada.	69
A.3. Exactitud, media y desviación estándar (Std) obtenida con <i>naive Bayes</i> para habla imaginada.	70
A.4. Exactitud, media y desviación estándar (Std) obtenidos con <i>random forest</i> usando los individuos de la BD2 para habla imaginada.	71
A.5. Exactitud, media y desviación estándar (Std) obtenidos con <i>SVM</i> usando los individuos de la BD2 para habla imaginada.	72
A.6. Exactitud, media y desviación estándar (Std) obtenidos con <i>naive Bayes</i> usando los individuos de la BD2 para habla imaginada.	73

A.7. Exactitud, media y desviación estándar (Std) obtenida con <i>random forest</i> usando los individuos de la BD1 para actividad e inactividad lingüística.	74
A.8. Exactitud, media y desviación estándar (Std) obtenida con <i>SVM</i> usando los individuos de la BD1 para actividad e inactividad lingüística.	75
A.9. Exactitud, media y desviación estándar (Std) obtenida con <i>naive Bayes</i> usando los individuos de la BD1 para actividad e inactividad lingüística.	76
A.10. Exactitud, media y desviación estándar (Std) obtenidos <i>random forest</i> usando los individuos de la BD2 para actividad e inactividad lingüística.	77
A.11. Exactitud, media y desviación estándar (Std) obtenidos con <i>SVM</i> usando los individuos de la BD2 para actividad e inactividad lingüística.	78
A.12. Exactitud, media y desviación estándar (Std) obtenidos con <i>naive Bayes</i> usando los 20 individuos de la BD2 para actividad e inactividad lingüística.	79
A.13. Exactitud, media y desviación estándar (Std) obtenidos con <i>random forest</i> usando los individuos de la BD2 para actividad e inactividad lingüística. Análisis del impacto del número de canales y la frecuencia de muestreo.	80
A.14. Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD1.	81
A.15. Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD2.	82
A.16. Validación de las muestras de tamaño $\frac{1}{2}$ entre sujetos. . . .	83
A.17. Validación de las muestras de tamaño $\frac{1}{16}$ entre sujetos. . . .	84
A.18. Validación de las muestras de tamaño $\frac{1}{2}$ entre sujetos. . . .	85
A.19. Validación de las muestras de tamaño $\frac{1}{16}$ entre sujetos. . . .	86

A.20.Exactitud, media y desviación estándar (Std) obtenidos con <i>random forest</i> usando los 20 individuos de la BD2 para transferencia de clasificador usando actividad e inactividad lingüística con segmentos de $\frac{1}{2}$ de segundo. Análisis del impacto del número de canales y la frecuencia de muestreo.	87
A.21.Exactitud, media y desviación estándar (Std) obtenidos con <i>random forest</i> usando los individuos de la BD2 para transferencia de clasificador usando actividad e inactividad lingüística con segmentos de $\frac{1}{16}$ de segundo. Análisis del impacto del número de canales y la frecuencia de muestreo.	88

Capítulo 1

Introducción

La organización mundial de la salud (OMS) define el término de discapacidad como todas las deficiencias, limitaciones y las restricciones de la participación en situaciones vitales, por ejemplo, la comunicación hablada.

Datos del Instituto Nacional de Estadística y Geografía (INEGI)¹ del 2015 muestran que el 6% de la población mexicana sufre alguna discapacidad, de las cuales el 18% sufre alguna discapacidad para hablar o comunicarse; entre ellas la esclerosis lateral amiotrófica, esclerosis múltiple, lesiones de médula espinal o cerebral, entre otras. Las personas que sufren dichas discapacidades tienen varios riesgos al tratar de convivir día a día con las personas que las rodean, no pueden desempeñar algunas tareas y no pueden comunicar lo que piensan o sienten.

Una interfaz cerebro-computadora(BCI), es un sistema de comunicación que monitorea la actividad cerebral y traduce determinadas características, correspondientes a las intenciones del usuario, en comandos de control de un dispositivo. Este tipo de sistemas ha motivado diversas investigaciones como las relacionadas con el habla imaginada también referida como habla no pronunciada que se refiere a la pronunciación interna o imaginada de palabras pero sin emitir sonidos ni articular gestos.

Las BCI se pueden dividir en invasivas y no invasivas. Las primeras, requieren de procedimientos quirúrgicos, con ellas se obtienen señales

¹Estadísticas a propósito del día internacional de las personas con discapacidad realizado por el INEGI en México (2015)

cerebrales muy nítidas gracias a que las medidas no son atenuadas por el cráneo y el cuero cabelludo. Por otro lado, las BCI no invasivas no requieren ningún tipo de cirugía, sin embargo, las señales que se obtiene son más débiles. Dentro de las no invasivas las basadas en electroencefalogramas (EEG) son las más utilizadas debido a su costo relativamente bajo. Sin embargo, captan ruido de otras actividades fisiológicas como la actividad cardiaca, la respiración, el movimiento de la lengua, entre otras. Además, también captan ruido proveniente de la corriente alterna y de los electrodos [9].

A los mecanismos neurológicos o procesos empleados por el usuario de la BCI para generar las señales de control, se les denomina fuentes electrofisiológicas. Las más empleadas son imágenes motoras, visuales evocados, corticales lentos y potenciales P300. Para este tipo de BCI existen dos problemas principales: el largo periodo de entrenamiento, pues las fuentes electrofisiológicas son generadas por el usuario de manera poco consiente; y las bajas tasas de comunicación, una sola palabra o menos procesada por minuto, lo que hace imposible la comunicación de manera natural [30].

1.1. Problemática

Los problemas mencionados anteriormente han generado una serie de investigaciones [40, 39, 8, 38, 31] que buscan utilizar los potenciales relacionados con la producción de habla. En estos trabajos la fuente electrofisiológica es el habla no pronunciada, habla interna o habla imaginada que se refiere a imaginar la dicción de una palabra sin emitir ni articular sonidos [44]. El estudio de este tipo de interfaces se conoce como interfaces de habla silenciosa, cuya finalidad es desarrollar sistemas capaces de permitir la comunicación hablada que toman lugar cuando emitir una señal acústica audible es imposible [10]. Hay pocos trabajos en esta área y los que existen manejan vocabularios reducidos o incluso se limitan únicamente al reconocimiento de sílabas o fonemas. Además, el problema de reconocimiento en línea no ha sido abordado y no hay esfuerzos en ese sentido.

El propósito de la presente investigación es analizar si es posible distinguir entre actividad e inactividad cerebral lingüística, esto con el fin

de orientar los trabajos hacia el reconocimiento del inicio y terminación de dicha actividad de manera automática y así avanzar en la implementación de una BCI en línea que sea capaz de reconocer habla imaginada.

Saber en dónde inicia y termina dicha actividad permitirá sentar las bases para el reconocimiento del habla imaginada en tiempo real. Para crear el método mencionado se probaron diferentes formas de obtener características para clasificar una señal de electroencefalograma (EEG) eficientemente, considerando principalmente la transformada discreta de wavelet y medidas estadísticas.

Elegidas las características útiles se experimentó con diferentes mecanismos de segmentación de la señal, variando el tamaño de la ventana para encontrar el segmento más pequeño que aún contiene información que permita clasificar de manera adecuada las señales de EEG para actividad e inactividad lingüística.

De acuerdo a lo expuesto anteriormente, a continuación se enuncian los objetivos de la presente investigación.

1.2. Objetivo

Proponer y evaluar una representación de la señal de EEG con el propósito de distinguir entre actividad e inactividad lingüística para interfaces cerebro-computadora basadas en habla imaginada.

1.3. Objetivos específicos

- Determinar las características apropiadas para clasificación de actividad e inactividad lingüística.
- Definir el tamaño mínimo de segmento para identificar los estados de actividad e inactividad lingüística.

El presente documento, está organizado de la siguiente manera: en el capítulo dos se enuncian los conceptos teóricos que serán utilizados en el resto del documento para posteriormente, en el capítulo tres analizar algunos trabajos relacionados con habla imaginada y mostrar los resultados que obtuvieron, además, de mencionar las técnicas para

preprocesamiento y extracción de características. En el capítulo cuatro se describe detalladamente el método propuesto para el preprocesamiento y obtención de características de una señal para la tarea de clasificación por clase (y multiclase). En el capítulo cinco se muestran los experimentos realizados y los resultados que se obtuvieron en cada uno de ellos, para finalmente dar una conclusión de los resultados obtenidos y el aprendizaje que se obtuvo en la tarea de caracterización y clasificación de habla imaginada. Después de presentar los resultados, se analiza hacia donde se orientan las investigaciones referentes al habla imaginada y las posibles aplicaciones útiles.

Capítulo 2

Marco teórico

En este capítulo se describen brevemente los conceptos básicos que son usados a lo largo del documento y que son necesarios para comprender la parte teórica de los experimentos realizados.

2.1. El cerebro

En los seres humanos, el cerebro pesa entre 1.3 y 1.6 kg compuesto por grasas y tejidos gelatinosos, es la más compleja de todas las estructuras vivas conocidas. Hasta un billón de células nerviosas trabajan unidas para coordinar las actividades físicas y los procesos mentales. La corteza cerebral (es decir, la superficie del cerebro) alberga 15,000 a 33,000 millones de neuronas [14]. El cerebro es el responsable de la complejidad, origen y funcionamiento del pensamiento, memoria, emociones y lenguaje. Sus partes se pueden clasificar en tres principales, la primera se puede observar en la figura 2.1.

- Corteza cerebral
 - Lóbulo occipital: información visual
 - Lóbulo parietal: órganos de sensación, coordina el equilibrio
 - Lóbulo temporal: procesa información de sonidos y funciones asociadas al lenguaje

- Lóbulo frontal: área de broca, encargada de producción lingüística y oral
- Núcleo estriado: asociado con los movimientos
- Rinencéfalo: percepción olfativa

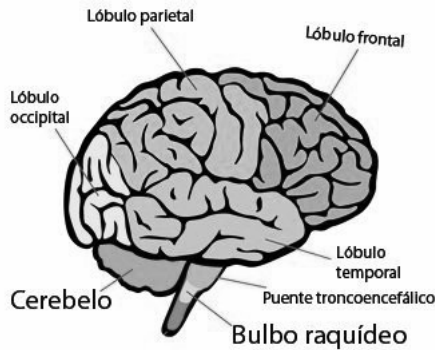


Figura 2.1: Partes del cerebro.

2.1.1. El cerebro y el lenguaje

La función del lenguaje depende del trabajo de un grupo de zonas corticales y subcorticales. El dominio del lenguaje está relacionada a la lateralidad [35]. Aproximadamente, el 95 % de los diestros tiene localizado el lenguaje en el hemisferio izquierdo y sólo un 5 % en el hemisferio derecho. En cambio, los zurdos, el 70 % lo tiene representado en el hemisferio izquierdo, un 15 % en el hemisferio derecho y un 15 % en forma bilateral [35]. Otros factores que influyen en la representación del lenguaje son:

- Sexo: las mujeres lo tienden a tener en forma bilateral. En cambio, en los hombres está más lateralizado a la izquierda [34]

- **Escolaridad:** los sujetos diestros con mayor escolaridad tienen el lenguaje más lateralizado a izquierda. Por el contrario, los sujetos con baja escolaridad tienden a tener una representación bilateral del lenguaje; por lo tanto, la incidencia de afasia cruzada es mayor en estos sujetos [34, 33].
- **Edad:** se ha planteado que al momento de nacer el lenguaje está representado en forma bilateral. Lo primero que se lateraliza a la izquierda es la región anterior, la que está relacionada con la expresión, y en etapas más tardías lo hace la región posterior, relacionada con la comprensión. Esto explicaría la razón por la cual, en las personas más jóvenes la afasia más frecuente es la de Broca y en los sujetos más añosos, la de Wernicke [32].

Lo anterior es muy importante, pues en la literatura relacionada con clasificación de habla imaginada se presentan mejores resultados en hombres que además sean diestros. Para el caso de las mujeres resulta más difícil que los sensores hagan contacto correctamente.

Normalmente para habla imaginada se consideran todos los electrodos que disponga el Electroencefalograma, sin embargo, se han realizado pruebas solo tomando en cuenta las áreas que están relacionadas con el lenguaje. Estas áreas se describen a continuación:

Área de Wernicke: Se localiza en el lóbulo temporal izquierdo, como se muestra en la figura 2.2. Las funciones de esta área son la comprensión auditiva y el procesamiento de la selección del léxico (eje paradigmático).

Área de Broca: Se localiza en el lóbulo frontal izquierdo. Las funciones de esta área son: la formulación verbal(morfosintaxis) que corresponde a la expresión y comprensión de estructuras sintácticas, y cumple un rol en el procesamiento de verbos. Además, tiene una participación en la planificación y programación motora para la articulación del habla. Esta área está relacionada con procesos de secuenciación.

En la figura 2.2 pueden observarse diferentes áreas de la corteza cerebral, donde las abreviaturas significan:

B: Área de Broca; W: Área de Wernicke; SM: Circunvolución supramarginal; A: Circunvolución angular; FE: Función ejecutiva; PS: Procesamiento sintáctico; F: Fusiforme; M Semántica: Memoria

semántica; FLS-II: Fascículo longitudinal superior 2º componente; FLS-III: Fascículo longitudinal superior 3er componente; FLS-IV (FA): Fascículo longitudinal superior 4to componente (fascículo arqueado); FLS-V: Fascículo longitudinal superior 5to componente; FU: Fascículo uncinado; FLI: Fascículo longitudinal inferior; FFOI: Fascículo fronto-occipital inferior.

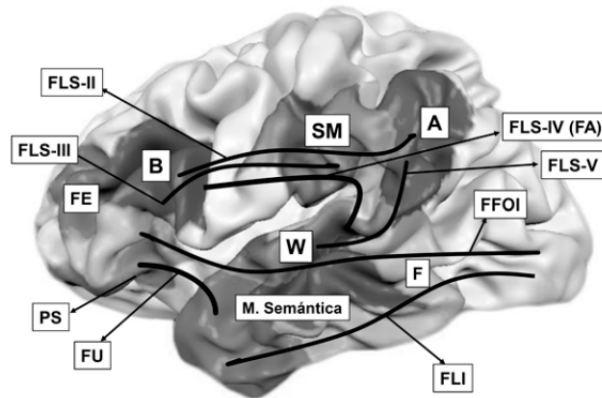


Figura 2.2: Áreas corticales, mapa de Brodmann [35].

2.2. Electroencefalograma

La exploración neurofisiológica basada en el registro de actividad bioeléctrica cerebral en condiciones basales de reposo, en vigilia o sueño, y durante diversas activaciones es llamada Electroencefalografía. La actividad eléctrica(onda cerebral) producida en el cerebro de un sujeto sano se clasifica en:

- Delta(0 a 4 Hz): están asociadas con etapas de sueño profundo. Se presentan en las etapas tres y cuatro, en casos de daño cerebral y coma.
- Theta(4 a 8 Hz): están asociadas con las fases de sueño uno y dos. Se generan tras la interacción entre el lóbulo temporal y frontal.

- Alpha(8 a 12 Hz): se origina en el lóbulo occipital durante períodos de relajación con los ojos cerrados, pero todavía despierto. Estas ondas se atenúan al abrir los ojos, con la somnolencia y el sueño. Se piensa que representan la actividad de la corteza visual en un estado de reposo.
- Beta(12 a 30 Hz): están asociadas con etapas de sueño nulo, donde se está despierto y consciente, y las ondas son más frecuentes en comparación con las ondas delta, alfa y theta.
- Gamma(30 a 100 Hz): su presentación más habitual es a 40 Hz. Se ha teorizado que las ondas gamma podrían estar implicadas en el proceso de percepción consciente.

Además se sabe que las frecuencias del habla imaginada llegan hasta 60 Hz, es decir que, prácticamente involucra toda la actividad eléctrica clasificada anteriormente. Como se verá más adelante, las frecuencias de las señales de EEG más allá de 60 Hz deben ser removidas, pues no son de interés para este trabajo.

La técnica de EEG se basa en colocar electrodos de superficie sobre el cuero cabelludo según las posiciones y mediciones del Sistema Internacional 10-20, como se muestra en la figura 4.2. El registro de la señal de EEG es una señal que representa el estado funcional del cerebro al detectar con alto grado de sensibilidad cambios eléctricos debidos a la oxigenación, la presencia de actividad paroxística cortical, la acción de anestésicos y otras drogas, así como cambios por manipulación mecánica del tejido cerebral [22, 15, 16]. Dichas señales están clasificadas como no estacionarias [25], es decir, cambian rápidamente con el tiempo y entre sesiones.

2.3. Interfaces cerebro-computadora

Una Interfaz Cerebro Computadora (BCI, por sus siglas en inglés), detecta e interpreta las señales del cerebro y utiliza los resultados para comunicar la intención de un usuario [45], convirtiéndose así en una nueva modalidad de interacción hombre-máquina que permite a los usuarios utilizar sus pensamientos para controlar diversos dispositivos externos. La clasificación de las BCI se especifica a continuación.

- Invasivas: el registro de señales se realiza en forma intracraneal o intracortical. Estos capturan directamente los potenciales de acción de grupos neuronales en la corteza cerebral [19]. Entre ellos Electrocartogramas, Potenciales de campo locales(LFP, por sus siglas en inglés), Unidades simples de actividad(SUA, por sus siglas en inglés)
- No Invasivas: utilizan electrodos de superficie para registrar de Electroencefalograma (EEG) y Magnetoencefalogramas (MEG).

En la figura 2.3 se describen los bloques funcionales de una BCI, en el cual un usuario controla un dispositivo (en este caso una silla de ruedas motorizada). El usuario puede también monitorear el estado del dispositivo, a través de realimentaciones que le permiten determinar el resultado de sus esfuerzos por controlarlo. Con el avance en los algoritmos de aprendizaje automático y tecnologías de procesamiento digital, la investigación reside en la exploración de técnicas de extracción de características y clasificación de las señales registradas [19].

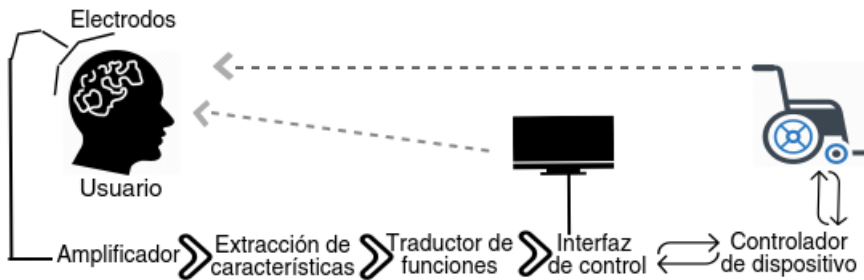


Figura 2.3: Modelo funcional de un sistema BCI no invasivo, adaptado de [3].

Las fuentes electrofisiológicas se refieren a los mecanismos o procesos neurológicos usadas por un usuario de una BCI para generar señales de control [2]. Wolpaw [45] separó estas fuentes en 5 categorías basadas en los mecanismos neuronales y la tecnología de grabación que utilizan. Estas categorías son:

- Actividad sensoriomotora.

- Potenciales P300.
- Potenciales evocados visuales(VEP).
- Potenciales corticales lentos(SCP).
- Actividad de la célula neural(ANC).

Luego Bashashati [3] agregó 2 categorías más:

- Respuesta a tareas mentales.
- Múltiples neuromecanismos.

2.4. Procesamiento digital de señales

El procesamiento digital de señales(PDS) es la representación, transformación y manipulación matemática de una señal de información para modificarla o mejorarla en algún sentido.

Existen dos tipos de señales, estacionarias y no estacionarias. Las primeras son ampliamente conocidas y tratadas con la Transformada de Fourier, las segundas, son de rápida transitoriedad. En el caso de las segundas, las señales pueden ser, por ejemplo, las capturadas con un Electroencefalograma durante el habla imaginada, del presente trabajo. Para las señales de naturaleza no estacionaria es común utilizar la transformada discreta de wavelet (DWT, por sus siglas en inglés), con base en una función conocida como wavelet madre que puede representar una señal en términos de versiones trasladadas y dilatadas de dichas funciones, descomponiendo una señal de EEG (en este caso) finita, en diferentes componentes de frecuencia.

2.4.1. Transformada discreta wavelet

A diferencia de la transformada de Fourier donde las funciones base son senoides, la transformada wavelet discreta(DWT) está basada en pequeñas ondas llamadas wavelets. Las wavelets analizan a las señales a partir de una función conocida como wavelet madre, mediante sus traslaciones y

dilataciones. Existen varias funciones wavelet madre, la usada en este trabajo es la biortogonal 2.2.

Esta transformada es eficiente en el análisis local de señales estacionarias y no estacionarias y de rápida transitoriedad. Mapea la señal en una representación de tiempo-escala. La DWT provee análisis con ventanas dilatadas. El análisis de las frecuencias de mayor rango se realiza usando ventanas angostas y el análisis de las frecuencias de menor rango se hace utilizando ventanas anchas.

En seguida se especifica el funcionamiento de la DWT. Para una señal s de longitud N , la DWT consiste a lo más de $\log_2 N$ etapas. El primer paso produce, a partir de la señal s , dos conjuntos de coeficientes: coeficientes de aproximación CA_1 , y coeficientes de detalle CD_1 .

Estos vectores son obtenidos con la convolución de s con el filtro **Lo_D** pasa-bajo(low-pass) para la aproximación. Con el filtro pasa-alto(high-pass) **Hi_D** para detalles, seguido del submuestreo (downsampling). Esto se ilustra en la figura 2.4

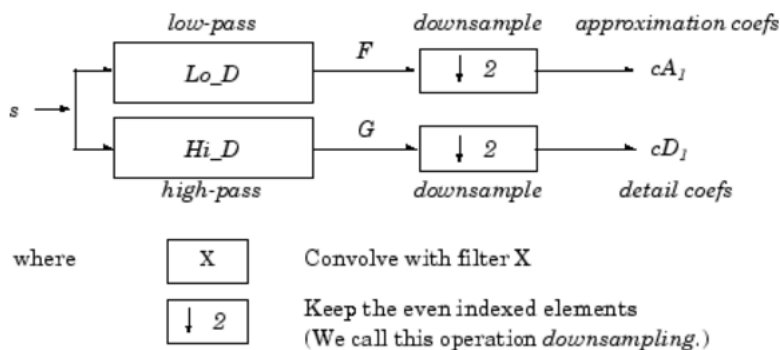


Figura 2.4: Funcionamiento de DWT [26]

La longitud de cada filtro es de $2N$. Si $n = \text{length}(s)$, las señales F y G son de longitud $n + 2N - 1$ y los coeficientes CA_1 y CD_1 son de tamaño $\text{floor}(\frac{n-1}{2}) + N$.

El siguiente paso divide el coeficiente de aproximación CA_j en dos partes usando el mismo esquema anterior pero reemplazando s por CA_j y ahora produciendo CA_{j+1} y CD_{j+1} , como se muestra en la figura 2.5.

En resumen, la descomposición wavelet de la señal s en el nivel j , tiene

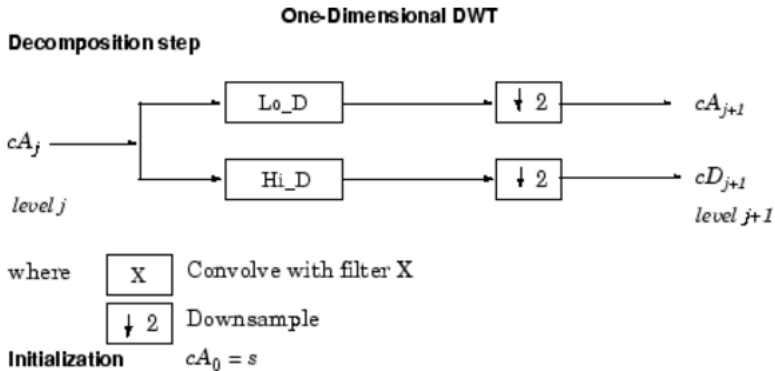


Figura 2.5: Funcionamiento de DWT [26] en el nivel de descomposición j .

la siguiente estructura $[cA_j, cD_j, \dots, cD_1]$. Por ejemplo, para el nivel de descomposición $j = 3$, el árbol que se va creando es el que se muestra en la figura 2.6.

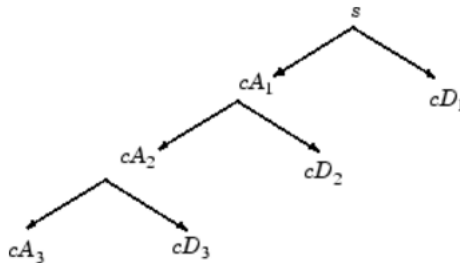


Figura 2.6: Árbol creado en el nivel de descomposición $j=3$.

Wavelet biortogonal 2.2

La calidad de la información obtenida de un análisis donde se aplica la DWT depende de la wavelet madre utilizada. En la bibliografía no se muestra un método o algoritmo para hacer la selección de una wavelet madre adecuada basado en las propiedades de cada una, sino que sólo emplean algunos criterios como:

- Determinar visualmente la semejanza que tiene la forma de la wavelet madre y la señal a analizar
- Ensayar distintas wavelets o métodos para cada caso de estudio y a partir de los resultados escoger una.
- Usar db4 por ser la más empleada en el análisis de señales

Las wavelet madre permiten analizar con buena resolución en frecuencia componentes que oscilan lentamente, sus armónicos y sus subarmónicos. También permiten analizar con buena resolución en tiempo los componentes que oscilan rápidamente, que normalmente corresponden a eventos transitorios [20], tales como las señales cerebrales. En el caso particular de habla imaginada, se han usado distintas wavelet madre y con la que se obtienen mejores resultados es con la wavelet madre biortogonal 2.2 (bior2.2).

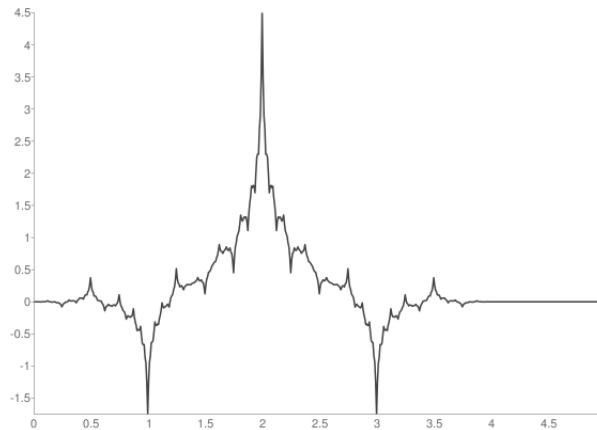


Figura 2.7: Wavelet biortogonal 2.2.

2.4.2. Distribución de la energía

Al aplicar la transformada DWT bior2.2 con un nivel de descomposición j se obtiene una estructura con vectores de coeficientes de aproximación(CA_j) y detalle(CD_j): $[CA_j, CD_j, CD_{j-1}, \dots, CD_1]$. Sin

embargo, el número de coeficientes wavelet en cada nivel de descomposición variará dependiendo del tamaño de la ventana análisis. Para tratar este problema se puede calcular la distribución de varios tipos de energía (instantánea, jerárquica, relativa y teager) [12] en cada nivel. De esta manera se logra normalizar los tamaños de las señales analizadas para unificar el número de coeficientes calculados. A continuación se muestran las formulas de los diferentes tipos de energía.

Sea $w_j(r)$ el coeficiente wavelet en el tiempo r y banda de frecuencia j . Si N es la longitud de la ventana de análisis, $w_j(r)$ tiene $N_j = N/2^j$ muestras en el nivel de descomposición j :

- Energía instantánea: este tipo de energía refleja la amplitud de la señal

$$f_j = \log_{10} \left(\frac{1}{N_j} \sum_{r=1}^{N_j} (w_j(r))^2 \right) \quad (2.1)$$

- Energía jerárquica: provee una resolución jerárquica en tiempo y le da más importancia al centro de la ventana de análisis

$$f_j = \log_{10} \left(\frac{1}{N_j} \sum_{r=(N_j-N_j)/2}^{(N_j+N_j)/2} (w_j(r))^2 \right) \quad (2.2)$$

Donde además, N_j 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 y es un parámetro robusto para el reconocimiento de habla ya que atenúa el ruido auditivo

$$f_j = \log_{10} \left(\frac{1}{N_j} \sum_{r=1}^{N_j-1} |(w_j(r))^2 - w_j(r-1) * w_j(r+1)| \right) \quad (2.3)$$

- Energía relativa: la distribución de energía puede ser considerada como una densidad tiempo-escala y provee información para caracterizar la distribución de energía de la señal en diferentes bandas de frecuencia [23]

$$f_j = \frac{E_j}{E_{total}} \quad (2.4)$$

Donde E_j representa la energía en el j -ésimo nivel de descomposición y E_{total} representa la energía total de los coeficientes wavelet de una señal dada. La energía en el j -ésimo nivel de descomposición E_j se define como:

$$E_j = \begin{cases} \sum_k |d_{j,k}|^2 & \text{Si } j \leq N \\ \sum_k |a_k|^2 & \text{en caso contrario} \end{cases} \quad (2.5)$$

Y la energía total E_{total} como sigue:

$$E_{total} = \sum_{j=1}^{N+1} E_j \quad (2.6)$$

2.5. Valores estadísticos

Los datos pueden ser analizados mediante métodos simples o también pueden ser complejos. Estos métodos, por ejemplo, pueden ser con el uso de gráficos estadísticos de tendencia central y dispersión. El análisis con valores estadísticos tiene el propósito de reducir el volumen de información.

Los estadísticos simples permiten estimar y describir el comportamiento de un conjunto de datos. Los más comunes son: el promedio, la media aritmética, el rango de variación, la desviación estándar y el coeficiente de variación, que se utilizan en el análisis de datos cuantitativos.

A continuación se detallan algunas medidas de interés para este trabajo.

- **La media aritmética:** es una medida de tendencia central que ayuda a caracterizar y permite relacionar un atributo con un valor central. Se obtiene a partir de la suma de todos sus valores dividida entre el total de números sumados. Para n números $\{x_1, x_2, \dots, x_n\}$, la media aritmética se define como:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.7)$$

- **Varianza:** es una medida de dispersión y sirve para identificar a la media de las desviaciones cuadráticas de una variable respecto a su

media aritmética.

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n \left(X_i - \bar{X} \right)^2 \quad (2.8)$$

- **Desviación estándar:** es la medida de dispersión más común, que indica qué tan dispersos están los datos con respecto a la media. Mientras mayor sea la desviación estándar, mayor será la dispersión de los datos. Se define como la raíz cuadrada de la varianza. Para analizar un conjunto de datos, no basta con conocer las medidas de tendencia central, sino que además, se necesita conocer la desviación que presentan los datos respecto de la media aritmética, con el fin de tener una visión de los mismos más de acuerdo con la realidad. Esta medida se calcula como sigue.

$$\sigma = \sqrt{\sigma^2} \quad (2.9)$$

- **Curtois:** este coeficiente analiza el grado de concentración que presentan los valores alrededor de la zona central de la distribución, de manera que a mayor grado de curtosis, más apuntada será la forma de la curva. Se mide promediando la cuarta potencia de la diferencia entre cada elemento del conjunto y la media, dividido entre la desviación estándar elevado también a la cuarta potencia.

Sea el conjunto $X = (x_1, x_2, \dots, x_N)$, el coeficiente de curtosis se calcula como:

$$curtosis = \frac{\sum_{i=1}^N (x_i - \bar{X})^4}{N * \sigma^4} - 3 \quad (2.10)$$

- **Oblicuidad:** es una medida de que tan asimétrica es una distribución alrededor de la media. También se le conoce como tercer momento de la distribución y se calcula de la siguiente forma.

$$Oblicuidad = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{X}}{\sigma} \right)^3 \quad (2.11)$$

Otras medidas estadísticas de interés en este trabajo han sido omitidas por su amplio conocimiento. Entre ellas se encuentra el mínimo, máximo, la suma y la mediana. Estas medidas también son importantes y aportan mucha información para caracterizar un conjunto de datos.

2.6. Clasificación

El aprendizaje automático es una rama de la inteligencia artificial que tiene como objetivo eliminar la necesidad del conocimiento de un experto de los procesos de análisis de datos. Los problemas que se tratan con aprendizaje automático se dividen en dos: problemas de clasificación y de regresión. En este documento se tratan sólo problemas de clasificación, esto consiste en asignar una clase, es decir que predice a que clase pertenece un conjunto de datos, en estos problemas los valores son discretos.

En [25] se hace un análisis sobre los problemas que afectan a los clasificadores utilizados en BCI que se basan en EEG. Se analizan dos problemas en la clasificación que son especialmente relevantes en investigaciones en BCI.

El primer problema es denominado “maldición de la dimensionalidad”, este se debe a que la dimensión del vector de características es mucho mayor que el número de ejemplos para entrenamiento del clasificador. Este problema se puede resolver mediante selección de características, que implica seleccionar un subconjunto mínimo, con M características de 2.12

$$S = (S_1, \dots, S_M) \tag{2.12}$$

Del conjunto de características original 2.13

$$F = (F_1, \dots, F_N) \tag{2.13}$$

donde $M \prec N$ y $S \subset F$, de tal manera que el espacio de características es reducido y el desempeño de la clasificación mejora.

El segundo problema trata la varianza y el sesgo. La varianza representa la sensibilidad al conjunto de datos de entrenamiento. El sesgo representa la divergencia entre la salida estimada y la mejor salida. Para obtener el error de clasificación mas bajo, el sesgo y la varianza deben ser bajos. Además,

se pueden usar varias técnicas, conocidas como técnicas de estabilización, que sirven para reducir la varianza.

2.6.1. Random Forest

Este algoritmo de clasificación es una combinación de árboles predictores tal que cada árbol depende del valor de un vector aleatorio muestreado independientemente y con la misma distribución para todos los árboles del bosque [4].

El esquema del algoritmo *random forest* es:

- Aleatoriamente se crea (seleccionando con reemplazado) el conjunto de entrenamiento de igual tamaño que el conjunto original. Al seleccionarse aleatoriamente con reemplazo no todos los datos de conjunto general estarán en el conjunto de entrenamiento. La probabilidad de que un dato particular esté en el conjunto de entrenamiento es aproximadamente 66 %.
- Los datos que no forman parte del conjunto de entrenamiento forman el conjunto de validación o *out of bag data* (OOB data).
- En cada punto de división del árbol o nodo, la búsqueda de la mejor variable para dividir los datos no se realiza sobre todas las variables sino sobre un subconjunto, m , de las mismas. La elección del subconjunto de variables se realiza de forma aleatoria.
- Se busca la mejor división de los datos de entrenamiento teniendo en cuenta solo las m variables seleccionadas. Para esta tarea se debe implementar una función objetivo. Habitualmente ésta es la entropía o el índice de Gini.
- Los anteriores procesos son repetidos varias veces, de forma que se tiene un conjunto de árboles de decisión entrenados sobre diferentes conjuntos de datos y de atributos.
- Una vez el algoritmo entrenado, la evaluación de cada nueva entrada es realizado con el conjunto de árboles. La categoría final de la clase (clasificación) es realizado por el voto mayoritario del conjunto de árboles, y en caso de regresión por el valor promedio de los resultados.

El principal problema que tienen los árboles de decisión es que son muy sensibles a pequeñas variaciones en los datos de entrada que pueden encaminar al árbol de decisión por un camino diferente, dando lugar a una clasificación muy diferente. Además, se debe considerar que el algoritmo *random forest* utiliza dos parámetros: el número de árboles y el número de predictores a utilizar en cada partición de cada uno de los árboles. Sin embargo, una de las grandes ventajas de este algoritmo es su baja sensibilidad a estos parámetros, por lo que los valores por defecto suelen producir buenos resultados [24].

2.6.2. Máquina de vectores de soporte

La máquina de vectores de soporte (SVM, por sus siglas en inglés), es una técnica de clasificación que ha mostrado tener mejor desempeño que las formas de aprendizaje tradicional como las redes neuronales (donde existen mínimos locales) [21]. Las SVM originariamente fueron pensadas para resolver problemas de clasificación binaria, sin embargo, actualmente se utilizan para resolver otros tipos de problemas (Regresión, agrupamiento, multclasificación).

La figura 2.8 muestra el problema clásico de la separación de dos clases perfectamente separables por una región de decisión lineal. En dicha figura H1 y H2 definen dos hiperplanos, la distancia que separa estos hiperplanos es conocida como *margen*. El hiperplano óptimo es aquel que maximiza el margen y minimiza el riesgo empírico, este debe tener buenas propiedades de generalización para ayudar a garantizar que es el mejor. Los ejemplos más cercanos dentro y fuera de clase que se encuentran en estos dos hiperplanos son llamados *vectores de soporte* [17].

Los problemas de clasificación normalmente involucran datos que no se pueden separar linealmente. Para resolver este problema, las entradas son mapeadas a un espacio de dimensiones mayor, conocido como espacio de características, donde la separación lineal si es posible. A este proceso se le conoce como *truco del kernel*, esto permite que algoritmos lineales se puedan aplicar a problemas no lineales.

Los *kernels* más comunes son:

- Lineal: $K(x_i, x_j) = X_i^T x_j$

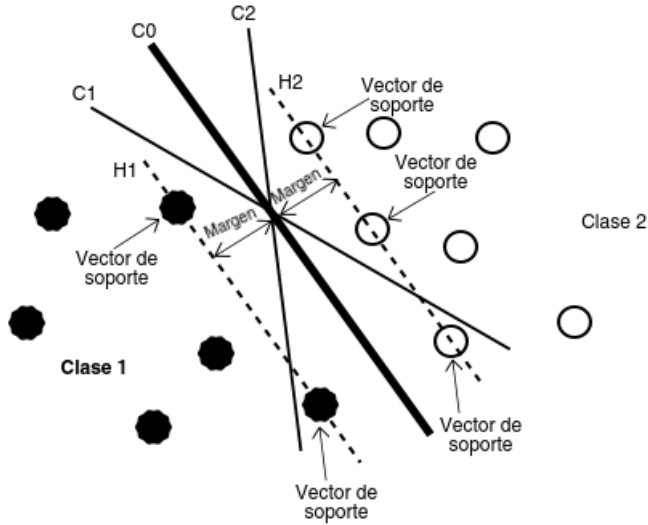


Figura 2.8: Hiperplano óptimo (CO) que maximiza el margen entre los hiperplanos H1 y H2.

- Gausiano: $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$
- Polinómico: $K(x_i, x_j) = (X_i^T x_j + 1)^n$

Para los problemas de clasificación multiclase se utiliza el esquema *one-vs-All* [36]. El esquema consiste en etiquetar la i -ésima clase como positiva y el resto como negativas, luego repetir el proceso iterando la i -ésima clase positiva para todas las clases.

2.6.3. Naive Bayes

Un clasificador bayesiano se basa en la aplicación del teorema de Bayes para predecir la probabilidad condicional de cada clase C , como el producto de la probabilidad a priori de la clase por la probabilidad condicional de los atributos (A) dada la clase C , dividido por la probabilidad de los atributos:

$$P(C|A) = \frac{P(C)P(A|C)}{P(A)} \quad (2.14)$$

El clasificador bayesiano simple, *naive Bayes*, es relativamente sencillo de implementar y es muy usado en tareas de clasificación debido a que produce resultados comparables con los obtenidos por otros métodos más sofisticados [27]. Este clasificador supone que los valores de los atributos son condicionalmente independientes entre sí:

$$P(a_1, a_2, \dots, a_n | c) = \prod_{i=1}^n P(a_i | c) \quad (2.15)$$

Dado un ejemplo x representado por k valores el clasificador *naive Bayes* se basa en encontrar la hipótesis más probable que describa a ese ejemplo. Si la descripción de ese ejemplo viene dada por los valores (a_1, a_2, \dots, a_n) , la hipótesis más probable será aquella que cumpla:

$$c_{MAP} = \operatorname{argmax}_{c \in C} P(c | a_1, a_2, \dots, a_n) \quad (2.16)$$

Es decir, la probabilidad de que conocidos los valores que describen a ese ejemplo, éste pertenezcan a la clase c (donde c es el valor de la función de clasificación $f(x)$ en el conjunto finito C). Luego por el teorema de Bayes:

$$\begin{aligned} c_{MAP} &= \operatorname{argmax}_{c \in C} P(c | a_1, a_2, \dots, a_n) \\ &= \operatorname{argmax}_{c \in C} \frac{P(a_1, a_2, \dots, a_n | c) P(c)}{P(a_1, a_2, \dots, a_n)} \\ &= \operatorname{argmax}_{c \in C} P(a_1, a_2, \dots, a_n | c) P(c) \end{aligned} \quad (2.17)$$

Aunque *naive Bayes* funciona bien en muchos dominios, en ocasiones el rendimiento decrece debido a que los atributos no son condicionalmente independientes como inicialmente se asume.

2.6.4. Validación cruzada

En la etapa de clasificación se utiliza la validación cruzada con K-pliegues (k-fold cross validation). En donde el conjunto de datos es dividido en K partes y de las K partes obtenidas, una de las particiones se usa como datos de prueba y los restantes $K - 1$ particiones son usadas como datos de entrenamiento.

Este proceso se repite K veces (K -pliegues) y en cada repetición cada una de las K partes se utilizan como datos de validación y las restantes para entrenamiento. Luego los resultados de los K -pliegues son promediados. La validación cruzada con 10 pliegues es la más común.

2.7. Medidas de desempeño para los clasificadores

El criterio de evaluación es un factor clave a la hora de medir el desempeño de un clasificador y la más utilizada es la **exactitud** (accuracy), ésta se define como el número de decisiones tomadas correctamente dividida por el número total de casos presentados al clasificador. En muchos casos la exactitud aporta poca información o no es suficiente. En esta sección se describen algunas medidas para evaluar la eficacia de un clasificador con clases balanceadas y con clases que no lo están.

En un problema de dos clases, la matriz de confusión contiene los resultados de los objetos clasificados correctamente e incorrectamente en cada clase, como se muestra en la tabla 2.1.

Tabla 2.1: Matriz de confusión para dos clases

	Positiva real	Negativa real
Positiva predicha	Verdaderos positivos(VP)	Falsos positivos(FP)
Negativa predicha	Falsos negativos(FN)	Verdaderos negativos(VN)

Se pueden obtener cuatro métricas de la tabla 2.1 para medir el rendimiento de clasificación para cada una de las clases. Donde:

- $VP_{rate} = \frac{TP}{VP+FN}$ corresponde a la fracción de objetos clasificados correctamente en la clase positiva
- $VN_{rate} = \frac{VN}{FP+VN}$ corresponde a la fracción de objetos clasificados correctamente en la clase negativa

- $FP_{rate} = \frac{FP}{FP+VN}$ corresponde a la fracción de objetos clasificados incorrectamente en la clase positiva
- $FN_{rate} = \frac{FN}{VP+FN}$ corresponde a la fracción de objetos clasificados incorrectamente en la clase negativa

A continuación se detallan las medidas de interés en este trabajo. Donde la *Medida F* resulta de gran interés ya que es una buena forma de validar la clasificación con clases desbalanceadas.

La **exactitud** es la fracción de observaciones clasificadas correctamente como positivas y las clasificadas correctamente como negativas, sobre todas las predicciones clasificadas como positivas y negativas.

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.18)$$

La **precisión** es la fracción de observaciones clasificadas correctamente como positivas, sobre todas las predicciones clasificadas como positivas.

$$precision = \frac{VP}{VP + FP} \quad (2.19)$$

La **sensibilidad** o **recuerdo** es la fracción de observaciones clasificadas correctamente como positivas, sobre todas las observaciones positivas.

$$recuerdo = \frac{VP}{VP + FN} \quad (2.20)$$

La **medida F** es el significado armónico entre precisión y recuerdo. Tratando de mejorar una medida, frecuentemente, se deteriora la otra, por lo que una manera de combinar ambas medidas es mediante ésta métrica.

$$medidaF = \frac{2 * precision * recuerdo}{precision + recuerdo} \quad (2.21)$$

Capítulo 3

Estado del arte

Las fuentes electrofisiológicas se refieren a los mecanismos o procesos neurológicos usadas por un usuario de una BCI para generar señales de control [2]. Wolpaw [45] separó estas fuentes en cinco categorías basadas en los mecanismos neuronales y la tecnología de grabación que utilizan. Estas categorías son: actividad sensoriomotora, potenciales P300, potenciales evocados visuales(VEP, por sus siglas en inglés), potenciales corticales lentos(SCP, por sus siglas en inglés) y actividad de la célula neural(ANC, por sus siglas en inglés). Luego Bashashati [2] agrego dos categorías más: respuesta a tareas mentales y Múltiples neuromecanismos.

Sin embargo dichas fuentes electrofisiológicas no son fáciles de generar y existen dos problemas principales [30].

- Largo periodo de entrenamiento que va desde algunos días, semanas o incluso meses; pues las fuentes electrofisiológicas son generadas por el usuario de manera poco consiente.
- Bajas tasas de comunicación, una sola palabra o menos procesada por minuto.

A pesar de tantos años de desarrollo, estos problemas hacen imposible la comunicación de manera natural. Es por esto que Desain [11] dice que aunque prometedor, todavía se está lejos de lo que se necesita para un control rápido y fiable de las interfaces y que en lugar de tratar de mejorar los métodos que usan las fuentes electrofisiológicas anteriores, se deben

aprovechar otras tareas mentales. Entre ellas menciona el habla imaginada o interna, de esta manera los métodos de análisis también pueden enfocarse en puntos de tiempo e intervalos de tiempo específicos. El habla imaginada como fuente electrofisiológica tiene algunas ventajas respecto a las otras, pues necesita poco entrenamiento; sin embargo, es una tarea compleja, pues está ligada a múltiples artefactos tales como movimiento de músculos y pulsos de arterias en el cuero cabelludo.

A continuación se presenta una revisión de los trabajos realizados en la tarea de reconocimiento de habla no pronunciada o imaginada. En algunos de ellos se han reportado experimentos enfocados a reconocer sólo sílabas o fonemas para el control de algún dispositivo, y otros consideran un enfoque orientado a reconocer palabras imaginadas. Cabe señalar que no todos estos trabajos intentan detectar el inicio del habla imaginada.

En [39] se hizo una aproximación hacia detectar el inicio del habla imaginada vs. estados de inactividad. Para esto se realizaron tres tareas para obtener información lingüística con cuatro sujetos (tres hombres y una mujer) entre 20-25 años. En este trabajo se imaginó sílaba *um* y se hicieron 6 experimentos con cada uno de los participantes y se capturó la señal de 64 canales colocados de acuerdo al sistema internacional 10-20 usando Biosemi Actiview con una frecuencia de muestreo de 256 Hz. La tarea *overt* consistía en producir un tono alto que se podía escuchar claramente. La tarea *I_Overt* implicaba dos elementos, el primero respirar hacia afuera y el segundo tensar las cuerdas vocales pero sin emitir sonidos de tonos altos. La tercer tarea *cover* consistió en imaginar el proceso que conduce a la producción del sonido alto, los sujetos fueron instruidos para imaginar la pronunciación y escuchar el sonido. Cada experimento fue elegido aleatoriamente, es decir: *over-covert-covert-I_Overt...* etc. para cada participante.

Para experimentar que tan lejos se encontraba la actividad, se tomaron segmentos de ventana de 1,1.5,2,y 3 segundos después del marcador de inicio. Después a cada segmento se se aplicó un filtro pasa banda en el rango de 4-20 Hz para luego eliminar en ruido ambiental. Además utilizaron patrones de espaciado común (CSP) para selección de canales. Para extraer características de los canales se usó un modelo autoregresivo (AR). Para esto se usó el método Burg's en Matlab (arbug) que regresa un vector con características. Finalmente para la clasificación se usó análisis

de discriminación lineal (LDA) que busca una línea que maximiza la separación entre clases y minimiza el traslape entre ellas. Los resultados muestran que con la tarea tres (covert) se obtuvo la mejor clasificación, para cada individuo 82.4% (1.5 segundos) 81.20% (3 segundos) 91.24% (1.5 segundos) y 75.68% (2 segundos) de exactitud respectivamente.

En [8] los autores utilizaron vocales *a* y *u* de tres sujetos diestros entre 26-29 años; dos hombres y una mujer, para mostrar que es viable el control de prótesis utilizando la imaginación de dichas vocales y obtuvieron resultados de 68% y 78% de precisión para cada palabra. Su trabajo propone un *framework* que realiza el procedimiento de clasificación de las vocales casi instantáneamente, es decir, en cuestión de segundos. Cabe mencionar, que se aprovechó la activación de las vocales seleccionadas *a* y *u* con diferentes músculos. Para este trabajo se usó Biosemi ActiView capturando 64 canales EEG con una frecuencia de muestreo de 2048 Hz. Se hicieron ventanas de 500 milisegundos y se obtuvieron características con patrones espaciales comunes (common spatial pattern,csp) para luego clasificar con una máquina de vectores de soporte con la función de base radial.

Los trabajos antes mencionados hacen uso de habla imaginada, sin embargo, hay otros estudios [38] que además utilizan expresiones faciales y audio. En el experimento realizado en [38] participaron 12 sujetos con 27.4 años de edad en promedio; cuatro mujeres y ocho hombres que pronunciaron e imaginaron las silabas *iy/uw/piy/tiy/diy/m/n/pat/pot/knew/gnaw*. Se realizaron 132 repeticiones, después de cada 40 los participantes descansaban. En este trabajo se calcularon medidas estadísticas para cada ventana: media, mediana, desviación estándar, varianza, máximo, mínimo, entropía espectral, energía, curtosis y oblicuidad. Los resultados obtenidos promedio son de 59.60% de exactitud para */iy/* y de 79.16 para */uw/*.

Para la clasificación empleando *SVM* se combinó información de EEG, facial, y audio. Específicamente se realizaron las siguientes combinaciones de información.

- EEG
- Audio
- Facial

- EEG+audio
- EEG+facial
- EEG+audio+facial

Mientras que en [39, 8, 38] se utilizan sílabas o vocales, en [40, 18] se utilizan las palabras *arriba*, *abajo*, *izquierda*, *derecha* y *seleccionar*. Cabe mencionar que estas palabras se seleccionaron por su aplicación para control de prótesis o un cursor de una computadora.

En [18] se utilizó la transformada discreta de wavelet para extraer características y estas usarlas como entrada para cuatro clasificadores: *naive Bayes*, *SVM*, *random forest* y perceptrón multicapa para una base de datos de 120 ejemplos de las palabras *arriba* y *abajo* obtenidas de tres personas. En este trabajo se estudiaron más a profundidad los canales cercanos al área de Broca y Wernicke, que corresponden a los canales F7, FC5, T7 y P7 en el sistema internacional 10-20. Se muestra que la clasificación de las palabras imaginadas es independiente del clasificador, pues no se muestran cambios significativos entre los clasificadores usados. Además, se da una aproximación de cómo se podrían mejorar los resultados, aplicando análisis de componentes independientes y evaluando componentes que se obtengan mediante el exponente de Hurst. Además, en futuros experimentos [40, 41, 42] realizaron pruebas con las mismas palabras imaginadas pero con diferentes formas de caracterizar la señal, usando DWT con la wavelet madre daubechies y la biortogonal 2.2, primero tomando ventanas de 2 segundos y después usando el cálculo de la distribución de la energía instantánea, jerárquica, teager y relativa. En los mejores casos obtuvieron 70.94% de exactitud. Al igual que en [40], se utilizó el kit emotiv para obtener los datos con una frecuencia de muestreo de 128 Hz y con 14 canales.

En [31], se utilizan las palabras *alpha*, *bravo*, *charlie*, *delta* y *echo* del alfabeto radiofónico por su uso para eliminar ambigüedad en las comunicaciones. Se utilizaron 20 repeticiones de cada uno de los 18 participantes. Estos ejemplos fueron tomados de acuerdo a cuatro estados *Drugs*, *Health*, *Alertness* y *Nervousness*. Para obtener características se utilizó Double-Tree Complex Wavelet Transform (DTCWT), y para la etapa de clasificación offline se utilizó Janus recognition toolkit(Jrtk).

Para la selección de características en [7] se consideran diferentes técnicas y experimentos donde se aplican las mismas. De un vector con características, en algunos casos simplemente se selecciona un subconjunto representativo. También se muestra una comparación entre los clasificadores usados para tareas de clasificación de señales EEG. Los métodos de este trabajo, no se aplican a habla imaginada, sin embargo, las técnicas que se usan para clasificar y obtener características muestran buenos resultados en las tareas desempeñadas, tales como el control de prótesis de mano para amputados, movimientos imaginados, sistemas de respuesta si y no, etc.

Como puede observarse, la mayoría de los trabajos relacionados se enfocan en el uso de sílabas, fonemas o incluso vocales. Este trabajo se enfoca en proponer un método para la detección del inicio y final de la actividad lingüística para cuando se desean reconocer palabras o incluso frases. Por la naturaleza no estacionaria de las señales de EEG y por las grandes dimensiones se debe tener especial atención en la extracción de características y así orientar la investigación al reconocimiento en línea de habla imaginada.

Capítulo 4

Método propuesto

Para lograr el objetivo planteado y encontrar características adecuadas para representar la señal de EEG, a continuación en la figura 4.1 se muestra el método seguido en los experimentos que se exponen en el siguiente capítulo. Primero se muestran las dos bases de datos utilizadas, luego para realizar el preprocesamiento de las señales de EEG se usa el método de referencia promedio común (CAR, por sus siglas en inglés). Para la extracción de características se usan dos maneras, la primera es usando la DWT bior2.2 y luego a cada vector de coeficientes de aproximación y detalle calcular la distribución de la energía (instantánea, jerárquica, relativa y teager). La segunda manera de extraer características es usando medidas estadísticas consta de dos conjuntos, como se verá más adelante, el primero con 15 coeficientes y el segundo con 9.

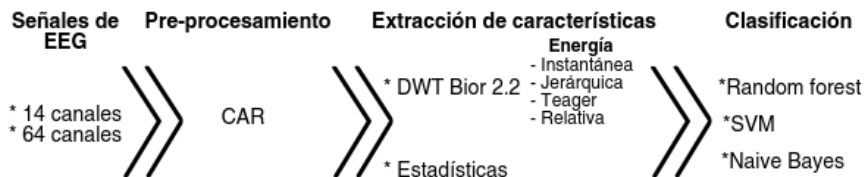


Figura 4.1: Modelo general del método.

4.1. Infraestructura

Para realizar los experimentos y lograr los objetivos planteados se utilizaron dos bases de datos, la primera contiene la señal de EEG de 27 sujetos grabada con una frecuencia de muestreo de 128 Hz y la segunda con la señal de EEG de 20 sujetos grabada con una frecuencia de muestreo de 500 Hz. Cabe destacar que los protocolos de adquisición de señales de EEG son diferentes. A continuación se muestran a detalle como están formadas las bases de datos usadas y los protocolos de estimulación y adquisición.

4.1.1. Primera base de datos

Esta base de datos (BD1) contiene las señales de EEG de 27 sujetos sanos entre 12 y 30 años, donde 6 mujeres son diestras, 2 mujeres zurdas y 19 hombres diestros. Cada persona fue grabada mientras imaginaba 33 veces cinco palabras, que se muestran ordenadas en la tabla 4.1.

Para cada palabra analizada se obtuvo y guardó información de 14 canales de EEG organizados de acuerdo al sistema internacional 10-20 (ver figura 4.2), detallada en la Tabla 4.2. La frecuencia de muestreo de EMOTIV EPOC es de 128 Hz.

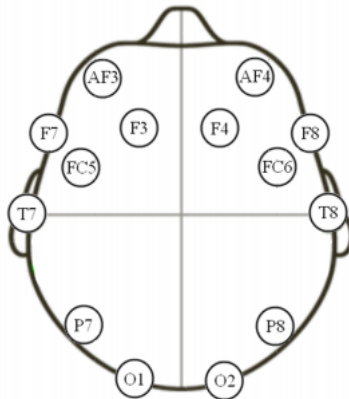


Figura 4.2: Localización de 14 electrodos de acuerdo al sistema internacional 10-20 para EEG.

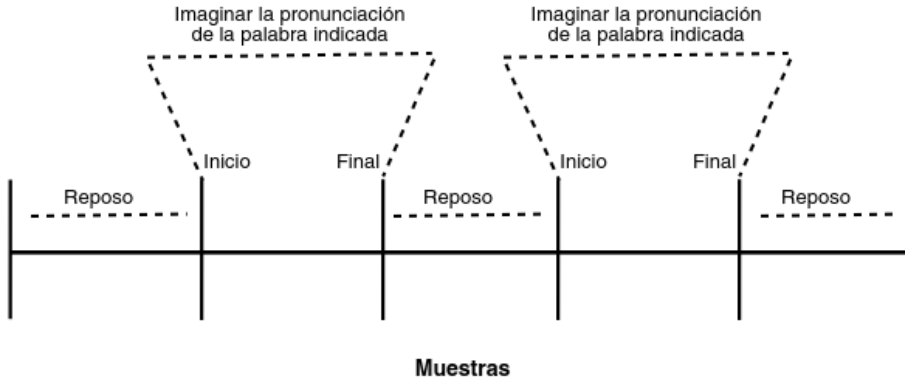


Figura 4.3: Diseño del protocolo para adquisición de la señal de EEG para la BD1 usando EMOTIV EPOC.

En la figura 4.3 se muestra el protocolo seguido para la adquisición de las señales de EEG. Los sujetos imaginaban una palabra y luego seguida una etapa de reposo para luego hacer otra instancia de la misma palabra imaginada. Así se realizó el proceso hasta completar las 33 instancias por cada sujeto.

Tabla 4.1: índice de palabras imaginadas y almacenadas en la BD1.

ÍndicePalabra	Nombre
1	Arriba
2	Abajo
3	Izquierda
4	Derecha
5	Seleccionar

4.1.2. Segunda base de datos

La segunda base de datos (BD2) contiene las señales de EEG de 20 sujetos hablantes nativos del idioma español que fueron grabados mientras imaginaban 40 repeticiones de 4 palabras (*arriba, abajo, derecha, izquierda*)

Tabla 4.2: Descripción de los canales almacenados en la BD1.

Columna	Etiqueta	Típo	Ubicación
1	AF3	Canal de EEG	Pre-frontal
2	F7	Canal de EEG	Frontal
3	F3	Canal de EEG	Frontal
4	FC5	Canal de EEG	Frontal-central
5	T7	Canal de EEG	Temporal
6	P7	Canal de EEG	Parietal
7	O1	Canal de EEG	Occipital
8	O2	Canal de EEG	Occipital
9	P8	Canal de EEG	Parietal
10	T8	Canal de EEG	Temporal
11	FC6	Canal de EEG	Fronto-central
12	F4	Canal de EEG	Frontal
13	F8	Canal de EEG	Frontal
14	AF4	Canal de EEG	Pre-frontal
15	GIRO X	Giroscopio en x	-
16	GIRO Y	Giroscopio en y	-
17	MARKER	Marcador	-

separadas por una pausa, así como se describe en [42]. En la figura 4.4 se muestra la localización de 64 canales de acuerdo al sistema internacional 10-20, los canales marcados con el círculo más oscuro se refiere a la correspondencia que existe con la localización de 14 canales del EMOTIV EPOC.

El protocolo de estimulación y adquisición de las señales de EEG de la BD2 se muestra en la figura 4.5, donde RN es el marcador final para $N = 1, 2, 3, 4$ correspondientes a las 4 palabras imaginadas: $R1$ para la palabra imaginada *arriba*, $R2$ para *abajo*, $R3$ para *derecha* y $R4$ para *izquierda*. Además, el marcador $\{R13 \text{ o } R14 \text{ o } R15\}$ es el final de la respuesta por teclado de la pronunciación imaginada e indica el momento en que el sujeto presionaba una flecha en el teclado. Esta flecha debería coincidir con la palabra que se solicitó imaginar. El tiempo máximo de espera para imaginar la palabra fue de 1,5s, a menos que el sujeto

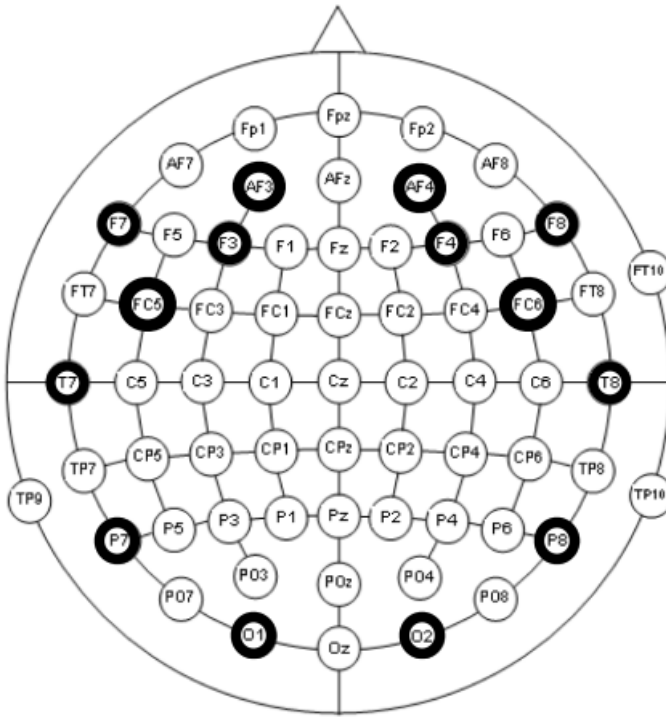


Figura 4.4: Localización de 64 electrodos de acuerdo al sistema internacional 10-20 para EEG.

presionara la tecla antes. Además, $\{R13 \text{ o } R14 \text{ o } R15\}$ es el inicio de una pausa terminando en el marcador $S4$. Para cada palabra imaginada se guardó información de 64 canales de EEG con una frecuencia de muestreo de 500 Hz.

Lo que se muestra en la figura 4.5, son las etapas seguidas para la estimulación y adquisición de la señal de EEG. Primero se presentó una etapa de pre-estímulo, luego mostrarle un objeto y pronunciar una frase relacionada con dicho objeto. Lo siguiente es esperar la respuesta que corresponde a la imaginación de una palabra para finalizar la instancia con una pausa. En la figura se muestra la presentación del objeto *casa* seguida de la frase *La casa estaba* y la imaginación de la palabra *arriba*,

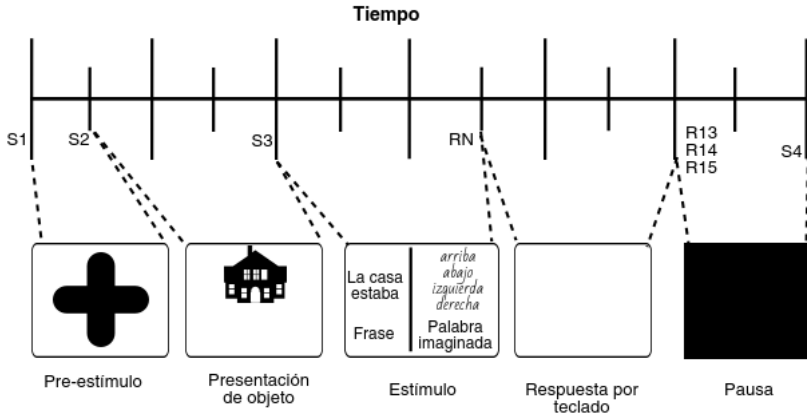


Figura 4.5: Diseño del protocolo para estimulación y adquisición de la señal de EEG para la BD2 usando Brain Vision.

abajo, izquierda o derecha, luego una respuesta por teclado cuando el sujeto presionaba una flecha que debería coincidir con la palabra imaginada. En el ejemplo sería la Palabra imaginada *arriba* y la tecla presionada \uparrow . En la tabla 4.3 se muestra el orden de las palabras almacenadas en la BD2.

Tabla 4.3: índice de palabras imaginadas y almacenadas en la BD2.

IndicePalabra	Nombre
1	Arriba
2	Abajo
3	Derecha
4	Izquierda

La tabla 4.4 presenta una descripción de las principales características de las bases de datos utilizadas.

Tabla 4.4: Comparación de las características de las bases de datos usadas.

	BD1	BD2
Num. sujetos	27	20
Num. palabras imaginadas	5	4
Num. repeticiones por palabra	33	40
Canales de EEG	14	64
Frecuencia de muestreo	128 Hz	500 Hz

4.2. Pre-procesamiento

El pre-procesamiento de los datos, se refiere a la selección, limpieza, reducción y transformación de las bases de datos con el fin de tener mejores datos con los cuales trabajar. Para pre-procesar las señales EEG se usa el método de referencia promedio común (CAR, por sus siglas en inglés). Este método mejora la relación señal-ruido de la señal de EEG y busca quitar lo que es común en todas las lecturas simultáneas de los electrodos. Esto se calcula mediante la siguiente formula:

$$V_i^{CAR} = V_i^{ER} - \frac{1}{n} \sum_{j=1}^n V_j^{ER} \quad (4.1)$$

donde V_i^{ER} es el potencial entre el i -ésimo electrodo y la referencia, y n es el número de electrodos.

4.3. Extracción de características

Para obtener características de las señales de EEG se calcularon coeficientes de energía con DWT o medidas estadísticas. En ambos casos fueron calculadas las características para cada uno de los canales y posteriormente concatenados, formando así un sólo vector de características para cada instancia o época. A continuación se describen las dos formas de obtener características detalladamente.

4.3.1. Transformada discreta wavelet

Se ha mencionado que las señales de EEG son no estacionarias ya que cambian rápidamente con el tiempo. Además, estas características contienen información del tiempo debido a que los patrones de actividad cerebral están generalmente relacionados a variaciones específicas del EEG en el tiempo. Por esa razón se debe usar una representación que considere estos detalles. Como se ha visto en el capítulo 2, en las señales EEG surge el problema que ha sido denominado "maldición de la dimensionalidad", es por esta razón que se debe transformar la señal en un nuevo conjunto de valores con las cuales sea más fácil tratar.

Las características con DWT bior2.2 se deben calcular con distintos niveles de descomposición. Para describir la descomposición se parte de conocer la frecuencia de muestreo (f_N) que es de 128 Hz para la BD1 y 500 Hz para la BD2. Luego, el teorema de Nyquist dice que la frecuencia de muestreo debe ser por lo menos 2 veces mayor a la frecuencia de la señal a muestrear, es decir:

$$f_N \geq 2f_s, \text{ donde } f_s \text{ es la frecuencia de la señal muestreada.} \quad (4.2)$$

Despejando f_s , se obtiene: $f_s \approx \frac{f_N}{2}$. Por lo tanto para la BD1 se tiene que $f_s \approx 64Hz$. y para la BD2 $f_s \approx 250Hz$.

Por otro lado, también es necesario observar el tamaño de las muestras para determinar el máximo número de niveles de descomposición que pueden ser calculados. La tabla 4.5 muestra los posibles niveles de descomposición para cada base de datos.

Para seleccionar el conjunto de coeficientes DWT más adecuado para la señal de EEG se consideró lo planteado en [31, 5] donde menciona que las frecuencias de EEG llegan hasta 60 Hz y más allá de dichas frecuencias no tienen interés para nuestro análisis. Tomando en cuenta lo anterior se debe realizar la caracterización para los 27 sujetos de la BD1 y los 20 sujetos de la BD2 con la DWT bior 2.2 con 4 niveles de descomposición para la BD1 y con 5 y 6 para la BD2 para luego eliminar los coeficientes de detalle correspondientes a los rangos mayores a 60 Hz.

Tabla 4.5: Rangos de frecuencias(Hz) para el nivel de descomposición 4 de DWT para la BD1 y los niveles de descomposición 5 y 6 para la BD2.

BD1		BD2			
Nivel	Rango	Nivel	Rango	Nivel	Rango
D1	32-64	D1	125-250	D1	125-250
D2	16-32	D2	62.5-125	D2	62.5-125
D3	8-16	D3	31.25-62.5	D3	31.25-62.5
D4	4-8	D4	15.62-31.25	D4	15.62-31.25
A4	0-4	D5	7.81-15.62	D5	7.81-15.62
		A5	0-7.81	D6	3.9-7.81
				A6	0-3.9

4.3.2. Medidas estadísticas

Otra opción para la caracterización de la señal fue motivada por lo planteado en [46]. En dicho trabajo se caracterizó la señal al calcular un conjunto de valores estadísticos. Inicialmente tomamos 15 características estadísticas por canal. Es decir, para cada uno de los canales se calcularon las 15 características y luego se concatenaron, para tener un vector de características por cada instancia.

Las características de interés para este trabajo constan de 2 conjuntos: el primero es de 15 valores estadísticos y el segundo de 9. El primer conjunto(C1) son las características usadas por [46] que consta de 9 valores estadísticos y 6 combinaciones de algunos de ellos. En el segundo(C2) conjunto sólo usamos los 9 valores sin considerar las combinaciones, como se muestra a continuación.

- C1: media, media⁶, absoluto de la media, máximo, absoluto del máximo, mínimo, absoluto del mínimo, máximo+mínimo, máximo-mínimo, desviación estándar, varianza, curtosis, oblicuidad, suma y mediana.
- C2: media, máximo, mínimo, desviación estándar, varianza, curtosis, oblicuidad, suma y mediana.

4.4. Clasificación

El concepto de clasificar se refiere a la acción de organizar u ordenar. El propósito de este trabajo es organizar un conjunto de datos en clases que contengan elementos en común. Esto se detalla a continuación.

Una vez obtenidos los vectores de características se realiza la etapa de clasificación con los algoritmos *random forest*, *SVM* y *naive Bayes* con 10-pliegues. Estos algoritmo fueron seleccionados porque son los que muestra mejores resultados de acuerdo a la bibliografía descrita en el estado del arte.

Cada vector de características está asociado a una etiqueta de clase(arriba, abajo, izquierda, derecha y selección). Dicha etiqueta de clase es usada durante el entrenamiento del clasificador y posteriormente sirve como parámetro para comparar la salida del clasificador. Finalmente, el número de ejemplos clasificados correcta e incorrectamente, sirven para evaluar al clasificador con las medidas descritas en la sección 2.6.

Para comprobar si existe información única en los estados de reposo y en las demás clases, se plantean diferentes experimentos que ayudan a comprender que aunque el cerebro sea entrenado para estar en estado de reposo puede ser que exista información importante que relaciona una repetición de una palabra con la siguiente repetición(ver capítulo 5). Esto es muy relevante ya que se espera comprobar si se pueden separar las clases que contienen información lingüística con las que supuestamente no la contienen.

Capítulo 5

Experimentos y resultados

En el presente capítulo se presentan los resultados de los experimentos realizados calculando características con la distribución de la energía de la DWT bior2.2 y medidas estadísticas para luego clasificar usando el algoritmo de *random forest*, *SVM* y *naive Bayes*. Primero se presenta la clasificación de palabras imaginadas (Habla imaginada) para comprobar si existe información que pueda distinguir entre las diferentes clases (arriba, abajo, izquierda, derecha y selección en el caso de la BD1 y arriba, abajo, derecha e izquierda en el caso de la BD2). Luego las palabras y los estados de reposo o pausa son separados en actividad e inactividad. Después se comparan segmentos de la señal con diferentes tamaños con el fin de encontrar el tamaño más pequeño que aún contiene suficiente información para distinguir entre actividad e inactividad. En el apéndice A se presentan los resultados a detalle para cada uno de los sujetos de las dos bases de datos al realizar los experimentos que a continuación se exponen.

5.1. Clasificación de palabras imaginadas

Después de aplicar el método descrito anteriormente se entrenaron tres clasificadores: *random forest*, *SVM* y *naive Bayes* con las cinco palabras de la BD1 y las 4 de la BD2, esto se realiza para cada uno de los sujetos.

En la figura 5.1 se muestra la exactitud promedio obtenida utilizando la BD1 para la clasificación multiclase de habla imaginada. Aquí se puede

observar que al utilizar la DWT y luego calcular la distribución de la energía instantánea, además, el clasificador que resultó mejor es *random forest* que en promedio muestra una exactitud de 0,77. Otro clasificador que dio buenos resultados es *SVM* con un promedio de exactitud de 0,76 también usando DWT y calculando la distribución de la energía instantánea.

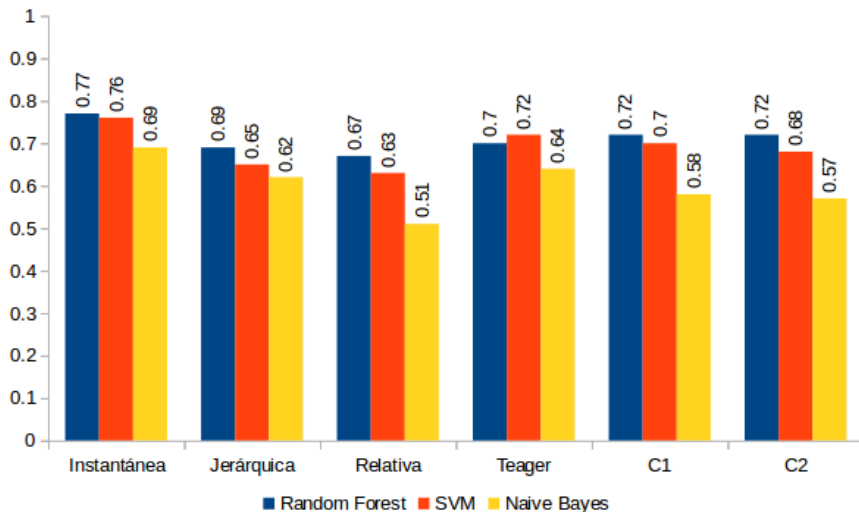


Figura 5.1: Promedios de exactitud con los tres clasificadores para los individuos de la BD1, características DWT bior2.2 con 4 niveles de descomposición y estadísticos (EST) para habla imaginada.

Para realizar este experimento en la BD2 se utilizó sólo la parte de habla imaginada, según la figura 4.5. Cabe destacar que por el diseño del protocolo de estimulación y adquisición, no es posible saber exactamente en donde inició la pronunciación imaginada de las palabras. Por lo que para los fines de clasificación se consideró que la palabra imaginada se encontraba $\frac{1}{2}$ segundo antes de mandar el marcador de final *RN*.

Como ya se vio en el método seguido del capítulo 4, para la BD2 se utilizó la DWT con 5 y 6 niveles de descomposición para luego eliminar los primeros *n* niveles que corresponden a rangos de frecuencia más allá de 60 Hz. Luego a estos vectores se les calculó la distribución de la energía.

En la figura 5.2 se puede ver el promedio de exactitud al realizar

clasificación multiclase. Aunque la exactitud está por encima de azar [13], comparados con los resultados de la BD1 son considerablemente más bajos.

Los resultados indican que la clasificación multiclase no muestra buenos resultados porque posiblemente las palabras imaginadas tienen características en común. Esto se puede entender mejor al realizar clasificación de todas las palabras imaginadas como una sola clase, como se verá más adelante.

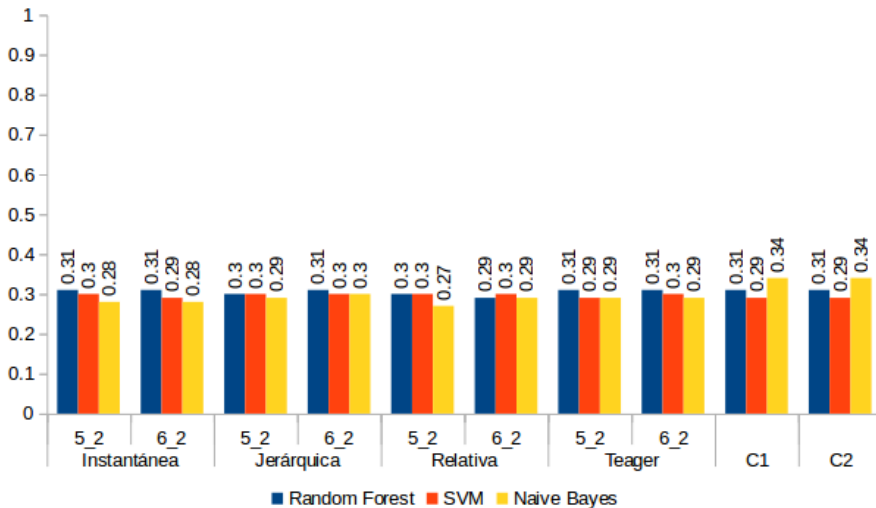


Figura 5.2: Promedios de exactitud con los tres clasificadores para los individuos de la BD2 en habla imaginada.

5.2. Clasificación actividad e inactividad lingüística

El reconocimiento de habla imaginada es muy importante para el desarrollo de una BCI, sin embargo, en una aplicación en línea esto posiblemente no sería de gran utilidad pues entre palabras imaginadas pudiera existir otro tipo de actividad diferente a palabras imaginadas. Es por ello que se debe estudiar el comportamiento y realizar clasificación

de los que pudieran ser otros tipos de actividad cerebral y las palabras imaginadas considerando estos como sólo dos clases.

Para este experimento las señales de EEG fueron separadas en 2 clases, un conjunto de palabras imaginadas consideradas la clase de *actividad lingüística*; y los estados de reposo o pausa como ejemplos de otra actividad cerebral. Estos últimos denominados por simplicidad de *inactividad lingüística*.

Para la BD1 la clase actividad contiene 165 instancias que corresponden a las 33 repeticiones de cada una de las 5 palabras. La clase inactividad contiene 165 instancias que corresponden a los estados de reposo que existen entre la repetición de una palabra y otra. Eso para cada uno de los 27 sujetos.

La figura 5.3 muestra los resultados de clasificar la actividad e inactividad cerebral lingüística con el algoritmo de *random forest*, *SVM* y *naive Bayes*. Para la caracterización se utilizó la DWT calculando la distribución de la energía como en el experimento anterior y medidas estadísticas. Los resultados muestran que al calcular características con medidas estadísticas las señales parecen distinguirse de una mejor manera. Al usar el conjunto de características C2 y con el algoritmo *random forest* se obtiene una exactitud promedio de 0,83. En general, con todas las formas de obtener características y con los tres algoritmos de clasificación se obtienen resultados por encima del azar.

Para la BD2 la clase considerada como actividad contiene 160 instancias que corresponden a las 40 repeticiones de cada una de la frase y las 4 palabras imaginadas entre los marcadores *S3* a *RN* según la figura 4.5 correspondiente al protocolo de adquisición de la señal. La clase inactividad contiene 160 muestras que corresponden a 40 instancias de el pre-estímulo, 40 instancias de la presentación del objeto, 40 instancias de la respuesta por teclado y finalmente 40 instancias de la pausa.

Los resultados utilizando la BD2 son particularmente interesantes, pues como se vio en la clasificación de habla imaginada los resultados parecían indicar que no existía información suficiente para distinguir entre clases. Al separar en las clases de en actividad y considerar la inactividad sí existe información que logra distinguir con una exactitud promedio muy por arriba del azar, en el peor caso se obtiene una exactitud de 0,84 con *naive Bayes*.

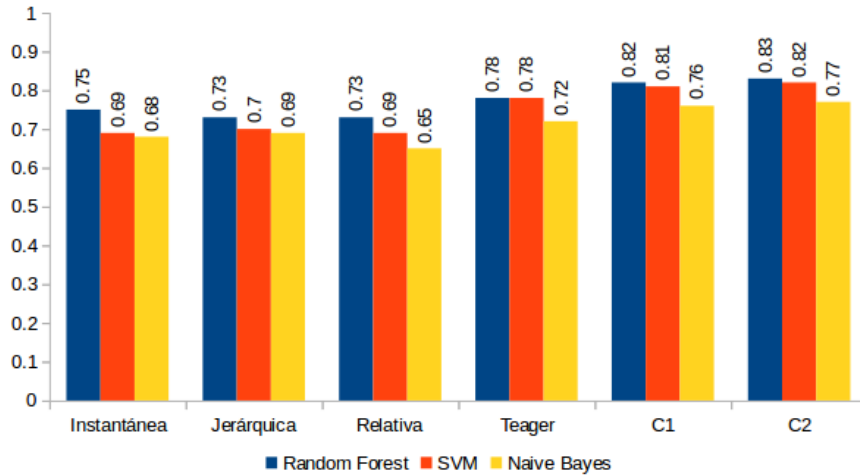


Figura 5.3: Promedios de exactitud con los tres clasificadores usando la BD1 para distinguir actividad e inactividad lingüística.

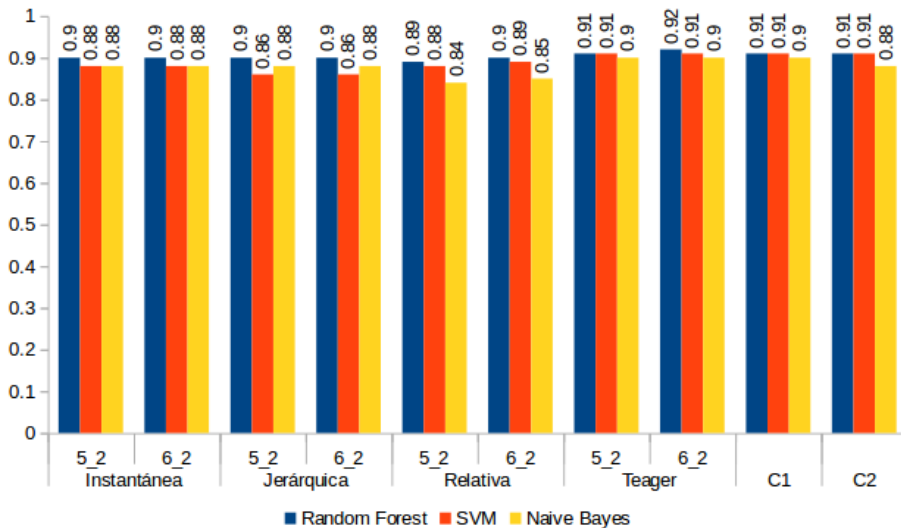


Figura 5.4: Promedios de exactitud con los tres clasificadores usando la BD2 para distinguir actividad e inactividad lingüística.

Lo planteado anteriormente es muy interesante pues muestra que es posible distinguir con buenos resultados la actividad e inactividad cerebral lingüística. Esto permitirá que al orientar futuros trabajos hacia una BCI en línea un proceso inicial logre encontrar en donde se trata de actividad y que luego un proceso posterior puede utilizar solo estos segmentos para realizar clasificación multiclase.

128128

5.2.1. Análisis sobre el número de canales y frecuencia de muestreo.

El experimento anterior mostró que al usar la BD2 se obtienen mejores resultados respecto a usar la BD1, concluyendo que dichos resultados están relacionados con la frecuencia de muestreo y el número de canales, que son diferentes en ambas bases de datos. En la BD2 se cuenta con 64 canales localizados de acuerdo al sistema internacional 10-20, con una frecuencia de muestreo de 500 Hz. Aquí surge la siguiente pregunta *¿Qué tanto afecta esta diferencia de características en las bases de datos?*

Para este experimento sólo se consideró la BD2. Para poder analizar el impacto de usar menos canales que los 64 de BRAIN VISION, se utilizaron los mismos canales del EMOTIV EPOC (ver figura 4.4), esto se puede realizar ya que las dos bases de datos fueron capturadas siguiendo el sistema internacional 10-20 y por lo tanto hay correspondencia de canales. Después, para el caso de la frecuencia de muestreo original de 500 Hz, la señal de EEG fue submuestreada para lograr tener una frecuencia de muestreo de 128 Hz.

Por los resultados obtenidos, en la clasificación de actividad e inactividad, sólo se utilizó *random forest* y luego para obtener características se usaron las medidas estadísticas C2.

En la figura 5.5 se muestran los resultados obtenidos. Primero se muestran los resultados obtenidos en el experimento anterior utilizando todos los canales y la frecuencia de muestreo de BRAIN VISION (500 Hz.). Después, sólo se consideraron los 14 canales en común con EMOTIV EPOC utilizando la misma frecuencia de muestreo, esto último con el fin de analizar si la frecuencia de muestreo afecta a los resultados obtenidos.

En la figura 5.5 se puede observar que al utilizar los mismos canales que

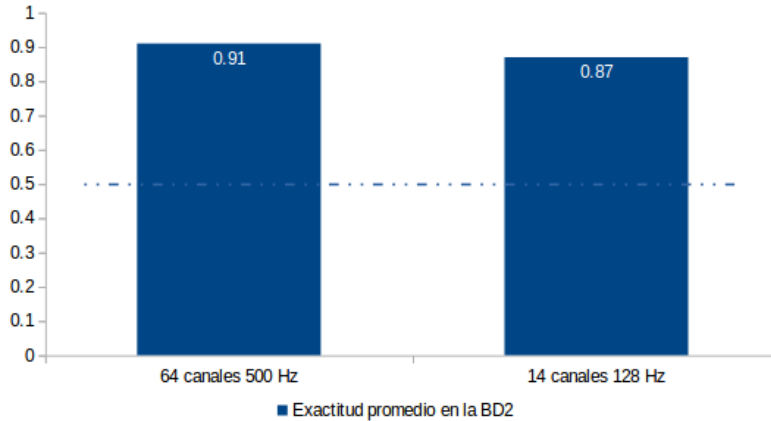


Figura 5.5: Promedios de exactitud con los tres clasificadores usando la BD1 para distinguir actividad e inactividad lingüística. Considerando 14 canales y una frecuencia de muestreo de 128 Hz.

se utilizaron en la BD1 los resultados son comparables con los resultados obtenidos usando los 64 canales. Sin embargo, también muestran que los resultados obtenidos si se ven beneficiados por el número de canales y la frecuencia de muestreo de BRAIN VISION. Los resultados a detalle de este experimento se pueden ver en la tabla A.13

Para comprobar que las modificaciones (selección de canales y submuestreo) realizadas se ven reflejadas en los resultados, se realizó una prueba estadística con la cual se comprobó que los resultados son estadísticamente significativos ($p\text{-valor} = 0,0103$ con $\alpha = 0,05$).

5.3. División en segmentos de señal de EEG que aún contienen información lingüística

Por los resultados de los experimentos anteriores, de aquí en adelante sólo se usará el algoritmo *random forest* y las características estadísticas C2 pues se obtienen los mejores resultados y el tiempo de procesamiento es mucho menor.

Para tener una respuesta más rápida se deben encontrar segmentos de

tamaño pequeño, para este propósito la señal de EEG de las dos bases de datos fueron divididas en segmentos de $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ y $\frac{1}{16}$. Para el tamaño de segmento de $\frac{1}{4}$ de segundo, para cada instancia se tomó segmentos de dicho tamaño desde el inicio y hasta el final de la señal. Por ejemplo, para una instancia de tamaño 256 se tomaron 8 segmentos de señal de EEG y se consideraron como actividad o inactividad lingüística según el fuera el caso.

Al dividir la actividad e inactividad en los segmentos antes mencionados, se debe tener cuidado al mezclar los datos para el entrenamiento y la validación ya que pudiera ser que algunos segmentos de una instancia de actividad quede tanto en entrenamiento como en la validación. Con lo anterior se obtendrían resultados erróneos pues los datos de entrenamiento ya contendrían partes de lo que se quiere clasificar, obteniendo así altas tasas de exactitud erróneamente.

Para asegurar que segmentos de las muestras divididas no estén tanto en el entrenamiento como en la validación, primero se separó el 20 % para validación y el 80 % para entrenar como se muestra en figura 5.6 para la BD1. Este proceso se realizó para formar 10 conjuntos de entrenamiento y 10 de validación en las dos bases de datos para hacer la validación cruzada con 10 pliegues.

La exactitud promedio luego de segmentar la señal de EEG, correspondiente a actividad e inactividad, se muestra en la figura 5.7. En dicha figura se puede observar que al dividir en segmentos cada vez más pequeños la exactitud decae. Para el caso de las dos bases de datos, al dividir la señal en segmentos de $\frac{1}{8}$ de segundo y de $\frac{1}{16}$ la exactitud ya no decayó. Los resultados parecen indicar que se puede identificar con exactitud muy por arriba del azar si se trata de actividad o de inactividad aún en segmentos muy pequeños.

En la figura 5.7 se muestra que para el caso de la BD2, al tomar segmentos de $\frac{1}{2}$, se obtiene peor exactitud que al tomar segmentos más pequeños de la señal. Esto pudiera estar relacionado con el protocolo de adquisición de la señal, pues como se mencionó, hay una frase antes de la palabra imaginada. Tomando en cuenta lo anterior y el tamaño promedio entre los marcadores $S3$ y RN se sabe que es 1.7 segundos, al tomar segmentos de $\frac{1}{2}$ de seg posiblemente las muestras contengan una parte de la frase y otras muestras partes de la palabra imaginada. Por lo que

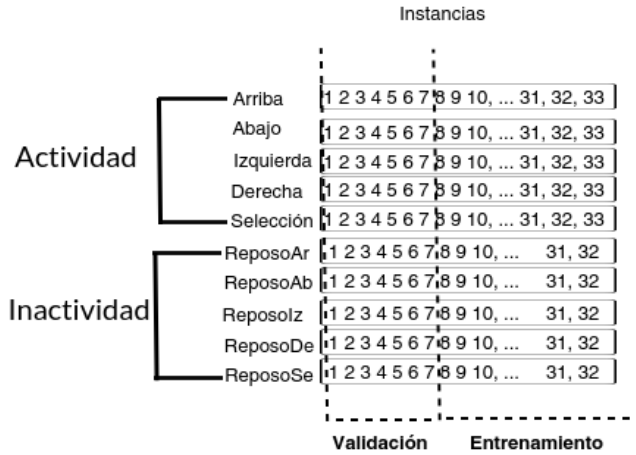


Figura 5.6: Procedimiento para separar las instancias de entrenamiento y validación en la BD1.

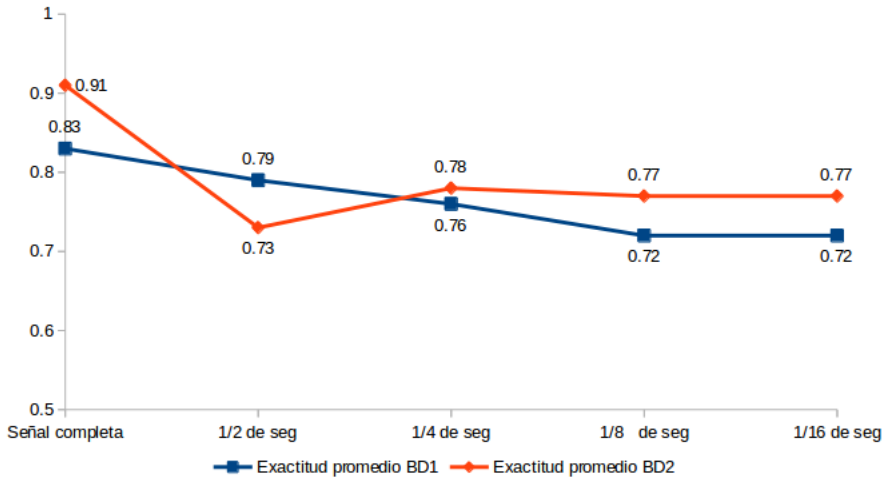


Figura 5.7: Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD1 y BD2.

al dividir en segmentos más pequeños este problema no se presenta pues, aunque se divida en partes de la frase y partes de habla imaginada, existen

más muestras con las cuales comparar.

5.4. Identificando periodos de actividad e inactividad lingüística de nuevos sujetos

Para el caso de crear una interfaz para un nuevo sujeto es necesario adquirir datos para entrenar el clasificador propuesto. El objetivo del presente experimento es observar si es posible reutilizar los datos ya adquiridos de otros sujetos para ser utilizados por nuevos sujetos.

Sin embargo, el primer paso para validar si existe información común en las instancias de los sujetos puede ser, crear un clasificador con ciertos sujetos para luego validar con uno nuevo que no esté en el entrenamiento, este enfoque es conocido como *transferencia de clasificador* [43]. Para realizar este experimento se construyó un clasificador con la información de 26 sujetos para luego validar con el restante. Por ejemplo, para validar al clasificador con el sujeto uno, el clasificador se construyó con los sujetos 2,3,...,27 en el caso de la BD1. Lo anterior se repite para cada uno de los sujetos.

Para ello se uso la señal dividida en segmentos de $\frac{1}{2}$ y $\frac{1}{16}$ de segundo de las dos bases de datos para comprobar si aún en segmentos más pequeños de la señal se logran distinguir las clases sin usar técnicas sofisticadas. En las tabla A.15 hasta la A.19 del apéndice se muestran los resultados (Exactitud, recuerdo y medida F) obtenidos al entrenar los clasificadores para luego validar con cada uno de los sujetos.

La figura 5.8 muestra una comparación de la exactitud promedio de la tabla A.14 (Exp 3) y la exactitud promedio obtenida al validar a los sujetos en un modelo que no los incluía en el entrenamiento (Exp 4) con los segmentos de señal antes mencionados.

En la figura 5.9 muestra una comparación de la exactitud promedio de la tabla A.14 (Exp3) y la exactitud promedio obtenida al validar a los sujetos en un modelo que no los incluía en el entrenamiento (Exp 4).

Aún sin utilizar ninguna técnica para hacer la transferencia del aprendizaje, los resultados obtenidos muestran que las instancias de los sujetos, sólo en la primer base de datos, contienen información común. Esto es muy interesante pues muestra que posiblemente al utilizar técnicas

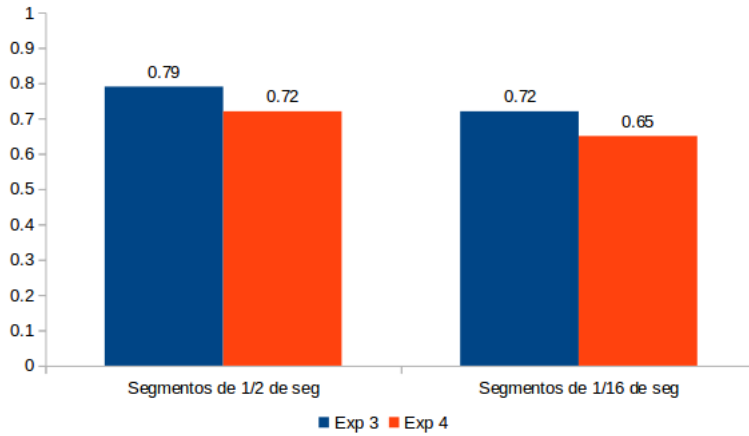


Figura 5.8: Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD1 al validar sujetos que no estaban en el entrenamiento.

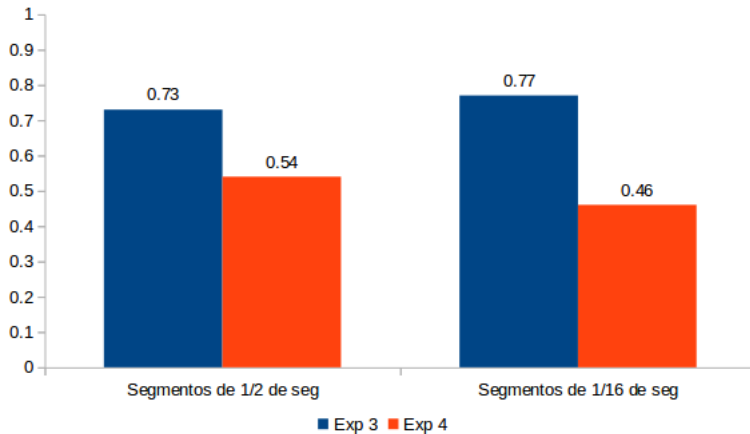


Figura 5.9: Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD2 al validar sujetos que no estaban en el entrenamiento.

adicionales de transferencia de aprendizaje se logren tasas de exactitud aún mejores.

Al analizar la figura 5.8 y 5.9 se puede notar que en la BD2 la

transferencia del clasificador funciona mejor. Esto puede estar relacionado a varios factores, por ejemplo, en la BD2 la frecuencia de muestreo y el número de canales es mayor que en la BD1. Aunque para clasificar actividad e inactividad lingüística, tener muchos canales sirvió para tener más información y lograr una mejor separación, para hacer transferencia de clasificador parece no ser tan bueno, pues el nuevo sujeto tiene que tener información en común de más canales. Para dar más información sobre lo anterior, se plantea el siguiente experimento que muestra una conclusión con evidencia experimental.

5.4.1. Análisis sobre el número de canales y frecuencia de muestreo en transferencia de clasificador

Para este experimento sólo se consideró la BD2. Para poder analizar el impacto de usar menos canales que los 64 de BRAIN VISION, se utilizaron los mismos canales del EMOTIV EPOC (ver figura 4.4), esto se puede realizar ya que las dos bases de datos fueron capturadas siguiendo el sistema internacional 10-20 y por lo tanto hay correspondencia de canales. Después, para el caso de la frecuencia de muestreo original de 500 Hz, la señal de EEG fue submuestreada para lograr tener una frecuencia de muestreo de 128 Hz.

Por los resultados obtenidos, en la clasificación de actividad e inactividad, sólo se utilizó *Random Forest* y luego para obtener características se usaron las medidas estadísticas C2. Los resultados de este experimento se muestran a continuación en la figura 5.10, los resultados a detalle se muestran en la tabla A.20 y A.21.

Como se puede observar en la figura 5.10, la exactitud al usar los 64 canales y frecuencia de muestreo de 500 Hz. está apenas por arriba de la cota del azar para esta experimento. Después al usar sólo los 14 canales iguales a los usados en la BD1 y haciendo submuestreo para obtener una frecuencia de muestreo de 128 Hz., los resultados muestran que al usar menos canales la exactitud no mejora (como se planteó en el experimento anterior.). Esto permite conjeturar que la transferencia de aprendizaje para el caso de clasificación de actividad e inactividad lingüística, no es posible, por lo menos sin hacer uso de técnicas adicionales que permitan calibración. En otras palabras, esto significa que existe una diferencia entre las señales

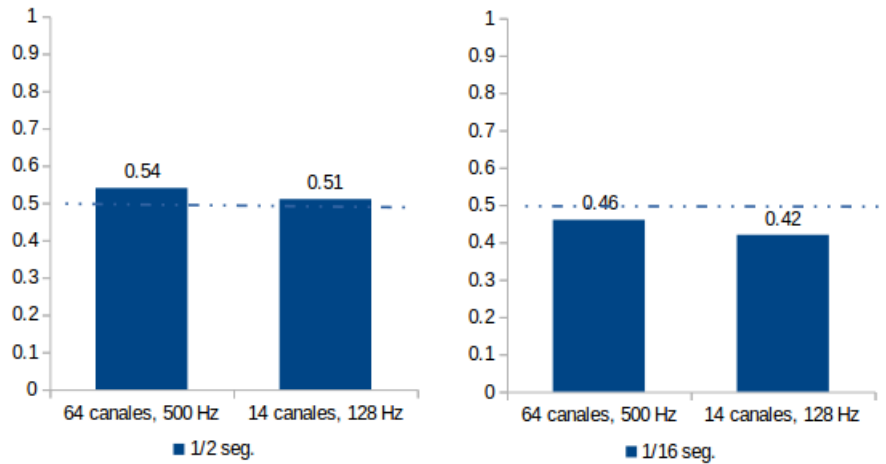


Figura 5.10: Exactitud promedio obtenida con segmentos de señal de $\frac{1}{2}$ y $\frac{1}{16}$ en la BD2 al validar sujetos que no estaban en el entrenamiento. Comparación del uso de 64 y 14 canales, con frecuencias de 500 y 128 Hz.

de EEG de los sujetos (al menos para el caso de actividad e inactividad lingüística) de ahí que esta diferencia podría aprovecharse para generar modelos biométricos orientados a reconocer sujetos para aplicaciones en sistemas de seguridad [6].

Capítulo 6

Conclusiones y trabajo futuro

El reconocimiento de habla imaginada, a partir de señales de EEG, ha sido estudiado en sistemas fuera de línea como una tarea de clasificación de distintas palabras de un vocabulario propuesto. En el primer experimento se realizó clasificación multiclase con distintas maneras de representación de la señal, con el fin de corroborar si esto era posible y analizar los resultados.

El algoritmo *random forest* presenta altas tasas de exactitud para la BD1 en la tarea de clasificación multiclase de habla imaginada, además, al calcular la distribución de la energía instantánea la señal de EEG se logró distinguir de mejor manera. Por lo menos para esta base de datos el algoritmo *naive Bayes*, es con el que se obtiene los peores resultados, estando aún arriba de la cota del azar.

Los resultados de la clasificación multiclase en la tarea de habla imaginada usando la BD2 son en el mejor caso 0,34. Esto pudiera ser por el protocolo de estimulación y adquisición de la señal de EEG, ya que las palabras imaginadas eran precedidas por una frase que no se marcó en donde terminaba y esto obligó a tomar el último segmento donde posiblemente se imaginó la palabra (Ver sección 5.1).

Para la parte medular de este trabajo se presentó un estudio para identificar la actividad e inactividad lingüística con el fin de dar el primer paso hacia el reconocimiento en línea. Los experimentos realizados a pesar de llevarse a cabo fuera de línea dan evidencia de la factibilidad

de hacer la tarea en línea. Se analizó si, independientemente de la palabra imaginada (actividad lingüística) es posible diferenciarse de otras actividades cerebrales (inactividad lingüística).

A pesar de que sólo se evaluó en dos bases de datos los resultados son alentadores indicando que es posible distinguir entre actividad o inactividad lingüística. Los resultados alcanzados están muy por arriba de la cota del azar, lo que permite conjeturar que es posible crear un proceso en línea que logre clasificar una señal como actividad o inactividad y posteriormente con otro clasificador identificar la palabra imaginada que finalmente se convertirá en un comando para fines de comunicación o control de dispositivos.

Al comparar los resultados de las tablas A.3 y A.4 se puede notar que siguiendo el mismo procedimiento en ambas bases de datos se obtiene mayor exactitud en la BD2. Son varios factores por los cuales la exactitud es mayor para la BD2. Como puede verse en la tabla 4.4 la BD2 comprende un número mayor de repeticiones de cada palabra, una frecuencia de muestreo mayor y se cuenta con información de un mayor número de canales (64 canales).

Por otra parte, este trabajo evaluó si la caracterización de las señales de EEG tiene algún efecto en la discriminación de habla imaginada y de actividad e inactividad lingüística. Las caracterizaciones estudiadas fueron características estadísticas obtenidas desde las señales en tiempo y energías calculadas a partir de la DWT. Usando los vectores estadísticos se obtuvieron resultados iguales o superiores que con las energías wavelet; con la ventaja de que éstos no requieren de una transformación a otro dominio, ni definición de parámetros adicionales a la señal.

El estudio de la distribución de la energía en los vectores de coeficientes, resultado de aplicar la DWT bior2.2, resultó de gran interés para comparar la tarea de clasificación de habla imaginada y la de clasificar actividad e inactividad lingüística. Para la primera tarea, los mejores resultados usando sólo la distribución de la energía se obtuvo con la energía instantánea. Esto posiblemente esta relacionado con el hecho de que la amplitud de las señales de EEG de una misma palabra en diferentes sujetos, es semejante y como se mencionó en el marco teórico, la energía instantánea hace precisamente este análisis. En la tarea de clasificación de actividad e inactividad lingüística, los mejores resultados se obtienen al usar la distribución de la energía teager

y esto puede ser que se relacione con el hecho de que para poder hacer una mejor discriminación, no es suficiente el análisis de la señal en amplitud, sino que además se debe hacer el análisis en la frecuencia. Al hacer el análisis en amplitud y frecuencia puede ser que las palabras imaginadas (actividad lingüística) tengan más información semejante que con el resto de actividad lingüística (inactividad lingüística).

Las señales de EEG se dividieron en segmentos de diferentes tamaños con el fin de dar evidencia experimental de la existencia de información que permite realizar la clasificación, al disminuir el tamaño del segmento de señal de EEG, la exactitud se ve afectada, decae entre 10% y 15% respecto a usar la señal completa. Sin embargo, la exactitud aún están por arriba de la cota del azar y esto permite suponer que se pueden tener respuestas rápidas a estímulos al usar una BCI. Estos resultados muestran la factibilidad de segmentar las señales para obtener una idea de donde inicia y termina la actividad lingüística para orientar los trabajos futuros hacia la construcción de una BCI en línea.

Por último, al evaluar a sujetos en clasificadores que no contenían información sobre ellos no se logró tener buenas tasas de exactitud. Esto parece indicar que las repeticiones de las palabras imaginadas entre sujetos no contienen información común y como se planteo en este experimento, esto puede ser usado como sistema de seguridad biométrico para identificar sujetos usando habla imaginada.

6.1. Trabajo futuro

Los experimentos expuestos fueron probados en dos bases de datos, es así que, sería interesante saber si el método seguido se puede generalizar a otras formas de estimulación y adquisición de las señales de EEG. En este sentido, también sería interesante reproducir los experimentos con otras bases de datos. En este trabajo se usó una sola forma de preprocesar la información, algunos algoritmos de clasificación necesitan datos tratados de una manera diferente y posiblemente esto permita distinguir las clases tanto de habla imaginada como de actividad e inactividad con exactitud mayor a la obtenida con el método propuesto de este trabajo.

La selección de los canales que aportan más información ha sido estudiada y se han aplicado diferentes mecanismos para identificarlos, como

trabajo futuro se plantea que se pueden buscar si los canales que resultan mejor para distinguir habla imaginada (clasificación multiclase) también son útiles para la clasificación de actividad e inactividad lingüística. De no ser así, identificar o dar evidencia de cuáles resultan mejor para la tarea. Esto es importante pues aporta ganancia a la hora de implementar una BCI en línea, pues se tendrían menos datos y más información (de acuerdo a los canales que sean seleccionados) para analizar e identificar cuando se trata de una u otra clase.

Como se vió, la actividad e inactividad lingüística se puede distinguir aun en segmentos pequeños. Esto permite conjeturar la posibilidad de tomar segmentos y evaluarlos en un clasificador con características como las usadas en este trabajo para así realizar clasificación en línea. Posiblemente será necesario plantear reglas para reevaluar los segmentos en donde no sea clara la clasificación pero en general se podría concluir que al tomar muestras lo suficientemente pequeñas se podría encontrar en donde inicia y donde termina la actividad cerebral lingüística, como se puede observar en la figura 6.1.

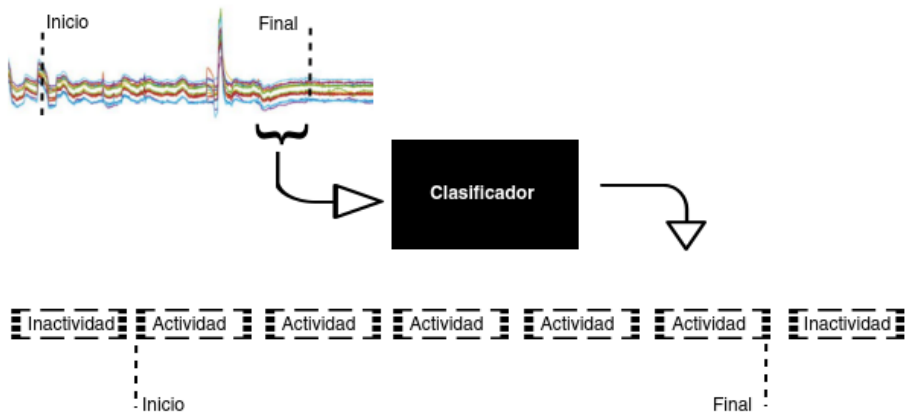


Figura 6.1: Posible proceso para detectar el inicio y final de la actividad cerebral lingüística.

En un caso hipotético, las muestras de inactividad que se encuentren precedidas por actividad y sucedidas por actividad, pudiera ser que hubieran sido clasificadas erróneamente y para esto es necesario fijar

CAPÍTULO 6. CONCLUSIONES Y TRABAJO FUTURO

algunas reglas que permitan asegurarlo. Esto podría resolverse planteando ventanas traslapadas y lo mismo se tendría que aplicar para muestras de actividad precedidas por inactividad y sucedidas por inactividad.

Apéndice A

Resultados a detalle de los experimentos

A continuación se presentan los resultados los experimentos realizados para cada sujeto a detalle. De la tabla A.1 a la tabla A.12, para la extracción de características se usa la DWT con diferentes niveles de descomposición y medidas estadísticas, para luego clasificar con *random forest*, *SVM*, y *naive Bayes*.

De la tabla A.14 en adelante solo se utiliza el conjunto de medidas estadísticas C2 y el algoritmo de *random forest*, como se explica en el capítulo 5.

Además, de la tabla A.16 en adelante, el encabezado EV, se refiere a utilizar al sujeto en cuestión como datos de validación y a los restantes como entrenamiento. Por ejemplo, en la tabla A.16, la exactitud 0,51 del sujeto 1, se obtiene al crear un clasificador con los sujetos 2 al 27 para luego utilizar los datos del primer sujeto como datos de validación.

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.1: Exactitud, media y desviación estándar (Std) obtenida con *random forest* para los individuos de la BD1 para habla imaginada.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.92	0.86	0.87	0.92	0.76	0.76
2	0.59	0.53	0.55	0.58	0.52	0.54
3	0.80	0.81	0.65	0.68	0.84	0.85
4	0.90	0.86	0.69	0.85	0.92	0.89
5	0.81	0.73	0.74	0.73	0.77	0.75
6	0.37	0.45	0.37	0.35	0.50	0.47
7	0.84	0.76	0.69	0.58	0.80	0.79
8	0.96	0.89	0.93	0.94	0.92	0.91
9	0.82	0.73	0.72	0.83	0.66	0.64
10	0.87	0.83	0.71	0.57	0.85	0.83
11	0.89	0.86	0.90	0.79	0.91	0.92
12	0.86	0.83	0.66	0.79	0.87	0.85
13	0.84	0.77	0.74	0.75	0.75	0.72
14	0.59	0.40	0.41	0.57	0.57	0.53
15	0.91	0.84	0.82	0.82	0.78	0.80
16	0.77	0.63	0.68	0.63	0.57	0.62
17	0.70	0.61	0.66	0.66	0.83	0.82
18	0.93	0.78	0.73	0.89	0.74	0.71
19	0.56	0.48	0.48	0.46	0.56	0.54
20	0.93	0.87	0.85	0.88	0.76	0.75
21	0.55	0.50	0.44	0.59	0.65	0.69
22	0.82	0.70	0.77	0.75	0.78	0.78
23	0.71	0.61	0.60	0.69	0.68	0.68
24	0.60	0.51	0.67	0.65	0.78	0.71
25	0.46	0.43	0.39	0.36	0.45	0.46
26	0.80	0.67	0.68	0.74	0.71	0.75
27	0.92	0.71	0.72	0.90	0.63	0.66
Media	0.77	0.69	0.67	0.70	0.72	0.72
Std	±0.16	±0.15	±0.15	±0.16	±0.13	±0.12

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.2: Exactitud, media y desviación estándar (Std) obtenida con *SVM* para los individuos de la BD1 para habla imaginada.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.90	0.79	0.85	0.89	0.71	0.67
2	0.63	0.44	0.41	0.57	0.41	0.35
3	0.77	0.66	0.53	0.64	0.80	0.81
4	0.83	0.69	0.63	0.78	0.88	0.88
5	0.85	0.72	0.66	0.84	0.75	0.70
6	0.36	0.43	0.41	0.37	0.45	0.44
7	0.86	0.80	0.69	0.62	0.77	0.81
8	0.96	0.89	0.95	0.96	0.90	0.90
9	0.84	0.66	0.64	0.79	0.57	0.52
10	0.75	0.66	0.71	0.63	0.88	0.88
11	0.92	0.84	0.77	0.92	0.87	0.83
12	0.88	0.72	0.58	0.80	0.85	0.82
13	0.86	0.78	0.78	0.74	0.77	0.76
14	0.55	0.48	0.37	0.62	0.51	0.48
15	0.89	0.79	0.80	0.86	0.77	0.79
16	0.75	0.57	0.69	0.73	0.65	0.64
17	0.73	0.57	0.60	0.75	0.79	0.79
18	0.90	0.78	0.50	0.88	0.70	0.67
19	0.59	0.55	0.56	0.51	0.49	0.51
20	0.89	0.82	0.81	0.92	0.79	0.72
21	0.54	0.52	0.47	0.56	0.68	0.62
22	0.78	0.65	0.72	0.68	0.72	0.72
23	0.67	0.61	0.52	0.65	0.69	0.64
24	0.61	0.51	0.62	0.71	0.70	0.68
25	0.45	0.40	0.46	0.43	0.48	0.45
26	0.76	0.60	0.53	0.72	0.71	0.64
27	0.89	0.71	0.62	0.89	0.52	0.55
Media	0.76	0.65	0.63	0.72	0.70	0.68
Std	±0.16	±0.14	±0.15	±0.15	±0.14	±0.15

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.3: Exactitud, media y desviación estándar (Std) obtenida con *naive Bayes* para habla imaginada.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.88	0.77	0.65	0.89	0.51	0.52
2	0.54	0.43	0.40	0.53	0.38	0.37
3	0.76	0.74	0.51	0.59	0.73	0.72
4	0.85	0.79	0.50	0.74	0.82	0.79
5	0.75	0.66	0.46	0.68	0.51	0.55
6	0.33	0.38	0.31	0.33	0.46	0.46
7	0.74	0.72	0.53	0.55	0.69	0.68
8	0.92	0.82	0.80	0.88	0.85	0.84
9	0.74	0.65	0.44	0.74	0.26	0.29
10	0.68	0.68	0.58	0.46	0.73	0.73
11	0.67	0.67	0.59	0.56	0.70	0.67
12	0.76	0.77	0.50	0.72	0.70	0.62
13	0.75	0.75	0.57	0.70	0.48	0.50
14	0.54	0.38	0.26	0.50	0.37	0.37
15	0.88	0.74	0.75	0.84	0.67	0.66
16	0.72	0.54	0.57	0.57	0.55	0.48
17	0.64	0.61	0.49	0.61	0.63	0.63
18	0.85	0.76	0.47	0.90	0.57	0.55
19	0.51	0.40	0.36	0.40	0.46	0.47
20	0.90	0.83	0.75	0.90	0.61	0.64
21	0.46	0.47	0.36	0.47	0.56	0.57
22	0.72	0.63	0.66	0.70	0.71	0.71
23	0.62	0.58	0.44	0.64	0.53	0.53
24	0.57	0.47	0.45	0.62	0.57	0.56
25	0.47	0.45	0.34	0.40	0.44	0.45
26	0.62	0.56	0.43	0.62	0.54	0.56
27	0.72	0.59	0.49	0.69	0.58	0.56
Media	0.69	0.62	0.51	0.64	0.58	0.57
Std	±0.15	±0.14	±0.13	±0.16	±0.14	±0.13

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.4: Exactitud, media y desviación estándar (Std) obtenidos con *random forest* usando los individuos de la BD2 para habla imaginada.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.36	0.40	0.31	0.37	0.31	0.33	0.31	0.29	0.27	0.23
2	0.31	0.36	0.29	0.34	0.34	0.27	0.33	0.37	0.36	0.33
3	0.27	0.30	0.29	0.32	0.31	0.31	0.29	0.30	0.25	0.25
4	0.37	0.33	0.32	0.33	0.30	0.26	0.30	0.34	0.35	0.36
5	0.33	0.25	0.33	0.28	0.29	0.25	0.33	0.27	0.30	0.29
6	0.40	0.31	0.28	0.33	0.34	0.37	0.33	0.34	0.33	0.33
7	0.24	0.24	0.23	0.23	0.21	0.22	0.22	0.24	0.27	0.30
8	0.25	0.33	0.39	0.32	0.31	0.28	0.33	0.28	0.29	0.31
9	0.34	0.31	0.31	0.32	0.29	0.26	0.31	0.29	0.32	0.36
10	0.29	0.31	0.27	0.33	0.31	0.37	0.30	0.33	0.26	0.28
11	0.27	0.22	0.28	0.27	0.23	0.22	0.26	0.25	0.33	0.33
12	0.36	0.34	0.40	0.34	0.36	0.31	0.37	0.34	0.43	0.37
13	0.32	0.31	0.30	0.32	0.30	0.24	0.36	0.33	0.32	0.24
14	0.29	0.36	0.32	0.36	0.35	0.32	0.38	0.43	0.30	0.29
15	0.33	0.32	0.32	0.36	0.29	0.30	0.35	0.33	0.35	0.36
16	0.26	0.25	0.26	0.27	0.25	0.25	0.28	0.29	0.26	0.24
17	0.27	0.30	0.35	0.28	0.36	0.28	0.32	0.29	0.30	0.28
18	0.37	0.35	0.39	0.31	0.36	0.36	0.37	0.35	0.34	0.35
19	0.37	0.37	0.29	0.30	0.36	0.35	0.36	0.34	0.30	0.32
20	0.29	0.26	0.28	0.32	0.26	0.29	0.29	0.30	0.33	0.28
Media	0.31	0.31	0.30	0.31	0.30	0.29	0.31	0.31	0.31	0.31
Std	±0.04	±0.04	±0.04	±0.03	±0.04	±0.04	±0.04	±0.04	±0.04	±0.04

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.5: Exactitud, media y desviación estándar (Std) obtenidos con *SVM* usando los individuos de la BD2 para habla imaginada.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.40	0.42	0.34	0.36	0.35	0.38	0.33	0.41	0.29	0.29
2	0.28	0.27	0.32	0.27	0.19	0.32	0.31	0.32	0.29	0.29
3	0.29	0.29	0.29	0.33	0.27	0.30	0.27	0.29	0.29	0.29
4	0.29	0.30	0.26	0.29	0.23	0.23	0.28	0.26	0.29	0.29
5	0.29	0.30	0.36	0.36	0.29	0.26	0.28	0.31	0.29	0.29
6	0.29	0.30	0.26	0.39	0.33	0.26	0.27	0.30	0.29	0.29
7	0.25	0.22	0.25	0.21	0.21	0.24	0.25	0.23	0.29	0.29
8	0.28	0.25	0.28	0.24	0.29	0.30	0.28	0.30	0.29	0.29
9	0.32	0.28	0.30	0.22	0.36	0.33	0.26	0.25	0.29	0.29
10	0.36	0.31	0.33	0.29	0.35	0.34	0.32	0.34	0.29	0.29
11	0.28	0.25	0.27	0.23	0.20	0.21	0.31	0.31	0.29	0.29
12	0.39	0.36	0.35	0.35	0.32	0.30	0.37	0.34	0.29	0.29
13	0.27	0.34	0.31	0.40	0.26	0.32	0.24	0.30	0.29	0.29
14	0.26	0.26	0.31	0.26	0.29	0.32	0.31	0.30	0.29	0.29
15	0.25	0.28	0.31	0.32	0.29	0.31	0.31	0.32	0.29	0.29
16	0.30	0.22	0.29	0.32	0.29	0.28	0.22	0.22	0.29	0.29
17	0.29	0.31	0.29	0.23	0.31	0.29	0.26	0.23	0.29	0.29
18	0.32	0.26	0.30	0.32	0.41	0.38	0.37	0.33	0.29	0.29
19	0.35	0.34	0.35	0.35	0.37	0.36	0.35	0.37	0.29	0.29
20	0.32	0.32	0.25	0.29	0.31	0.33	0.26	0.29	0.29	0.29
Media	0.30	0.29	0.30	0.30	0.30	0.30	0.29	0.30	0.29	0.29
Std	±0.04	±0.05	±0.03	±0.06	±0.06	±0.05	±0.04	±0.05	±0.00	±0.00

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.6: Exactitud, media y desviación estándar (Std) obtenidos con *naive Bayes* usando los individuos de la BD2 para habla imaginada.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.33	0.33	0.32	0.31	0.28	0.31	0.34	0.35	0.34	0.34
2	0.24	0.26	0.24	0.25	0.12	0.16	0.24	0.25	0.34	0.34
3	0.30	0.29	0.32	0.31	0.28	0.30	0.27	0.27	0.34	0.34
4	0.24	0.24	0.28	0.29	0.23	0.25	0.22	0.20	0.34	0.34
5	0.25	0.22	0.27	0.30	0.26	0.24	0.27	0.25	0.34	0.34
6	0.34	0.34	0.28	0.35	0.31	0.31	0.30	0.32	0.34	0.34
7	0.22	0.26	0.26	0.24	0.20	0.21	0.26	0.28	0.34	0.34
8	0.30	0.28	0.29	0.29	0.29	0.30	0.30	0.25	0.34	0.34
9	0.30	0.30	0.29	0.28	0.40	0.39	0.22	0.29	0.34	0.34
10	0.42	0.38	0.39	0.40	0.33	0.31	0.45	0.43	0.34	0.34
11	0.17	0.12	0.17	0.14	0.18	0.18	0.17	0.18	0.34	0.34
12	0.32	0.34	0.36	0.39	0.27	0.30	0.39	0.33	0.34	0.34
13	0.22	0.23	0.24	0.29	0.18	0.18	0.25	0.22	0.34	0.34
14	0.31	0.33	0.37	0.30	0.28	0.33	0.36	0.39	0.34	0.34
15	0.29	0.29	0.30	0.29	0.34	0.34	0.30	0.28	0.34	0.34
16	0.22	0.28	0.27	0.30	0.24	0.29	0.25	0.29	0.34	0.34
17	0.30	0.30	0.28	0.30	0.37	0.46	0.32	0.32	0.34	0.34
18	0.28	0.27	0.32	0.33	0.36	0.30	0.33	0.29	0.34	0.34
19	0.30	0.33	0.35	0.34	0.30	0.30	0.33	0.33	0.34	0.34
20	0.29	0.30	0.28	0.29	0.28	0.29	0.30	0.34	0.34	0.34
Media	0.28	0.28	0.29	0.30	0.27	0.29	0.29	0.29	0.34	0.34
Std	±0.05	±0.06	±0.05	±0.05	±0.07	±0.07	±0.06	±0.06	±0.00	±0.00

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.7: Exactitud, media y desviación estándar (Std) obtenida con *random forest* usando los individuos de la BD1 para actividad e inactividad lingüística.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.72	0.73	0.73	0.76	0.82	0.83
2	0.88	0.89	0.88	0.90	0.98	0.99
3	0.57	0.55	0.56	0.67	0.65	0.66
4	0.63	0.66	0.68	0.74	0.79	0.79
5	0.96	0.96	0.94	0.97	0.99	1.00
6	0.80	0.79	0.83	0.90	0.94	0.95
7	0.92	0.92	0.91	0.92	0.91	0.91
8	0.73	0.73	0.78	0.90	0.92	0.93
9	0.65	0.60	0.51	0.66	0.75	0.77
10	0.85	0.85	0.90	0.98	1.00	1.00
11	0.62	0.61	0.63	0.64	0.72	0.74
12	0.77	0.78	0.77	0.77	0.81	0.84
13	0.48	0.50	0.49	0.46	0.54	0.57
14	0.47	0.50	0.49	0.57	0.60	0.63
15	0.80	0.80	0.81	0.90	0.96	0.96
16	0.68	0.67	0.69	0.75	0.78	0.76
17	0.53	0.55	0.51	0.52	0.53	0.52
18	0.44	0.50	0.47	0.48	0.41	0.42
19	0.90	0.91	0.89	0.99	1.00	1.00
20	0.77	0.75	0.78	0.81	0.93	0.93
21	0.83	0.85	0.83	0.88	0.94	0.94
22	0.83	0.81	0.85	0.93	0.98	0.98
23	0.52	0.52	0.47	0.57	0.51	0.53
24	0.79	0.78	0.81	0.86	0.92	0.92
25	0.93	0.91	0.92	0.98	0.99	1.00
26	0.75	0.75	0.75	0.77	0.85	0.87
27	0.75	0.73	0.72	0.85	0.97	0.97
Media	0.75	0.73	0.73	0.78	0.82	0.83
Std	±0.15	±0.14	±0.16	±0.16	±0.18	±0.17

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.8: Exactitud, media y desviación estándar (Std) obtenida con *SVM* usando los individuos de la BD1 para actividad e inactividad lingüística.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.67	0.68	0.70	0.75	0.80	0.79
2	0.84	0.85	0.65	0.92	0.97	0.97
3	0.43	0.53	0.62	0.63	0.64	0.63
4	0.53	0.59	0.70	0.73	0.72	0.71
5	0.95	0.95	0.88	0.97	0.99	0.99
6	0.69	0.71	0.77	0.89	0.91	0.92
7	0.93	0.92	0.91	0.94	0.91	0.91
8	0.67	0.72	0.72	0.89	0.90	0.92
9	0.67	0.63	0.52	0.66	0.72	0.75
10	0.70	0.68	0.83	0.97	1.00	1.00
11	0.57	0.55	0.52	0.68	0.70	0.69
12	0.77	0.79	0.78	0.79	0.82	0.82
13	0.46	0.51	0.44	0.49	0.57	0.57
14	0.50	0.51	0.52	0.51	0.48	0.49
15	0.79	0.76	0.78	0.93	0.96	0.96
16	0.62	0.61	0.67	0.77	0.74	0.76
17	0.51	0.58	0.54	0.57	0.54	0.54
18	0.45	0.48	0.50	0.50	0.44	0.47
19	0.90	0.88	0.88	0.99	1.00	1.00
20	0.78	0.80	0.72	0.85	0.92	0.93
21	0.82	0.84	0.79	0.88	0.94	0.94
22	0.78	0.77	0.76	0.94	0.99	0.99
23	0.54	0.53	0.62	0.48	0.60	0.58
24	0.76	0.74	0.73	0.88	0.90	0.91
25	0.83	0.85	0.82	0.91	0.96	0.97
26	0.78	0.76	0.69	0.81	0.85	0.85
27	0.69	0.63	0.65	0.86	0.95	0.97
Media	0.69	0.70	0.69	0.78	0.81	0.82
Std	±0.15	±0.14	±0.13	±0.16	±0.17	±0.17

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.9: Exactitud, media y desviación estándar (Std) obtenida con *naive Bayes* usando los individuos de la BD1 para actividad e inactividad lingüística.

Sujeto	DWT				EST	
	Instantánea	Jerárquica	Relativa	Teager	C1	C2
1	0.73	0.76	0.65	0.71	0.61	0.62
2	0.66	0.82	0.50	0.67	0.91	0.92
3	0.59	0.54	0.60	0.62	0.61	0.63
4	0.66	0.68	0.51	0.73	0.77	0.77
5	0.80	0.81	0.80	0.81	0.97	0.97
6	0.73	0.72	0.76	0.80	0.88	0.91
7	0.91	0.91	0.90	0.91	0.87	0.88
8	0.70	0.71	0.63	0.84	0.82	0.83
9	0.64	0.61	0.51	0.65	0.60	0.61
10	0.68	0.70	0.69	0.92	0.98	0.98
11	0.53	0.56	0.44	0.56	0.64	0.67
12	0.77	0.76	0.75	0.77	0.78	0.79
13	0.51	0.51	0.50	0.50	0.53	0.54
14	0.53	0.46	0.45	0.54	0.51	0.53
15	0.72	0.76	0.71	0.77	0.89	0.92
16	0.63	0.64	0.66	0.72	0.74	0.76
17	0.56	0.54	0.52	0.56	0.46	0.47
18	0.47	0.47	0.53	0.42	0.43	0.72
19	0.87	0.90	0.85	0.97	0.99	0.99
20	0.77	0.77	0.69	0.77	0.82	0.85
21	0.79	0.83	0.77	0.85	0.90	0.91
22	0.73	0.75	0.74	0.78	0.89	0.90
23	0.56	0.58	0.61	0.58	0.62	0.61
24	0.65	0.68	0.61	0.76	0.79	0.80
25	0.83	0.86	0.79	0.88	0.92	0.93
26	0.66	0.71	0.61	0.68	0.76	0.76
27	0.69	0.69	0.63	0.72	0.91	0.91
Media	0.68	0.69	0.65	0.72	0.76	0.77
Std	±0.11	±0.13	±0.12	±0.14	±0.17	±0.17

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.10: Exactitud, media y desviación estándar (Std) obtenidos *random forest* usando los individuos de la BD2 para actividad e inactividad lingüística.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.89	0.88	0.89	0.89	0.88	0.90	1.00	0.99	1.00	1.00
2	0.88	0.90	0.89	0.89	0.86	0.87	0.87	0.87	0.85	0.86
3	0.89	0.88	0.89	0.89	0.86	0.86	0.87	0.88	0.88	0.88
4	0.90	0.89	0.87	0.89	0.85	0.86	0.89	0.90	0.83	0.84
5	0.87	0.89	0.86	0.89	0.88	0.88	0.89	0.88	0.88	0.88
6	0.89	0.89	0.90	0.89	0.87	0.88	0.88	0.89	0.85	0.84
7	0.86	0.88	0.87	0.90	0.85	0.86	0.89	0.88	0.86	0.88
8	0.91	0.92	0.91	0.92	0.92	0.94	0.95	0.95	0.98	0.98
9	0.89	0.90	0.90	0.89	0.88	0.91	0.91	0.92	0.89	0.92
10	0.90	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.89	0.92
11	0.92	0.91	0.89	0.91	0.90	0.91	0.93	0.93	0.93	0.93
12	0.90	0.91	0.91	0.91	0.88	0.89	0.91	0.89	0.89	0.89
13	0.90	0.90	0.90	0.90	0.89	0.90	0.91	0.92	0.91	0.93
14	0.93	0.94	0.92	0.92	0.95	0.94	0.93	0.94	0.96	0.96
15	0.88	0.89	0.89	0.89	0.86	0.89	0.90	0.91	0.89	0.88
16	0.91	0.90	0.91	0.89	0.88	0.89	0.88	0.89	0.86	0.88
17	0.88	0.90	0.89	0.89	0.90	0.93	0.93	0.94	0.94	0.94
18	0.94	0.94	0.90	0.92	0.94	0.94	0.94	0.94	0.96	0.96
19	0.91	0.91	0.91	0.91	0.93	0.92	0.93	0.93	0.95	0.96
20	0.93	0.94	0.92	0.93	0.92	0.92	0.95	0.96	0.94	0.94
Media	0.90	0.90	0.90	0.90	0.89	0.90	0.91	0.92	0.91	0.91
Std	±0.02	±0.02	±0.02	±0.01	±0.03	±0.03	±0.03	±0.03	±0.05	±0.05

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.11: Exactitud, media y desviación estándar (Std) obtenidos con *SVM* usando los individuos de la BD2 para actividad e inactividad lingüística.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.89	0.88	0.84	0.84	0.88	0.87	0.87	0.88	0.86	0.85
2	0.86	0.87	0.87	0.85	0.76	0.78	0.84	0.86	0.83	0.84
3	0.88	0.86	0.86	0.83	0.87	0.85	0.88	0.86	0.88	0.87
4	0.84	0.85	0.83	0.80	0.87	0.89	0.90	0.90	0.83	0.84
5	0.86	0.85	0.84	0.88	0.84	0.86	0.90	0.90	0.92	0.93
6	0.86	0.86	0.83	0.81	0.84	0.87	0.88	0.88	0.84	0.85
7	0.81	0.84	0.80	0.80	0.87	0.88	0.90	0.89	0.87	0.87
8	0.91	0.91	0.89	0.90	0.89	0.94	0.93	0.93	0.97	0.97
9	0.83	0.85	0.80	0.82	0.85	0.87	0.90	0.90	0.92	0.92
10	0.89	0.90	0.85	0.85	0.85	0.88	0.94	0.95	0.91	0.92
11	0.92	0.92	0.91	0.90	0.90	0.91	0.93	0.91	0.93	0.92
12	0.87	0.86	0.87	0.85	0.89	0.88	0.92	0.92	0.90	0.91
13	0.91	0.93	0.91	0.91	0.86	0.88	0.92	0.94	0.92	0.91
14	0.94	0.93	0.94	0.91	0.94	0.96	0.93	0.96	0.96	0.96
15	0.84	0.86	0.83	0.86	0.87	0.89	0.90	0.91	0.91	0.91
16	0.87	0.88	0.86	0.89	0.90	0.90	0.91	0.90	0.91	0.91
17	0.87	0.89	0.84	0.85	0.89	0.94	0.94	0.94	0.95	0.94
18	0.91	0.92	0.88	0.92	0.91	0.93	0.94	0.93	0.95	0.95
19	0.91	0.91	0.88	0.87	0.94	0.94	0.93	0.91	0.95	0.94
20	0.94	0.93	0.88	0.91	0.94	0.94	0.97	0.96	0.96	0.96
Media	0.88	0.88	0.86	0.86	0.88	0.89	0.91	0.91	0.91	0.91
Std	±0.04	±0.03	±0.04	±0.04	±0.04	±0.04	±0.03	±0.03	±0.04	±0.04

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.12: Exactitud, media y desviación estándar (Std) obtenidos con *naive Bayes* usando los 20 individuos de la BD2 para actividad e inactividad lingüística.

Sujeto	DWT								EST	
	Instantánea		Jerárquica		Relativa		Teager		C1	C2
	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂	5 ₂	6 ₂		
1	0.89	0.89	0.88	0.88	0.81	0.81	0.99	0.99	0.90	0.92
2	0.85	0.86	0.85	0.86	0.71	0.74	0.88	0.89	0.90	0.87
3	0.89	0.89	0.90	0.90	0.84	0.84	0.90	0.89	0.90	0.89
4	0.85	0.85	0.84	0.86	0.79	0.82	0.89	0.88	0.90	0.88
5	0.86	0.84	0.86	0.87	0.81	0.87	0.90	0.90	0.90	0.87
6	0.88	0.87	0.88	0.87	0.87	0.86	0.89	0.89	0.90	0.87
7	0.86	0.87	0.86	0.87	0.85	0.86	0.90	0.90	0.90	0.87
8	0.88	0.89	0.88	0.89	0.84	0.87	0.90	0.90	0.90	0.89
9	0.88	0.89	0.89	0.90	0.83	0.84	0.90	0.90	0.90	0.90
10	0.88	0.88	0.87	0.89	0.85	0.86	0.90	0.90	0.90	0.89
11	0.87	0.88	0.87	0.88	0.77	0.78	0.90	0.90	0.90	0.88
12	0.89	0.89	0.88	0.88	0.86	0.87	0.90	0.89	0.90	0.88
13	0.86	0.86	0.85	0.85	0.87	0.89	0.85	0.84	0.90	0.88
14	0.89	0.89	0.88	0.88	0.88	0.89	0.90	0.90	0.90	0.92
15	0.88	0.89	0.88	0.90	0.86	0.86	0.90	0.90	0.90	0.87
16	0.89	0.89	0.89	0.89	0.87	0.87	0.90	0.90	0.90	0.88
17	0.89	0.89	0.89	0.89	0.85	0.90	0.90	0.91	0.90	0.89
18	0.88	0.90	0.89	0.92	0.81	0.89	0.88	0.92	0.90	0.85
19	0.88	0.88	0.87	0.88	0.85	0.85	0.90	0.90	0.90	0.85
20	0.90	0.92	0.90	0.91	0.89	0.88	0.90	0.90	0.90	0.89
Media	0.88	0.88	0.88	0.88	0.84	0.85	0.90	0.90	0.90	0.88
Std	±0.01	±0.02	±0.02	±0.02	±0.04	±0.04	±0.03	±0.03	±0.00	±0.02

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.13: Exactitud, media y desviación estándar (Std) obtenidos con *random forest* usando los individuos de la BD2 para actividad e inactividad lingüística. Análisis del impacto del número de canales y la frecuencia de muestreo.

Sujeto	64 canales, 500 Hz	14 canales, 128 Hz
1	1.00	1.00
2	0.86	0.85
3	0.88	0.83
4	0.84	0.79
5	0.88	0.82
6	0.84	0.81
7	0.88	0.79
8	0.98	0.97
9	0.92	0.87
10	0.92	0.80
11	0.93	0.95
12	0.89	0.87
13	0.93	0.92
14	0.96	0.93
15	0.88	0.80
16	0.88	0.78
17	0.94	0.87
18	0.96	0.92
19	0.96	0.92
20	0.94	0.84
Media	0.91	0.87
Std	±0.05	±0.07

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.14: Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD1.

Sujeto	$\frac{1}{2}$ seg	$\frac{1}{4}$ de seg	$\frac{1}{8}$ de seg	$\frac{1}{16}$ de seg
1	0.69	0.67	0.66	0.67
2	0.84	0.82	0.82	0.81
3	0.61	0.59	0.61	0.59
4	0.58	0.58	0.58	0.58
5	0.93	0.92	0.91	0.91
6	0.81	0.78	0.76	0.76
7	0.88	0.85	0.83	0.80
8	0.79	0.77	0.76	0.76
9	0.66	0.64	0.62	0.61
10	0.85	0.83	0.82	0.81
11	0.64	0.62	0.61	0.60
12	0.79	0.77	0.73	0.72
13	0.55	0.52	0.51	0.52
14	0.53	0.55	0.53	0.53
15	0.85	0.81	0.53	0.77
16	0.70	0.68	0.67	0.67
17	0.55	0.55	0.54	0.55
18	0.52	0.49	0.49	0.49
19	0.83	0.81	0.78	0.78
20	0.81	0.77	0.75	0.74
21	0.84	0.82	0.80	0.78
22	0.79	0.78	0.76	0.75
23	0.53	0.83	0.53	0.51
24	0.82	0.76	0.74	0.74
25	0.81	0.78	0.74	0.72
26	0.81	0.74	0.69	0.66
27	0.77	0.73	0.72	0.72
Media	0.79	0.76	0.72	0.72
Std	± 0.13	± 0.12	± 0.12	± 0.11

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.15: Segmentos de señal de diferentes tamaños para evaluar la muestra más pequeña que aún contiene información, usando la BD2.

Sujeto	$\frac{1}{2}$ de seg	$\frac{1}{4}$ de seg	$\frac{1}{8}$ de seg	$\frac{1}{16}$ de seg
1	0.78	0.78	0.78	0.77
2	0.76	0.79	0.79	0.79
3	0.73	0.79	0.79	0.79
4	0.64	0.55	0.55	0.57
5	0.79	0.79	0.79	0.79
6	0.70	0.64	0.66	0.64
7	0.56	0.78	0.78	0.78
8	0.76	0.76	0.74	0.73
9	0.54	0.71	0.65	0.58
10	0.55	0.63	0.63	0.59
11	0.79	0.78	0.77	0.77
12	0.68	0.74	0.68	0.67
13	0.79	0.79	0.80	0.80
14	0.79	0.79	0.79	0.79
15	0.79	0.79	0.79	0.79
16	0.72	0.66	0.69	0.68
17	0.61	0.79	0.79	0.79
18	0.57	0.54	0.60	0.60
19	0.80	0.79	0.80	0.79
20	0.62	0.68	0.64	0.62
Media	0.73	0.78	0.77	0.77
Std	± 0.09	± 0.08	± 0.08	± 0.09

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.16: Validación de las muestras de tamaño $\frac{1}{2}$ entre sujetos.

Sujeto	A.14	EV		
	Exactitud	Exactitud	Recuerdo	MedidaF
1	0.69	0.51	0.52	0.51
2	0.84	0.83	0.83	0.83
3	0.61	0.62	0.54	0.53
4	0.58	0.50	0.44	0.44
5	0.93	0.87	0.88	0.87
6	0.81	0.78	0.73	0.75
7	0.88	0.76	0.74	0.72
8	0.79	0.72	0.70	0.70
9	0.66	0.60	0.61	0.61
10	0.85	0.80	0.81	0.80
11	0.64	0.61	0.61	0.60
12	0.79	0.72	0.72	0.72
13	0.55	0.54	0.53	0.52
14	0.53	0.51	0.50	0.50
15	0.85	0.81	0.77	0.78
16	0.70	0.72	0.69	0.70
17	0.55	0.50	0.50	0.50
18	0.52	0.50	0.50	0.50
19	0.83	0.84	0.80	0.81
20	0.81	0.76	0.76	0.76
21	0.84	0.79	0.79	0.79
22	0.79	0.75	0.75	0.75
23	0.53	0.55	0.55	0.55
24	0.82	0.81	0.77	0.78
25	0.81	0.74	0.75	0.74
26	0.81	0.72	0.72	0.72
27	0.77	0.65	0.68	0.63
Media	0.79	0.72	0.72	0.72
Std	± 0.13	± 0.12	± 0.12	± 0.13

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.17: Validación de las muestras de tamaño $\frac{1}{16}$ entre sujetos.

Sujeto	A.14	EV		
	Exactitud	Exactitud	Recuerdo	MedidaF
1	0.67	0.47	0.47	0.47
2	0.81	0.80	0.81	0.80
3	0.59	0.59	0.55	0.55
4	0.58	0.50	0.47	0.47
5	0.91	0.85	0.87	0.85
6	0.76	0.75	0.70	0.72
7	0.80	0.65	0.65	0.64
8	0.76	0.70	0.69	0.69
9	0.61	0.56	0.57	0.56
10	0.81	0.78	0.78	0.78
11	0.60	0.55	0.56	0.54
12	0.72	0.64	0.64	0.64
13	0.52	0.52	0.51	0.50
14	0.53	0.48	0.47	0.44
15	0.77	0.72	0.70	0.71
16	0.67	0.70	0.67	0.67
17	0.55	0.51	0.50	0.50
18	0.49	0.50	0.50	0.50
19	0.78	0.78	0.74	0.75
20	0.74	0.68	0.69	0.68
21	0.78	0.73	0.74	0.73
22	0.75	0.71	0.71	0.71
23	0.51	0.54	0.55	0.54
24	0.74	0.72	0.71	0.72
25	0.72	0.69	0.70	0.69
26	0.66	0.63	0.62	0.62
27	0.72	0.59	0.63	0.59
Media	0.72	0.65	0.65	0.64
Std	± 0.11	± 0.11	± 0.11	± 0.11

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.18: Validación de las muestras de tamaño $\frac{1}{2}$ entre sujetos.

Sujeto	A.14	EV		
	Exactitud	Exactitud	Recuerdo	MedidaF
1	0.78	0.33	0.57	0.42
2	0.76	0.42	0.65	0.51
3	0.73	0.44	0.66	0.53
4	0.64	0.44	0.66	0.53
5	0.79	0.78	0.68	0.55
6	0.70	0.77	0.68	0.58
7	0.56	0.43	0.66	0.52
8	0.76	0.45	0.67	0.54
9	0.54	0.47	0.69	0.56
10	0.55	0.46	0.68	0.55
11	0.79	0.74	0.72	0.65
12	0.68	0.42	0.65	0.51
13	0.79	0.73	0.68	0.60
14	0.79	0.46	0.68	0.55
15	0.79	0.46	0.68	0.55
16	0.72	0.45	0.67	0.54
17	0.61	0.78	0.67	0.55
18	0.57	0.48	0.69	0.56
19	0.80	0.46	0.69	0.55
20	0.62	0.79	0.70	0.62
Media	0.73	0.54	0.67	0.55
Std	± 0.09	± 0.16	± 0.03	± 0.05

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.19: Validación de las muestras de tamaño $\frac{1}{16}$ entre sujetos.

Sujeto	A.14	EV		
	Exactitud	Exactitud	Recuerdo	MedidaF
1	0.77	0.34	0.58	0.42
2	0.79	0.42	0.65	0.51
3	0.79	0.44	0.66	0.53
4	0.57	0.45	0.67	0.54
5	0.79	0.46	0.68	0.55
6	0.64	0.44	0.66	0.53
7	0.78	0.44	0.66	0.52
8	0.73	0.45	0.67	0.54
9	0.58	0.47	0.68	0.55
10	0.59	0.46	0.68	0.55
11	0.77	0.47	0.68	0.55
12	0.67	0.43	0.65	0.52
13	0.80	0.42	0.65	0.51
14	0.79	0.46	0.68	0.55
15	0.79	0.46	0.68	0.55
16	0.68	0.46	0.67	0.55
17	0.79	0.46	0.68	0.55
18	0.60	0.47	0.68	0.55
19	0.79	0.46	0.68	0.55
20	0.62	0.43	0.66	0.52
Media	0.77	0.46	0.67	0.55
Std	± 0.09	± 0.03	± 0.02	± 0.03

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.20: Exactitud, media y desviación estándar (Std) obtenidos con *random forest* usando los 20 individuos de la BD2 para transferencia de clasificador usando actividad e inactividad lingüística con segmentos de $\frac{1}{2}$ de segundo. Análisis del impacto del número de canales y la frecuencia de muestreo.

Sujeto	64 canales, 500 Hz	14 canales, 128 Hz
1	0.33	0.50
2	0.42	0.26
3	0.44	0.26
4	0.44	0.25
5	0.78	0.74
6	0.77	0.72
7	0.73	0.75
8	0.45	0.57
9	0.47	0.28
10	0.46	0.28
11	0.74	0.71
12	0.42	0.25
13	0.73	0.42
14	0.46	0.71
15	0.46	0.27
16	0.45	0.75
17	0.78	0.76
18	0.48	0.29
19	0.46	0.75
20	0.79	0.76
Media	0.54	0.51
Std	± 0.16	± 0.22

APÉNDICE A. RESULTADOS A DETALLE DE LOS
EXPERIMENTOS

Tabla A.21: Exactitud, media y desviación estándar (Std) obtenidos con *random forest* usando los individuos de la BD2 para transferencia de clasificador usando actividad e inactividad lingüística con segmentos de $\frac{1}{16}$ de segundo. Análisis del impacto del número de canales y la frecuencia de muestreo.

Sujeto	64 canales, 500 Hz	14 canales, 128 Hz
1	0.34	0.29
2	0.42	0.40
3	0.44	0.42
4	0.45	0.42
5	0.46	0.43
6	0.44	0.42
7	0.44	0.42
8	0.45	0.43
9	0.47	0.44
10	0.46	0.43
11	0.47	0.43
12	0.43	0.41
13	0.42	0.40
14	0.46	0.43
15	0.46	0.44
16	0.46	0.43
17	0.46	0.43
18	0.47	0.44
19	0.46	0.43
20	0.43	0.41
Media	0.46	0.42
Std	± 0.03	± 0.03

Referencias

- [1] Torres-García Alejandro Antonio, Reyes-García Carlos Alberto, and Villaseñor-Pineda Luis. Toward a silent speech interface based on unspoken speech. *Proceedings of BIOSIGNALS (BIOSTEC)*, pages 370–373, February 2012.
- [2] Ali Bashashati, Mehrdad Fatourech, Rabab K Ward, and Gary E Birch. A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. *Journal of Neural engineering*, 4, 2007.
- [3] Jaimie F Borisoff, Steven G Mason, Ali Bashashati, and Gary E Birch. Brain-computer interface design for asynchronous control applications: improvements to the lf-asd asynchronous brain switch. *IEEE Transactions on Biomedical Engineering*, 2004.
- [4] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [5] Katharine Brigham and BVK Vijaya Kumar. Imagined speech classification with eeg signals for silent communication: A preliminary investigation into synthetic telepathy. *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on, IEEE*, pages 1–4, 2010.
- [6] Katharine Brigham and BVK Vijaya Kumar. Subject identification from electroencephalogram (eeg) signals during imagined speech. *Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, 4:1–8, 2010.

- [7] Charles DaSalla, Makoto Sato, and Yasuharu Koike. Emg vowel recognition using a support vector machine. *International symposium on measurement, analysis and modeling of human functions*, 2007.
- [8] Charles S DaSalla, Hiroyuki Kambara, Makoto Sato, and Yasuharu Koike. Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Networks*, 22(9):1334–1339, November 2009.
- [9] José Luis Sánchez de la Rosa. Métodos para el procesamiento y análisis estadístico multivariante de señales multicanal: aplicación al estudio del eeg. Master’s thesis, Universidad de La Laguna, España, 1993.
- [10] Bruce Denby, Thomas Schultz, Kiyoshi Honda, Thomas Hueber, Jim M Gilbert, and Jonathan S Brumberg. Silent speech interfaces. *Speech Communication*, 52(4):270–287, 2010.
- [11] Peter Desain, Jason Farquhar, Pim Haselager, Christian Hesse, and RS Schaefer. What bci research needs. *Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems*, 2008.
- [12] Emmanuel Didiot, Irina Illina, Dominique Fohr, and Odile Mella. A wavelet-based parameterization for speech/music discrimination. *Computer Speech and Language, Elsevier*, 24, 2010.
- [13] Thomas G Dietterich. Ensemble methods in machine learning. *MCS: International Workshop on Multiple Classifier Systems*, 2000.
- [14] DP elvig, Henning Pakkenberg, AK Stark, and Bente Pakkenberg. Neocortical glial cell numbers in human brains. *Proceedings of The First Joint BMES/EMBS conference Serving Humanity, Advancing Technology*, 29:1754–1762, april 2007.
- [15] Caupolicán Muñoz Gamboa and Joel Jiménez Cruz. Potenciales evocados i: Introducción. *Revista Mexicana de Ingeniería Biomédica*, 12, 2001.
- [16] Caupolicán Muñoz Gamboa and Joel Jiménez Cruz. Potenciales evocados ii: potenciales exógenos. *Revista Mexicana de Ingeniería Biomédica*, 13, 2002.

-
- [17] Aravind Ganapathiraju, Jonathan E Hamaker, and Joseph Picone. Applications of support vector machines to speech recognition. *IEEE Transactions on Signal Processing*, 52(8):2348–2355, 2004.
- [18] AA Torres García, CA Reyes García, and L Villaseñor Pineda. Hacia la clasificación de habla no pronunciada mediante electroencefalogramas (eeg). *XXXIV Congreso Nacional de Ingeniería Biomédica, Ixtapa-Zihuatanejo, Guerrero, Mexico*, 2011.
- [19] Gerardo G Gentiletti, Carolina B Tabernig, and Rubén C Acevedo. Interfaces cerebro computadora: Definición, tipos y estado actual. *Bioengineering Solutions for Latin America Health*, 2007.
- [20] E Gómez-Luna, D Silva, and G Aponte. Selección de una wavelet madre para el análisis frecuencial de señales eléctricas transitorias usando wpd. *Ingeniare. Revista chilena de ingeniería*, 21, 2013.
- [21] Steve R Gunn et al. Support vector machines for classification and regression. *ISIS technical report*, 14:85–86, 1998.
- [22] Josefina Gutiérrez. Análisis de señales en el neuromonitoreo. *Revista Mexicana de Ingeniería Biomédica*, 22(2):67–77, 2001.
- [23] Isabelle Guyon, Steve Gunn, Masoud Nikravesh, and Lofti A Zadeh. *Feature extraction: foundations and applications*, volume 207. Springer, 2008.
- [24] Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R News*, 2(3):18–22, 2002.
- [25] Fabien Lotte, Marco Congedo, Anatole Lécuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4, 2007.
- [26] Stephane G Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 2, 1989.
- [27] Tom Mitchell. *Machine Learning*. McGraw Hill, 1997.

- [28] The University of Toronto. The kara one database: Phonological categories in imagined and articulated speech, 2012.
- [29] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.
- [30] Gert Pfurtscheller. Brain-computer interface-state of the art and future prospects. *Signal Processing Conference, 2004 12th European*, 4:509–510, 2004.
- [31] A Porbadnigk and T Schultz. Eegbased speech recognition: Impact of experimental design on performance. Master’s thesis, Institut für Theoretische Informatik Universität Karlsruhe, 2008.
- [32] Chapey R. Language intervention strategies in aphasia and related neurogenic communication disorders. *Lippincott Williams and Wilkins*, 5, 2008.
- [33] González R. Neuropsicología clínica y demencias: trabajos seleccionados. *Ediciones de la Sociedad de Neurología, Psiquiatría y Neurocirugía de Chile*, 2008.
- [34] González R. Trastorno del lenguaje y habla. *Neurología Fundamental*, 5, 2011.
- [35] Andrea Hornauer-Hughes Rafael González V. Cerebro y lenguaje. *Departamento de Neurología y Neurocirugía, HCUCH, Hosp Clín Univ Chile*, 2014.
- [36] Ryan Rifkin and Aldebaro Klautau. In defense of one-vs-all classification. *Journal of Machine Learning Research*, 5(Jan):101–141, 2004.
- [37] Chang Su Ryu, Yoonseon Song, Done-Sik Yoo, Sangsup Choi, Sung Sill Moon, and Jin-Hun Sohn. Eeg-based discrimination between yes and no. *Proceedings of The First Joint BMES/EMBS conference Serving Humanity, Advancing Technology*, January 1991.

-
- [38] Zhao S. and Rudzicz F. Classifying phonological categories in imagined and articulated speech. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 22(1):992–996, april 2015.
- [39] YoungJae Song and Francisco Sepulveda. Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state. *IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, 1(1):568–571, October 2014.
- [40] AA Torres-García, CA Reyes-García, L Villaseñor-Pineda, and JM Ramírez-Cortés. Análisis de señales electroencefalográficas para la clasificación de habla imaginada. *Revista mexicana de ingeniería biomédica*, 72(1):23–39, Febrero 2013.
- [41] Alejandro A Torres-García, Carlos A Reyes-García, Luis Villaseñor-Pineda, and Gregorio García-Aguilar. Implementing a fuzzy inference system in a multi-objective eeg channel selection model for imagined speech classification. *Expert Systems With Applications*, 59, 2016.
- [42] A. A Torres-García. *Análisis y clasificación de electroencefalogramas (EEG) registrados durante el habla imaginada*. PhD thesis, Instituto Nacional de Astrofísica, Óptica y Electrónica, 2016.
- [43] Peitao Wang, Jun Lu, Bin Zhang, and Zeng Tang. A review on transfer learning for brain-computer interface classification. *5th International Conference on Information Science and Technology (IC1ST)*, pages 315–322, 2015.
- [44] Marek Wester and T Schultz. Unspoken speech - speech recognition based on electroencephalography. Master's thesis, Institut für Theoretische Informatik Universität Karlsruhe, 2006.
- [45] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113, 2002.
- [46] Shunan Zhao and Frank Rudzicz. Classifying phonological categories in imagined and articulated speech. *International Conference on*

Acoustics, Speech and Signal Processing (ICASSP), *IEEE*, 22:992–996, 2015.



BUAP