



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**RECONOCIMIENTO DEL HABLA SILENCIOSA CON SEÑALES
ELECTROENCEFALOGRÁFICAS (EEG) PARA INTERFACES
CEREBRO-COMPUTADOR**

Luis Carlos Sarmiento Vela

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Mecánica y Mecatrónica
Doctorado en Ingeniería - Ingeniería Mecánica y Mecatrónica
Bogotá, Colombia
2015

**RECONOCIMIENTO DEL HABLA SILENCIOSA CON SEÑALES
ELECTROENCEFALOGRÁFICAS (EEG) PARA INTERFACES
CEREBRO-COMPUTADOR**

Luis Carlos Sarmiento Vela

Tesis o trabajo de investigación presentada como requisito parcial para optar al
título de:

Doctor en Ingeniería - Ingeniería Mecánica y Mecatrónica

Director

Carlos Julio Cortés Rodríguez. PhD

Codirector

Andrés Tovar Pérez. PhD

Línea de investigación
Biomecánica Aplicada

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Mecánica y Mecatrónica
Doctorado en Ingeniería - Ingeniería Mecánica y Mecatrónica
Bogotá, Colombia
2015

Agradecimientos

Agradezco la asesoría y colaboración del profesor Carlos Julio Cortés, director del proyecto, ya que sin su colaboración no hubiera sido posible el desarrollo de este trabajo de tesis. Quiero resaltar su dedicación a la parte investigativa y sus grandes aportes como docente.

Igualmente quiero destacar al profesor Andrés Tovar por sus valiosos aportes durante el desarrollo de esta tesis y su fuerte vocación en el campo de la docencia y en la parte investigativa.

Agradezco al profesor Jan Bacca por sus aportes en el procesamiento de señales, al profesor Rubén Guevara por sus aportes en la parte de estadística funcional y al Dr. Pablo Lorenzana por todo su apoyo desde el área de la neurofisiología para el desarrollo de esta tesis.

De igual manera quiero resaltar el grupo de profesores del Doctorado que me brindaron todo su apoyo para la realización de este proyecto.

Quiero agradecer a Amparo Ramírez y Wilmar Hernández por su colaboración para la asistencia técnica en la grabación de los datos en el laboratorio de Electrofisiología Clínica de la Universidad Nacional de Colombia.

También quiero agradecer a toda mi familia por el constante apoyo.

Resumen

Las interfaces cerebro computador tienen relevancia médica en el tratamiento de individuos que sufren de parálisis motora o amputaciones de miembro superior o miembro inferior; sin embargo, la dificultad para extraer y procesar con exactitud, particularmente las señales cerebrales de habla silenciosa, limita en gran medida su aplicación. Esta tesis presenta dos novedosos sistemas de procesamiento de señales basados en electroencefalografía con la habilidad de clasificar vocales y sílabas con habla silenciosa. Una de las metodologías se basa en las características de la entropía de la información con la dimensión de regularización, y el otro en las características de datos funcionales en el espacio de Hilbert L^2 , utilizando los datos de la posición de los electrodos. Dentro de las ventajas de los métodos desarrollados en comparación con otros métodos de BCI, pueden establecerse los siguientes: no requieren de procesos de entrenamiento dispendiosos como en el caso de la imaginería motora; no requieren de un proceso de atención riguroso como ocurre utilizando potenciales evocados visuales de estado estable (*steady-state visual evoked potential* - SSVEP) o imaginería motora; no requieren de un estímulo externo como en el caso de SSVEP o P300; y no requieren de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imaginería motora. Adicionalmente, utilizan señales cerebrales que están relacionadas con el área de lenguaje (vocales y sílabas) y tienen la posibilidad de utilizar la innumerable cantidad de palabras (léxico) de un lenguaje. La relevancia de esta tesis está en aportar dos metodologías novedosas de habla silenciosa con EEG, como una opción importante donde las BCIs mejoren su desempeño para controlar dispositivos como: *spellers*, sillas de ruedas, prótesis y robots, entre otros.

En esta tesis, una máquina de soporte vectorial para clasificación multiclase es implementada usando la estrategia uno contra el resto (1-*rest*) y uno contra uno (1-1) con una función kernel de base radial. Los parámetros óptimos son calculados con un algoritmo genético. Los resultados son demostrados con la clasificación de cinco vocales (/a/, /e/, /i/, /o/, /u/) y cinco sílabas (/fa/, /pa/, /ma/, /la/ /ra/), usando habla silenciosa con señales electroencefalográficas. El desempeño de las metodologías propuestas medidas en términos de exactitud (*accuracy*) son los siguientes:

Con la metodología basada en vector de características con entropía de la información y dimensión de regularización, se seleccionaron dos algoritmos SVM de clasificación multiclase (1-*rest*) y (1-1). Los mejores resultados de clasificación fueron obtenidos con el clasificador (1-1) para vocales y sílabas con habla silenciosa. En el caso de vocales con habla silenciosa se alcanzó una exactitud (media estadística) de 69.83 %, y en sílabas con habla silenciosa una exactitud (media estadística) de 66.89 %.

Para la metodología basada en vector de características con datos funcionales, aplicado a vocales y sílabas con habla silenciosa, se seleccionó el algoritmo SVM de clasificación multiclase (1-1). Para el caso de vocales con habla silenciosa se utilizaron ritmos δ , $\delta\theta$ y $\delta\theta\alpha$. Los mejores resultados de exactitud fueron obtenidos con los ritmos $\delta\theta\alpha$ con una media estadística de 71.92 %. En el caso de sílabas con habla silenciosa se utilizaron ritmos δ , con los cuales se alcanzó una exactitud (media estadística) de 67.13 %.

De los resultados de la clasificación se concluye que la exactitud alcanzada para vocales y sílabas con habla silenciosa, utilizando la metodología basada en vector de características con datos funcionales, es más alta que aquella alcanzada con la metodología basada en vector de características con entropía de la información y dimensión de regularización.

Abstract

Brain-computer interfaces have medical relevance in the treatment of individuals suffering from motor paralysis or amputation of upper limb or lower limb; however, the difficulty to extract and accurately process, particularly silent speech brain signals, greatly limit its application. This thesis presents two novel processing systems based on electroencephalography with the ability to classify vowels and syllables with silent speech signals. One methodology is based on the features of the information entropy with dimension regularization, and the other one is based on the features of functional data in the Hilbert space L^2 , using the position data of the electrodes. Among the advantages of the developed methodologies in comparison to other BCI methods, the following can be established: they do not require training wasteful processes as in the case of motor imagery; they do not require a rigorous attention such as using (*steady-state visual evoked potential* - SSVEP) or motor imagery; they do not require an external stimulus such as in the case of SSVEP or P300; and they do not require cognitive tasks that generate cognitive or muscle fatigue as in the case of motor imagery. In addition to this, they use brain signals that are related to the language area (vowels and syllables) and they have the possibility to work with the countless number of words (vocabulary) of a language. The relevance of this thesis is to provide two novel methodologies of silent speech with EEG, as an important option where BCIs can improve their performance to control devices such as: *spellers*, wheelchairs, prostheses and robots, among others.

In this thesis, a support vector machine for multiclass classification was implemented using the one against-rest (1-rest) and one against-one (1-1) strategy with a radial basis function kernel. The optimal parameters are calculated with a genetic algorithm. The results are demonstrated with the classification of five vowels (/a/, /e/, /i/, /o/, /u/) and five syllables (/ fa/, /pa/, /ma/, /la/ /ra/) using silent speech with electroencephalographic signals. The performance of the proposed methodologies measured in terms of accuracy is as follows:

In regard to the feature vector based information entropy and dimension regularization methodology, two multiclass SVM classification algorithms (1-rest) and (1-1) were selected. The best results were obtained with the classifier (1-1) to vowels and syllables with silent speech. In the case of silent speech vowels an accuracy (statistical average) of 69.83% was reached, and for silent speech syllables an accuracy (statistical average) of 66.89% was reached.

For the feature vector based functional data methodology, the multiclass SVM classification algorithm (1-1) was selected to vowels and syllables with silent speech. In the case of silent speech vowels δ , $\delta\theta$ and $\delta\theta\alpha$ rhythms were used. The best results of accuracy were obtained with $\delta\theta\alpha$ rhythms, with a statistical average of 71.92%. In the case of silent speech syllables δ rhythms were used, where an accuracy (statistical average) of 67.13% was reached.

From the classification results it can be concluded that the accuracy to the feature vector based functional data methodology to vowels and syllables with silent speech, is higher than the feature vector based information entropy and dimension regularization methodology.

Índice general

1. Introducción	1
1.1. Justificación	1
1.2. Objetivos	4
2. Referentes teóricos	5
2.1. El procesamiento del lenguaje en los seres humanos	5
2.1.1. El cerebro humano	5
2.1.2. Modelos neurológicos del lenguaje en seres humanos	10
2.2. Interfaces cerebro computador	17
2.2.1. Arquitectura de BCIs	18
2.2.2. Métodos para capturar señales cerebrales	19
2.2.3. BCIs invasivos	24
2.2.4. Estrategias mentales no invasivas	25
2.3. El habla silenciosa (<i>silent speech</i>) (SS)	30
2.3.1. Captura del movimiento de puntos fijos en el aparato articulador usando sensores de articulografía electromagnética (EMA)	31
2.3.2. Caracterización del tracto vocal usando ultrasonido (US) e imágenes ópticas de la lengua y los labios	31
2.3.3. Electromiografía superficial (sEMG) basada en reconocimiento del habla	32
2.3.4. Habla silenciosa con BCIs utilizando electrocorticografía	33
2.3.5. Habla silenciosa con BCIs empleando electroencefalografía (EEG) . .	33
2.4. Entropía de la información	35
2.5. Dimensión de regularización	38
2.6. Análisis de datos funcionales	40
3. Modelos propuestos basados en habla silenciosa	45
3.1. Protocolo experimental	45
3.1.1. Adquisición de datos para vocales con SS	47
3.1.2. Adquisición de datos para sílabas con SS	48
3.2. Metodología basada en vector de características con entropía de la información y la dimensión de regularización	49
3.2.1. Filtraje	50
3.2.2. Consideración de artefactos	51
3.2.3. Remoción artefactos de parpadeo	52
3.2.4. Clasificador multiclase con máquinas de soporte vectorial	57
3.3. Metodología basada en vector de características con datos funcionales	62
3.3.1. Filtraje	63

3.3.2.	Remoción artefactos de parpadeo	64
3.3.3.	Transformación a densidad espectral de potencia	65
3.3.4.	Transformación posición de los electrodos y PSD a datos funcionales	67
4.	Resultados y discusión	73
4.1.	Resultados de señales EEG para vocales con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización	73
4.1.1.	Resultados por sujeto	73
4.1.2.	Resultados por localización	76
4.1.3.	Resultados de la clasificación	82
4.2.	Resultados de señales EEG para sílabas con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización	84
4.2.1.	Resultados por sujeto	84
4.2.2.	Resultados por localización	87
4.2.3.	Resultados de la clasificación	93
4.3.	Comparación entre los resultados de la metodología basada en vector de características con entropía de la información y dimensión de regularización para vocales y sílabas con habla silenciosa	94
4.4.	Discusión sobre señales EEG para vocales y sílabas con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización	96
4.4.1.	Determinación del método de clasificación	96
4.4.2.	Desempeño de vector de características en función de la metodología desarrollada	97
4.4.3.	Criterios para la selección de la metodología desarrollada de habla silenciosa	98
4.5.	Resultados de señales EEG para vocales con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales	100
4.5.1.	Resultados por sujeto	100
4.5.2.	Resultados por localización	105
4.5.3.	Resultados utilizando datos funcionales	115
4.5.4.	Resultados de la clasificación	126
4.6.	Resultados de señales EEG para sílabas con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales	127
4.6.1.	Resultados por sujeto	127
4.6.2.	Resultados por localización	129
4.6.3.	Resultados utilizando datos funcionales	132
4.6.4.	Resultados de la clasificación	137
4.7.	Comparación entre los resultados de la metodología basada en vector de características con datos funcionales para vocales y sílabas con habla silenciosa	137
4.8.	Discusión sobre señales EEG para vocales y sílabas con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales	140
4.8.1.	Determinación del método de clasificación	140
4.8.2.	Desempeño del vector de características en función de la metodología desarrollada	141

4.8.3.	Criterio para la selección de la metodología desarrollada de habla silenciosa	142
4.8.4.	Viabilidad de las metodologías propuestas de habla silenciosa aplicadas a BCIs	144
4.8.5.	Proyección de mejoras a las metodologías desarrolladas de habla silenciosa	146
5.	Conclusiones	147
5.1.	Metodología basada en vector de características con entropía de la información y dimensión de regularización	148
5.2.	Metodología basada en vector de características con datos funcionales	149
5.3.	Recomendaciones y proyecciones	151
5.4.	Publicación de resultados y difusión	153
A.	Tablas de la experimentación	155

Índice de figuras

2.1. Sistema nervioso central. Adaptado de [40]	6
2.2. Cubierta protectora del encéfalo. Adaptado de [40]	6
2.3. Planos de referencia para el cerebro	7
2.4. Vista lateral izquierda de la corteza cerebral. Circunvolución frontal superior (CFS), circunvolución frontal medial (CFM), circunvolución frontal inferior (CFI), circunvolución temporal superior (CTS), circunvolución temporal medial (CTM), circunvolución temporal inferior (CTI), circunvolución parietal superior (CPS), circunvolución parietal inferior (CPI), circunvolución supra-marginal (CS), circunvolución angular (CA), circunvolución precentral (Cpre), circunvolución postcentral (Cpos), pars opercular (PO), pars triangular (PT), pars orbitaria (Por). Surco temporal superior (STS) y surco temporal inferior (STI).	8
2.5. Lóbulos del cerebro	8
2.6. Áreas de Brodmann de la corteza cerebral	9
2.7. Modelo de Broca. Color rojo área de Broca.	11
2.8. Modelo de Wernicke. Color rojo, área de Broca; color azul, área de Wernicke.	11
2.9. Modelo de Litchtheim. Afasia de conducción (AC), afasia transcortical motora (ATM) y afasia transcortical sensorial (ATS)	12
2.10. Modelo neurológico de Geshwind	13
2.11. Diagrama modelo cognitivo de Benedet. Adaptado de [52]	14
2.12. Diagrama de bloques modelo de comprensión oral de Hickok y Poeppel. Adaptado de [54]	15
2.13. Modelo neurológico de comprensión oral de Hickok y Poeppel. Tomado de [54]	16
2.14. Arquitectura general de una BCI. Adaptado de [24].	18
2.15. Esquema de una neurona. Adaptado de [61]	20
2.16. Potencial de acción en la membrana de una neurona. Adaptado de [62].	20
2.17. Comportamiento de los canales de Na^+ , K^+ ante un potencial de acción. Tomado de [39].	21
2.18. Generación de una señal EEG. Adaptado de [39].	22
2.19. Colocación de los electrodos EEG en el sistema 10-20. Adaptado de [26].	23
3.1. Sujeto usando el EEG <i>neuroheadset</i> en la experimentación de vocales con habla silenciosa	46
3.2. Sujeto usando el EEG <i>neuroheadset</i> en la experimentación de sílabas con habla silenciosa	46
3.3. Arreglo de 21 electrodos cubriendo el área de Broca, Wernicke y corteza auditiva primaria para vocales con SS	47

3.4. Arreglo de 21 electrodos cubriendo el área de Broca, Wernicke y corteza auditi- va primaria para sílabas con SS	47
3.5. Intervalos seleccionados para el procesamiento de vocales con SS	48
3.6. Intervalos seleccionados para el procesamiento de sílabas con SS	49
3.7. Diagrama del modelo propuesto con entropía de la información y dimensión de regularización para vocales con SS	50
3.8. Diagrama del modelo propuesto con entropía de la información y dimensión de regularización para sílabas con SS	51
3.9. Diagrama del modelo propuesto con datos funcionales para vocales con SS . .	63
3.10. Diagrama del modelo propuesto con datos funcionales para sílabas con SS . .	64
4.1. Valor promedio por sujeto y por vocal con habla silenciosa (SS) de la entropía de la informacion	74
4.2. Diagrama de dispersión del valor promedio de entropía para cada vocal con habla silenciosa (SS)	75
4.3. Valor promedio por sujeto y por vocal con habla silenciosa (SS) de la dimensión de regularización	75
4.4. Diagrama de dispersión del valor promedio de la dimensión de regularización para cada vocal con habla silenciosa (SS)	76
4.5. Representación espacial del valor promedio por electrodo de la entropía de la información, realizando tareas de habla silenciosa (SS) con vocales	79
4.6. Valor promedio de entropía por electrodo en vocales con habla silenciosa (SS)	80
4.7. Valor promedio de dimensión de regularización por electrodo en vocales con habla silenciosa (SS)	81
4.8. Representación espacial del valor promedio por electrodo de la dimensión de regularización en vocales con habla silenciosa (SS)	82
4.9. Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-rest) . .	83
4.10. Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1) . . .	84
4.11. Valor promedio por sujeto y por sílaba con habla silenciosa (SS) de la entropía de la informacion	85
4.12. Diagrama de dispersión del valor promedio de entropía para cada sílaba con habla silenciosa (SS)	85
4.13. Valor promedio por sujeto y por sílaba con habla silenciosa (SS) de la dimen- sión de regularización	86
4.14. Diagrama de dispersión del valor promedio de dimensión de regularización por sílaba con habla silenciosa (SS)	87
4.15. Representación espacial del valor promedio por electrodo de la entropía para sílabas con habla silenciosa (SS)	90
4.16. Valor promedio de entropía por electrodo para sílabas con habla silenciosa (SS)	91
4.17. Representación espacial del valor promedio por electrodo de la dimensión de regularización para sílabas con habla silenciosa (SS)	92
4.18. Valor promedio de dimensión de regularización por electrodo para sílabas con habla silenciosa (SS)	93
4.19. Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-rest) . .	94
4.20. Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-1)	94
4.21. Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos δ	101

4.22. Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos δ	102
4.23. Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos $\delta\theta$	102
4.24. Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos $\delta\theta$	103
4.25. Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos $\delta\theta\alpha$	104
4.26. Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos $\delta\theta\alpha$	105
4.27. Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos δ	108
4.28. Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos δ	109
4.29. Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos $\delta\theta$	110
4.30. Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos $\delta\theta$	111
4.31. Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos $\delta\theta\alpha$	113
4.32. Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos $\delta\theta\alpha$	114
4.33. Error medio cuadrático (MSE) relacionado al número de bases para vocales con habla silenciosa (SS), utilizando ritmos $\delta\theta\alpha$	115
4.34. Datos funcionales con ritmos δ . La curva en color negro corresponde a la media.120	
4.35. Datos funcionales con ritmos δ . La curva en color negro corresponde a la media.121	
4.36. Datos funcionales con ritmos $\delta\theta$. La curva en color negro corresponde a la media.122	
4.37. Datos funcionales con ritmos $\delta\theta$. La curva en color negro corresponde a la media.123	
4.38. Datos funcionales con ritmos $\delta\theta\alpha$. La curva en color negro corresponde a la media.	124
4.39. Datos funcionales con ritmos $\delta\theta\alpha$. La curva en color negro corresponde a la media.	125
4.40. Clasificación de vocales con habla silenciosa utilizando (SS) SVM (1-1) - ritmos δ 126	
4.41. Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1)- ritmos $\delta\theta$	127
4.42. Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1) - ritmos $\delta\theta\alpha$	127
4.43. Valor promedio del PSD estimado por sujeto y por sílaba con habla silenciosa (SS) - ritmos δ	128
4.44. Diagrama de dispersión del valor promedio de PSD estimado por sílaba con habla silenciosa (SS) - ritmos δ	129
4.45. Valor promedio de PSD estimado por electrodo para sílabas con habla silenciosa (SS) - ritmos δ	131
4.46. Representación espacial del valor promedio por electrodo del PSD estimado para sílabas con habla silenciosa (SS) - ritmos δ	132
4.47. Datos funcionales para sílabas con ritmos δ . La curva en color negro corresponde a la media.	135

4.48. Datos funcionales para sílabas con ritmos δ . La curva en color negro corresponde a la media.	136
4.49. Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-1) - ritmos δ	137

Índice de tablas

3.1.	Clasificación multiclase usando el método <i>one against-rest</i> para vocales con SS	58
3.2.	Clasificación multiclase usando el método <i>one against-rest</i> para sílabas con SS	58
3.3.	Clasificación multiclase usando el método <i>one against-one</i>	61
3.4.	Valores óptimos de la distancia angular Θ_i^* para vocales con habla silenciosa, usando un algoritmo genético. E es el electrodo. Θ_i^* está en grados	69
3.5.	Valores óptimos de la distancia angular Θ_i^* para sílabas con habla silenciosa, usando un algoritmo genético. E es el electrodo. Θ_i^* está en grados	70
4.1.	MSE variando el número de bases para vocales con habla silenciosa (SS) - ritmos cerebrales δ , $\delta\theta$ y $\delta\theta\alpha$	116
4.2.	MSE variando el número de bases para sílabas con habla silenciosa (SS) - ritmos cerebrales δ	133
A.1.	Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para vocales con SS - SVM (1-1)	155
A.2.	Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para vocales con SS - SVM (1-rest)	156
A.3.	Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos δ	156
A.4.	Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos $\delta\theta$	157
A.5.	Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos $\delta\theta\alpha$	157
A.6.	Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para sílabas con SS - SVM (1-1)	158
A.7.	Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para sílabas con SS - SVM (1-rest)	158
A.8.	Metodología basada en vector de características con datos funcionales - Exactitud para sílabas con SS - ritmos δ	158

Capítulo 1

Introducción

1.1. Justificación

Existe un conjunto de enfermedades que afectan las neuronas motoras, generando en los individuos pérdidas en movimientos motores como el habla, la respiración, caminar y agarrar objetos [1]. En este grupo de enfermedades se encuentran las siguientes patologías: la esclerosis lateral amiotrófica, la esclerosis lateral primaria, la atrofia muscular progresiva y la atrofia muscular espinal. La esclerosis lateral amiotrófica (ALS) o enfermedad de Lou Gehrig, afecta progresivamente las señales motoras de todos los músculos voluntarios [2]. La esclerosis lateral primaria (PLE) afecta las neuronas motoras superiores y contribuye a la pérdida del control voluntario de los músculos del cuerpo. La atrofia muscular progresiva afecta a las neuronas motoras inferiores, comprometiendo, inicialmente, el movimiento de las manos y, de manera progresiva, el resto de los movimientos del cuerpo [3]. La atrofia muscular espinal (SMA) es una enfermedad que afecta las neuronas de la médula espinal generando pérdida de movilidad en piernas y brazos [4].

Del conjunto de enfermedades que afectan al ser humano, el 6.2% está relacionado con desordenes neurológicos. De acuerdo con las estadísticas de la Organización Mundial de la Salud referidas al 2005, la población afectada por cada 10000 habitantes debido a desordenes neurológicos fue: lesiones neurológicas, 425.4; enfermedades cerebrovasculares, 176.8; neuroinfecciones, 98.4; esclerosis múltiple, 20 [5]. Esto sugiere el desarrollo de investigaciones desde el área de la medicina y la ingeniería que mejoren la calidad de vida de estas personas, brindándoles así la posibilidad de recuperar parcial o totalmente los movimientos motores que han perdido.

Otra de las causas para la pérdida del movimiento motor está relacionada con las amputaciones de miembros superiores o miembros inferiores. Dentro de las causas más comunes para amputaciones están la diabetes, los traumas, la falta de flujo sanguíneo y las infecciones, entre otros [6]. En Estados Unidos, por ejemplo, hay aproximadamente 1.9 millones de personas que han perdido alguna extremidad. Una estimación sugiere que una de cada 200 personas tiene algún tipo de amputación. De los casos de amputación, el 82% corresponde a amputaciones relacionadas a problemas vasculares. Dentro de los casos de amputación traumática, el 68.6% corresponde a extremidades superiores [7], [8].

Cuando una persona ha sufrido una amputación de mano, pierde una parte del cuerpo importante para la interacción con el entorno, transfiriendo todos los procesos de agarre a la otra mano disponible [9]. Sin embargo, el cerebro sigue enviando la información de la intención de movimiento de la mano amputada, pero los nervios eferentes se encuentran aislados en el

nivel de la amputación, generando la problemática del miembro fantasma [10].

Como una forma de suplir la problemática de personas con amputación de miembro superior se han desarrollado prótesis mioeléctricas, que capturan las señales eléctricas generadas por músculos residuales en forma no invasiva [9]. Este tipo de bioseñales tienden a ser utilizadas en uno o dos músculos, posibilitando procesos de prensión, generalmente de agarre, siendo necesario un largo proceso de entrenamiento para adaptar el músculo a la tarea requerida. La posibilidad de utilizar más señales musculares se dificulta en función de la carga cognitiva que el sujeto tiene que realizar para desarrollar las tareas de prensión. Otra opción para el control de prótesis de miembro superior es la utilización de neuroprótesis [11]. Esta metodología, de carácter invasivo, requiere la utilización de sensores electrónicos que se implantan directamente sobre el nervio objetivo. Como una ventaja, este procedimiento tiende a utilizar los mismos nervios que producían el movimiento de los procesos de prensión antes de la amputación. Como elemento desfavorable, se encuentra el requerimiento de hacer una o más cirugías y la posibilidad de rechazo por parte del organismo.

En un trabajo previo, se desarrolló una prótesis de 10 grados de libertad para personas amputadas de mano y muñeca [12]. Esta prótesis se controló con señales mioeléctricas y señales de voz. Dentro de las conclusiones de esa tesis, se encontró que para el caso de estudio las señales de voz eran más eficientes y eficaces que las señales mioeléctricas. Pero, entonces, ¿cuál es el tipo de señal más adecuada para controlar una prótesis?. Lee y Roberson sugieren que controlar una prótesis con el pensamiento es el medio más natural para controlar una prótesis [13].

Para afrontar la pérdida de movilidad en el aparato motor debido a enfermedades o amputaciones, desde la ingeniería se han desarrollado interfaces cerebro-computador que buscan solucionar esta problemática, siendo la atención selectiva y la imaginaria motora dos de los métodos más explorados. En el caso de la atención selectiva, el sujeto requiere focalizar su atención en un estímulo determinado, siendo el potencial P300 y el potencial evocado visual de estado estable (*steady-state visual evoked potentials* - SSVEP) dos patrones cerebrales representativos.

En el caso de los potenciales evocados P300, una de las ventajas que posee este método aplicado a BCIs es el no requerir de un entrenamiento dispendioso y el poder utilizar este método con P300 en pocos minutos. Dentro de las limitaciones que presenta se encuentra el requerir de un estímulo externo (fuentes luminosas, auditivas o hápticas), y la fatiga que produce el hecho de prestar atención a fuentes luminosas, auditivas o hápticas por un largo lapso de tiempo. Se han realizado algunas aplicaciones como aquella hecha por Rebsamen et al. [14], quienes desarrollaron un prototipo de silla de ruedas basado en P300 habilitado para navegar en un ambiente, como una oficina típica o un hospital. Una de las desventajas de este método es el tiempo de respuesta.

En el caso de SSVEP, una de las ventajas que posee este método aplicado a BCIs es que no requiere de un entrenamiento dispendioso. Dentro de las limitaciones que presenta se encuentra el requerir de un estímulo externo (fuentes luminosas), la cantidad limitada de fuentes luminosas en relación con al rango de frecuencias y la fatiga que produce mirar una fuente luminosa que parpadea por largo tiempo. Se han realizado aplicaciones con SSVEP como los de Mandel et al., [15] que han desarrollado una interfaz cerebro computador (BCI) para que un sujeto pueda manejar una silla de ruedas a través de 4 señales de potenciales evocados visuales. Torres et al. [15] integraron una interface cerebro-computador (BCI) basada en potenciales evocados visuales de estado permanente (SSVEPs) y una silla de ruedas robótica, con el objeto de ser comandada por este sistema. Pfurtscheller et al. han controlado

una prótesis de brazo con dos grados de libertad, utilizando señales EEG con SSVEP [16]. En este método se han colocado 4 leds en diferentes partes de la prótesis. Los movimientos que puede realizar son: abrir la mano (led a 8 Hz), cerrar la mano (led a 13 Hz), girar a la izquierda la muñeca (led a 6 Hz) y girar a la derecha la muñeca (led a 7 Hz).

La imaginería motora aplicada a BCIs tiene ventajas frente a los métodos de atención, como el hecho de no requerir de un estímulo externo. Dentro de las limitaciones que presenta, se encuentra la necesidad para los sujetos de una fase de entrenamiento, donde el tiempo de entrenamiento está en función de la cantidad y complejidad de las tareas. Otra limitación consiste en la fatiga cognitiva, que se incrementa con el número de tareas y el tiempo respectivo. Severens et al., desarrollaron un sistema con señales EEG para determinar si una persona está caminando o imaginando caminar, o si está quieta. Para este estudio se utilizaron 62 electrodos y se realizaron las tareas de imaginar la caminata y caminar. Los resultados muestran su aplicabilidad en procesos de rehabilitación motora para miembros inferiores [17]

Para aumentar el desempeño de estos métodos se ha buscado realizar combinaciones entre los métodos de atención selectiva e imaginería motora. Duan et al., desarrollaron un BCI basado en SSVEP e imaginería motora para controlar un robot. Con tres estímulos visuales podían controlar avance, girar a la izquierda y girar a la derecha. Con la señal de imaginería motora el robot agarraba el objeto [18]. Yin et al., desarrollaron un sistema híbrido entre potenciales P300 y SSVEP aplicado a *spellers*. La combinación de P300 con SSVEP permite incrementar el número de bits por segundo, comparado con cada uno de los métodos aplicado a *spellers* [19]. Li et al., desarrollaron un sistema híbrido entre potenciales P300 y SSVEP para el control de una silla de ruedas. La utilización al mismo tiempo de estas dos señales cerebrales permite desarrollar BCIs con mayor transferencia de bits en comparación con cada uno de los métodos por separado [20]. Zhang et al., desarrollaron un control con BCI para que una persona pueda navegar utilizando silla de ruedas. En este caso, la persona puede seleccionar por imaginería motora o P300 el destino a donde quiere ir de un conjunto previo de posibilidades. Posteriormente, un sistema automático de navegación lo conduce al lugar seleccionado [21].

En la última década se ha buscado desarrollar nuevos métodos de BCIs que puedan realizar una amplia cantidad de tareas y que no requieran de: un estímulo externo, procesos cognitivos de atención con alta concentración, tiempos de respuesta factibles para manejar dispositivos y tiempos de entrenamiento dispendiosos, entre otros.

Una reciente línea de investigación en BCIs es el habla silenciosa con señales EEG. Dentro de las características que presenta esta nueva línea de BCIs están: el no ser invasiva; el no requerir de procesos de entrenamiento dispendiosos como en el caso de la imaginería motora [22]; el no requerir un proceso de atención riguroso como en SSVEP [23], P300 [14] o imaginería motora [24]; el no requerir de un estímulo externo como en el caso de SSVEP [25] o P300 [26]; y el no requerir de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imaginería motora [27]. Adicionalmente, utiliza señales cerebrales que están relacionadas con el área de lenguaje y tiene la posibilidad de utilizar la innumerable cantidad de palabras (léxico) de un lenguaje. A pesar de estas ventajas, aún son limitadas sus aplicaciones. Una de las causas está en que áreas como la neurociencia aún intentan entender los procesos neuronales cerebrales relacionados con el lenguaje. A pesar de esta dificultad, se han logrado algunos avances en el reconocimiento de vocales ([28], [29], [30]), sílabas ([31], [32]) y palabras ([33],[34]); pero aún requiere de una mayor cantidad de investigaciones para su aplicación práctica en BCIs, siendo un área abierta de investigación.

De esta manera, esta tesis doctoral busca responder la siguiente pregunta de investigación:

¿Qué señales cerebrales EEG, desarrollos numéricos y protocolos experimentales permiten el reconocimiento del habla silenciosa como base para el desarrollo de interfaces cerebro-computador?

Para responder la pregunta de investigación se plantean las siguientes hipótesis:

1. La identificación de señales cerebrales EEG a través del desarrollo de un algoritmo de clasificación basado en vector de características con entropía de la información y dimensión de regularización permiten el reconocimiento del habla silenciosa.

2. La identificación de señales cerebrales EEG a través del desarrollo de un algoritmo de clasificación basado en vector de características con datos funcionales permiten el reconocimiento del habla silenciosa.

1.2. Objetivos

Como objetivo general se busca establecer los fundamentos teóricos, desarrollos numéricos y protocolos experimentales que permitan el reconocimiento del habla silenciosa a través de señales electroencefalográficas (EEG) como base para el desarrollo de interfaces cerebro-computador.

Para el logro del objetivo general, esta propuesta de tesis doctoral propone tres objetivos específicos:

- Determinar los parámetros de las señales EEG que permitan la identificación de fonemas, con habla silenciosa, que incluyan vocales y sílabas.
- Establecer un algoritmo computacional que permita el procesamiento y clasificación de señales EEG con habla silenciosa.
- Comprobar la exactitud del algoritmo para el BCI con habla silenciosa a través de protocolos experimentales.

Capítulo 2

Referentes teóricos

Este capítulo describe los referentes teóricos de la tesis doctoral. Primero presenta los principales modelos neurológicos del lenguaje en los seres humanos, posteriormente muestra las principales características de las interfaces cerebro-computador, luego hace una descripción del habla silenciosa y finalmente presenta una fundamentación teórica de la entropía de la información, la dimensión de regularización y el análisis de datos funcionales.

2.1. El procesamiento del lenguaje en los seres humanos

El lenguaje humano es una de las características que más nos diferencia de otros animales [35]. El lenguaje, permite transmitir una infinidad de mensajes basados en un conjunto específico de signos [36]. Los signos de un lenguaje permiten conformar estructuras lingüísticas de nivel sintáctico, como sílabas, palabras, oraciones, textos, etcétera [37]. En el ser humano, el lenguaje está ligado a procesos de alto nivel tales como: la percepción, la atención, la memoria, el raciocinio y las emociones, garantizando así el desarrollo de la personalidad y la formación de la conducta con base en la comunicación como función primordial [38].

Actualmente, una de las líneas de investigación de la neurociencia está enfocada en el desarrollo de modelos que intentan explicar cómo el cerebro procesa la información cognitiva del lenguaje. De esta forma, se busca interpretar con mayor objetividad conductas verbales como el habla, la comprensión, la escritura, la lectura y la repetición de palabras [39].

2.1.1. El cerebro humano

El sistema nervioso humano está compuesto por células especializadas llamadas neuronas, que tienen como objetivo recibir la información, procesarla y responder a esos estímulos activando músculos u hormonas [41]. El sistema nervioso tiene dos partes principales, que son el sistema nervioso periférico y el sistema nervioso central. El sistema nervioso central está compuesto de la médula espinal y el encéfalo (Fig. 2.1). El encéfalo se encuentra en la cavidad craneana y se prolonga hasta la médula espinal; está cubierto por tres capas o meninges denominadas duramadre, aracnoides y piamadre (Fig. 2.2). Anatómicamente, el encéfalo se divide en tres partes: prosencéfalo, mesencéfalo y rombencéfalo (Fig. 2.1). El rombencéfalo se compone del bulbo raquídeo, la protuberancia y el cerebelo; y el prosencéfalo se compone del diencéfalo (parte interna central de los dos hemisferios) y el cerebro [40]. Para

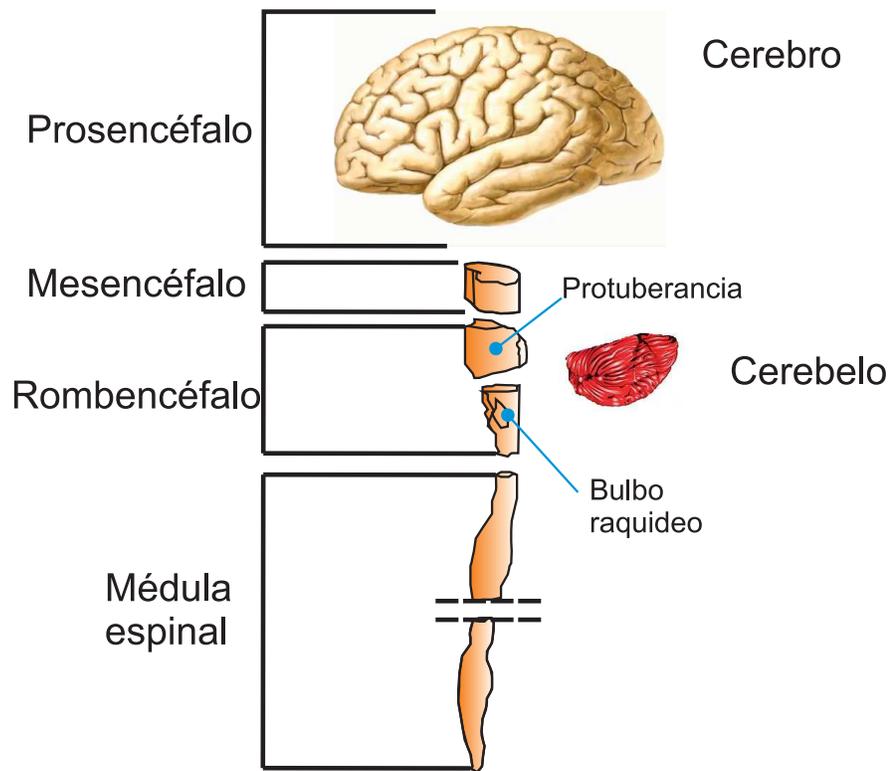


Figura 2.1: Sistema nervioso central. Adaptado de [40]

el estudio neuroanatómico, el cerebro se divide en planos denominados: plano sagital o medial, plano coronal o frontal y plano horizontal (Fig. 2.3).

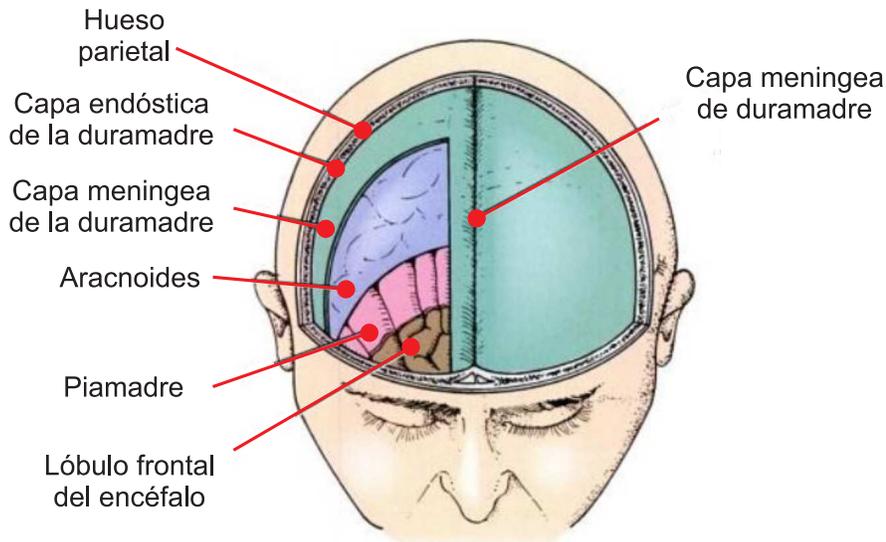


Figura 2.2: Cubierta protectora del encéfalo. Adaptado de [40]

El cerebro está compuesto por dos hemisferios (izquierdo y derecho) separados por un cuerpo calloso. La corteza de cada hemisferio cerebral posee una serie de pliegues denominados circunvoluciones o giros, que se encuentran separados por fisuras o surcos (Fig. 2.4)

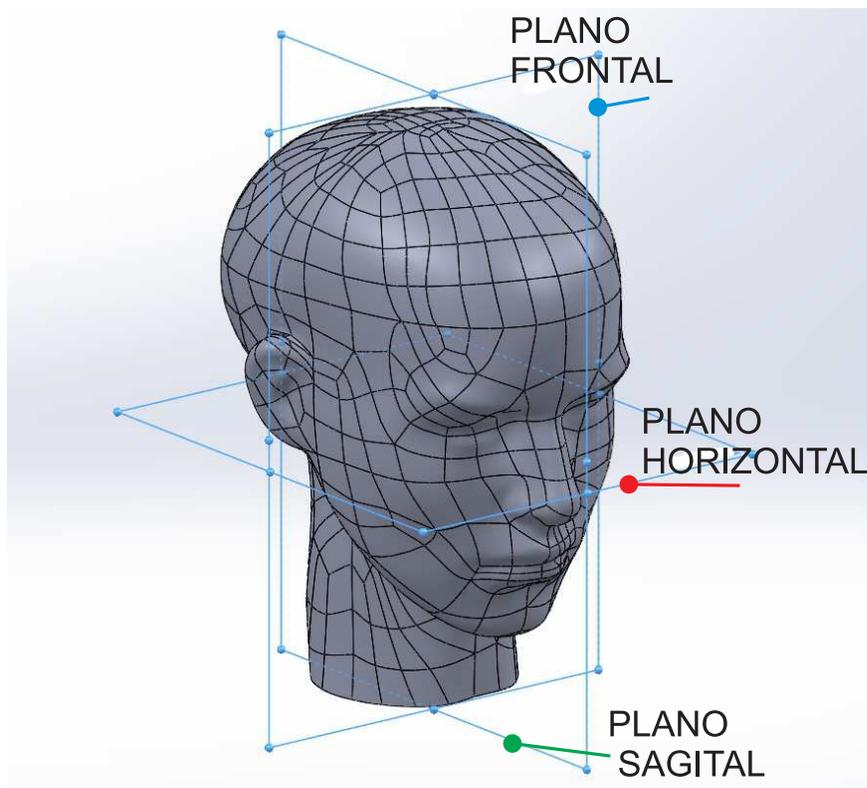


Figura 2.3: Planos de referencia para el cerebro

[42]. Dentro de las circunvoluciones se encuentran: la circunvolución frontal superior (CFS), la circunvolución frontal medial (CFM), la circunvolución frontal inferior (CFI), la circunvolución temporal superior (CTS), la circunvolución temporal medial (CTM), la circunvolución temporal inferior (CTI), la circunvolución parietal superior (CPS), la circunvolución parietal inferior (CPI), la circunvolución supramarginal (CS), la circunvolución angular (CA), la circunvolución precentral (Cpre), la circunvolución postcentral (Cpos), pars opercular (PO), pars triangular (PT), pars orbitaria (Por). De forma complementaria, se especifica el surco temporal superior (STS) y el surco temporal inferior (STI).

Las subdivisiones de cada hemisferio se basan en varios surcos de la corteza cerebral y se denominan lóbulos. Para referenciar partes de la corteza cerebral se utiliza, complementariamente, las áreas de Brodmann, que están numeradas del 1 al 51 por la citoarquitectura o disposición de las neuronas que conforman la corteza cerebral (Fig. 2.6). A continuación se presentan los cinco lóbulos del cerebro donde el nombre es dado por el hueso del cráneo debajo del cual se ubican.

El lóbulo frontal está en la parte posterior de la cisura de Rolando y se subdivide a la vez en: área precentral, área motora suplementaria, campo ocular, área motora del lenguaje de Broca y corteza prefrontal (Fig. 2.5). El área precentral se ubica principalmente sobre la circunvolución precentral. El área precentral se subdivide a la vez en área motora primaria (área 4 de Brodmann (BD)) y el área premotora (6 BD y partes de 8, 44 y 45 BD) (Fig. 2.6). La estimulación del área motora primaria produce el movimiento de algunos músculos del cuerpo ubicados en el hemisferio opuesto al del estímulo. La función del área premotora es almacenar programas de actividad motora con aprendizaje previo [43].

El área motora suplementaria está ubicada principalmente sobre la circunvolución frontal

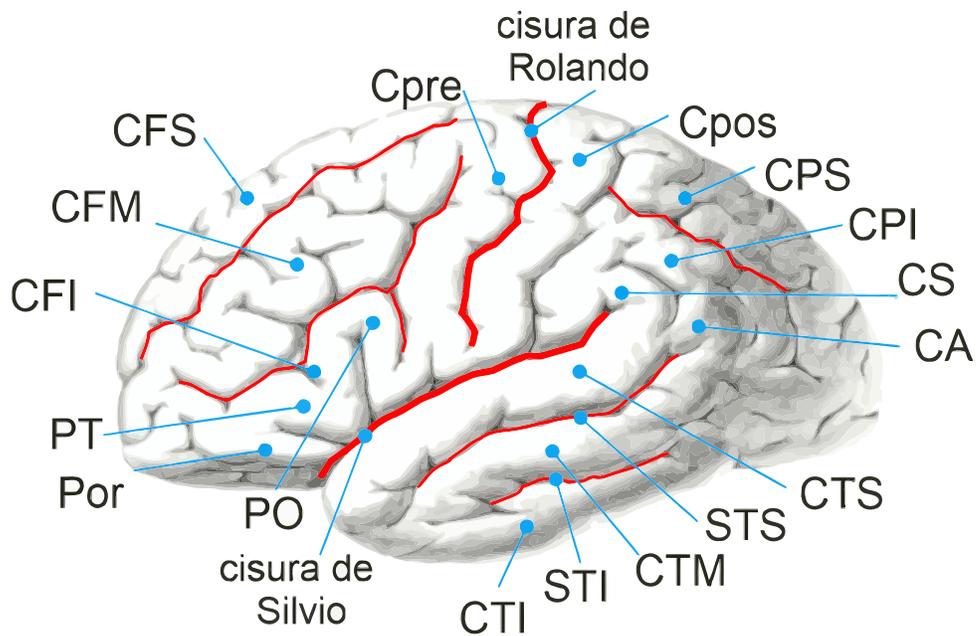


Figura 2.4: Vista lateral izquierda de la corteza cerebral. Circunvolución frontal superior (CFS), circunvolución frontal medial (CFM), circunvolución frontal inferior (CFI), circunvolución temporal superior (CTS), circunvolución temporal medial (CTM), circunvolución temporal inferior (CTI), circunvolución parietal superior (CPS), circunvolución parietal inferior (CPI), circunvolución supramarginal (CS), circunvolución angular (CA), circunvolución precentral (Cpre), circunvolución postcentral (Cpos), pars opercular (PO), pars triangular (PT), pars orbitaria (Por). Surco temporal superior (STS) y surco temporal inferior (STI).

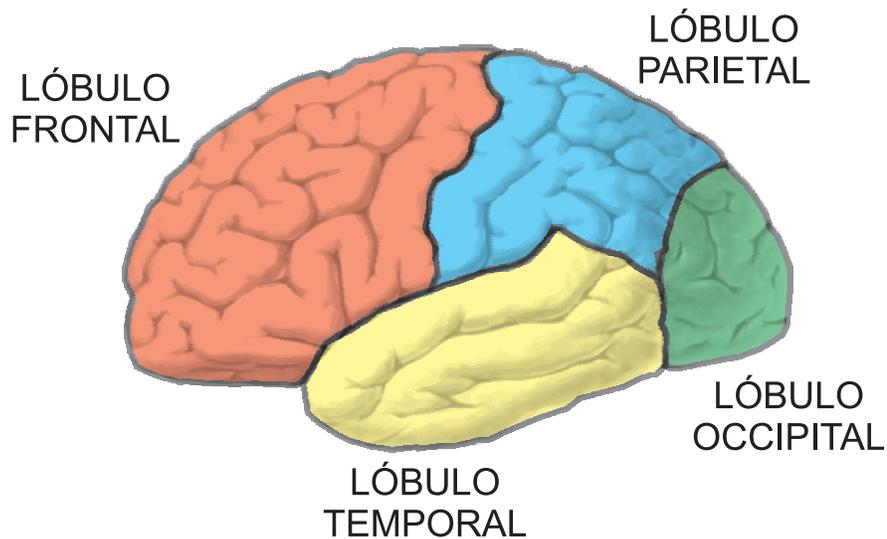


Figura 2.5: Lóbulos del cerebro

media y también está relacionada con el movimiento motor [43].

El campo ocular frontal se encuentra sobre partes de la circunvolución frontal media relacionada con las áreas 6,8 y 9 BD (Fig. 2.6). La función principal de esta área es controlar el movimiento voluntario de los ojos.

La corteza prefrontal está ubicada sobre las áreas 9, 10, 11 y 12 de Brodmann (Fig. 2.6), principalmente en la parte anterior del cerebro, sobre partes de la circunvolución superior, circunvolución media y circunvolución inferior, entre otras. A la vez, está relacionada con el cerebelo, el hipotálamo y el cuerpo estriado, entre otros. Esta área es fundamental para determinar procesos cognitivos relacionados con la personalidad [40].

El área motora del lenguaje de Broca está ubicada principalmente sobre la circunvolución frontal inferior (áreas 44 y 45 BD), y su función está relacionada con la producción del habla.

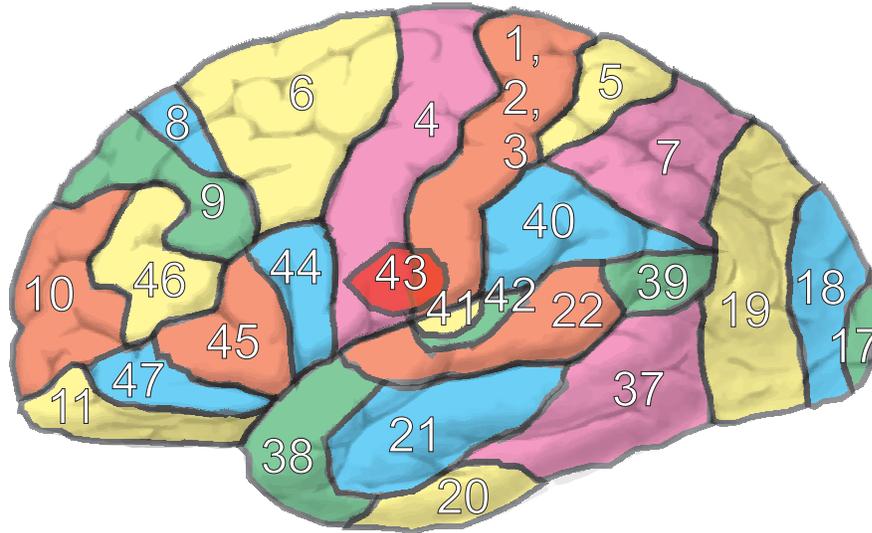


Figura 2.6: Áreas de Brodmann de la corteza cerebral

El lóbulo parietal posee tres áreas: la somatoestésica primaria, la somatoestésica secundaria y la somatoestésica de asociación (Fig. 2.5). Las áreas somatoestésica primaria y secundaria están ubicadas sobre las áreas 1, 2 y 3 de Brodmann, localizadas principalmente sobre la circunvolución postcentral (Fig. 2.6). Esta área recibe la información sensorial de las diferentes partes del cuerpo como los dermatomas. La proyección de una parte del cuerpo en esta área está en relación directa con la cantidad de receptores sensitivos que posea un área determinada del cuerpo, destacándose la cara, los labios y el pulgar. El área somatoestésica de asociación, por su parte, está ubicada sobre las áreas 5 y 7 de Brodmann (Fig. 2.6), y está relacionada con la habilidad cognitiva de representación espacial e integrada con otras áreas sensitivas que le permiten a una persona el reconocer la forma de un objeto con los ojos cerrados [42].

El lóbulo occipital está conformado por el área visual primaria y el área visual secundaria (Fig. 2.5). El área visual primaria está localizada sobre el área 17 de Brodmann y está conectada con fibras provenientes de los nervios ópticos (Fig. 2.6). El lóbulo occipital izquierdo recibe información del nervio óptico derecho, y el lóbulo occipital derecho recibe información del nervio óptico izquierdo. El área visual secundaria rodea al área visual primaria y está localizada sobre las áreas 17 y 18 de Brodmann (Fig. 2.6). La función de esta área es integrar información de la corteza visual primaria con experiencias previas que estén almacenadas en la memoria. Tanto en la corteza visual primaria como secundaria se integra información visual de luminancia y de croma [43].

El lóbulo temporal está conformado por el área auditiva primaria, el área auditiva secundaria y el área de Wernicke (Fig. 2.5). El área auditiva primaria está ubicada sobre las

áreas 21 y 41 de Brodmann, donde se procesa la información auditiva proveniente de células sensitivas del órgano de Corti (oído interno); al igual que en el lóbulo occipital, esta recibe información del hemisferio opuesto. El área auditiva secundaria está ubicada sobre el área 22 de Brodmann y se cree que está involucrada en la interpretación, por parte del cerebro, de los sonidos y procesos de asociación auditiva con otras áreas del cerebro (Fig. 2.6). El área de Wernicke está relacionada con áreas de la circunvolución temporal superior y con el proceso de comprensión del lenguaje [40].

2.1.2. Modelos neurológicos del lenguaje en seres humanos

El cerebro humano es el sistema más complejo que se haya descubierto en el universo, pues contiene billones de neuronas y trillones de conexiones o sinápsis [44]. El lenguaje es una de las actividades cognitivas más complejas que esta realiza. Para el estudio de los procesos de lenguaje en el cerebro humano, desde la neurociencia se han desarrollado diferentes modelos basados en estudios invasivos y no invasivos.

Los primeros estudios sobre las bases neurológicas del lenguaje partían de pacientes con algún tipo de trastorno del lenguaje (afasia). Al sujeto se le hacía un seguimiento durante su vida, y al fallecer se le practicaba una autopsia. Allí se observaba la zona del cerebro que estaba dañada y se hacía la relación del área cerebral con la afasia.

Después de la segunda mitad del siglo *XX* se desarrollaron métodos no invasivos con los cuales fue posible estudiar a los sujetos en vida realizando tareas de lenguaje. Dentro de los métodos no invasivos se encuentran: la tomografía por emisión de positrones (PET), la resonancia magnética funcional (fMRI), las técnicas electroencefalográficas, la magnetoencefalografía y la estimulación magnética intracraneal (TMS), entre otras [45].

Un primer estudio con bases neurológicas del lenguaje fue realizado por Marc Darx en 1836, en el que se describe una correlación entre daños en el hemisferio izquierdo y la pérdida del habla. Este estudio lo realizó con 20 pacientes [46].

Modelo de Broca

Este modelo fue propuesto por Paul Broca en 1861 al estudiar a un paciente con problemas de producción del habla. En este caso el paciente sólo podía pronunciar la palabra "tan", pero podía comprender lo que le decían. En las observaciones del cerebro de este paciente, después de fallecido, y de otros seis sujetos, Broca encontró daños similares en la circunvolución frontal inferior del hemisferio izquierdo[47]. Estas observaciones condujeron a Paul Broca a desarrollar un primer modelo del habla, donde infirió que la circunvolución frontal inferior del hemisferio izquierdo está relacionada con la producción del habla (Fig. 2.7) [46].

Modelo de Wernicke

Este modelo fue propuesto por Carl Wernicke en 1874 cuando analizó una serie de pacientes que hablaban pero no comprendían lo que decían. De los exámenes post mortem, se encontraron lesiones en la parte posterior de la circunvolución temporal superior en el hemisferio izquierdo. De esta forma, Wernicke correlacionó esta área del cerebro con la comprensión del habla. El modelo que propuso Wernicke tenía dos centros cerebrales, uno encargado de comprender el habla (área de Wernicke) y otro encargado de la producción del habla (área de Broca) (Fig. 2.8). Adicionalmente, supuso que debían existir unas fibras nerviosas para

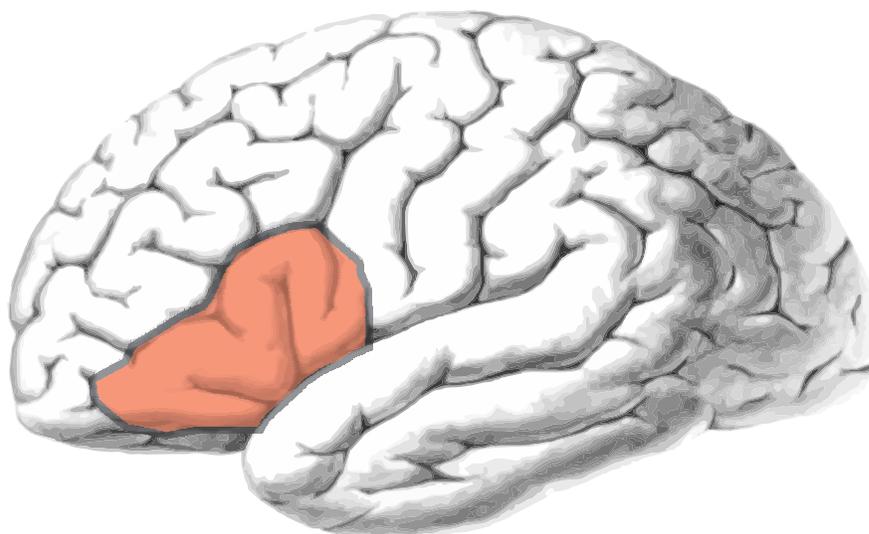


Figura 2.7: Modelo de Broca. Color rojo área de Broca.

comunicar estas dos regiones, y que estas al dañarse podrían impedir la repetición de una palabra (afasia de conducción) [48].

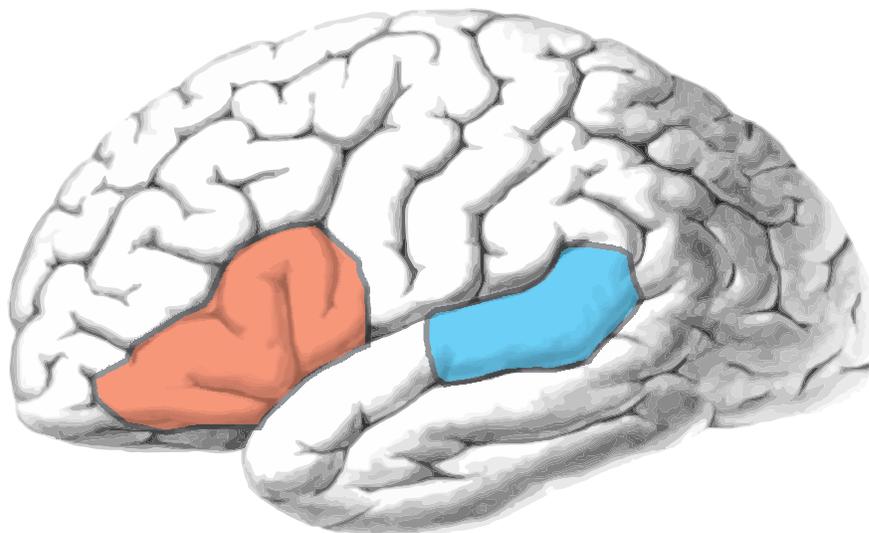


Figura 2.8: Modelo de Wernicke. Color rojo, área de Broca; color azul, área de Wernicke.

Modelo de Litchtheim

En 1885, Litchtheim analizó un paciente que comprendía las palabras, podía pronunciarlas, pero no podía realizar procesos de repetición de estas. En este sentido, concluyó que la hipótesis de Wernicke sobre la afasia de conducción era cierta. El modelo que propuso Litchtheim, aún vigente hoy en día, está compuesto por tres centros en el hemisferio izquierdo (el auditivo, el motor y el conceptual) junto con sus conexiones [47]. El centro auditivo está relacionado con la comprensión del habla, el centro motor con la producción del habla y el centro conceptual con el almacenamiento del significado de componentes del lenguaje (Fig. 2.9). Aunque el modelo de Litchtheim no determinó en qué lugar se encontraba el centro

conceptual, permitió modelar otras afasias como la afasia transcortical motora (ATM) y la afasia transcortical sensorial (ATS). La ATM consiste en un daño en la conexión entre el centro conceptual y el centro motor, donde el sujeto tiene las características de la afasia de Broca pero puede repetir palabras. En el caso de la ATS se presenta un daño entre el centro auditivo y el centro conceptual, presentando características de la afasia de Wernicke pero con la posibilidad que los sujetos puedan repetir palabras [48].

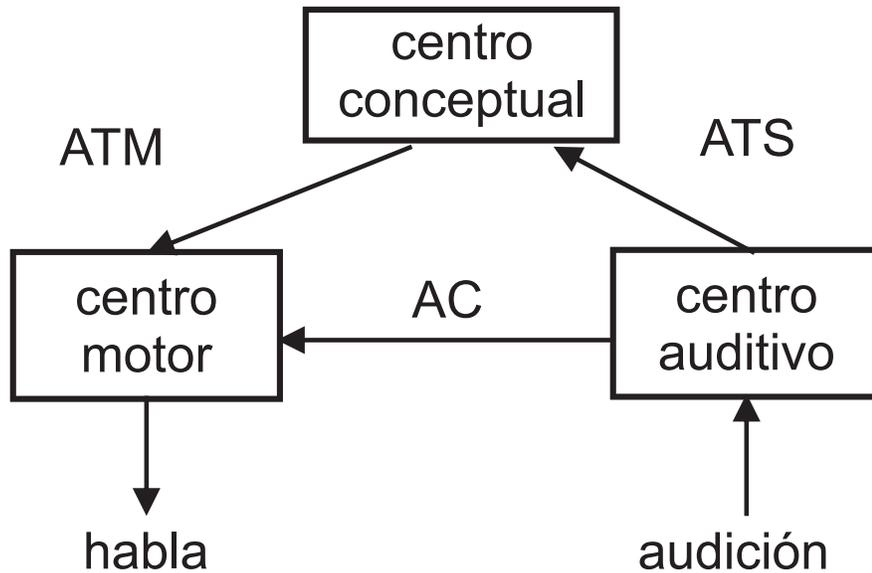


Figura 2.9: Modelo de Litchtheim. Afasia de conducción (AC), afasia transcortical motora (ATM) y afasia transcortical sensorial (ATS)

Modelo globalista

El modelo globalista, cuyo principal representante fue Lashley, surge a inicios del siglo XX y tiene su auge en la primera mitad del siglo XX, desplazando el modelo de Litchtheim. Este modelo supone que todas las tareas del lenguaje están repartidas en el cerebro y que no existe un lugar en particular para una tarea del lenguaje (principio de equipotencialidad). El modelo de Lashley relaciona una afasia, exclusivamente, con el tamaño del área afectada, sin tener en cuenta un lugar en particular [47].

Modelo de Geshwind

El modelo de Geshwind, desarrollado en 1965, retomó los trabajos de Wernicke y Litchtheim. El modelo de Geshwind se basa en centros de lenguaje y conexiones entre esos centros conformados por fibras o tractos nerviosos; plantea que hay tres centros en el hemisferio izquierdo (Fig. 2.10): el centro de producción del lenguaje (área de Broca), el centro de comprensión del lenguaje (área de Wernicke) y el centro conceptual que lo ubica en la circunvolución supramarginal y circunvolución angular, ambas pertenecientes a la zona posterior e inferior del lóbulo parietal [49].

Con este modelo es posible relacionar las afasias de: Wernicke, Broca, AC, ATM, ATS, global y anómica. La afasia global involucra al mismo tiempo la afasia de Broca con la afasia

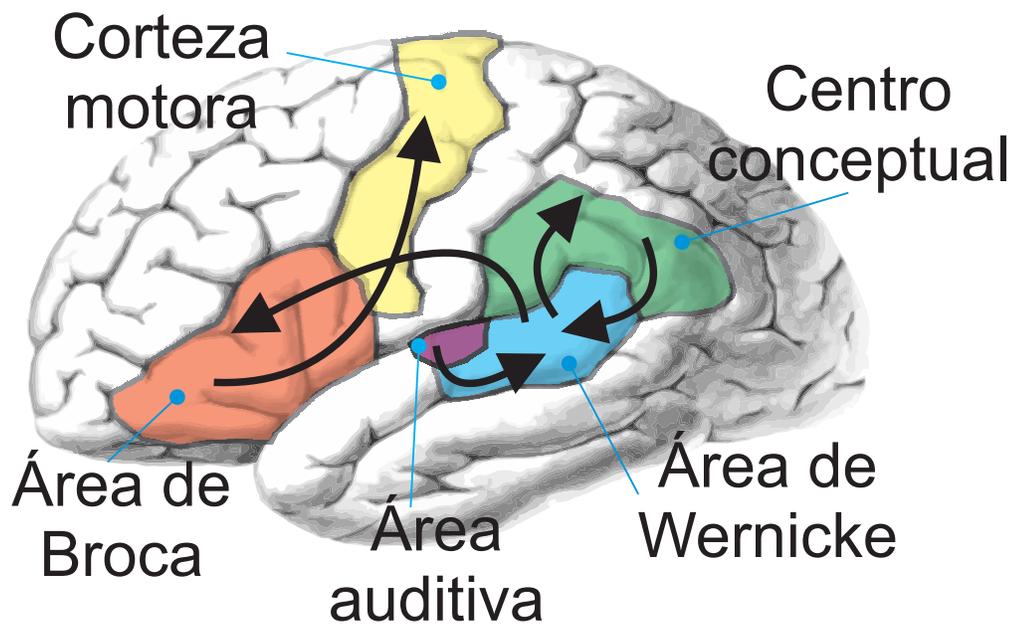


Figura 2.10: Modelo neurológico de Geshwind

de Wernicke, y la afasia anómica está relacionada con la dificultad para recuperar el léxico (palabras).

La siguiente es una descripción del modelo de Geshwind cuando una persona escucha una palabra y desea repetirla (Fig. 2.10). Primero, la información auditiva es recibida en el cerebro por parte del área auditiva primaria; posteriormente, la información pasa al área de Wernicke donde la persona comprende las señales que está escuchando; esta información se relaciona con el centro conceptual donde las palabras cobran un significado, a continuación se envía la información del área de Wernicke al área de Broca a través del fascículo arqueado y finalmente la información se envía al área motora para iniciar el proceso de articulación vocal [50].

Este modelo aún se utiliza en la actualidad, pero presenta diferentes limitaciones; por ejemplo, existen diferentes afasias como la afasia mixta, que no puede modelarse con el modelo de Geshwind. Las técnicas de neuroimágenes actuales muestran que existen más áreas que se activan en el hemisferio izquierdo que las propuestas en el modelo de Geshwind. Estas técnicas también muestran que el hemisferio derecho y partes subcorticales del cerebro también se activan durante tareas de procesamiento del lenguaje. En las técnicas de neuroimágenes también se ha detectado que las áreas de Wernicke y de Broca no son fijas para todos los individuos y presentan desfases con relación a la ubicación en la corteza cerebral planteada por Geshwind [51].

Modelo cognitivo de Benedet

El modelo cognitivo de Benedet representa los procesos cerebrales del lenguaje como flujos de información entre bloques. Este modelo tiene tres entradas perceptuales de información: escuchar una palabra, ver un objeto o ver una palabra en forma escrita [52].

Cuando se trata de escuchar una palabra y pronunciarla en voz alta, se genera la siguiente secuencia (Fig. 2.11): primero, se genera una conversión de la señal acústica a información de fonemas. La información de los fonemas puede tener una vía semántica o asemántica.

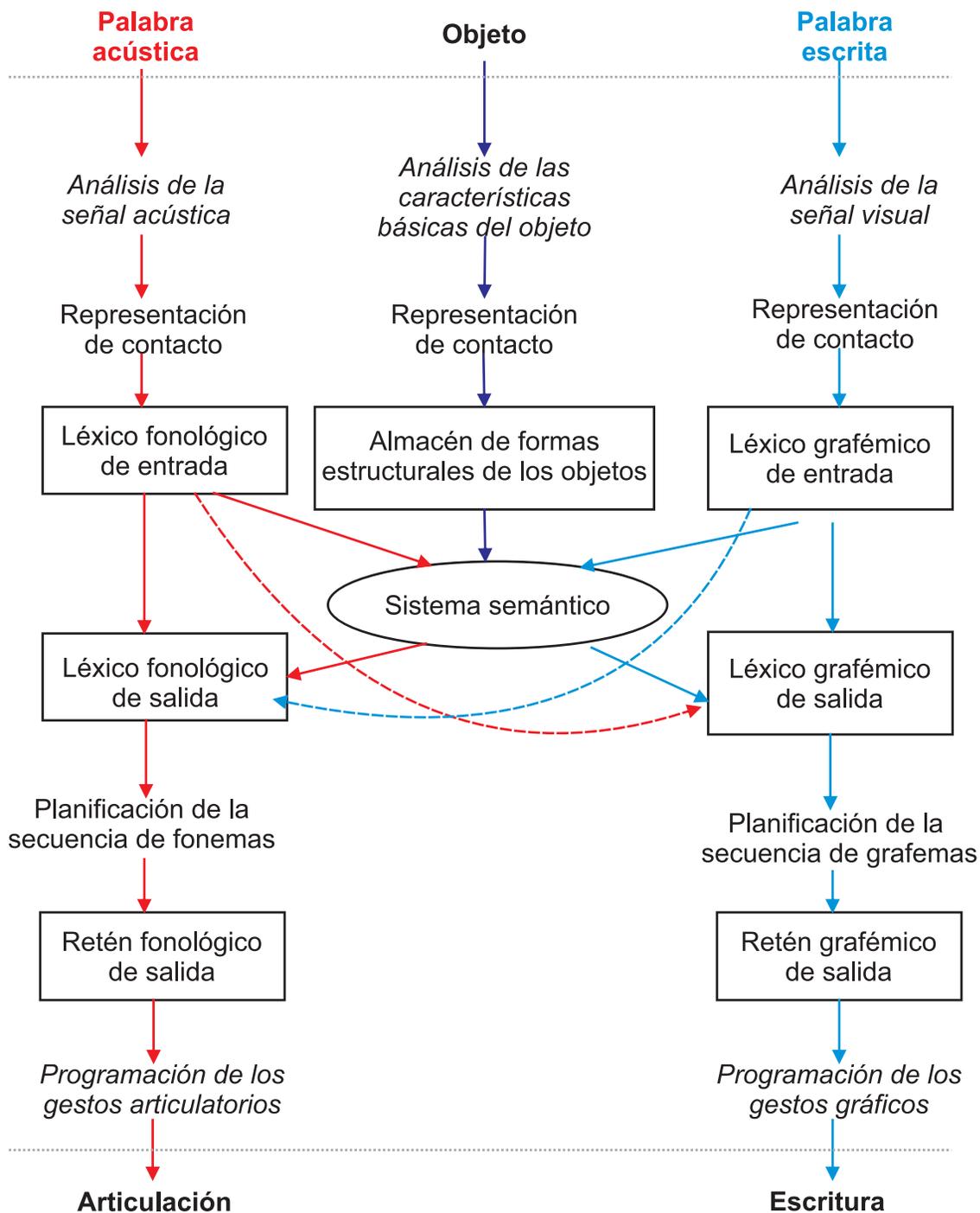


Figura 2.11: Diagrama modelo cognitivo de Benedet. Adaptado de [52]

En el caso de la vía semántica, el fonema que se está percibiendo tiene un significado (vía semántica). Un ejemplo de casos asemánticos puede ser las seudopalabras, que son fragmentos discursivos que no tienen significado (vía directa entre el léxico fonológico de entrada y de salida). Después de que el cerebro determina los fonemas que va a vocalizar, se produce una planificación de la secuencia de fonemas. Estos fonemas se almacenan en una memoria de trabajo denominada retén fonológico de salida. Posteriormente se programan los gestos articulatorios para producir los fonemas en voz alta [52].

Cuando se trata de leer una palabra sin pronunciarla y posteriormente escribirla, entonces se genera la siguiente secuencia (Fig. 2.11): en primer lugar, se compara de forma visual cada grafema de la palabra con el léxico de grafemas que el sujeto posee. La información de los grafemas puede tener una vía semántica o asemática. En el caso de la vía semántica, el grafema que se está percibiendo tiene un significado (vía semántica). En el caso en que el grafema no tenga un significado, continúa por la vía directa entre el léxico grafémico de entrada y de salida. A la salida del léxico grafémico se planea la secuencia de grafemas a escribir. Posteriormente, cada grafema pasa al retén grafémico de salida relacionado con la memoria de trabajo. Finalmente, se programan los gestos gráficos que van a producir la escritura.

En el caso en que el sujeto ve un objeto y quiere pronunciar el respectivo nombre, se generan los siguientes pasos (Fig. 2.11): primero se hace un análisis de las características visuales del objeto. Posteriormente, se compara con el almacén de formas que el sujeto ha aprendido. Si el sujeto lo reconoce y encuentra un significado, relaciona el léxico fonológico propio de ese objeto. Después de que el cerebro determina los fonemas que va a vocalizar, se produce una planificación de la secuencia de fonemas. Estos fonemas se almacenan en una memoria de trabajo denominado retén fonológico de salida. Posteriormente, se programan los gestos articulatorios para producir los fonemas en voz alta [53].

En el caso en que el sujeto ve un objeto y quiere escribir la respectiva palabra, se genera la siguiente secuencia (Fig. 2.11): primero se hace un análisis de las características visuales del objeto. Posteriormente, se compara con el almacén de grafemas que el sujeto ha aprendido. Si el sujeto lo reconoce y encuentra un significado, relaciona el léxico fonológico propio de ese objeto. A la salida del léxico grafémico de salida, se planea la secuencia de grafemas a escribir. Luego cada grafema pasa al retén grafémico de salida relacionado con la memoria de trabajo y, finalmente, se programan los gestos gráficos que van a producir la escritura [52].

Modelo de comprensión oral de Hickok y Poeppel

Hickok y Poeppel (2007) proponen un modelo de la estructura funcional de la corteza cerebral en procesos de comprensión oral, donde la persona escucha una palabra y tiene que repetirla. En este modelo hay muchas corrientes de información y múltiples áreas del cerebro (bloques) que trabajan juntas para el procesamiento del lenguaje [54].

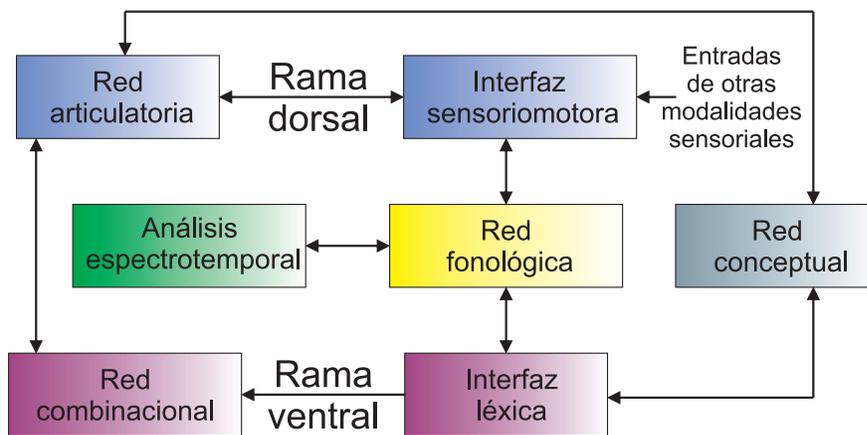


Figura 2.12: Diagrama de bloques modelo de comprensión oral de Hickok y Poeppel. Adaptado de [54]

En este modelo, un primer bloque denominado análisis espectral temporal recibe la información sonora en el área auditiva primaria (hemisferios izquierdo y derecho) de la corteza cerebral (Fig. 2.12),(Fig. 2.13). En estas áreas existen dos niveles de codificación, el primer nivel responde a una amplia gama de señales, incluido el ruido, y el segundo nivel responde a señales con estructura temporal (ritmos musicales, señales de voz, entre otros). Ambas regiones se encuentran cerca a la circunvolución de Helch que se ubica en BD 41 y 42. En las figuras (Fig. 2.12), (Fig. 2.13) el área de análisis espectral temporal está representado en color verde [54].

El siguiente bloque se denomina red fonológica el cual tiene como objetivo reconocer, de la señales generadas en el bloque de análisis espectral temporal, elementos específicos del lenguaje como los fonemas (Fig. 2.12), (Fig. 2.13). La red fonológica está ubicada en la circunvolución temporal superior y la mitad posterior de la cisura temporal superior de ambos hemisferios. En las figuras (Fig. 2.12), (Fig. 2.13) la red fonológica está representado en color amarillo. Desde este bloque las señales divergen en dos vías: una vía ventral relacionada con el significado de los sonidos y una vía dorsal que relaciona los sonidos con las respectivas articulaciones.

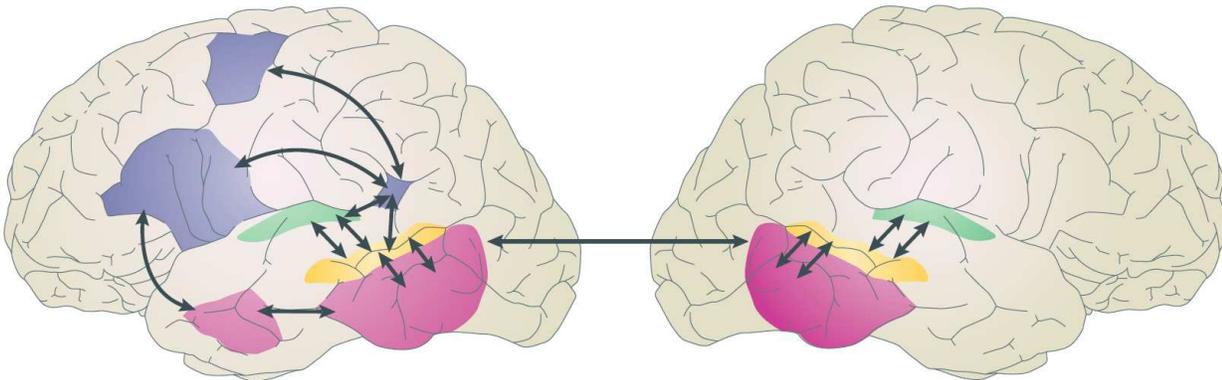


Figura 2.13: Modelo neurológico de comprensión oral de Hickok y Poeppel. Tomado de [54]

Las señales de la red fonológica que van por la vía ventral se dirigen al bloque de interfaz léxica (Fig. 2.12), (Fig. 2.13). En este lugar se le asigna una representación conceptual a los fonemas que han sido reconocidos en la red fonológica. Una persona que no tenga problemas de afasias puede reconocer las palabras de un léxico determinado; por ejemplo, si una persona de habla hispana escucha la palabra "gato". En este caso la persona diferencia esa palabra del resto de su léxico. La ubicación de esta red está principalmente en la parte posterior de la circunvolución temporal media y en la parte posterior del surco temporal inferior de los dos hemisferios cerebrales [54].

El cerebro puede no solamente reconocer y diferenciar palabras aisladas, pero también puede realizar procesos donde se estructuran palabras para formar sistemas lingüísticos de mayor complejidad. En el bloque red combinacional permite formar sistemas complejos del lenguaje con carácter sintáctico. Con esta red, una persona puede comprender oraciones, textos, etcétera. La ubicación de esta red está principalmente en la parte anterior de la circunvolución temporal media y en la parte anterior del surco temporal inferior. La red combinacional en este modelo está ubicada en el hemisferio izquierdo. En las figuras (Fig. 2.12), (Fig. 2.13) la red de interfaz léxica y la red combinacional están representadas en color magenta [55].

Las señales de la red fonológica que van por la vía dorsal, se dirigen al bloque de interfaz sensoriomotora (Fig. 2.12), (Fig. 2.13). Este bloque es una vía de comunicación desde los procesos de percepción auditiva de la corteza auditiva primaria hacia el lóbulo parietal y la producción del habla (lóbulo frontal). El bloque de interfaz sensoriomotora se encuentra localizado en una parte profunda de la cisura de Silvio, entre el lóbulo parietal y temporal. Este bloque también posee otra entrada de información relacionada con procesos de percepción no auditivos, como el visual o el táctil, y aquellos que se integran con la información verbal. Este bloque es importante porque permite formar la representación motora de los fonemas que conoce un hablante y palabras nuevas que no se encuentren en su léxico [56].

Cuando una persona ha codificado la información de una palabra y la desea expresar en voz alta, requiere transformar esta información en una codificación para que el sistema articulatorio pueda recibir la información requerida y para que el aparato fonador pueda producir la articulación requerida. El bloque red articulatoria se encarga de tomar la información de la interface sensoriomotora y enviarla a la circunvolución precentral, para que sea producida la señal de voz por el aparato fonador. Este bloque se encuentra ubicado en el hemisferio izquierdo y principalmente en la parte posterior de la circunvolución frontal inferior, la corteza premotora y la parte anterior de la ínsula, un lugar del cerebro ubicada en la parte profunda de la cisura de Silvio. En las figuras (Fig. 2.12) y (Fig. 2.13) la interfaz sensoriomotora y la red articulatoria están representadas en color azul [54].

Finalmente, el bloque de red conceptual implica la relación del área del lenguaje con una amplia red de circuitos neuronales en el cerebro, donde la información de otras áreas se integra con la información del lenguaje. En las figuras (Fig. 2.12) y (Fig. 2.13) la red conceptual está representada en color gris.

Modelo neurológico sobre ubicación de la percepción del habla

En 2004, Indefrey y Cutler desarrollaron un estudio sobre localización cortical de tareas lingüísticas auditivas de tipo pasivo. Esta investigación se desarrolló con 1058 personas utilizando información hemodinámica de neuroimágenes [57].

En este estudio se encontró que el procesamiento de la información fonológica se halla en la parte posterior de la corteza temporal superior, relacionada con el área de Wernicke. Adicionalmente, se encontró que la información fonológica, en la corteza cerebral, tiene una estructura jerárquica. Mientras más compleja sea la información lingüística, más tenderá a ubicarse en la parte anterior y ventral de los lóbulos temporales. El estudio se realizó con los siguientes elementos lingüísticos: tonos, seudopalabras, palabras y oraciones. En resumen, las oraciones se ubicaron en la parte anterior de la corteza temporal superior, mientras los tonos tendieron a ubicarse en la parte posterior, relacionada con la corteza temporal superior [47].

2.2. Interfaces cerebro computador

En esta sección se hace una presentación sobre las interfaces cerebro-computador *brain-computer interface* (BCIs). Se busca especificar en que consiste las BCIs, que señales cerebrales son las más utilizadas en este campo y que desarrollos se han alcanzado.

2.2.1. Arquitectura de BCIs

Las interfaces cerebro-computador (BCIs) también denominadas interfaz hombre máquina, son sistemas invasivos o no invasivos que usan señales cerebrales para controlar computadores o dispositivos de *hardware* [58]. En este sentido, una interfaz cerebro-computador es un canal de salida artificial que permite una comunicación directa entre el cerebro y un computador o máquina, que puede aceptar comandos cerebrales sin requerir movimientos del sistema muscular [24].

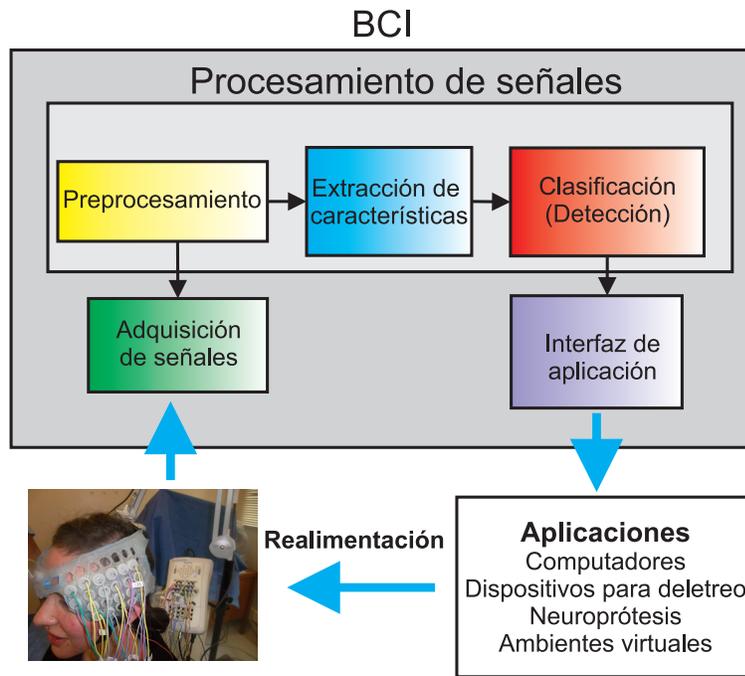


Figura 2.14: Arquitectura general de una BCI. Adaptado de [24].

El diagrama básico de una BCI consta de: un bloque de adquisición de las señales cerebrales, un bloque de preprocesamiento, un bloque de extracción de características, un bloque de clasificación y un bloque interface de salida [24] (Fig. 2.14). El bloque de adquisición de las señales se encarga de capturar las señales del cerebro y transformar esa información en un conjunto de valores discretizados. En el caso de electroencefalografía, se trata con valores en μV , y en el caso de imágenes de resonancia magnética funcional fMRI, se trata con valores en voxels. El bloque de preprocesamiento es un algoritmo que busca filtrar la mayor cantidad de ruido (artefactos) posible de las señales cerebrales. El bloque de extracción de características es un algoritmo que toma los parámetros más relevantes de las señales cerebrales para ser utilizados como elementos representativos de cada clase. El bloque de clasificación es un algoritmo que busca separar las características de una clase (patrón) con respecto a las otras clases o patrones. Finalmente, la interfaz convierte los resultados de la clasificación en señales apropiadas, dependiendo del dispositivo o máquina que se desee controlar. El cerebro percibe el resultado del BCI generando un proceso de realimentación.

2.2.2. Métodos para capturar señales cerebrales

Existen diferentes técnicas para la captura de las señales neuronales. En el campo de las neuroimágenes, están: PET, MEG, rTMS, NIRS, ES, fMRI, EEG, entre otras. La PET (*positron emission tomography*) consiste en inyectar en forma intravenosa trazadores radioactivos, y se basa en procesos de contraste debido a la utilización del oxígeno en diferentes áreas del cerebro. La MEG (*magnetoencephalography*) se basa en la utilización de grandes campos magnéticos, utilizando superconductores; de esta forma, se mide la actividad biomagnética en el cerebro. La rTMS (*repetitive transcranial magnetic stimulation*) utiliza cambios rápidos del campo magnético a específicas áreas del cerebro, inhibiendo o excitando grupos de neuronas. NIRS (*near-infrared spectroscopy*) es un método óptico que se basa en los cambios de absorción del oxígeno en el cerebro, basado en propiedades ópticas de la hemoglobina (deoxihemoglobina y oxihemoglobina). La ES (*electrocortical stimulation*) aplica corriente eléctrica sobre la corteza cerebral para generar pequeñas interrupciones temporales. Dos métodos ampliamente utilizados y de carácter no invasivo son el fMRI y las señales EEG [59].

fMRI

Las imágenes de resonancia magnética funcional (*functional magnetic resonance imaging* - fMRI) es una técnica de neuroimágenes que permite observar la hemodinámica cerebral de forma no invasiva. El método consiste en colocar a una persona en un gran campo magnético del orden de 1.5 Teslas. Cuando se coloca el cerebro dentro de un campo magnético de gran intensidad, los espines de los átomos de hidrógeno se orientan en una sola dirección. En las regiones del cerebro donde se está produciendo mayor actividad neuronal, las respectivas neuronas tienen una mayor irrigación vascular [60]. En este caso se tiene una mayor concentración de oxihemoglobina, incrementando el nivel de oxígeno en la sangre, y una disminución de la deoxihemoglobina. El efecto de la dependencia del nivel de oxígeno en la sangre se denomina efecto BOLD, por sus siglas en inglés. Debido a que la deoxihemoglobina tiene diferentes propiedades magnéticas, sirve como elemento de contraste. Una ventaja del fMRI es la buena resolución espacial (3 mm) y el no ser invasivo. Dentro de las limitaciones se encuentran: la baja resolución temporal (1 - 5 s), el alto costo del equipo, el requerir de un alto campo magnético, el no ser portable, el ser sensible al movimiento de la persona, el hecho de que el campo magnético produce un ruido entre 93 a 98 dB puede afectar estudios relacionados con el área del lenguaje [59].

Electroencefalografía

Las primeras investigaciones de electroencefalografía (EEG) en seres humanos fueron realizadas por Berger en 1920, utilizando un galvanómetro de alta sensibilidad. El primer reporte de una señal EEG fue realizado por Berger en 1929, donde mostró las ondas cerebrales alfa. Para este trabajo utilizó solamente un canal entre el lóbulo occipital y el lóbulo frontal [24]. También realizó la primera grabación de una señal EEG durante 3 minutos. Berger también realizó el primer registro de una señal EEG en periodo de sueño. A principio de los años 30, Toennies desarrolló el primer amplificador para señales cerebrales. Kornmuller realizó la primera grabación simultánea de varios electrodos cubriendo la corteza cerebral. Fischer y Lowenbach realizaron el primer estudio de epilepsia utilizando señales EEG. Gray fue el pionero en aplicación de la electroencefalografía clínica y desde ese momento se ha expandido su aplicación para detectar diferentes enfermedades como: traumatismos craneocerebrales,

tumores cerebrales, desórdenes cerebrovasculares, epilepsia, síndrome de Down, glaucomas, etc [24].

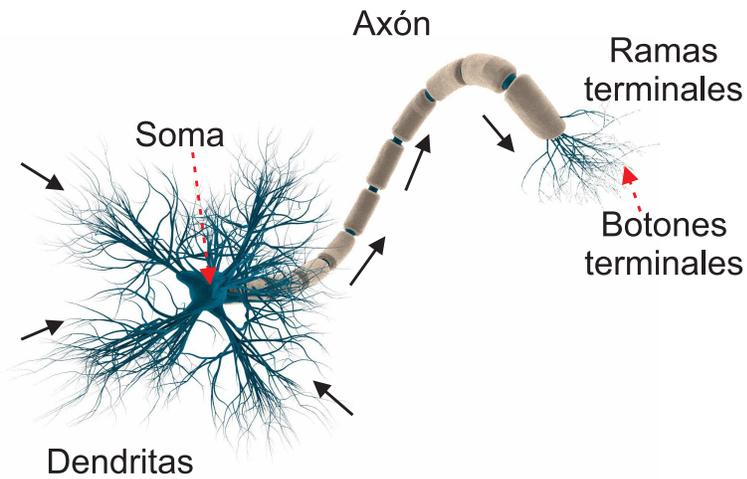


Figura 2.15: Esquema de una neurona. Adaptado de [61]

Potencial de acción generado en las neuronas El cerebro es un sistema de alta complejidad compuesto por millones de neuronas. Cada neurona está conformada por un cuerpo celular (soma), dendritas, axones y terminaciones nerviosas (Fig. 2.15) [39]. La entrada de información a la neurona se realiza a través de las dendritas, las cuales reciben información de en promedio 10000 neuronas. Con relación a la membrana celular, el exterior de la neurona tiene mayor concentración de iones Na^+ y el interior mayor concentración de iones K^+ . Con esta concentración, el potencial de membrana es -70 mV , donde la parte interna se carga negativamente y la parte externa positivamente [39], [61]. En el momento en que ocurre una despolarización de la neurona debido a la información recibida en las dendritas, se produce un potencial de acción que viaja a través del axón (Figs. 2.16 y 2.17) [62].

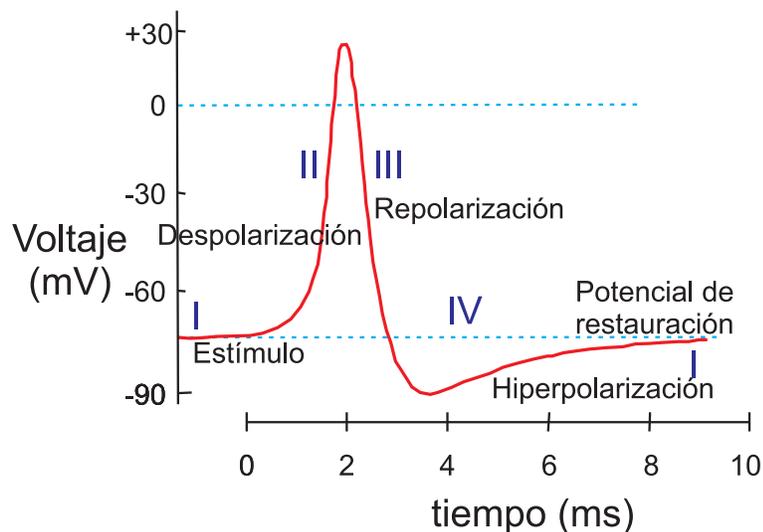


Figura 2.16: Potencial de acción en la membrana de una neurona. Adaptado de [62].

El mecanismo que produce el potencial de acción está relacionado con la bomba de sodio potasio. En el momento que la despolarización alcanza un umbral entre -70 mV a -55 mV , los canales de sodio dejan entrar los iones Na^+ , invirtiendo la polaridad de la membrana [61]. Esta fase se llama despolarización y alcanza un valor promedio de 30 mV (Figs. 2.16 y 2.17) [62].

Cuando se ha producido la despolarización, se cierran los canales de sodio y se abren los canales de potasio [61]. De esta forma salen los iones K^+ , produciéndose la repolarización de la membrana. Cuando se alcanza el valor de -90 mV este proceso se denomina hiperpolarización, y sirve para que en esa etapa la neurona no reciba otro estímulo [39]. Posteriormente la bomba de sodio potasio restablece el potencial de membrana a -70 mV (Figs. 2.16 y 2.17).

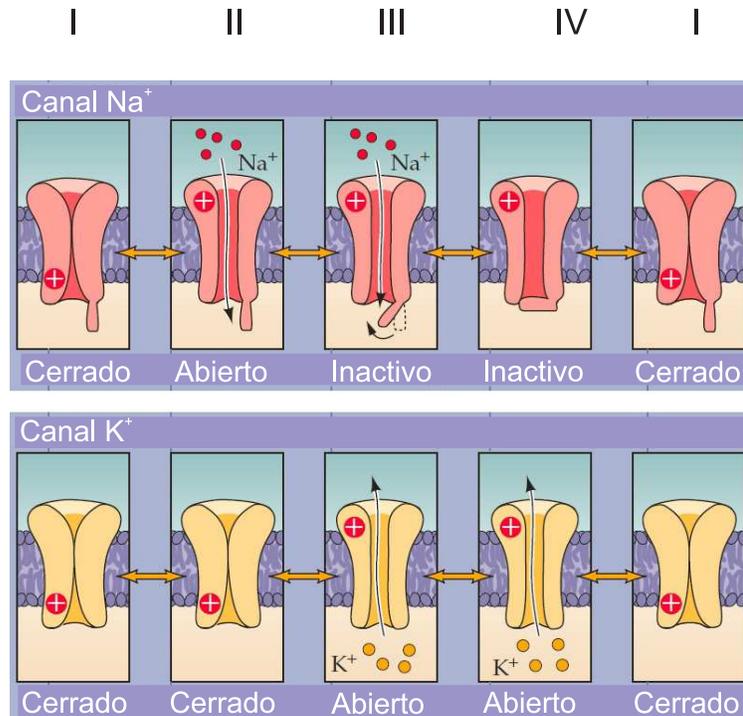


Figura 2.17: Comportamiento de los canales de Na^+ , K^+ ante un potencial de acción. Tomado de [39].

El potencial de acción viaja a través del axón a una velocidad de entre 1 y 100 m/s , y transmite la información a otras neuronas a través de un proceso químico denominado sinapsis. La sinapsis puede ser excitadora si produce una disminución de la permeabilidad de la membrana postsináptica o inhibidora, si aumenta la permeabilidad de la membrana postsináptica. La suma de las sinapsis excitadoras e inhibidoras que ingresan por las dendritas de una neurona establece si la neurona conduce o no conduce [62].

Generación de las señales EEG La corteza cerebral está conformada principalmente por neuronas piramidales [1], y las señales EEG son producidas por la sincronización local de miles de neuronas piramidales [39]. Esta sincronización se presenta por la sumatoria de sinapsis excitadoras en las dendritas de miles de neuronas piramidales. Esto genera en cada neurona un dipolo conformado entre la carga del soma y la carga de las dendritas (Fig. 2.18). De esta manera, la corriente en el cerebro está conformada por el bombeo de iones Na^+ , K^+ , Ca^{++} y Cl^- a través de la membrana de las neuronas [62].

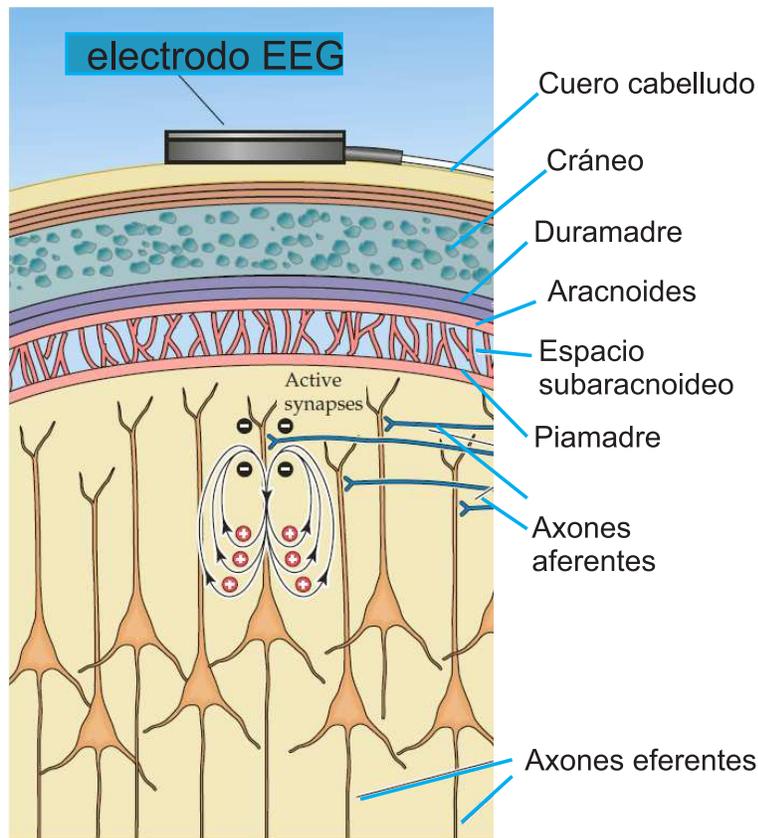


Figura 2.18: Generación de una señal EEG. Adaptado de [39].

En consecuencia, las señales electroencefalográficas (EEG) consisten en un conjunto de ondas eléctricas generadas por la corteza cerebral (Fig. 2.18) [39]. Cuando se mide con electrodos superficiales, el rango de voltajes está entre 0 a $300 \mu V$, con frecuencias que oscilan entre 0.5 y $100 Hz$. Dentro de las ventajas que tienen este tipo de señales, están la buena resolución temporal (ms), el hecho de que las señales están listas para ser capturadas, que la grabación es sencilla y el costo de los equipos es bajo. Una desventaja de las señales EEG es su baja resolución espacial (cm) [62].

Ritmos cerebrales Las señales cerebrales se agrupan en cinco ritmos: delta (δ), teta (θ), alfa (α), beta (β) y gama (γ) [62].

Los ritmos delta (δ) son ondas cerebrales en el rango de $0.5-4 Hz$. Se presentan con más frecuencia en niños y adolescentes. De igual forma, están asociadas a estados de sueño y se atenúan con abrir los ojos y estados de somnolencia [63].

Los ritmos teta (θ) son ondas cerebrales en el rango de $4-8 Hz$. Estas ondas se presentan con más frecuencia en niños y tienden a disminuir con la adultez. En los adultos tienen mayor actividad en F_z y C_z . Los ritmos (θ) se han asociado a estados de meditación profunda [62].

Los ritmos alfa (α) son ondas cerebrales en el rango de $8-13 Hz$. Se presentan en todas las etapas de la vida y tienden a tener la misma intensidad. Se presentan principalmente cuando la persona está despierta, en reposo y con los ojos cerrados. Los ritmos (α) presentan la mayor amplitud en los lóbulos occipitales y desaparecen cuando la persona está dormida [63].

Los ritmos beta (β) son ondas cerebrales en el rango de 13-30 Hz . La presencia de estos ritmos en personas adultas es normal y se asocia con estados de pensamiento o procesos de atención. Se presenta con mayor intensidad en el lóbulo frontal y el lóbulo parietal. Los niveles de ritmos beta (β) pueden aumentar en periodos de pánico.

Los ritmos gama (γ) son ondas cerebrales por encima de 30 Hz . Se presentan con menor intensidad en comparación con los otros ritmos cerebrales. Estas ondas se han asociado con procesos de imaginación motora [62].

Convención para la colocación de los electrodos La Federación Internacional de Sociedades para la Electroencefalografía y Neurofisiología Clínica ha recomendado el sistema 10-20 para la colocación de los electrodos EEG. De igual manera, en el campo de las BCIs se toma como referencia el sistema 10-20 para la colocación de los 21 electrodos (Fig. 2.19) [62].

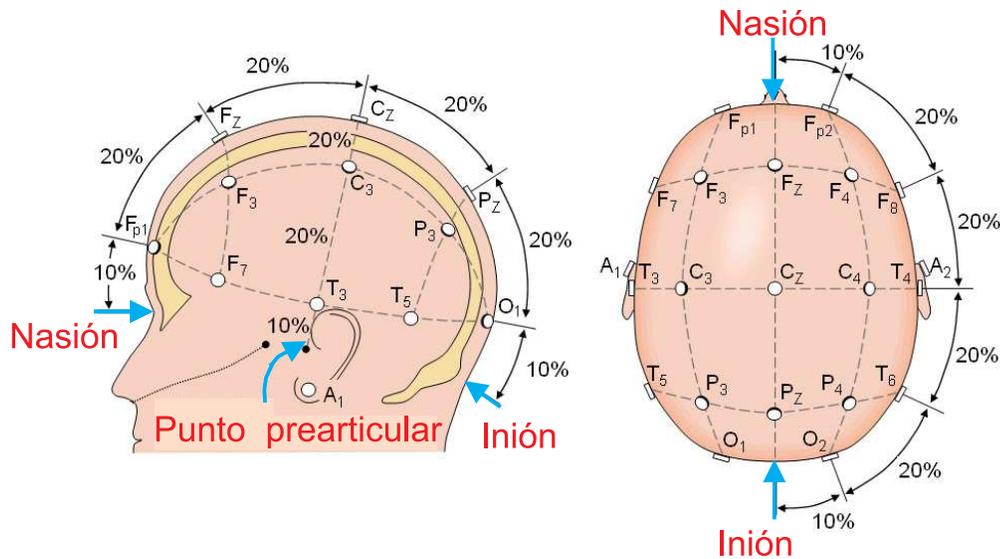


Figura 2.19: Colocación de los electrodos EEG en el sistema 10-20. Adaptado de [26].

El sistema 10-20 toma como referencia cuatro puntos: el nasión, el inión y los puntos prearticulares. El nasión es una depresión que se ubica entre la nariz y la frente. El inión es una protuberancia que se encuentra en la parte posterior de la cabeza entre cuello y cabeza. Los puntos prearticulares se ubican en la oreja izquierda y derecha, específicamente en una depresión arriba del cartílago que cubre las orejas (Fig. 2.19) [64].

Los electrodos se nombran de acuerdo con la ubicación craneal. El valor F con el hueso frontal, el valor T con el hueso temporal, el valor P con el hueso parietal y el valor O con el hueso occipital. El subíndice z especifica el plano sagital o línea media que divide el hemisferio izquierdo y derecho; los subíndices impares relacionan el hemisferio izquierdo y los subíndices pares relacionan el hemisferio derecho [62].

Después de ubicar los puntos de referencia se procede a ubicar los electrodos. En primer lugar, se toma como 100% la longitud sobre la superficie de la cabeza entre el nasión y el inión. Con relación al nasión, F_z está al 30%, C_z está al 50% y P_z está al 70%. De igual manera, se toma la longitud sobre la superficie de la cabeza entre los dos puntos articulares, y ese valor ahora corresponde al 100%. Ahora, con relación al punto articular de la izquierda,

T_3 está al 10 %, C_3 está al 30 %, C_z está al 50 %, C_4 está al 70 % y T_4 está al 90 %. De esta forma se continúa la colocación del resto de los electrodos (Fig. 2.19) [64].

Existen dos electrodos adicionales llamados $A1$ y $A2$ que se ubican en el lóbulo de la oreja izquierda y derecha, y un electrodo denominado GND que tiende a ubicarse sobre el cráneo.

2.2.3. BCIs invasivos

La electrooculografía (ECoG) consiste en la colocación de electrodos directamente sobre la corteza cerebral [65]. Estos estudios se han realizado en animales y personas con algún tipo de patología como la epilepsia. Dentro de las ventajas que tiene este método está el poder grabar información de la corteza cerebral con buena resolución espacial, amplio rango de frecuencias cerebrales y una mejor relación señal a ruido comparada con las señales EEG [66]. Dentro de las desventajas se encuentra: el requerir de una cirugía para la colocación de los electrodos, el posible rechazo del sistema inmunológico, problemas éticos y financieros, entre otros.

Existen varios trabajos relevantes que utilizan ECoG. El trabajo desarrollado por Ryu et al., [67] se centró en la utilización de un teclado por medio de un sistema BCI, implantando una matriz de electrodos sobre la corteza premotora de un mono, logrando operar un teclado a una velocidad de 6.5 *bits* por segundo. Serruya et al., [68] implantaron una matriz de electrodos sobre la corteza motora de monos, tomando un conjunto determinado de neuronas de la corteza cerebral y decodificando las señales para mover el cursor de un computador en un ángulo visual de 14 grados.

Chapin et al., [69] entrenaron varios ratones para mover una palanca con el fin de obtener alimento. A estos ratones, se les implantaron algunos electrodos para controlar un robot. Por medio de este dispositivo se consiguió que los ratones logaran controlar el robot y así obtener su alimento. Fetz [65] también logró que ratones movieran un robot utilizando señales del área cortical y del tálamo. Taylor et al., [66] implantaron electrodos sobre la corteza motora de monos logrando un movimiento tridimensional de una prótesis al usar algoritmos de control que detectan cambios en el movimiento con retroalimentación visual. Weesberg et al., [70] implantaron electrodos sobre la corteza cerebral de algunos primates en la corteza motora primaria, la corteza premotora y el área parietal. Con estas señales, los primates lograron manejar un robot con trayectorias tridimensionales utilizando algoritmos lineales y no lineales de clasificación.

Georgopoulos et al., [71] implantaron electrodos sobre la corteza motora de primates, en especial en el área del brazo; encontraron que la dirección para controlar el movimiento del brazo en los primates se puede predecir por la acción de una población de neuronas corticales motoras. La sumatoria de los pesos de un grupo determinado de neuronas permite predecir la dirección del movimiento asimilando esta información como un vector que indica la dirección. Georgopoulos et al., [72] implantaron electrodos sobre el área del brazo en la corteza motora de primates, y utilizaron los vectores de análisis para predecir si alguno de los dedos o la muñeca están activados. Con este método no se logró caracterizar un movimiento particular para los dedos. Poliakov y Schieber [73] estudiaron la relación del área de la mano en la corteza motora durante el movimiento individual de los dedos en primates. Desarrollaron grupos funcionales de neuronas con criterios de vecindad y aplicaron análisis de cluster para encontrar relaciones de vecindad, hallando que la corteza motora utiliza diferentes regiones para el control de un sólo dedo.

2.2.4. Estrategias mentales no invasivas

Existen dos estrategias no invasivas utilizadas en el campo de BCIs. La primera se denomina atención selectiva y la segunda se fundamenta en la imaginación motora.

Atención selectiva

Desde la neurociencia la atención es una actividad del cerebro que consiste en procesar en forma diferencial fuentes simultáneas de información [61]. Desde la ciencia cognitiva la atención se define como la capacidad de filtrar y seleccionar uno o varios estímulos [53]. En esta misma corriente, la atención se considera un proceso cognitivo donde diferentes zonas cerebrales interactúan para determinar un foco de atención durante un lapso de tiempo, con la capacidad de generar nuevos focos de atención, de acuerdo a los requerimientos del individuo [74]; también se considera un proceso cognitivo que comienza cuando el sujeto se encuentra ante diferentes estímulos del ambiente y genera una clasificación de los estímulos de acuerdo con su prioridad o necesidad, para posteriormente realizar un proceso de toma de decisiones [75].

La atención se puede clasificar en atención focalizada, atención dividida y atención sostenida. En la atención focalizada el cerebro selecciona un estímulo que a su parecer tiene la mayor importancia. En el caso de la visión, el cerebro tiene un área visual en donde la información que proviene del entorno es priorizada y genera un máximo interés. Para el cerebro, este tipo de atención es primordial porque le permite atender la gran complejidad de estímulos presentes en el ambiente [53].

La propuesta de atención dividida entra en antagonismo con la propuesta de atención focalizada. En ella se propone que el cerebro es capaz de procesar varios estímulos a la vez y el proceso cognitivo de atención puede ser sobrellevado por el sujeto, siendo procesado de forma no consciente por el individuo. Esta teoría se basa en la arquitectura paralela del cerebro [53].

En la atención sostenida se presentan la conciencia y el intervalo de tiempo como variables características. En este tipo de atención el cerebro es consciente de una tarea u objetivo a realizar durante un intervalo de tiempo, y centra todo su esfuerzo en atender los estímulos que lleven a desarrollar la tarea [53].

Existen diferentes zonas cerebrales relacionadas con el proceso de atención dependiendo en gran medida de si el estímulo es visual, auditivo o somatosensitivo, entre otros. Sin embargo, independiente del estímulo, la corteza cerebral del área parietal siempre está relacionada con los procesos de atención [61].

El proceso cognitivo de atención es utilizado para BCIs y el tipo de estímulo puede ser auditivo, visual o somatosensorial. Cada uno de estos estímulos puede ser utilizado como un comando para controlar un dispositivo. El sujeto requiere focalizar su atención en un estímulo determinado. El potencial P300 y el potencial evocado visual de estado estable (*steady-state visual evoked potentials* - SSVEP) son dos patrones cerebrales que produce el cerebro relacionados con los procesos de atención.

P300 Los potenciales de eventos relacionados (*event-related potentials* - ERP) son señales EEG que se producen en la corteza cerebral en respuesta a eventos cognitivos, afectivos o sensoriales [26]. Estas señales son difíciles de detectar y se caracterizan por la amplitud de la onda, por el tiempo en el que se produce el máximo de la onda (latencia) y por la

distribución de esta señal en la corteza cerebral. Los ERP se clasifican como positivos o negativos de acuerdo con la latencia máxima o mínima que generen. Si la latencia es positiva se especifica con la letra P, y si es negativa se especifica con la letra N. En ambos casos, después de la respectiva letra se especifica el tiempo de latencia con el valor en milisegundos. Los ERP son señales dinámicas y varían de acuerdo con el tipo de estímulo [26].

Los ERP relacionados con procesos de atención de eventos inesperados poseen dos componentes: la componente negativa (*mismatch negative* - MMN) y la componente P300. El MMN consiste en una respuesta del cerebro que se produce durante los primeros 10 *ms* después que este detecta que ha habido un cambio en el entorno (eventos homogéneos); esta señal se genera en el cerebro cerca de la corteza auditiva primaria. La componente P300 relaciona procesos cognitivos tales como la focalización de la atención, la actualización del contexto, la modulación de respuesta y la resolución de respuesta [26].

Las señales cerebrales P300 pertenecen al conjunto de señales ERP que se generan ante la atención a estímulos de eventos inesperados. Se caracterizan por su amplitud, latencia y distribución; en la corteza cerebral tienen un valor máximo (pico) que se produce aproximadamente 300 milisegundos después de haberse generado un proceso de atención cognitiva (evento inesperado). Las señales cerebrales P300 tienen diferentes características dependiendo del estímulo. Una señal P300 generada por un estímulo visual es diferente a una señal P300 generada por un estímulo auditivo [26].

La señal P300 posee dos componentes principales, la P3a y la P3b. La componente P3a está relacionada con la orientación del cerebro al estímulo novedoso y se localiza en la región frontal y central de la corteza cerebral. La componente P3b de la señal P300 se detecta en la región temporal y parietal de la corteza cerebral. Esta componente tiene una mayor amplitud en P_z .

Una de las ventajas que posee este método aplicado a BCIs es el no requerir de un entrenamiento dispendioso y el hecho de que en pocos minutos ya puede utilizarse. Dentro de las limitaciones que presenta el P300 se encuentra el requerir de un estímulo externo (fuentes luminosas, auditivas o hápticas), y la fatiga que produce el poner atención a una fuente luminosa, auditiva o háptica por largo tiempo.

Una de las primeras aplicaciones de P300 a BCIs fue el desarrollo de un *speller* para el deletreo de caracteres. En este caso, el estímulo que se le presenta al sujeto es información visual a través de una pantalla. En ella se muestra una matriz de 6x6 que contiene las letras del alfabeto, los números del 1 al 9 y un símbolo de espacio que sirve de separación entre las posibles palabras de una oración. Cuando se quiere que una letra resalte en la pantalla, cambia su luminancia de gris a blanco durante un intervalo de tiempo. El software de la matriz P300 inicia en el instante que lo decide el experimentador y finaliza cuando el sujeto determina que ha completado la tarea cerebral. En forma aleatoria se determina qué carácter es el que se ilumina; en su base de datos almacena la fila y columna de ese carácter y el tiempo en que se generó el evento respectivo. Esta información sirve para determinar en forma sincrónica con la captura de la señal EEG, el instante que el sujeto genera la señal P300 [76].

Kindermans et al., innovaron una nueva metodología basada en un enfoque probabilístico para la utilización de *spellers* con P300. Este enfoque tiene en cuenta, principalmente, la distribución gaussiana de los potenciales de eventos relacionados y la distribución probabilística del lenguaje a ser utilizado en el *speller* [77]. Speier et al., utilizaron un modelo oculto de Markov (HMM) y el algoritmo de Viterbi para aumentar el desempeño de *spellers* con potenciales P300. En este caso, cada uno de los caracteres del P300 es modelado como un estado

oculto y la señal EEG es la salida visible. Con esta información se utilizó el algoritmo de Viterbi para incorporar la información del lenguaje [78]. Zeyl et al., mejoraron el desempeño de un *speller* que utiliza señales P300. La mejora del desempeño está fundamentada en combinar la confianza de un *speller* con potenciales P300 y en el cálculo del potencial de error relacionado (ErrP). La confianza se mejoró utilizando un estimador bayesiano y el ErrP se basó en árboles de decisión [79].

Thurling et al., desarrollaron un nuevo método de BCI para el control de juegos por computador. En este caso utilizaron potenciales de eventos relacionados con estímulos táctiles que generan potenciales P300. La ventaja de este método está en el incremento del número de bits por segundo transmitidos por el respectivo BCI [80]. Spuler et al., aplicaron análisis de correlación canónica para el desarrollo de filtros espaciales; estos fueron aplicados para mejorar el desempeño en la separación de potenciales P300 [81].

En el área de las interfaces cerebro computador, utilizando potenciales evocados P300 (EP), Rebsamen et al., [14] desarrollaron un prototipo de silla de ruedas habilitado para navegar en un ambiente como una oficina típica o un hospital. El controlador para silla de ruedas (BCW) está basado en el sistema P300 que, para este, caso es lento, pero tiene la ventaja de ser bastante seguro. Li et al., desarrollaron un sistema híbrido entre potenciales P300 y SSVEP para el control de una silla de ruedas. La utilización simultánea de estas dos señales cerebrales permite desarrollar BCIs con mayor transferencia de *bits* comparado con cada uno de los métodos por separado [20]. Zhang et al., desarrollaron un control con BCI para que una persona pueda navegar utilizando silla de ruedas. En este caso, la persona puede seleccionar de un conjunto previo de posibilidades y por medio de imaginación motora o P300 el destino a donde quiere ir. Posteriormente, un sistema automático de navegación la conduce al lugar seleccionado [21].

Long et al., desarrollaron una BCI para controlar un cursor en forma horizontal y vertical. La metodología que utilizaron fue una combinación entre potenciales P300 e imaginación motora [82]. Xu et al., utilizaron una variación del *speller* clásico basado en el cambio de tamaño de la matriz que ve el sujeto y un modelo probabilístico. La matriz utilizada en los *speller* es de 6x6 y se modificó por dos matrices con tamaño de 3x3. La transición de una matriz a otra está dada por funciones de probabilidad propias de frecuencia en donde los caracteres aparecen en un lenguaje [83].

SSVEP El potencial evocado visual de estado estable (*steady-state visual evoked potential* - SSVEP) es una respuesta del cerebro ante la atención a un estímulo visual [25]. Los estímulos visuales corresponden a diferentes fuentes estables (lumínicas) con frecuencias fundamentales comprendidas entre 6 y 30 *Hz*. El método SSVEP aplicado a BCI requiere que el sujeto fije su mirada (atención) en una fuente de luminancia. De acuerdo con la frecuencia de la fuente luminosa que el sujeto esté observando, se produce en los lóbulos occipitales una frecuencia fundamental de la misma frecuencia de la fuente que está observando, y una serie de armónicos. De esta manera se puede relacionar el poner atención a una fuente luminosa, con el control de un dispositivo o máquina [24].

Una de las ventajas que posee este método aplicado a BCIs es que no requiere de un entrenamiento dispendioso y en pocos minutos puede aprender a utilizar el método SSVEP. Dentro de las limitaciones que presenta se encuentra el requerimiento de un estímulo externo (fuentes luminosas), la cantidad limitada de fuentes luminosas relacionado al rango de frecuencias que el cerebro puede detectar y la fatiga que produce mirar una fuente luminosa que parpadea por largo tiempo.

En el área de las interfaces cerebro computador con potenciales evocados visuales de estado sostenido (SSVEP), se presentan trabajos como los de Mandel et al. [15], quienes desarrollaron una interfaz cerebro computador (BCI) para que un sujeto pueda manejar una silla de ruedas por medio de 4 señales de potenciales evocados visuales. Pfurtscheller [16] desarrolló un sistema basado en potenciales evocados visuales para el control de una prótesis de mano que posee dos grados de libertad y tiene la capacidad de clasificar cuatro patrones cerebrales. Torres et al., [15] integraron una interface cerebro-computador (BCI) basada SSVEP para controlar una silla de ruedas. Lee et al., controlaron un carro pequeño utilizando señales SSVEP; para las direcciones (izquierda, derecha y adelante) utilizaron tres fuentes luminosas que oscilaban a 13 Hz , 14 Hz y 15 Hz . La señal fue capturada del electrodo O_z [84]. Chumerin et al., emplearon un sistema SSVEP aplicado a la navegación dentro de un juego. En este caso, con los estímulos visuales el sujeto podía realizar cuatro tareas de navegación: adelante, atrás, izquierda y derecha [23].

Shyu et al., innovaron en la aplicación de los sistemas SSVEP variando el ancho del pulso y la frecuencia para mejorar el desempeño. En ese trabajo encontraron que las mejores frecuencias para el estímulo visual estaban entre 24 a 36 Hz [85]. Han et al., presentaron una nueva aproximación al SSVEP reconociendo diferentes patrones en los estímulos visuales, donde los patrones visuales tienen diferentes configuraciones basados en modelos de tablero de damas [86].

Duan et al., desarrollaron un BCI basado en SSVEP e imaginación motora para controlar un robot. Con tres estímulos visuales podían controlar avance, girar a la izquierda y girar a la derecha; Con la señal de imaginación motora el robot agarraba el objeto [18]. Yin et al., desarrollaron un sistema híbrido entre potenciales P300 y SSVEP aplicado a *spellers*. La combinación de P300 con SSVEP permite incrementar el número de bits por segundo, comparado con cada uno de los métodos aplicado a *spellers* [19].

Imaginación motora

La imaginación es un proceso cognitivo en donde una persona transforma una información o un conjunto de observaciones desarrolladas intrínsecamente, generando una representación interna (mental) de esa información [27]. En el caso de la imaginación motora, esta consiste en hacer una simulación interna de un movimiento motor sin generar el movimiento de alguna parte del cuerpo. Existe relación entre el tiempo que gasta una persona realizando una actividad motora y su respectiva imaginación. Se han desarrollado varias teorías sobre cómo el cerebro realiza el proceso de imaginación motora, una de estas supone que el cerebro crea una plantilla donde modela los movimientos que tiene que desarrollar, sin requerir la activación de la corteza motora primaria [22].

En los procesos de imaginación motora o movimiento motor se activa la corteza motora (circunvolución precentral) y/o la corteza somatosensorial (circunvolución postcentral). En procesos de imaginación motora se ha encontrado que las ondas (μ) están dentro del rango de las ondas α y las ondas (γ). Estas ondas cerebrales están presentes durante este tipo de tareas cognitivas [24]. A los procesos de imaginación motora cuando las frecuencias cerebrales disminuyen en una determinada banda, se les denomina eventos relacionados con la desincronización (*event-related desynchronization* - ERD). Al mismo tipo de tareas cognitivas en que las frecuencias cerebrales aumentan en una determinada banda, se les denomina eventos relacionados con la sincronización (*event-related synchronization* - ERS). Un evento ERD/ERS se produce cuando en razón a un proceso de imaginación motora, una zona del cerebro produce

un ERD y otra zona del cerebro produce un ERD [25].

La corteza motora y la corteza somatosensorial tienen una organización topológica donde el hemisferio izquierdo controla el lado derecho del cuerpo y el hemisferio derecho controla el lado izquierdo. La representación de las piernas se encuentra cerca de C_z , la mano está cerca de C_3 y C_4 , y la boca cerca de T_3 y T_4 . Las señales ERD/ERS están relacionadas con el tipo de imaginación que se utilice. La imaginación o movimiento de los pies produce una variación de potencial en C_z . El movimiento o imaginación de la mano izquierda o derecha produce variación de los potenciales en C_3 para la mano derecha y C_4 para la mano izquierda. De esta manera, es difícil distinguir con señales EEG entre la imaginación o movimiento de los dedos, o entre la imaginación o movimiento del pie derecho e izquierdo [25].

Una de las ventajas que posee la imaginación motora aplicada a BCIs es que no requiere de un estímulo externo. Dentro de las limitaciones que presenta se encuentra el requerir de una fase de entrenamiento donde el tiempo de entrenamiento está en función de la cantidad y complejidad de las tareas. Otra limitación consiste en la fatiga cognitiva que se incrementa con el número de tareas y el tiempo respectivo.

Existen diferentes investigaciones que utilizan la habilidad cognitiva de imaginación mental para su utilización en interfaces cerebro computador: Matsunaga et al. [87], por ejemplo, encontraron que existe una relación entre la imaginación de los movimientos de los miembros superiores o inferiores y las señales μ y β . Ramos et al., [88] contrastaron las señales EEG con las señales EMG para el control de una prótesis de mano de Otto Bock, siendo la estrategia de control tareas imaginarias motoras (señales μ). Millán et al. [89] presentaron una metodología con imaginación motora para controlar un robot móvil a partir de señales EEG, utilizando tres electrodos de carácter no invasivo. Polak y Kostov [90] reportan un trabajo con imaginación motora para el control de una articulación artificial de muñeca. Erfanian y Gerivany [91] describen el desarrollo de un sistema computacional para predecir el movimiento voluntario (agarrar, sostener y soltar) de una mano, o las dos al mismo tiempo a través de señales EEG. Chen, Ju y Lin desarrollaron una interfaz cerebro computador que percibe el movimiento voluntario del dedo pulgar [92], [93].

Erfanian y Gerivany [91] desarrollaron un sistema para detectar tres movimientos de la mano por medio de EEG (agarrar, relajarse y sostener). En este caso utilizaron tres electrodos (Fz, C3 y F3), un filtro adaptativo, redes neuronales, funciones de base radial y sistemas estocásticos. Carrant et al. [94] utilizaron 4 electrodos EEG buscando contrastar cuatro tareas cognitivas e imaginando los siguientes ambientes: navegación espacial alrededor de un ambiente familiar, escuchar una melodía de familiar, imaginación de la mano izquierda y derecha abriendo y cerrando una mano. Se utilizó modelamiento de autorregresión y clasificadores con regresión logística.

Jia et al., [95] desarrollaron un sistema para controlar con EEG el desplazamiento de un *slider* en forma continua, y para ello se utilizó el fenómeno de desincronización de eventos relacionados (ERD), que se produce en procesos de imaginación motora relacionados con el movimiento rápido de los dedos. Los canales utilizados fueron C_3 y C_4 . Akrami et al., [96] desarrollaron un método para diferenciar la imaginación del movimiento de la mano izquierda y derecha utilizando electrodos en C_3 y C_4 ; se utilizaron clasificadores lineales y no lineales.

Wang et al., [97] utilizaron un algoritmo basado en aproximaciones de entropía para detectar tareas cognitivas como rotación, multiplicación, conteo, deletreo y la imaginación del movimiento de la mano izquierda o derecha. El resultado se tomó utilizando 64 electrodos, encontrándose relevancia significativa con el movimiento imaginativo de la mano izquierda o derecha en términos de la atenuación de señales en el hemisferio opuesto, denominado sincro-

nización de eventos relacionados (ERS). Morash et al. [98] colocaron 24 electrodos EEG sobre la corteza sensoromotora con el objetivo de clasificar diferentes tareas de imaginación motora utilizando la mano izquierda, la mano derecha, la lengua y el pie derecho. El procedimiento consistía en primero mover la parte respectiva del cuerpo, luego descansar un intervalo de tiempo y posteriormente imaginarse el movimiento respectivo. El cambio de secuencia estaba fijado por una imagen que aparecía en la pantalla de un computador. En este caso se utilizó un clasificador estocástico.

Li, Yang y Yuang [99] desarrollaron un clasificador que mejora significativamente la respuesta para señales EEG resultado de un proceso de imaginación mental. Se utilizó una combinación de un modelo autorregresivo (AR) y una máquina de soporte vectorial (SVM). Huang et al. [15] desarrollaron un método basado en BCIs para controlar el movimiento de un cursor con movimiento en un espacio bidimensional. El sistema consistía en imaginar el desplazamiento de la mano izquierda o de la mano derecha para controlar el movimiento del cursor en la pantalla. El método novedoso consistió en utilizar en forma combinada la desincronización de eventos relacionados (ERD) y al mismo tiempo la sincronización de eventos relacionados (ERS) observados en el hemisferio opuesto a la imaginación del respectivo miembro. Las señales fueron tomadas de cuatro electrodos EEG, en las posiciones FC_5 , T_7 , T_8 y FC_6 .

Severens et al., desarrollaron un sistema con señales EEG para determinar si una persona está caminando o imaginando caminar, o si está quieta. Para este estudio se utilizaron 62 electrodos y se realizaron las tareas de imaginar la caminata y caminar. Los resultados muestran la posibilidad de aplicar esto en procesos de rehabilitación motora para miembros inferiores [17]. Corralejo et al., desarrollaron un clasificador para mejorar el desempeño de las tareas de imaginación motora. El algoritmo utiliza análisis discriminante lineal, junto con información espectral, temporal y espacial. En este estudio se clasificó la imaginación de la mano derecha, la mano izquierda, los pies y la lengua [100]. Aghaei et al., proponen el desarrollo de un nuevo algoritmo de clasificación denominado patrones espacio espectrales comunes (SCSSP); este método es una generalización del algoritmo binario de clasificación de patrones espaciales comunes (CSP). El algoritmo reporta eficiencia computacional en el proceso de clasificación [101]. Baali et al., presentan un nuevo algoritmo para clasificar procesos de imaginación motora basados en una transformación ortogonal denominada descomposición de valores singulares basado en predicción lineal (LP-SVD). Para probar el algoritmo se utilizó la base de datos de la III competencia internacional de BCIs (2005) [102].

2.3. El habla silenciosa (*silent speech*) (SS)

Una primera definición afirma que el habla silenciosa es una área reciente de investigación, que corresponde a sistemas que habilitan la comunicación del habla cuando una señal acústica audible no está habilitada [103]. Otra, especifica que el habla silenciosa corresponde a la producción natural del habla sin la producción de sonidos [104]. Y una más, la define como la acción de imaginar la vocalización mientras el mismo sujeto permanece en silencio [30]. Las señales que se pueden trabajar para el habla silenciosa están relacionadas con el aparato articulador de la voz, las redes neuronales (nervios) o las señales cerebrales de la corteza motora [103]. De esta forma, el habla silenciosa tiene la factibilidad de aplicarse en diferentes campos como: personas que no pueden producir la voz (mudas), discapacitados que hayan sufrido de una laringectomía, procesos de comunicación de personas sanas en ambientes con alto ruido ambiental y comunicación entre personas con dispositivos celulares sin producir la

voz, entre otras [103].

En este capítulo se presentan diferentes metodologías de habla silenciosa, dentro de las cuales se han explorado distintas estrategias en función de la forma del aparato articulador o aparato fonador, de las señales nerviosas relacionadas a la producción del habla y de las señales cerebrales relacionadas con la corteza cerebral. Entre dichas estrategias se encuentran las siguientes: captura del movimiento de puntos fijos en el aparato articulador usando sensores de articulografía electromagnética (EMA); caracterización del tracto vocal usando ultrasonido (US) e imágenes ópticas de la lengua y labios; electromiografía superficial (sEMG) basada en reconocimiento del habla; habla silenciosa con BCIs utilizando electrocorticografía (ECoG); y habla silenciosa con BCIs empleando electroencefalografía (EEG).

2.3.1. Captura del movimiento de puntos fijos en el aparato articulador usando sensores de articulografía electromagnética (EMA)

El método de captura del movimiento de puntos fijos en el aparato articulador usando sensores de articulografía electromagnética, busca identificar el habla silenciosa producida por una persona, sin generar ningún sonido, cuando realiza el movimiento del aparato articulador. Este método de habla silenciosa requiere la utilización de bobinas o imanes, y de sensores de campo magnético [105]. Los primeros sistemas generan los campos magnéticos a través de bobinas; en trabajos más recientes, los campos magnéticos son generados por varios imanes. Esos imanes o bobinas se ubican en posiciones cartesianas con la condición que su posición varíe con respecto al aparato articulador [106]. La ubicación de estos imanes o bobinas tiende a estar en la lengua, los labios y los dientes. Los sensores magnéticos se ubican sobre una estructura (gafas) y su condición primaria es servir como sistema de referencia para las bobinas o imanes.

Para poder decodificar el habla silenciosa, en un estudio se utilizaron seis sensores magnéticos duales que buscaban detectar las variaciones de campo magnético generadas por imanes ubicados en labios, lengua y dientes [107]. En este trabajo se buscó determinar si era posible decodificar el habla silenciosa con la metodología EMA. La experimentación se desarrolló utilizando trece fonemas y nueve palabras; cada una de ellas se repitió diez veces [107]. Para la clasificación de los respectivos fonemas y palabras se utilizaron algoritmos de programación dinámica, obteniendo así una precisión superior al 90%. Este trabajo puede ser aplicado en personas a las que no les sea posible generar los sonidos del habla.

2.3.2. Caracterización del tracto vocal usando ultrasonido (US) e imágenes ópticas de la lengua y los labios

Otra forma de habla silenciosa es la caracterización del tracto vocal usando señales de ultrasonido e imágenes ópticas de la lengua y los labios, con el objetivo de identificar el habla sin la producción de sonidos [103]. Este sistema utiliza como entradas de información una cámara ultrasónica colocada en el plano sagital, debajo de la barbilla, y una cámara para luz visible que se ubica en frente de los labios del sujeto. El objetivo de la cámara ultrasónica es detectar los movimientos de la lengua; el de la cámara de luz visible, detectar el movimiento de los labios. Este método puede ser aplicado a personas que tengan patologías del habla o a aquellas que hayan sufrido de una laringectomía.

Un estudio desarrollado con estos dos tipos de cámaras fue aplicado para decodificar el habla sin la producción de sonidos. En este caso se utilizó un sintetizador como etapa de salida del decodificador [108]. Otro estudio relacionado, se ha centrado en decodificar estas señales utilizando frecuencias espectrales, e implementado redes neuronales artificiales durante la etapa de clasificación (*Backpropagation*) [109]. Otra alternativa para decodificar el habla silenciosa utilizando este tipo de cámaras se basa en desarrollar un método global para detectar el contorno de la lengua, denominado descomposición *EigenTongues*, junto a la detección de la forma de músculos relacionados a la lengua, el hueso hioides y el tendón corto [110]. Otro estudio que se vale de esta metodología, implementó un sistema de habla silenciosa basado en imágenes de la lengua y los labios, utilizando como principio el reconocimiento visual de fonemas; en este caso, se alcanzó una precisión del 60 % con un clasificador basado en modelos ocultos de Markov (HMM) [111].

2.3.3. Electromiografía superficial (sEMG) basada en reconocimiento del habla

Las señales electromiográficas (EMG) son señales producidas por las neuronas motoras para activar los músculos del cuerpo [112]. La célula fundamental de los músculos se llama fibra muscular y cada fibra muscular contiene miles de miofibrillas en donde se encuentra la unidad contráctil del músculo [61]. Un músculo recibe varias fibras nerviosas motoras. La unión entre una sola neurona motora alfa y las fibras musculares que inerva se llama unidad motora. Existe una variedad en la cantidad de fibras a inervar, desde 10 fibras musculares en el globo ocular (movimientos finos) hasta 200 fibras musculares en los músculos de las extremidades (movimientos bruscos). Las neuronas motoras o eferentes tienen su origen en la corteza cerebral motora [39], y se interconectan con otras neuronas motoras del tronco cerebral como el tálamo y el bulbo raquídeo. Estas señales viajan por las astas anteriores de la médula espinal terminando en un actuador como los músculos [61].

Las señales mioeléctricas generadas en los músculos son de naturaleza aleatoria y sus valores de voltaje se encuentran entre 0 y 6 mV . Las frecuencias generadas por una señal mioeléctrica se encuentran en el rango de 0 a 500 Hz y con mayor intensidad entre 50 y 150 Hz [113].

Existen dos formas de capturar las señales EEG: a través de electrodos tipo aguja, que tienen un carácter invasivo, y electrodos superficiales (sEMG), que tienen la ventaja de no ser invasivos y se pueden colocar sobre el músculo del que se requiere capturar la bioseñal [113]. En el caso de reconocimiento del habla con sEMG se busca identificar el habla silenciosa o el habla pronunciada a través de electrodos colocados sobre los músculos relacionados con la laringe [103]. La aplicabilidad de este sistema involucra a personas que tienen patologías relacionadas con la producción de la voz o personas sanas que requieran comunicarse en ambientes con alto nivel de ruido.

En un estudio de reconocimiento del habla con sEMG se utilizaron tres electrodos ubicados sobre la laringe, logrando la clasificación de vocales japonesas [114]. En otro estudio, se logró clasificar diez palabras habladas utilizando información de algunos músculos de la cabeza y el cuello [115]. De igual manera, se han desarrollado pruebas de esta metodología en ambientes con niveles de ruido de 95 dB , alcanzando precisiones del 74 % para 15 palabras. En un estudio más, se logró clasificar seis palabras no audibles con sEMG, utilizando electrodos sobre el área de la laringe [116].

El reconocimiento del habla con sEMG presenta algunas dificultades; entre ellas se encuentran: limitaciones para el reconocimiento del habla fluida o continua [117]; limitaciones para el desarrollo con esta metodología de diccionarios de reconocimiento basados en unidades menores que las palabras [118]; y desarrollo de sistemas de reconocimiento del habla basados en sEMG independientes del sujeto, entre otras.

2.3.4. Habla silenciosa con BCIs utilizando electrocorticografía

Las BCIs invasivas (ECoG) consisten en la captura de señales cerebrales ubicando electrodos directamente sobre la corteza cerebral. Este método requiere de una cirugía para la colocación de los electrodos donde el sujeto corre el riesgo de sufrir una infección o tener una respuesta de rechazo por parte del sistema inmunológico. Las BCIs invasivas con habla silenciosa han tendido a utilizar ritmos gama para los procesos de reconocimiento. En el caso de reconocimiento de vocales y consonantes con habla silenciosa, han optado por utilizar ritmos de gama alta (70 - 170 Hz.) [119]. En el caso de reconocimiento de sílabas con habla silenciosa [120] y de palabras con habla silenciosa [121], se han utilizado igualmente ritmos de gama alta, de 70 a 170 Hz.

De esta forma, las BCIs que utilizan habla silenciosa con métodos invasivos han optado por utilizar algoritmos para los procesos de clasificación, tales como máquinas de soporte vectorial (*support vector machine* - SVM) y clasificadores Bayesianos. Para el caso de reconocimiento de vocales y consonantes con habla silenciosa, se ha utilizado el algoritmo de Naive Bayes con una precisión promedio de 60 % [119]. En el caso de reconocimiento de sílabas con habla silenciosa, se ha utilizado un clasificador Bayesiano [120] y un clasificador basado en la amplitud de las ondas gama [121].

2.3.5. Habla silenciosa con BCIs empleando electroencefalografía (EEG)

Una nueva y promisoriosa línea de investigación en BCIs es el habla silenciosa (*silent speech* - SS), que utiliza señales electroencefalográficas (EEG) [103] en los siguientes rangos de frecuencia: δ (de 0.5 a 4 Hz), θ (de 4 a 8 Hz), α (de 8 a 13 Hz), β (de 13 a 30 Hz) y γ ($>$ 30 Hz) [26]. Dentro de las características que presenta esta nueva línea de BCI están: el no ser invasiva; el no requerir de procesos de entrenamiento dispendiosos como en el caso de la imaginería motora [22]; el no requerir un proceso de atención riguroso como en SSVEP [23], P300 [14] o imaginería motora [24]; el no requerir de un estímulo externo como en el caso de SSVEP [25] o P300 [26]; el no requerir de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imaginería motora [27]. Adicionalmente, utiliza señales cerebrales que están relacionadas con el área de lenguaje y tiene la posibilidad de utilizar la innumerable cantidad de palabras (léxico) de un lenguaje. Entre las desventajas que presenta está el hecho de que la neurociencia aún trata de entender los modelos neuronales del lenguaje, ya que los mecanismos no están completamente entendidos; y que la caracterización de las señales cerebrales relacionadas con el habla silenciosa aún es un área de investigación abierta.

Hasta la fecha se han reportado un limitado número de reconocimiento de patrones, identificando algunas vocales, algunas sílabas, algunas palabras y un limitado conjunto de oraciones. Por medio del habla silenciosa se han alcanzado a reconocer: la vocal silenciosa /a/ [28]; las vocales silenciosas /a/ y /u/ [29],[122],[123]; la clase vocales silenciosas abiertas-medias, en contraste con la clase vocales silenciosas cerradas que fue clasificada previamente

por el autor de esta tesis doctoral [124]; y cinco vocales silenciosas en el idioma inglés [125] y en el idioma japonés [126],[127], [30]. De igual manera, con este mismo método se han logrado reconocer: las sílabas silenciosas /ba/ y /ku/ [128], [31], [32], y dos caracteres silenciosos chinos "zuo" y "yi" [129], [130]. Finalmente, también se han logrado reconocer: dos palabras silenciosas [33],[34]; cinco palabras silenciosas [131], [132] y siete palabras silenciosas utilizando EEG y magnetoencefalografía (MEG) [133].

Las investigaciones de habla silenciosa con métodos no invasivos, han optado por utilizar como vector de características transformaciones en el dominio del tiempo, la frecuencia, o el tiempo-frecuencia. En vocales con habla silenciosa se han utilizado como vector de características: descriptores estadísticos [125]; transformada rápida de Fourier (FFT) [28]; patrones espaciales comunes (CSP) [29],[126], [127], [30]; periodogramas [124] y coeficientes cepstrales en las frecuencias de Mel (MFCC) [123]. Adicionalmente, el vector de características para sílabas con habla silenciosa ha utilizado: CSP [129],[130]; análisis de componentes independientes (ICA) con modelos de autorregresión (AR) [31]; espectro de Hilbert [32] y envoltura generada por el espacio de Hilbert [128]. De otra manera, se han utilizado los siguientes vectores de características para palabras con habla silenciosa: descriptores estadísticos [125]; FFT [133]; modelos ocultos de Markov (HMM) [134], [131] y transformadas de wavelet discretas (DWT) [132]. Finalmente, los vectores de características para oraciones con habla silenciosa han utilizado transformada de Fourier [135] [104] y análisis de componentes principales (PCA) con ICA [136].

Dentro de las investigaciones de habla silenciosa con métodos no invasivos se ha optado por utilizar ritmos cerebrales delta, teta, alfa, beta, gama y gama alta. Para reconocimiento de vocales con habla silenciosa se ha tendido a utilizar los ritmos: teta, alfa y beta [137]; delta, teta y alfa [124]; alfa, beta y gama [123]; delta, teta, alfa, beta y gama [29], [122], [30] y delta, teta, alfa, beta, gama y gama alta [126], [127]. En el caso de reconocimiento de sílabas con habla silenciosa, se han utilizado ritmos: delta, teta y alfa [32], y teta, alfa y beta [128], [31], [129], [130]. En el caso de reconocimiento de palabras con habla silenciosa, se han utilizado ritmos: beta [34]; teta, alfa y beta [132], y delta, teta, alfa, beta y gama [133], [134], [33]. Finalmente, en el caso de reconocimiento de oraciones con habla silenciosa se han utilizado ritmos delta, teta y alfa [135], y delta, teta, alfa y beta [136].

Una última etapa en el procesamiento de señales consiste en la clasificación de los patrones a reconocer; en el caso del habla silenciosa con EEG se han tendido a utilizar como clasificadores: SVM, Naive Bayes, HMM, análisis discriminante (DA) y mapas autoorganizativos (SOM), entre otros. Para reconocimiento de vocales con habla silenciosa se han utilizado algoritmos como: SVM, con el cual se ha alcanzado una precisión entre el 68 % y el 78 % para dos vocales [29]; la combinación de los algoritmos SVM con kernel gaussiano (SVM-G), máquinas de vectores relevantes con kernel gaussiano (RVM-G) y máquinas de vectores relevantes lineales (RVM-L) que han alcanzado una precisión entre el 57 % al 87 % para cinco vocales en idioma japonés [127]; el algoritmo de regresión logístico (SLR-VAR), que ha alcanzado una precisión del 49.9 % para dos vocales [122]; el Naive Bayes, con el cual se ha obtenido una precisión de entre el 70 % al 75 % para dos vocales [137]; el HMM con el algoritmo de los k vecinos más cercanos (k-NN), con el que se ha obtenido una precisión del 75 % para dos vocales [123]; y la red neuronal artificial *Backpropagation*, con la cual se ha alcanzado una precisión del 44 % para cinco vocales en idioma inglés [125].

De otra manera, en el caso del reconocimiento de sílabas con habla silenciosa se ha utilizado el algoritmo de producto interno generado por la envoltura de filtros, con el cual se ha obtenido una precisión del 75 % para dos sílabas [128]. Por otro lado, con el algoritmo k-NN

se ha alcanzado una precisión entre el 56 % y el 88 % para dos sílabas [31]; con el clasificador Bayesiano y el análisis discriminante lineal (LDA) se ha logrado una precisión entre el 54 % y el 73 % para dos sílabas [32]; y, por último, con el algoritmo de SVM se ha alcanzado, en un caso, precisión entre el 79.35 % y el 88.26 % para dos caracteres chinos [129] y, en otro caso, precisión entre 73.65 % y 95.76 % para otros dos caracteres chinos [130].

Finalmente, para el reconocimiento de palabras con habla silenciosa se han utilizado clasificadores basados en HMM con una precisión entre el 30 % y el 70 % para cinco palabras silenciosas [134], y de 45.5 % para cinco palabras silenciosas [131]; los clasificadores SVM, *random forest* (RF) y Naive Bayes, con una precisión menor al 42 % para cinco palabras silenciosas [132], y k-NN, con una precisión entre el 27.84 % y el 91.3 % para dos palabras silenciosas [34].

2.4. Entropía de la información

Esta sección tiene como propósito presentar los fundamentos teóricos de la entropía de la información, los cuales son utilizados en el desarrollo de la metodología basada en vector de características con entropía de la información y dimensión de regularización.

Dada una fuente de información, sin memoria, que tiene un número γ de símbolos diferentes, es posible estadísticamente representarla como una variable aleatoria discreta (X). De esta manera, cada realización u observación de X se representa por x_i , donde $i = 1, \dots, \gamma$ [138].

La probabilidad de cada valor x_i se denomina $z(x_i)$, y el vector de probabilidades de x_i es $\mathbf{Z}(X) = [z_1, \dots, z_\gamma]$, donde

$$\sum_{i=1}^{\gamma} z_i = 1 \quad (2.1)$$

De esta manera, la función masa de probabilidad de X con sus siglas en inglés (PMF) es representada por:

$$PMF(X) = \begin{pmatrix} x_i \\ z_i \end{pmatrix}. \quad (2.2)$$

Por consiguiente, la información $I(\cdot)$ contenida en X está condicionada por los niveles de incertidumbre z_i [138]. Si la información $I(x_i)$ es una medida de la incertidumbre con respecto a la realización u observación x_i , entonces se establece que

$$I(x_i) = \mathbf{F}(z_i) \quad (2.3)$$

indicando que la información de x_i está en función $\mathbf{F}(\cdot)$ de la incertidumbre z_i .

Para hallar la función $\mathbf{F}(\cdot)$, debe cumplirse que la información sea cero o tenga un valor positivo $\mathbf{F}(p) \geq 0$, y que sea aditiva [138].

De esta manera, cuando x_i esté compuesto por dos eventos independientes x_{ia} y x_{ib} , entonces

$$I(x_i) = I(x_{ia}) + I(x_{ib}). \quad (2.4)$$

Aplicando la ecuación (2.3) a la ecuación (2.4), se obtiene:

$$\mathbf{F}[z(x_i)] = \mathbf{F}[z(x_{ia})] + \mathbf{F}[z(x_{ib})] \quad (2.5)$$

dado que x_{ia} y x_{ib} son independientes, entonces

$$\mathbf{F}[z(x_{ia})z(x_{ib})] = \mathbf{F}[z(x_{ia})] + \mathbf{F}[z(x_{ib})] \quad (2.6)$$

Teniendo en cuenta la ecuación (2.3), la ecuación (2.6) y que la información es de carácter positivo, entonces [138]:

$$\mathbf{F}(z_i) = -\lambda \log(z_i) = I(x_i) \quad (2.7)$$

donde λ es una constante real positiva. La ecuación (2.7) se denomina la auto-información de x_i [138] que es una medida del contenido de información asociada con un evento en un espacio de probabilidad.

Existen diferentes formas para medir la información tales como el *nit*, el *dit* y el *bit*. De esta manera, 1 *nit* = $-\ln(1/e)$ y 1 *dit* = $-\log(1/10)$ [139].

En el caso de dos eventos equiprobables x_a y x_b ,

$$I(x_a)I(x_b) = -\lambda \log\left(\frac{1}{2}\right) = 1. \quad (2.8)$$

Para medir la información, Shanon utilizó un $\lambda = 1$ y el logaritmo en base 2, estableciendo el *bit* como la unidad de medida para la información [140]. De esta manera

$$I(x_a)I(x_b) = -\log_2\left(\frac{1}{2}\right) = 1 \text{ bit}. \quad (2.9)$$

Así, la información presente en x_i es:

$$I(x_i) = -\log_2(z_i). \quad (2.10)$$

En 1948 Shannon define la entropía de la información $H(\cdot)$ como el promedio de información por símbolo emitido, midiendo la incertidumbre de una fuente de información, dado por:

$$H(X) = \mathbf{E}[I(x_i)] = \sum_{i=1}^{\gamma} z_i I(x_i). \quad (2.11)$$

donde $\mathbf{E}[\cdot]$ es el valor esperado [138].

Remplazando la ecuación (2.10) en la ecuación (2.11), el resultado es:

$$H(X) = -\sum_{i=1}^{\gamma} z_i \log_2 z_i \quad (2.12)$$

De esta manera, la entropía es una medida de la razón de generación de información. Un alto valor de entropía corresponde a un incremento en la irregularidad o en la no predictibilidad. De otra manera, un bajo valor de entropía corresponde a una alta regularidad [141].

En el caso de la teoría de la información, el nombre de entropía se utilizó por la similitud con la interpretación probabilística de Boltzmann-Gibbs relacionada con la segunda ley de la termodinámica [139]:

$$S = -k_B \sum_{i=1}^n z_i \log(z_i) \quad (2.13)$$

donde S es la entropía termodinámica, k_B es la constante de Boltzmann y z_i son las posibilidades del sistema para estar en un micro estado i de n posibles estados. Kolgomorov demostró que la entropía de un sistema termodinámico S (Eq. (2.13)) es igual a la entropía de la información $H(\cdot)$ (Eq. (2.12)) [142].

La entropía tiene las siguientes propiedades [138]:

Es una función continua con respecto a cada variable z_i , dado que es la suma de funciones continuas.

Es aditiva, dado que está en función del valor medio de la auto-similaridad.

Tiene un valor máximo cuando los eventos del espacio muestral tienen igual probabilidad.

Es simétrico, dado que si los valores de x_i se intercambian el valor de la entropía no cambia.

El máximo valor de entropía se calcula maximizando la ecuacion 2.12 con la restricción 2.1, entonces:

$$\Psi(z_i) = - \sum_{i=1}^{\gamma} z_i \log_2 z_i + \lambda \left(\sum_{i=1}^{\gamma} z_i - 1 \right) \quad (2.14)$$

donde λ son los multiplicadores de Lagrange. De esta manera, la función máxima en términos de los multiplicadores de Lagrange es [138]:

$$\max (H(X)) = \max (\Psi(z_i)) \quad (2.15)$$

Para resolver la expresión 2.15 se calcula la derivada parcial:

$$\frac{\partial \Psi(z_i)}{\partial z_i} = 0 \quad (2.16)$$

para $\forall i = 1, \dots, \gamma$. Resolviendo la expresión (2.16) para z_i y z_j se obtiene [138]:

$$\frac{\partial \Psi(z_i)}{\partial z_i} = -\log_2 z_i - \log_2 e + \lambda = 0 \quad (2.17)$$

y

$$\frac{\partial \Psi(z_j)}{\partial z_j} = -\log_2 z_j - \log_2 e + \lambda = 0 \quad (2.18)$$

En consecuencia se deduce que $\log_2 z_i = \log_2 z_j$, entonces $z_i = z_j$. Por lo tanto para $i = 1, \dots, \gamma$:

$$z_1 = z_2 = z_3 = \dots = z_\gamma = \frac{1}{\gamma} \quad (2.19)$$

De modo que,

$$H_{\max}(X) = \log_2 \gamma \quad (2.20)$$

2.5. Dimensión de regularización

Esta sección tiene como propósito presentar los fundamentos teóricos de la dimensión de regularización, los cuales son utilizados en el desarrollo de la metodología basada en vector de características con entropía de la información y dimensión de regularización.

Inicialmente, Mandelbrot desarrolló la idea de los fractales cuando se propuso medir la costa de Inglaterra, percatándose que su longitud $N(\epsilon)$ aumentaba a medida que utilizaba escalas ϵ más pequeñas. De esta forma, Mandelbrot encontró la relación

$$N(\epsilon)\epsilon^\Lambda = \varsigma, \quad (2.21)$$

donde ς es una constante y Λ la variable que Mandelbrot denominó la dimensión fractal [143].

Los fractales tienen dos características, la auto-similaridad y la dimensión fractal. La auto-similaridad implica que cuando se observa un sistema u objeto a diferentes escalas, este tiende a tener la misma forma de la estructura original. Ejemplos clásicos de estructuras con auto-similaridad son: el conjunto de Cantor, la curva de Koch, el copo de nieve de Koch y el triángulo de Sierpinski, entre otros [144].

Desde la topología matemática, la dimensión (Λ) se define por [144]:

- La dimensión de un conjunto vacío es $\Lambda = -1$.
- Si κ es cualquier elemento que pertenezca al conjunto V donde $\kappa \in V$ y V es abierto en el espacio Δ , entonces existe un conjunto abierto U con frontera vacía que cumple $\kappa \in U \subset V$. En este caso el espacio topológico será $\Lambda = 0$.
- Un espacio topológico Δ tiene dimensión menor o igual a $\Lambda \Leftrightarrow$ para cada punto κ en Δ , y cada conjunto abierto $V \in \kappa \Rightarrow \exists U$ que es un conjunto abierto tal que $\kappa \in U \subset V$ y dimensión de frontera $U \leq \Lambda - 1$.

De esta manera, aplicando el concepto de espacio topológico, una línea tiene $\Lambda = 1$, un cuadrado tiene $\Lambda = 2$ y un cubo tiene $\Lambda = 3$. Cuando Hausdorff intentó hallar la dimensión topológica de ciertos cuerpos como el triángulo de Sierpinski, encontró que $1 \leq \Lambda \leq 2$. Es decir, $\Lambda \notin \mathbb{N}$ [144].

En este sentido, un fractal es un cuerpo o sistema que es autosimilar y donde la dimensión fractal es mayor a la dimensión topológica, siendo la dimensión fractal la capacidad de detectar dimensiones que no son enteras [144].

Diferentes métodos para medir la dimensión fractal se han desarrollado, como la dimensión de Hausdorff, la dimensión por conteo de cajas y la dimensión de regularización [145]. Hausdorff desarrolló un método basado en conceptos de topología general. La idea de Hausdorff consiste en utilizar una sucesión de discos cada vez más pequeños, donde en cada cubrimiento se utilizan discos del mismo tamaño que cubren el objeto de interés [146]. Si \mathfrak{S} es una figura cerrada y acotada, entonces:

$$h^\Lambda(\mathfrak{S}) = \lim_{\epsilon \rightarrow 0} N(\epsilon)\epsilon^\Lambda \quad (2.22)$$

donde h^Λ es la medida Λ -dimensional de Hausdorff, ϵ es el diámetro de los discos y $N(\epsilon)$ es el número de discos pertenecientes a \mathfrak{S} cuando $\epsilon \rightarrow 0$ [143]. Hausdorff demostró que existe un

único valor de Λ donde $h^\Lambda \neq 0$ y $h^\Lambda \neq \infty$. La dimensión de Hausdorff presenta el inconveniente del costo, en memoria y tiempo de ejecución, para su implementación computacional [144].

De acuerdo con el planteamiento de Mandelbrot (Eq. 2.21), para cuerpos autosimilares la dimensión fractal se puede calcular con:

$$\Lambda = \frac{\ln(N)}{\ln(1/\epsilon)} \quad (2.23)$$

donde Λ es la dimensión fractal autosimilar, ϵ es la escala y N es el número de subdivisiones relacionadas a ϵ para un determinado grafo.

Una variante de este método, aplicable a objetos que no son completamente autosimilares, es la dimensión por conteo de cajas. Consiste en crear elementos que tengan vecindad topológica utilizando para el caso bidimensional una red de cuadrículas de igual dimensión, donde cada una de esas cuadrículas se denomina una caja [147]. De esta forma se cuenta el número de cajas que pertenecen al objeto que se quiere medir. Este proceso se repite reduciendo la escala o tamaño de las cajas. Así, si un objeto $\mathfrak{S} \in \mathbb{R}^2$ y el menor número de subconjuntos de diámetro ϵ que cubren a \mathfrak{S} se denomina $N_\epsilon(\mathfrak{S})$, entonces la dimensión por conteo de cajas se define como [144]:

$$\Lambda_B(\mathfrak{S}) = \lim_{\epsilon \rightarrow 0} \frac{\ln(N_\epsilon(\mathfrak{S}))}{\ln(1/\epsilon)} \quad (2.24)$$

donde $\epsilon \in \mathbb{R}^+$.

La dimensión de regularización fue desarrollada por Roueff y Veheh como un camino novedoso para evaluar la regularidad del grafo de una función \mathfrak{S} [148]. La aproximación está basada en la medición de la razón de crecimiento de la longitud de menores y menores versiones regularizadas de \mathfrak{S} . De esta manera, se genera un índice denominado dimensión de regularización [148].

Si se asume que \mathfrak{S} es irregular y tiene longitud infinita, entonces se puede hallar una versión suavizada \mathfrak{S}_σ , donde

$$\mathfrak{S}_\sigma = \mathfrak{S} * \eta_\sigma \quad (2.25)$$

Así, \mathfrak{S}_σ es la versión suavizada de la señal que resulta de la convolución entre \mathfrak{S} y η_σ , donde η_σ es un kernel gaussiano con ancho σ . En este caso, la familia de funciones gaussianas es representada por: [149]:

$$\eta_\sigma(t) = e^{\left(\frac{-t^2}{\sigma^2}\right)} \quad (2.26)$$

donde t es el tiempo y la r -ésima derivada es representada como:

$$\eta_\sigma^{(r)}(t) = \frac{d^{(r)}}{dt^{(r)}} \eta_\sigma(t) \quad (2.27)$$

Cuando se convoluciona \mathfrak{S} con η_σ , se obtiene una curva suavizada más regular con una longitud R_σ [149]. A medida que el ancho σ tiende a cero ($\sigma \rightarrow 0$), el kernel gaussiano tiende a la función impulso. En consecuencia, la curva suavizada se aproxima a \mathfrak{S} y la longitud de R_σ tiende a infinito ($R_\sigma \rightarrow \infty$).

De esta forma, la dimensión de regularización (Λ_{RD}) se representa con:

$$\Lambda_{RD} = 1 - \lim_{\sigma \rightarrow 0} \frac{\ln R_\sigma}{\ln(\sigma)} \quad (2.28)$$

donde R_σ es la longitud de la señal suavizada y σ es el ancho del kernel gaussiano.

De esta manera, Λ_{RD} mide la velocidad en que la longitud de la señal suavizada converge a infinito ($R_\sigma \rightarrow \infty$) cuando el ancho del kernel gaussiano tiende a cero ($\sigma \rightarrow 0$), [148]. Para hallar Λ_{RD} , se lleva a cabo una regresión lineal entre $\ln R_\sigma$ y $\ln(\sigma)$ [149]. El valor de la pendiente de la regresión lineal es usada para estimar Λ_{RD} .

Dentro de las ventajas que presenta la dimensión de regularización se encuentran: se puede escoger el kernel de regularización, las versiones suavizadas son adaptativas por construcción, el parámetro de suavisamiento puede ser variado en muy pequeños pasos y presenta tolerancia a niveles de ruidos [149]. En este estudio se seleccionó la dimensión de regularización como una de las características necesarias para la clasificación de señales EEG con habla silenciosa.

2.6. Análisis de datos funcionales

Esta sección tiene como propósito presentar los fundamentos teóricos del análisis de datos funcionales, los cuales son utilizados en el desarrollo de la metodología basada en vector de características con datos funcionales.

La reciente área de la estadística llamada análisis de datos funcionales, analiza curvas o superficies, mirandolas como realizaciones continuas en un espacio de Hilbert. Estas funciones pueden ser definidas en el tiempo, longitud de onda, posición, peso, etc [150]. En este sentido, una variable randómica es una variable funcional si toma valores en un espacio dimensional infinito llamado espacio funcional χ [151]. Una observación de χ es llamado un dato funcional x .

En análisis multivariado los datos se pueden representar en un arreglo matricial donde las filas n corresponden a los individuos u objetos y las columnas p a sus variables o características. En análisis multivariado se requiere que $n \geq p$ para poder hallar las matrices de varianza y covarianza [152]. En el caso de datos funcionales, no se tiene esta restricción y puede hacer análisis con $p \geq n$. Otra de la ventajas de los datos funcionales es que puede analizar la forma de las curvas o funciones y puede hallar descriptores relevantes, que se encuentran ocultos, para el análisis de datos multivariado [150].

El modelo matemático de los datos funcionales es un espacio de Hilbert L^2 , siendo un espacio de Hilbert \mathcal{H} un espacio con producto interno $\langle \cdot, \cdot \rangle$ con la propiedad de completéz. En este caso, la completéz es definida cuando para toda secuencia de Cauchy $\{x_n\}$ converge en norma a algún elemento $x \in \mathcal{H}$ [153]. Un espacio de Hilbert L^2 tiene las características de un espacio de Hilbert donde para dos funciones $f(x)$ y $g(x)$ [154]:

$$L^2(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{C} \mid \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx \leq \infty\} \quad (2.29)$$

La norma en este caso es definida como:

$$\|f\|_2 = \sqrt{\int_{-\infty}^{\infty} |f(x)|^2 dx} \quad (2.30)$$

y su producto interno es:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx \quad (2.31)$$

Para calcular una estimación de los datos funcionales, un conjunto de funciones base independientes entre sí $\phi_k(\cdot)$ es utilizado [155]. Seleccionando estas funciones base es posible aproximar funciones arbitrarias, calculando una combinación lineal de las respectivas bases [150] de esta forma:

$$\hat{x}(\Theta) = \sum_{k=1}^K c_k \phi_k(\Theta) \quad (2.32)$$

donde \hat{x} es una función estimada (curva), c_k son los coeficientes de las bases, Θ es el dominio de los datos funcionales y K es el número de bases. Para estimar \hat{x} es necesario seleccionar y calcular $\phi_k(\Theta)$, c_k y K .

Existen diferentes tipos de funciones base aplicables a datos funcionales [155]: exponenciales, polinomiales, poligonales, wavelets, Fourier, B-splines, entre otras. Las funciones base más utilizadas son las bases de Fourier, aplicadas a datos periódicos y las bases B-spline, aplicadas a datos no periódicos [150]. En este estudio, las señales cerebrales con habla silenciosa son representadas como una señal randómica, por lo tanto el B-spline es una función base susceptible [150], [155]. Las funciones B-spline son una serie de polinomios con grado m y orden $m + 1$ unidos en puntos llamados nudos, o en inglés knots [150]. De esta forma, el número de knots (*nbreaks*) es definido como [155]:

$$nbreaks = nbases - norden + 2 \quad (2.33)$$

donde el número de bases *nbases* y el orden (*norden*) son parámetros a estimar. Así, los knots para definir una curva son: $a_k = a_0, \dots, a_{nbreaks}$.

Para definir los B-splines se utiliza el algoritmo recursivo de Boor [156]. Las condiciones iniciales \aleph_k^0 de este algoritmo son:

$$\aleph_k^0 = \begin{cases} 1 & \text{si } \Theta \in [a_k, a_{k+1}) \\ 0 & \text{de otra manera} \end{cases} \quad (2.34)$$

De esta forma, las funciones B-spline son calculadas con:

$$\aleph_k^m(\Theta) = \zeta_k^{m-1} \aleph_k^{m-1}(\Theta) + (1 - \zeta_{k+1}^{m-1}) \aleph_{k+1}^{m-1}(\Theta) \quad (2.35)$$

donde $\aleph_k^m(\cdot)$ son las funciones B-spline, k es el número de bases, m es el grado de los B-spline y el valor ζ_k^{m-1} es:

$$\zeta_k^{m-1} = \frac{\Theta - a_k}{a_{k+m} - a_k} \quad (2.36)$$

Por lo tanto, los datos funcionales estimados con las funciones B-spline son representados por [150]:

$$\hat{x}(\Theta) = \sum_{k=1}^K c_k \aleph_k^m(\Theta) \quad (2.37)$$

La expresión (2.37) en forma vectorial se representa por:

$$\hat{x}(\Theta) = \mathbf{c}^T \Phi \quad (2.38)$$

donde \mathbf{c}^T es el vector transpuesto de coeficientes c_k y Φ es una matriz de B-splines dado por:

$$\Phi = \begin{pmatrix} \aleph_1^m(\Theta_1) & \dots & \aleph_K^m(\Theta_1) \\ \aleph_1^m(\Theta_2) & \dots & \aleph_K^m(\Theta_2) \\ \vdots & \ddots & \vdots \\ \aleph_1^m(\Theta_n) & \dots & \aleph_K^m(\Theta_n) \end{pmatrix} \quad (2.39)$$

con un tamaño $n \times K$, donde n es la cantidad de observaciones discretas y_j que conforman el vector \mathbf{y} . Para ajustar las observaciones discretas y_j se utiliza el modelo $y_j = x(\Theta_j) + \epsilon_j$ donde ϵ_j corresponde a las perturbaciones [150].

Para obtener los coeficientes c_k , inicialmente se requiere calcular $SE(\cdot)$ dado por:

$$SE(\mathbf{y}|\mathbf{c}) = \sum_{j=1}^n \left[y_j - \sum_{k=1}^K c_k \aleph_k^m(\Theta_j) \right]^2 \quad (2.40)$$

donde $SE(\mathbf{y}|\mathbf{c})$ es el error de los datos ajustados. En forma matricial $SE(\cdot)$ es dado por:

$$SE(\mathbf{y}|\mathbf{c}) = (\mathbf{y} - \Phi \mathbf{c})^T (\mathbf{y} - \Phi \mathbf{c}). \quad (2.41)$$

Si la ecuación (2.41) es propuesta en términos de la función norma entonces:

$$SE(\mathbf{y}|\mathbf{c}) = \|\mathbf{y} - \Phi \mathbf{c}\|^2 \quad (2.42)$$

Para solucionar la ecuación (2.42) es necesario realizar la primera derivada de $SE(\mathbf{y}|\mathbf{c})$ con respecto a \mathbf{c} ; por lo tanto

$$2\Phi\Phi^T \mathbf{c} - 2\Phi^T \mathbf{y} = 0 \quad (2.43)$$

La solución de la ecuación (2.43) que permite encontrar el valor óptimo de c_k es:

$$\mathbf{c} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y} \quad (2.44)$$

Los nuevos valores $\hat{\mathbf{y}}$ que se obtienen con los coeficientes óptimos c_k son:

$$\hat{\mathbf{y}} = \Phi \mathbf{c} = \Phi (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y} \quad (2.45)$$

El último paso de este proceso es encontrar el número óptimo de funciones bases K para estimar la función \hat{x} [157]. Para encontrar este valor es necesario hallar el mínimo error cuadrático $MSE[\cdot]$ usando:

$$MSE[\hat{x}(\Theta)] = E [(\hat{x}(\Theta) - x(\Theta))^2] \quad (2.46)$$

donde $E[\cdot]$ corresponde al valor esperado y el valor mínimo $MSE[\hat{x}(\Theta)]$ es hallado por validación cruzada.

Adicionalmente, para calcular la primera derivada: $\hat{x}(\Theta)$ se utiliza

$$D\widehat{x}(\Theta) = \sum_{k=1}^K c'_k D\aleph_k^m(\Theta) = (\mathbf{c}')^T \mathbf{\Phi}' \quad (2.47)$$

donde $D\widehat{x}(\Theta)$ corresponde a la primera derivada de $\widehat{x}(\Theta)$; c'_k son los coeficientes a estimar de la primera derivada y $D\aleph_k^m(\Theta)$ corresponde a la primera derivada de las funciones B-spline [158]. La matriz $\mathbf{\Phi}'$, en este caso, es representada con la siguiente ecuación:

$$\mathbf{\Phi}' = \begin{pmatrix} D\aleph_1^m(\Theta_1) & \dots & D\aleph_K^m(\Theta_1) \\ D\aleph_1^m(\Theta_2) & \dots & D\aleph_K^m(\Theta_2) \\ \vdots & \ddots & \vdots \\ D\aleph_1^m(\Theta_n) & \dots & D\aleph_K^m(\Theta_n) \end{pmatrix} \quad (2.48)$$

En forma similar a c_k , la estimación por mínimos cuadrados fue utilizada para obtener los coeficientes c'_k . Por lo tanto,

$$\mathbf{c}' = \left((\mathbf{\Phi}')^T \mathbf{\Phi}' \right)^{-1} (\mathbf{\Phi}')^T \mathbf{y} \quad (2.49)$$

Finalmente, para hallar la media de las funciones con datos funcionales se utiliza la siguiente expresión [150]:

$$\bar{x}(\Theta) = \frac{\sum_{k=1}^{N_T} x_k(\Theta)}{N_T} \quad (2.50)$$

donde $\bar{x}(\cdot)$ corresponde a la curva media, $x_k(\cdot)$ corresponde a cada una de las curvas funcionales, Θ corresponde al dominio de las curvas funcionales, k es la referencia a cada una de las curvas y N_T es el número de curvas.

Capítulo 3

Modelos propuestos basados en habla silenciosa

En este capítulo se presenta los dos métodos desarrollados para caracterizar y clasificar el habla silenciosa utilizando señales EEG. El primer método de habla silenciosa se caracteriza por utilizar la entropía de la información y la dimensión de regularización. El segundo método se caracteriza por utilizar la densidad espectral de potencia (periodograma) y la información espacial de los electrodos para hacer una transformación de esas señales cerebrales en datos funcionales. Cada uno de estos métodos fue desarrollado para vocales y sílabas con habla silenciosa.

3.1. Protocolo experimental

La experimentación se desarrolló en el laboratorio de Electrofisiología Clínica de la Universidad Nacional de Colombia, en condiciones controladas de luminosidad y mínimo ruido ambiental. A continuación se explica el protocolo de la experimentación.

Los datos usados en esta investigación fueron obtenidos de 20 sujetos (17 hombres y 3 mujeres) para la experimentación de vocales con habla silenciosa (Fig. 3.1), y de 10 sujetos (6 hombres y 4 mujeres) para sílabas con habla silenciosa (Fig. 3.2). Durante la recolección de los datos, los sujetos estuvieron confortablemente sentados, con sus ojos cerrados, portando un EEG *neuroheadset* donde se acomodaban 21 electrodos. Las señales cerebrales se grabaron en un cuarto aislado de ruido (ASTM STC 63) con control de luminosidad ($3 \pm 0.2 \text{ lm/m}^2$). Una fuente luminosa (80 lm/m^2) fue colocada a una distancia de 1 ± 0.1 metros de los ojos del sujeto para indicar el momento en que debía comenzar y terminar de pensar en la vocal o sílaba específica con SS.

El EEG *neuroheadset* está fabricado en silicona, lo cual permite posicionar 21 electrodos $E1, \dots, E21$ en un arreglo matricial con una distancia entre ellos de 16 mm. Este arreglo incluye un electrodo de referencia posicionado en la mitad de la frente y un electrodo de tierra (GND) posicionado en el lóbulo de la oreja izquierda. El punto nasión (Fig. 2.19) en la cabeza del sujeto y el área preauricular de la oreja izquierda (Fig. 2.19), son usadas como puntos de referencia para ajustar el *neuroheadset* a los diferentes sujetos. Un gel conductor es aplicado para limpiar el cuero cabelludo antes de colocar cada electrodo. La posición de los electrodos fue proyectada para cubrir las áreas de Wernicke y Broca en el hemisferio izquierdo. Con cada sujeto se buscó que coincidiera la posición del electrodo $E4$ relacionado

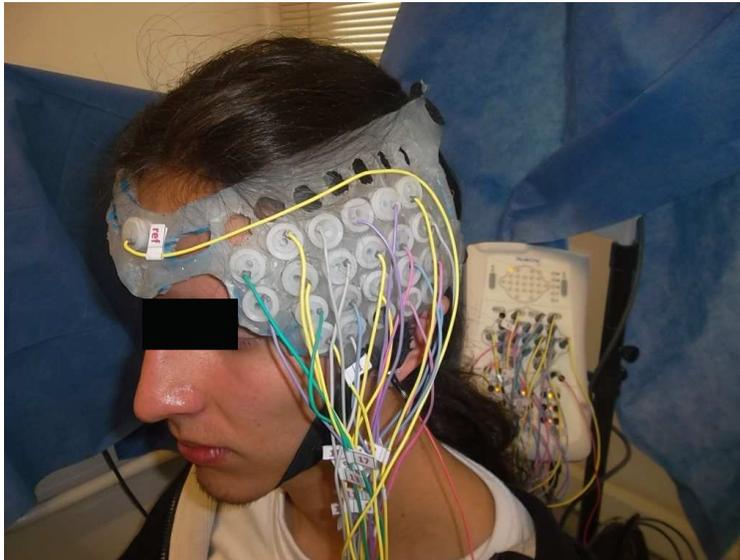


Figura 3.1: Sujeto usando el EEG *neuroheadset* en la experimentación de vocales con habla silenciosa

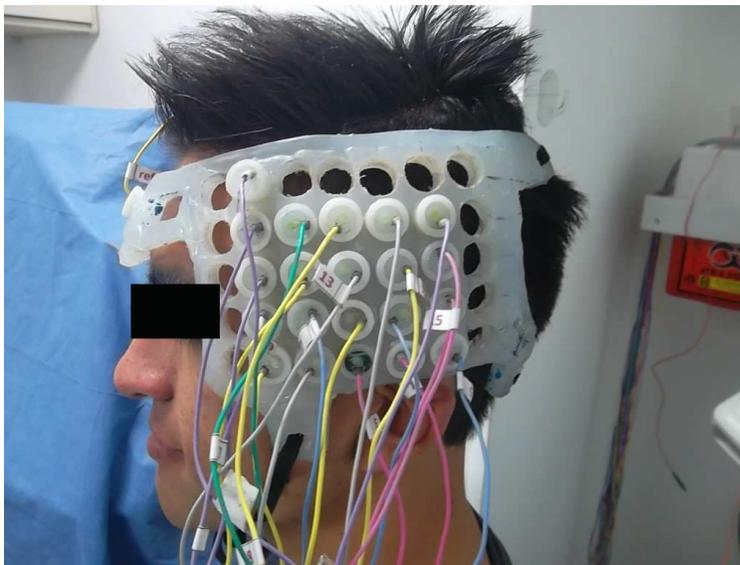


Figura 3.2: Sujeto usando el EEG *neuroheadset* en la experimentación de sílabas con habla silenciosa

al *neuroheadset* con la posición T_3 del sistema internacional para la colocación de electrodos 10-20, como puede verse en (Fig. 3.3) y (Fig. 3.4).

Los electrodos se ubicaron en el hemisferio izquierdo, de acuerdo a los modelos neurológicos sobre el lenguaje de: Broca [46], Wernicke [48], Litchtheim [48], Geshwind [49] y el modelo de Hickok y Poeppel [54]. De esta manera se busca que los electrodos cubran el área de Broca (Brodmann 44, 45, 6, 43, 46 y 47) y el área de Wernicke (Brodmann 39, 40, 22, 38, 42 y 21) (Fig. 2.6). Para vocales con habla silenciosa, los electrodos E1, E6, E7, E11, E12, E13, E14, E17, E18, E19 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E15, E16 y E21 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.3). Para sílabas con habla silenciosa, los electrodos E1, E6, E7, E11, E12,

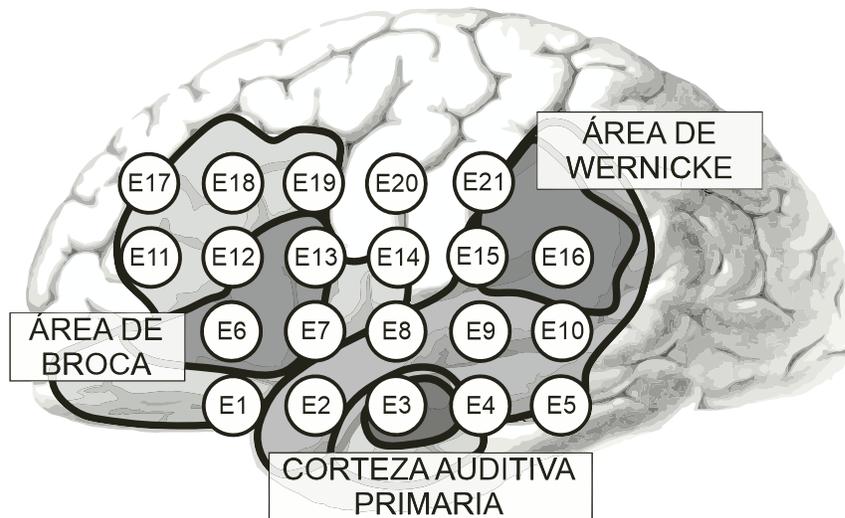


Figura 3.3: Arreglo de 21 electrodos cubriendo el área de Broca, Wernicke y corteza auditiva primaria para vocales con SS

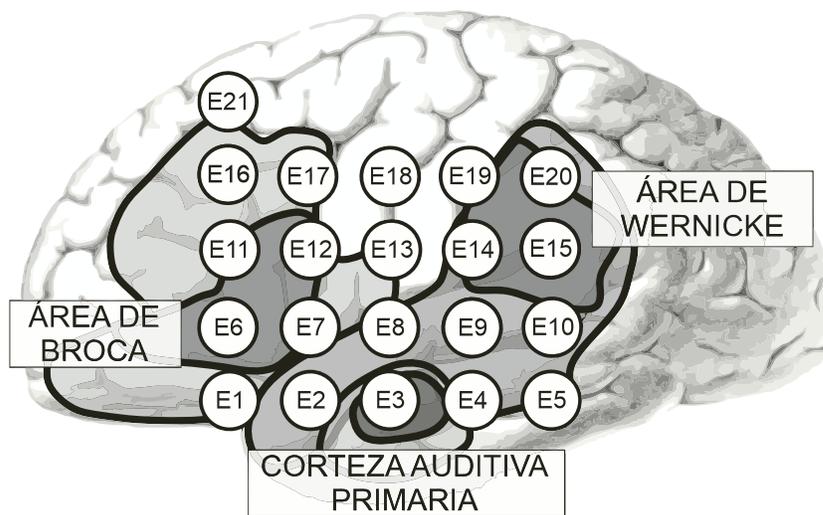


Figura 3.4: Arreglo de 21 electrodos cubriendo el área de Broca, Wernicke y corteza auditiva primaria para sílabas con SS

E13, E16, E17, E21 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E14, E15, E19 y E20 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.4).

3.1.1. Adquisición de datos para vocales con SS

Antes de que la señal sea grabada, se le explica a cada sujeto que mientras la fuente luminosa esté encendida, debe pensar en la respectiva vocal de forma continua con habla silenciosa; y que una vez la fuente luminosa se apague, ellos deben dejar de pensar en la vocal y pasar a un estado de relajación (Fig. 3.5(a)). Durante la grabación, la fuente luminosa está encendida por tres segundos y apagada por otros tres. El procedimiento es repetido 10

veces para cada una de las 5 vocales. Entre cada tarea de vocales, los sujetos descansan 5 minutos para realizar el cambio a la siguiente vocal. Las vocales fueron grabadas en el siguiente orden: primero, únicamente /a/; segundo, únicamente /e/; tercero, únicamente /i/; cuarto, únicamente /o/; y quinto, únicamente /u/. Los intervalos de tiempo cuando los sujetos pensaron en las vocales fueron (3-6 segundos, 9-12 segundos, 15-18 segundos,..., 57-60 segundos); estos intervalos fueron seleccionados para el análisis (Fig. 3.5(b)). De acuerdo con el grado de abertura de la cavidad oral, /a/ es abierta, /e/ es media, /i/ es cerrada, /o/ es media y /u/ es cerrada [159].

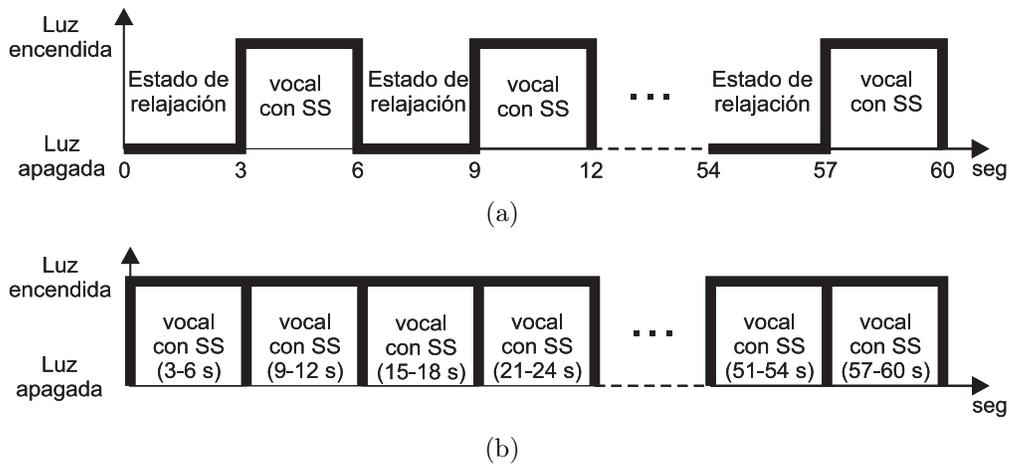


Figura 3.5: Intervalos seleccionados para el procesamiento de vocales con SS

Las señales fueron grabadas con un amplificador NicoletOne™ V32 (Viasys Healthcare, Inc. Conshohocken, Pennsylvania) con el software de adquisición Nicolet™ VEEG (Natus Neurology, Middleton, Wisconsin). Las grabaciones se desarrollaron en el laboratorio de Electrofisiología Clínica de la Universidad Nacional de Colombia (Bogotá, Colombia). Las grabaciones se realizaron a una frecuencia de muestreo de 500 Hz en configuración monopolar. Los datos fueron almacenados y exportados para ser procesados con el software Matlab™ (The MathWorks, Inc. Natick, Massachusetts) y el software R para computación estadística. Estos software permiten desarrollar funciones personalizadas para procesamiento de señales.

3.1.2. Adquisición de datos para sílabas con SS

Antes de iniciar la grabación de las señales cerebrales, se le explica a cada sujeto que cuando la fuente luminosa se apague, debe pensar en la respectiva sílaba una vez, luego dejar de pensar en la sílaba y a continuación pasar a un estado de relajación (Fig. 3.6(a)). Durante la grabación, la fuente luminosa está encendida por 100 ms y apagada por 3900 ms. El procedimiento es repetido 30 veces para cada una de las 5 sílabas. Entre cada tarea de sílabas, los sujetos descansan 5 minutos para realizar el cambio a la siguiente sílaba. Las sílabas fueron grabadas en el siguiente orden: primero, únicamente /fa/; segundo, únicamente /pa/; tercero, únicamente /ma/; cuarto, únicamente /la/; y quinto, únicamente /ra/. Cada vez que la fuente luminosa se apaga, se toma 200 ms para el registro y análisis de los datos (Fig. 3.6(b)). De acuerdo con el modo de articulación, /fa/ es fricativa, /pa/ es oclusiva, /ma/ es nasal, /la/ es lateral y /ra/ es vibrante [159].

Las señales, igualmente, fueron grabadas en el laboratorio de Electrofisiología Clínica de la Universidad Nacional de Colombia (Bogotá, Colombia) con un amplificador NicoletOne™

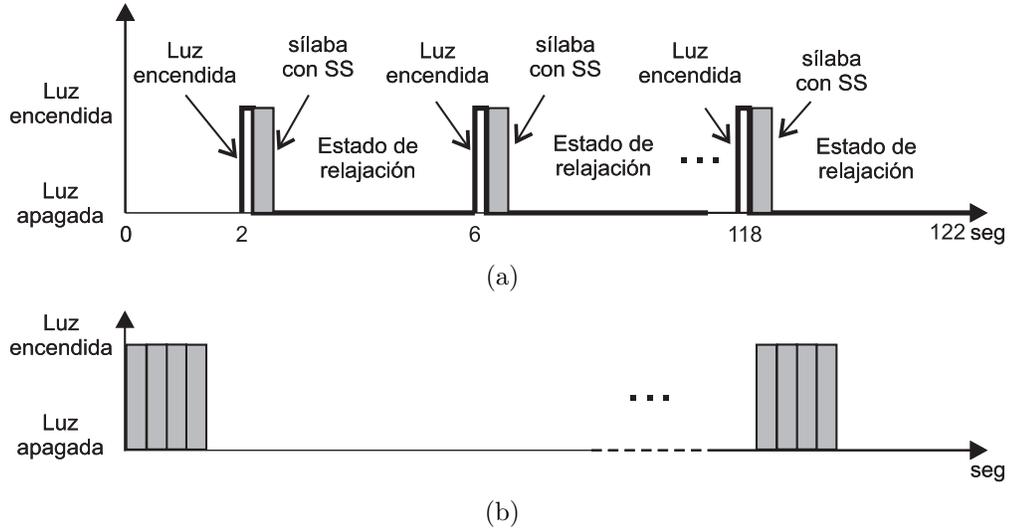


Figura 3.6: Intervalos seleccionados para el procesamiento de sílabas con SS

V32 (Viasys Healthcare, Inc. Conshohocken, Pennsylvania) con el software de adquisición NicoletTM VEEG (Natus Neurology, Middleton, Wisconsin). Las grabaciones se realizaron a una frecuencia de muestreo de 2000 Hz en configuración monopolar. Los datos fueron almacenados y exportados para ser procesados con el software MatlabTM (The MathWorks, Inc. Natick, Massachusetts) y el software R para computación estadística.

3.2. Metodología basada en vector de características con entropía de la información y la dimensión de regularización

El primer modelo desarrollado de habla silenciosa es un algoritmo que toma los registros de las grabaciones realizadas a los sujetos en tareas de vocales con habla silenciosa (Fig. 3.7) o sílabas con habla silenciosa (Fig. 3.8). Estas señales pasan a la etapa de filtraje donde se seleccionan las frecuencias entre 2 Hz a 50 Hz. Luego, los intervalos seleccionados donde los sujetos pensaron en la respectiva vocal o sílabas con SS, se denomina ($T_{s,v}^p$). Posteriormente, se pasa a un bloque de detección y eliminación de artefactos debidos al parpadeo ($V_{s,v}^p$). A continuación, viene un bloque de selección de intervalos en el cual se piensa en la vocal o sílabas con habla silenciosa ($W_{s,v}^p$). En seguida está el bloque de extracción de características, utilizando la entropía de Shannon y dimensión de regularización; con estos valores se implementa el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$. Después de esto, por medio de un proceso de validación cruzada se selecciona el conjunto de datos para entrenamiento y el conjunto de datos para prueba. Para el proceso de clasificación multiclase fue seleccionada una máquina de soporte vectorial (SVM) uno contra el resto (1-rest) y uno contra uno (1-1). Los parámetros del SVM fueron hallados a través de optimización con algoritmos genéticos. A continuación se explica cada uno de los bloques.

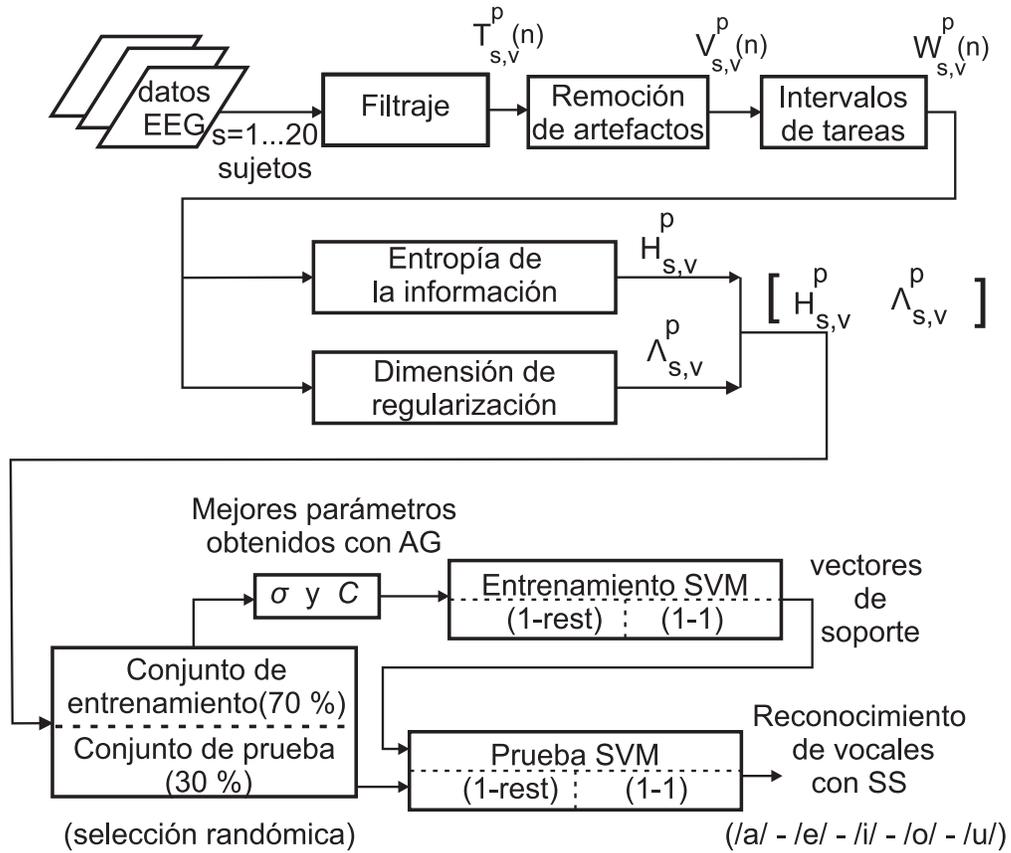


Figura 3.7: Diagrama del modelo propuesto con entropía de la información y dimensión de regularización para vocales con SS

3.2.1. Filtrado

En el filtrado de señales cerebrales (Fig. 3.7 y Fig. 3.8), dos filtros elípticos con respuesta de impulso infinita (IIR) son propuestos: un filtro pasa alto (HPF) y un filtro pasa bajo (LPF) [160]. En el caso de las vocales se utiliza una frecuencia de muestreo $F_s = 500 \text{ Hz}$; el filtro pasa alto tiene una frecuencia de corte de 2 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 2 dB y un orden de 6. El filtro pasa bajo tiene una frecuencia de corte de 50 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 1 dB y un orden de 6. En el caso de las sílabas se utiliza una frecuencia de muestreo $F_s = 2000 \text{ Hz}$; el filtro pasa alto tiene una frecuencia de corte de 2 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 2 dB y un orden de 6. El filtro pasa bajo tiene una frecuencia de corte de 50 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 1 dB y un orden de 11.

Para el caso de vocales con habla silenciosa (SS), $T_{s,v}^p(n)$ es la señal filtrada. El subíndice s se refiere al sujeto $s = 1, \dots, 20$, p se refiere al número de electrodos $p = 1, \dots, 21$ y v se refiere a la vocal con habla silenciosa $v = 1, \dots, 5$ (/a/, /e/, /i/, /o/ y /u/). La variable n denota las muestras de tiempo discreto. Los segmentos grabados proveen un total de 30000 (60 segundos) datos por vocal con habla silenciosa, por electrodo y por sujeto (Figs. 3.5).

Para el caso de las sílabas con SS, $T_{s,v}^p(n)$ es la señal filtrada. El subíndice s se refiere al sujeto $s = 1, \dots, 10$, p se refiere al número de electrodos $p = 1, \dots, 21$ y v se refiere a la sílaba con habla silenciosa $v = 1, \dots, 5$ (/fa/, /pa/, /ma/, /la/ y /ra/). La variable n denota las muestras de tiempo discreto. Los segmentos grabados proveen un total de 244000 (12

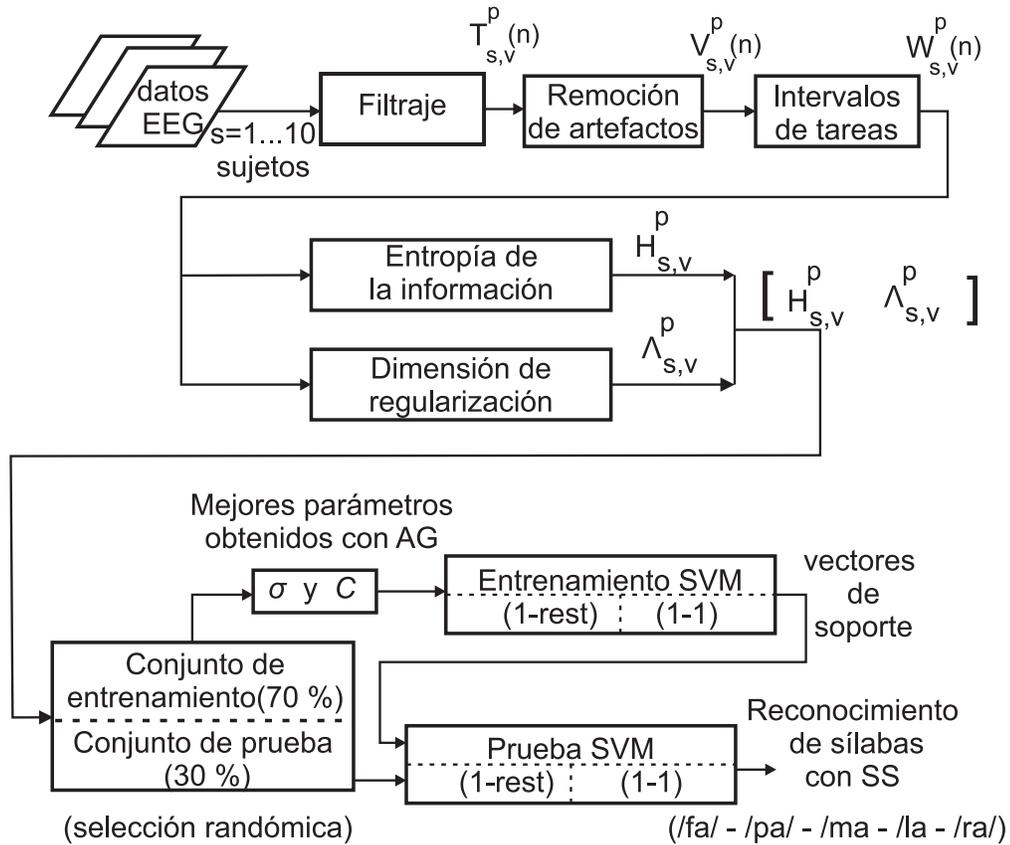


Figura 3.8: Diagrama del modelo propuesto con entropía de la información y dimensión de regularización para sílabas con SS

segundos) datos por sílaba con habla silenciosa, por electrodo, y por sujeto (Figs. 3.6).

3.2.2. Consideración de artefactos

Los siguientes artefactos fueron tomados en consideración para el diseño del sistema EEG con habla silenciosa: artefactos externos, artefactos cardiacos, artefactos por los electrodos y artefactos musculares [161].

Los artefactos externos son aquellos relacionados con las señales generadas fuera del cuerpo y que inciden sobre las señales EEG. Para minimizar estos artefactos, la experimentación se llevó a cabo en el laboratorio de Electrofisiología Clínica de la Universidad Nacional de acuerdo al ASTM STC 63 para aislamiento de cuartos relacionados con ruido. Adicionalmente el control de luminosidad estuvo en el rango de $3.0 \pm 0.2 \text{ lm/m}^2$. En el preprocesamiento de vocales y sílabas con habla silenciosa se tuvo en cuenta remover los 60 Hz utilizando LPF con frecuencias de corte menores a 50 Hz .

En otro caso, los artefactos relacionados con el electrocardiograma (ECG) son generados por la actividad eléctrica del corazón. El pulso normal en un adulto es de 60 latidos por minutos, en ingles *beats per minute* (bpm), a 100 bpm (1 Hz a 2 Hz). En el habla silenciosa con EEG se implementa un filtro (HPF) con una frecuencia de corte de 2 Hz para reducir esta clase de artefactos [64].

Los artefactos relacionados a los electrodos se fundamentan en el mal contacto o movimiento relativo de los electrodos con el cuero cabelludo [64]. En el caso de los artefactos

relacionados con el contacto de los electrodos se redujeron teniendo en cuenta que la impedancia de los electrodos fuera menor a $5k\Omega$. Adicionalmente, el movimiento relativo de los electrodos fue reducido con la utilización del EEG *neuroheadset*.

Adicionalmente, los artefactos musculares están relacionados con las señales electroencefalográficas (EMG) generadas por la activación de las células musculares. Las señales EMG están entre $6 Hz$ a $500 Hz$ con frecuencias de mayor potencia desde $50 Hz$ a $150 Hz$ [112]. En este estudio, se utilizaron ritmos cerebrales: δ (2 a $4 Hz$), $\delta\theta$ (2 a $8 Hz$), y $\delta\theta\alpha$ (2 a $13 Hz$). Para reducir los artefactos musculares se les solicitó a los sujetos en la experimentación no mover los miembros inferiores, miembros superiores o el cuello.

Finalmente, los artefactos oculares están relacionados con el parpadeo o movimiento de los ojos [64]. Para reducir el parpadeo se les pidió a los sujetos cerrar los ojos durante la experimentación.

3.2.3. Remoción artefactos de parpadeo

Los artefactos relacionados con el movimiento de los ojos (parpadeo) se detectaron a través de la evaluación de un valor de umbral U definido como el 170% de la raíz media cuadrática (RMS) de la señal EEG $T_{s,v}^p(n)$. Esto es:

$$U = 1.7 \times \sqrt{\frac{1}{N} \sum_{n=1}^N |T_{s,v}^p(n)|^2} \quad (3.1)$$

Para el caso de vocales con habla silenciosa, la longitud de los datos es $N = 30000$ ($60 s$). Si $T_{s,v}^p(n) \geq U$ en una ventana rectangular de ancho $N_i = 250$ ($0.5 s$), entonces el intervalo es considerado un artefacto de parpadeo. Para el caso de sílabas con habla silenciosa, la longitud de los datos es $N = 244000$ ($122 s$). Si $T_{s,v}^p(n) \geq U$ en una ventana rectangular de ancho $N_i = 400$ ($200 ms$), entonces el intervalo es considerado un artefacto de parpadeo.

Una vez que un artefacto de parpadeo es detectado, un algoritmo basado en wavelets es implementado para removerlo [162]. Con este algoritmo, la remoción del artefacto es desarrollada en tres pasos: (1) descomposición de la señal observable, (2) umbralización y (3) reconstrucción de la señal estimada.

En el primer paso, la señal observada $T_{s,v}^p(n)$ es expresada como $T_{s,v}^p(n) = \widehat{S}_d(n) + \varepsilon(n)$, donde $\widehat{S}_d(n)$ es la señal estimada sin ruido y $\varepsilon(n)$ es el ruido no observado (artefacto). Usando la transformada de wavelet discreta, en inglés *discrete wavelet transform* (DWT), la señal estimada puede ser expresada como $\widehat{S}_d(n) = \widehat{A}_d(n) + \widehat{D}_d(n)$, donde $\widehat{A}_d(n)$ es la señal aproximada y $\widehat{D}_d(n)$ es la señal detallada. La señal aproximada es definida como: $\widehat{A}_d(n) = \sum_k \widehat{\beta}_{j,k} \phi_{j,k}(n)$ y la señal detallada es definida como $\widehat{D}_d(n) = \sum_{1 \leq j \leq k} \sum_k \widehat{\alpha}_{j,k} \psi_{j,k}(n)$, donde j es el parámetro de escala, k es el parámetro de posición, $\widehat{\beta}_{j,k}$ son los coeficientes aproximados, $\widehat{\alpha}_{j,k}$ son los coeficientes detallados, $\phi_{j,k}(n)$ son la familia de wavelets, y $\psi_{j,k}(n)$ son la familia de wavelets ortogonales a $\phi_{j,k}(n)$ [163]. Este trabajo hace uso de un wavelet ortogonal Daubechies (*db3*), con un nivel de descomposición de seis para las vocales con habla silenciosa y de ocho para sílabas con habla silenciosa [163].

En el segundo paso, este trabajo implementa una umbralización universal [164] definida como:

$$T_i = \sqrt{2 \log(N_i)}, \quad (3.2)$$

donde N_i es la longitud del intervalo. Para vocales con habla silenciosa $N_i = 250$ (0.5 s) y para sílabas con habla silenciosa $N_i = 400$ (200 ms). Únicamente los coeficientes detallados que satisfacen $\hat{\alpha}_{j,k} \geq T_i$ son seleccionados, y todos los demás valores son asignados como cero.

Finalmente, en el paso tres, la señal estimada $\hat{S}_d(n)$ es reconstruida como:

$$\hat{S}_d(n) = \sum_k \hat{\beta}_{L,k} \phi_{L,k}(n) + \sum_{1 \leq j \leq L} \sum_k \hat{\alpha}_{j,k} \psi_{j,k}(n), \quad (3.3)$$

donde L es el nivel de descomposición del wavelet. En este caso, para vocales con habla silenciosa $L = 6$ y para sílabas con habla silenciosa $L = 8$ [163].

La señal resultante a la salida del bloque de remoción de artefactos es $V_{s,v}^p(n)$. Cuando un artefacto de parpadeo se detecta en la señal cerebral, entonces $V_{s,v}^p(n) = \hat{S}_d(n)$. En el caso de no detectar un artefacto de parpadeo, $V_{s,v}^p(n) = T_{s,v}^p(n)$.

El próximo paso es seleccionar todos los intervalos, donde el sujeto s hace las tareas de habla silenciosa ($W_{s,v}^p(n)$), (Figs. 3.5(a) y 3.5(b)) para las vocales con habla silenciosa y (Figs. 3.6(a) y 3.6(b)) para las sílabas con habla silenciosa. Los valores $W_{s,v}^p(n)$ son organizados en una matriz $\mathbb{W}_{s,v}$ donde cada columna de $\mathbb{W}_{s,v}$ contiene la información del electrodo p .

En el caso de vocales con habla silenciosa, esta información es almacenada desde $n = 0, \dots, 14999$. De esta forma,

$$\mathbb{W}_{s,v} = \begin{pmatrix} W_{s,v}^1(0) & \dots & W_{s,v}^{21}(0) \\ \vdots & \ddots & \vdots \\ W_{s,v}^1(14999) & \dots & W_{s,v}^{21}(14999) \end{pmatrix} \quad (3.4)$$

En el caso de sílabas con habla silenciosa, esta información es almacenada desde $n = 0, \dots, 11999$. De esta forma,

$$\mathbb{W}_{s,v} = \begin{pmatrix} W_{s,v}^1(0) & \dots & W_{s,v}^{21}(0) \\ \vdots & \ddots & \vdots \\ W_{s,v}^1(11999) & \dots & W_{s,v}^{21}(11999) \end{pmatrix} \quad (3.5)$$

Implementación entropía de la información

En este estudio, el algoritmo entropía de la información se desarrolla de la siguiente manera:

Se inicia calculando el valor máximo $nmax$ de los datos contenidos en las matrices $|\mathbb{W}_{s,1}|$, $|\mathbb{W}_{s,2}|$, $|\mathbb{W}_{s,3}|$, $|\mathbb{W}_{s,4}|$ y $|\mathbb{W}_{s,5}|$; donde s es un sujeto. En el caso de vocales con habla silenciosa, el valor $v = 1, \dots, 5$ corresponde a (/a/, /e/, /i/, /o/, /u/). En el caso de sílabas con habla silenciosa, el valor $v = 1, \dots, 5$ corresponde a (/fa/, /pa/, /ma/, /la/, /ra/).

Posteriormente se normaliza cada una de las matrices anteriores, donde

$$\underline{\mathbb{W}}_{s,v} = \frac{\mathbb{W}_{s,v}}{nmax} \quad (3.6)$$

para un sujeto s .

En esta parte, el objetivo es hallar la matriz de entropías $\mathbf{H}_{s,v}$ definida para un sujeto s y una vocal o sílaba de habla silenciosa v . Esta matriz está compuesta por los valores de entropía H_{s,v,n_1,n_2}^p , donde p es un electrodo y n_1 y n_2 corresponden a la fila máxima y mínima

de la respectiva matriz $\underline{\mathbb{W}}_{s,v}$. En el caso de vocales con habla silenciosa $A = 250$, entonces $n_2 = n_1 + A - 1$. De esta manera la matriz de entropías $\mathbf{H}_{s,v}$ corresponde a:

$$\mathbf{H}_{s,v} = \begin{pmatrix} H_{s,v,0,249}^1 & \cdots & H_{s,v,0,249}^{21} \\ \vdots & \ddots & \vdots \\ H_{s,v,14744,14999}^1 & \cdots & H_{s,v,14744,14999}^{21} \end{pmatrix} \quad (3.7)$$

donde el número de filas de $\mathbf{H}_{s,v}$ es $15000/A = 60$, y el número de columnas es 21, de manera que el tamaño de $\mathbf{H}_{s,v}$ para vocales con habla silenciosa es 60x21.

En el caso de sílabas con habla silenciosa $A = 400$, entonces $n_2 = n_1 + A - 1$. De esta manera la matriz de entropías $\mathbf{H}_{s,v}$ corresponde a

$$\mathbf{H}_{s,v} = \begin{pmatrix} H_{s,v,0,399}^1 & \cdots & H_{s,v,0,399}^{21} \\ \vdots & \ddots & \vdots \\ H_{s,v,11599,11999}^1 & \cdots & H_{s,v,11599,11999}^{21} \end{pmatrix} \quad (3.8)$$

donde el número de filas de $\mathbf{H}_{s,v}$ es $12000/A = 30$, y el número de columnas es 21, de manera que el tamaño de $\mathbf{H}_{s,v}$ para sílabas con habla silenciosa es 30x21.

Para hallar cada valor de entropía correspondiente a H_{s,v,n_1,n_2}^p se realizan los siguientes pasos:

El primer paso corresponde a la selección del número de muestras que van a conformar el espacio muestral. En este caso, se selecciona $\gamma = 100$.

El segundo paso define un vector \mathbf{d} , que contiene valores con un espaciado lineal entre $U_1 = -1$ hasta $U_2 = 1$, con $U_3 = \gamma + 1$ valores. De esta manera, el vector es $\mathbf{d} = [-1, -0.98, \dots, 0.98, 1]$ para vocales y sílabas con habla silenciosa. Cada componente de \mathbf{d} es representado por d_i , siendo $i = 1, \dots, U_3$.

El tercer paso establece un vector para almacenar el número de veces que se repite un evento de habla silenciosa, con relación a los valores del vector \mathbf{d} . Este vector se define inicialmente como

$$\mathbf{I} = [0, \dots, 0] \quad (3.9)$$

donde cada elemento de \mathbf{I} se define como I_k , siendo $k = 1, \dots, \gamma$.

El cuarto paso establece los valores para el vector \mathbf{I} de la siguiente forma:

Se asigna a un vector ϑ los valores de $\underline{\mathbb{W}}_{s,v}^p$, pertenecientes a la matriz $\underline{\mathbb{W}}_{s,v}$, comprendidos entre la fila n_1 y la fila n_2 del electrodo p . Ahora, se busca determinar la pertenencia de cada valor de ϑ con relación a los valores del vector \mathbf{d} . De esta manera, si el valor de ϑ está entre d_i y d_{i+1} , entonces $k = i$ y el valor de I_k se incrementa en 1.

El paso quinto consiste en determinar la probabilidad de los eventos generados en el espacio muestral \mathbf{I} . Si \mathbf{Z} es el vector de probabilidades, entonces

$$\mathbf{Z} = [Z_1, \dots, Z_\vartheta]. \quad (3.10)$$

Cada elemento de \mathbf{Z} se denomina Z_i siendo

$$Z_i = \frac{I_i}{\sum_{k=1}^{\gamma} I_k} \quad (3.11)$$

El último paso consiste en calcular el valor de entropía (Eq. 2.12) para cada valor de

$$H_{s,v,n_1,n_2}^p = - \sum_{k=1}^{\gamma} Z_i \log_2 Z_i \quad (3.12)$$

Implementación dimensión de regularización

En este estudio se desarrollaron los siguientes pasos para la implementación del algoritmo dimensión de regularización.

El primer paso consiste en definir para vocales con habla silenciosa el valor mínimo $U_1 = 5$ y máximo $U_2 = 200$ para el kernel gaussiano. En el caso de sílabas con habla silenciosa, el valor mínimo seleccionado es $U_1 = 5$ y máximo $U_2 = 350$ para el kernel gaussiano.

El segundo paso consiste en definir el número de tamaños intermedios para el kernel gaussiano $U_3 = 32$. Este valor es el mismo para vocales y sílabas con habla silenciosa.

El tercer paso consiste en generar un espaciamento logarítmico entre U_1 y U_2 , donde $\mathbf{d} = [5, 5.6318, \dots, 177.5623, 200]$ para vocales con SS y $\mathbf{d} = [5, 5.7344, \dots, 305.1749, 350]$ para sílabas con SS. Cada componente de \mathbf{d} se representa por d_i , siendo $i = 1, \dots, U_3$.

El cuarto paso, consiste en redondear los valores d_i . Por consiguiente, $\mathbf{D} = \text{round}(\mathbf{d})$ donde cada componente se representa por D_i .

El quinto paso consiste en hallar los $U_3 = 32$ kernel gaussianos. Estos 32 kernels son representados por:

$$\boldsymbol{\omega} = [\omega_1(n), \dots, \omega_{U_3}(n)] \quad (3.13)$$

donde cada kernel gaussiano de $\boldsymbol{\omega}$ se representa por $\omega_i(n)$ y se define como:

$$\omega_i(n) = e^{-\frac{1}{2} \left(\zeta \frac{n}{(D_i-1)/2} \right)^2} = e^{-\frac{n^2}{2\sigma_i^2}} \quad (3.14)$$

donde $\zeta = 2.5$ y el valor de la desviación estandar es:

$$\sigma_i = \frac{D_i - 1}{2\zeta}. \quad (3.15)$$

El sexto paso, consiste en hallar la primera derivada de los kernel gaussianos por diferencias finitas donde

$$\boldsymbol{\omega}' = [\omega'_1(n), \dots, \omega'_{U_3}(n)] \quad (3.16)$$

siendo $\boldsymbol{\omega}'$ el vector derivadas de kernel gaussiano y $\omega'_i(n)$ el i -ésimo kernel gaussiano. La derivada de cada kernel gaussiano se determina con:

$$\omega'_i(n) = \nabla_h[\omega'_i](n) = \omega_i(n) - \omega_i(n - h), \quad (3.17)$$

siendo $h = 1$ el valor entre muestra y muestra.

En el paso siete, se calcula la escala de los diferentes valores (anchos) de kernel gaussianos, donde

$$\boldsymbol{\sigma} = [\sigma_1, \dots, \sigma_{U_3}] \quad (3.18)$$

el cual se denomina vector de escalas y cada escala se representa por σ_i dado por la ecuación 3.15.

El octavo paso consiste en hallar el vector de convoluciones entre la señal cerebral y los kernel gaussianos dado por:

$$\mathbf{G} = [G_1(n), \dots, G_{U_3}(n)], \quad (3.19)$$

siendo \mathbf{G} el vector de convoluciones y $G_i(n)$ definido como:

$$G_i(n) = W_{s,v,n_1,n_2}^p * \omega'_i(n), \quad (3.20)$$

donde W_{s,v,n_1,n_2}^p corresponde al conjunto de valores entre la fila n_1 y la fila n_2 de la matriz $\mathbb{W}_{s,v}$. Cada elemento de la matriz es representado por $W_{s,v}^p$, para un electrodo p , una vocal o sílaba v y un sujeto s . De esta manera, cada valor $G_i(n)$ es la convolución entre los datos cerebrales W_{s,v,n_1,n_2}^p y la derivada del kernel gaussiano $\omega'_i(n)$. En este estudio el número de filas seleccionadas de $\mathbb{W}_{s,v}$ para cada convolución es $A = 250$ para vocales con habla silenciosa y $A = 400$ para sílabas con habla silenciosa, de tal manera que $n_2 = n_1 + A - 1$.

En el paso 9 se calcula la longitud de las convoluciones para cada valor $G_i(n)$, siendo

$$\mathbf{R} = [R_1, \dots, R_{U_3}] \quad (3.21)$$

donde \mathbf{R} representa el vector longitud de convoluciones y cada componente está definida como:

$$R_i = \sum_{k=2}^{U_3} |G_k(n) - G_k(n-1)|. \quad (3.22)$$

El paso 10, consiste en hallar la regresión lineal por mínimos cuadrados dado por [165]:

$$\ln(\mathbf{R}) = \beta_0 + \beta_1 \ln(\boldsymbol{\sigma}) \quad (3.23)$$

donde β_0 y β_1 son parámetros de la regresión lineal. El valor de β_0 es:

$$\beta_0 = \frac{\sum \ln(R_i) \sum (\ln(\sigma_i))^2 - \sum \ln(\sigma_i) \sum \ln(\sigma_i) \ln(R_i)}{N \sum (\ln(\sigma_i))^2 - (\sum \ln(\sigma_i))^2} \quad (3.24)$$

donde $N = U_3$ e $i = 1, \dots, U_3$. El valor del parámetro β_1 es:

$$\beta_1 = \frac{N \sum \ln(\sigma_i) \ln(R_i) - \sum \ln(\sigma_i) \sum \ln(R_i)}{N \sum (\ln(\sigma_i))^2 - (\sum \ln(\sigma_i))^2} \quad (3.25)$$

De esta forma, el valor de la dimensión de regularización (Eq. 2.28) se calcula con:

$$\Lambda_{s,v,n_1,n_2}^p = 1 - \beta_1 \quad (3.26)$$

donde s es el sujeto, v es la vocal o sílaba con habla silenciosa y n_1 hasta n_2 es el rango de filas de $\mathbb{W}_{s,v}$.

Finalmente, la matriz resultante de dimensión de regularización para todos los valores de $\mathbb{W}_{s,v}$ con habla silenciosa es:

$$\mathbf{\Lambda}_{s,v} = \begin{pmatrix} \Lambda_{s,v,0,249}^1 & \dots & \Lambda_{s,v,0,249}^{21} \\ \vdots & \ddots & \vdots \\ \Lambda_{s,v,14744,14999}^1 & \dots & \Lambda_{s,v,14744,14999}^{21} \end{pmatrix} \quad (3.27)$$

donde el tamaño de la matriz $\Lambda_{s,v}$ es 60x21.

Para el caso de sílabas con habla silenciosa, la matriz resultante de dimensión de regularización para todos los valores de $\mathbb{W}_{s,v}$ es

$$\Lambda_{s,v} = \begin{pmatrix} \Lambda_{s,v,0,399}^1 & \cdots & \Lambda_{s,v,0,399}^{21} \\ \vdots & \ddots & \vdots \\ \Lambda_{s,v,11599,11999}^1 & \cdots & \Lambda_{s,v,11599,11999}^{21} \end{pmatrix} \quad (3.28)$$

donde el tamaño de la matriz $\Lambda_{s,v}$ es 30x21.

3.2.4. Clasificador multiclase con máquinas de soporte vectorial

Una máquina de soporte vectorial, o en inglés *support vector machine* (SVM), es una máquina de aprendizaje supervisada en la cual un conjunto de vectores característicos (datos de entrenamiento) es transformado en un vector de más alta dimensión $\Omega(\cdot)$ con datos separables linealmente y máximo margen de separación entre clases [166]. Dentro de los clasificadores multiclase que utilizan SVM, se encuentran los métodos *one against-rest* (1-rest) y *one against-one* (1-1). Si hay M clases, el método *one against-rest* descompone el problema en M problemas binarios, mientras el método *one against-one* descompone el problema en $M(M - 1)/2$ problemas binarios. La metodología basada en vector de características con entropía de la información y la dimensión de regularización hace uso de los dos métodos de clasificación multiclase [167].

Clasificador multiclase SVM con *one against-rest*

El clasificador con el método *one against-rest*, o uno contra todos, (1-rest) toma las señales de $\mathbf{H}_{s,v}$, resultado de la entropía de información, y de $\Lambda_{s,v}$, resultado de la dimensión de regularización; con esta señales se implementa el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \Lambda_{s,v}]$ por sujeto s y por vocal o sílaba v . Con los valores $\Xi_{s,v}$ se estructura Ξ_s , los datos para el clasificador, donde:

$$\Xi_s = \begin{pmatrix} \Xi_{s,1} \\ \Xi_{s,2} \\ \Xi_{s,3} \\ \Xi_{s,4} \\ \Xi_{s,5} \end{pmatrix} \quad (3.29)$$

cada número $v = 1, \dots, 5$ representa las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) y las sílabas con habla silenciosa (/fa/, /pa/, /ma/, /la/, /ra/). El tamaño de Ξ_s es 300x42 para vocales con habla silenciosa y 150x42 para sílabas con habla silenciosa.

Los datos Ξ_s se dividen en dos conjuntos usando validación cruzada como sigue: 70 % de los datos (P_k) son seleccionados para entrenar el SVM (1-rest), y el restante 30 % (T_k) es usado para evaluación (Figs. 3.7 y 3.8). Con el método (1-rest) se descompone el problema en M problemas binarios, donde M es el número de clases (vocales o sílabas con SS) a reconocer. En este caso se desea clasificar 5 vocales o 5 sílabas con SS, y por lo tanto es necesario hallar 5 hiperplanos de la forma $(w^m)^T \Omega(P_k) + b^m = 0$ con máximo margen de separación, donde $m = 1, \dots, M$, w^m es el vector de parámetros de pesos y b^m es el vector de parámetros de offset. El conjunto de datos para entrenamiento es $(P_1, y_1), \dots, (P_l, y_l)$, $\Xi_{s,k} \in \mathbb{R}^J$, donde

J es el número de electrodos, l es el número de muestras, y cada número $k = 1, \dots, l$, y $y_k \in \{1, \dots, 5\}$ representa las cinco vocales con habla silenciosa o las cinco sílabas con habla silenciosa. El problema de optimización se establece como:

$$\begin{aligned}
&\text{encontrar} && w^m, b^m, \varepsilon^m \\
&\text{minimizar} && \frac{1}{2}(w^m)^T w^m + C \sum_{k=1}^l \varepsilon_k^m \\
&\text{sujeto a} && (w^m)^T \Omega(P_k) + b^m \geq 1 - \varepsilon_k^m, \text{ if } y_k = m \\
&&& (w^m)^T \Omega(P_k) + b^m \geq -1 + \varepsilon_k^m, \text{ if } y_k \neq m \\
&&& \varepsilon_k^m \geq 0, \quad C > 0
\end{aligned} \tag{3.30}$$

donde C es el parámetro de penalización y ε_k^m las variables relajadas (*slack*) (Figs. 3.7 y 3.8). El objetivo de minimizar la primera parte de la función objetivo es encontrar el máximo margen entre las clases, mientras el objetivo de minimizar la segunda parte de la función es encontrar el más pequeño error de las variables *slack* [167]. En (3.30), si la clase $y_k = m$ entonces $Y_k^m = +1$; de otra manera, $Y_k^m = -1$. En otras palabras, una muestra es positiva ($Y_k^m = +1$) si pertenece a la clase; de otra forma, es negativa ($Y_k^m = -1$). Los valores de la muestra para las cinco vocales con SS son mostrados en las tablas 3.1 y 3.2..

Tabla 3.1: Clasificación multiclase usando el método *one against-rest* para vocales con SS

m	Y_1^m	Y_2^m	Y_3^m	Y_4^m	Y_5^m
/a/	1	-1	-1	-1	-1
/e/	-1	1	-1	-1	-1
/i/	-1	-1	1	-1	-1
/o/	-1	-1	-1	1	-1
/u/	-1	-1	-1	-1	1

Tabla 3.2: Clasificación multiclase usando el método *one against-rest* para sílabas con SS

m	Y_1^m	Y_2^m	Y_3^m	Y_4^m	Y_5^m
/fa/	1	-1	-1	-1	-1
/pa/	-1	1	-1	-1	-1
/ma/	-1	-1	1	-1	-1
/la/	-1	-1	-1	1	-1
/ra/	-1	-1	-1	-1	1

De esta manera, la función objetivo Eq. (3.30) es cuadrática mientras que las restricciones son lineales. Los resultados de la programación cuadrática están en relación con un problema de optimización convexa que puede ser solucionado utilizando el método de los multiplicadores de Lagrange [166]. El Lagrangiano de Eq. (3.30) es:

$$\begin{aligned}
L^m &= \frac{1}{2} \|w^m\|^2 + C \sum_{k=1}^l \varepsilon_k^m - \sum_{k=1}^l \mu_k^m \varepsilon_k^m \\
&\quad - \sum_{k=1}^l \lambda_k^m [Y_k^m (w^m \cdot \Omega(P_k) + b^m) - 1 + \varepsilon_k^m],
\end{aligned} \tag{3.31}$$

donde λ_k^m representa los multiplicadores de Lagrange, y μ_k^m los parámetros relacionados con ε_k^m . El primer y segundo término están relacionados con la función objetivo, el tercer término está relacionado con la condición no negativa de los valores ε_k^m y el cuarto término

está relacionado con las restricciones. La condición necesaria de primer orden para los resultados de optimización es:

$$\frac{\partial L^m}{\partial w^m} = w^m - \sum_{k=1}^l \lambda_k^m Y_k^m \Omega(P_k) = 0, \quad (3.32)$$

$$\frac{\partial L^m}{\partial b^m} = \sum_{k=1}^l \lambda_k^m Y_k^m = 0, \quad (3.33)$$

$$\frac{\partial L^m}{\partial \varepsilon_k^m} = C - \lambda_k^m - \mu_k^m = 0. \quad (3.34)$$

El Lagrangiano es transformado a un problema dual L_d^m reemplazando (3.32), (3.33) y (3.34) en (3.31), dado por:

$$L_d^m = \sum_{k=1}^l \lambda_k^m - \frac{1}{2} \sum_{k,r=1}^l \lambda_k^m \lambda_r^m Y_k^m Y_r^m \Omega(P_k) \cdot \Omega(P_r) \quad (3.35)$$

para $r = 1, \dots, l$, la cual es una función más simple en términos de los multiplicadores de Lagrange y de los datos de entrenamiento.

Para solucionar L_d^m es necesario conocer la función de transformación no lineal $\Omega(\cdot)$ y evitar la maldición de la dimensionalidad [168]. Estos dos problemas se pueden resolver utilizando el teorema de Mercer [169], que permite usar la función kernel definida como un producto punto de los datos de entrenamiento en una alta dimensión, y así evitar problemas como la maldición de la dimensionalidad usando los datos de entrenamiento de la dimensión original, $K(P_k, P_r) = \Omega(P_k) \cdot \Omega(P_r)$. El teorema de Mercer es satisfecho si $K(P_k, P_r)$ es semidefinido positivo [166]. Reemplazando la función kernel en (3.35), resulta:

$$L_d^m = \sum_{k=1}^l \lambda_k^m - \frac{1}{2} \sum_{k,r=1}^l \lambda_k^m \lambda_r^m Y_k^m Y_r^m K(P_k, P_r). \quad (3.36)$$

Una función kernel que satisface el teorema de Mercer es la función de base radial (RBF):

$$K(P_k, P_r) = \exp\left(\frac{-\|P_k - P_r\|^2}{2\sigma^2}\right), \quad (3.37)$$

donde σ es un parámetro a ser encontrado. Los parámetros óptimos σ y C fueron encontrados con algoritmos genéticos para cada sujeto y cada vocal o sílaba con SS (ver algoritmo 1).

El algoritmo genético presenta las siguientes características [170]: posee 100 individuos para el proceso de búsqueda; cada individuo está compuesto por un cromosoma binario (0 o 1), donde cada cromosoma está compuesto de 22 lugares (locus) (11 locus para σ y 11 locus para C); la función *fitness* toma agrupaciones de dos clases, donde los valores de la población se representan como cromosomas (σ y C). Con esos datos se realiza el entrenamiento con SVM y se halla la sumatoria en el error de la clasificación; con los valores menores de *fitness*, se seleccionan los 50 más aptos; el cruce se realiza probabilísticamente entre parejas de padres, intercambiando la información de los cromosomas referidos a un locus seleccionado al azar. De esta manera se generan los individuos hijos; el valor de la probabilidad de mutación es

0.01 [171]. El resultado de la búsqueda con algoritmos genéticos halla los valores óptimos de σ y C .

Finalmente, usando el método *one against-rest*, los datos de prueba son asignados a la clase con el máximo argumento m^* . De esta forma, el valor seleccionado para la clasificación de la vocal o sílaba con SS es:

$$m^* = \operatorname{argmax}_{m=1,\dots,5} \sum_{k=1}^{l_2} ((w^m)^T \Omega(T_k) + b^m) \quad (3.38)$$

donde l_2 son la cantidad de datos para prueba.

Algorithm 1: Algoritmo genético

Data: Clases SVM

Data: Ítems a entrenar (P_k)

Data: Ítems a evaluar (T_k)

Result: Mejor σ , Mejor C

```

1 begin
2   Seleccionar el número de cromosomas crom;
3   Seleccionar el tamaño de la población ind;
4   Seleccionar la probabilidad de mutación pm;
5   Generar randómicamente una población pob (tamaño ind);
6   Seleccionar el límite de generaciones GG;
7   for  $SS \in GG$  do
8     Aplicar el fitness para cada individuo;
9     if desviación estándar del fitness es < un error permisible then
10      | break for;
11     Seleccionar los mejores individuo padr;
12     Generar nuevos individuo usando cruce hij;
13     Aplicar la probabilidad de mutación para cada nuevo individuo;
14     Generar una nueva población con los mejores nuevos individuos;
15 return Mejor  $\sigma$  , Mejor  $C$ 
```

Clasificador multiclase SVM con *one against-one*

El clasificador con el método *one against-one*, o uno contra uno (1-1), toma las señales de $\mathbf{H}_{s,v}$, resultado de la entropía de información, y $\mathbf{\Lambda}_{s,v}$, resultado de la dimensión de regularización; con estas, señales se implementa el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$ por sujeto s y por vocal o sílaba v . Con los valores $\Xi_{s,v}$ se estructura Ξ_s ; y estos serán los datos para el clasificador, donde:

$$\Xi_s = \begin{pmatrix} \Xi_{s,1} \\ \Xi_{s,2} \\ \Xi_{s,3} \\ \Xi_{s,4} \\ \Xi_{s,5} \end{pmatrix} \quad (3.39)$$

Cada número $v = 1, \dots, 5$ representa las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) y las sílabas con habla silenciosa (/fa/, /pa/, /ma/, /la/, /ra/). El tamaño de Ξ_s es 300x42 para vocales con habla silenciosa y 150x42 para sílabas con habla silenciosa.

Los datos Ξ_s se dividen en dos conjuntos usando validación cruzada como sigue: 70% de los datos (P_k) son seleccionados para entrenar el SVM (1-1), y el restante 30% (T_k) es usado para evaluación (Figs. 3.7 y 3.8). Con el método (1-1) se descompone el problema en $M(M-1)/2$ problemas binarios, donde M es el número de clases (vocales o sílabas con SS) a reconocer. En este caso se desea clasificar 5 vocales o 5 sílabas con SS, y por lo tanto resulta necesario hallar 10 hiperplanos de la forma $(w^m)^T \Omega(P_k) + b^m = 0$ con máximo margen de separación donde $m = 1, \dots, M$, w^m es el vector de parámetros de pesos, y b^m es el vector de parámetros de offset. El conjunto de datos para entrenamiento es $(P_1, y_1), \dots, (P_l, y_l)$, $\Xi_{s,k} \in \mathbb{R}^J$, donde J es el número de electrodos, l es el número de muestras, $k = 1, \dots, l$, y en $y_k \in \{1, \dots, 5\}$ cada número representa las cinco vocales con habla silenciosa o las cinco sílabas con habla silenciosa. El problema de optimización se establece como:

$$\begin{aligned}
&\text{encontrar} && w^{hq}, b^{hq}, \varepsilon^{hq} \\
&\text{minimizar} && \frac{1}{2}(w^{hq})^T w^{hq} + C \sum_{k=1}^l \varepsilon_k^{hq} \\
&\text{sujeito a} && (w^{hq})^T \Omega(P_k) + b^{hq} \geq 1 - \varepsilon_k^{hq}, \text{ if } y_k = h \\
&&& (w^{hq})^T \Omega(P_k) + b^{hq} \geq -1 + \varepsilon_k^{hq}, \text{ if } y_k = q \\
&&& \varepsilon_k^{hq} \geq 0, C > 0
\end{aligned} \tag{3.40}$$

donde h y q son la h -ésima y q -ésima clase, C es el parámetro de penalidad y ε_k^{hq} las variables relajadas (*slack*). Se busca minimizar el primer término de la función objetivo para encontrar el máximo margen entre las clases (Eq. (3.61)), mientras que la idea de minimizar la segunda parte de la función objetivo es encontrar el más pequeño error de las variables *slack* [167]. Los valores de las clases para (y_h, y_k) se exponen en la tabla 3.3.

Tabla 3.3: Clasificación multiclase usando el método *one against-one*

y_h	$y_1 (+1)$	$y_1 (+1)$	$y_1 (+1)$	$y_1 (+1)$	$y_2 (+1)$	$y_2 (+1)$	$y_2 (+1)$	$y_3 (+1)$	$y_3 (+1)$	$y_4 (+1)$
y_q	$y_2 (-1)$	$y_3 (-1)$	$y_4 (-1)$	$y_5 (-1)$	$y_3 (-1)$	$y_4 (-1)$	$y_5 (-1)$	$y_4 (-1)$	$y_5 (-1)$	$y_5 (-1)$

La función objetivo Eq. (3.61) es cuadrática mientras que las restricciones son lineales. El resultado de la programación cuadrática corresponde a un problema de optimización convexa para ser solucionado utilizando el método de los multiplicadores de Lagrange [166]. El Lagrangiano de Eq. (3.61) es:

$$\begin{aligned}
L^{hq} = & \frac{1}{2} \|w^{hq}\|^2 + C \sum_{k=1}^l \varepsilon_k^{hq} - \sum_{k=1}^l \mu_k^{hq} \varepsilon_k^{hq} - \\
& \sum_{k=1}^l \lambda_k^{hq} \left[Y_k^{hq} (w^{hq} \Omega(P_k) + b^{hq}) - 1 + \varepsilon_k^{hq} \right],
\end{aligned} \tag{3.41}$$

donde λ_k^{hq} representa los multiplicadores de Lagrange, Y_k^{hq} están relacionadas con las clases (y_k^h, y_k^q) y μ_k^{hq} son parámetros relacionados con ε_k^{hq} . El primer y segundo término están relacionados con la función objetivo (Eq. (3.41)), el tercero con la condición no negativa en los valores ε_k^{hq} y el cuarto con las restricciones. La condición necesaria para el resultado de optimización es:

$$\frac{\partial L^{hq}}{\partial w^{hq}} = w^{hq} - \sum_{k=1}^l \lambda_k^{hq} Y_k^{hq} \Omega(P_k) = 0, \tag{3.42}$$

$$\frac{\partial L^{hq}}{\partial b^{hq}} = \sum_{k=1}^l \lambda_k^{hq} Y_k^{hq} = 0, \quad (3.43)$$

$$\frac{\partial L^{hq}}{\partial \varepsilon_k^{hq}} = C - \lambda_k^{hq} - \mu_k^{hq} = 0. \quad (3.44)$$

El Lagrangiano es transformado al Lagrangiano dual L_d^{hq} reemplazando (3.42), (3.43) y (3.44) en (3.41), dado por:

$$L_d^{hq} = \sum_{k=1}^l \lambda_k^{hq} - \frac{1}{2} \sum_{k,r=1}^l \lambda_k^{hq} \lambda_r^{hq} Y_k^{hq} Y_r^{hq} \Omega(P_k) \cdot \Omega(P_r) \quad (3.45)$$

para $r = 1, \dots, l$, que es la función más simple en términos de los multiplicadores de Lagrange y los datos de entrenamiento.

Con el propósito de solucionar L_d^{hq} , se requiere conocer la función de transformación no lineal $\Omega(\cdot)$ y evitar la maldición de la dimensionalidad [168]. Estos dos problemas pueden ser solucionados aplicando el teorema de Mercer [169]. Este teorema permite usar la función kernel definida como un producto punto de los datos de entrenamiento en una alta dimensión, y evitar la maldición de la dimensionalidad usando los datos de entrenamiento de la dimensión original, $K(P_k, P_r) = \Omega(P_k) \cdot \Omega(P_r)$. El teorema de Mercer es satisfecho si $K(P_k, P_r)$ es semidefinido positivo [166]. Reemplazando la función kernel en (3.45), resulta en:

$$L_d^{hq} = \sum_{k=1}^l \lambda_k^{hq} - \frac{1}{2} \sum_{k,r=1}^l \lambda_k^{hq} \lambda_r^{hq} Y_k^{hq} Y_r^{hq} K(P_k, P_r). \quad (3.46)$$

Una función kernel que satisface el teorema de Mercer es la función de base radial (RBF):

$$K(P_k, P_r) = \exp\left(\frac{-\|P_k - P_r\|^2}{2\sigma^2}\right), \quad (3.47)$$

donde σ es un parámetro a ser encontrado. Los parámetros óptimos C y σ fueron seleccionados con un algoritmo genético para cada sujeto y cada vocal (ver algoritmo 1).

Finalmente, si se desea conocer la clase a la cual pertenece T_k , entonces debe aplicarse la ecuación $sign((w^{hq})^T \Omega(T_k) + b^{hq})$. En el caso en que T_k pertenezca a la m -ésima clase, entonces el voto es para la m -ésima clase. La clase ganadora es la que tenga la mayor votación.

3.3. Metodología basada en vector de características con datos funcionales

El segundo modelo de habla silenciosa desarrollado en este estudio, es un algoritmo que toma los registros de las grabaciones realizadas a los sujetos en tareas de vocales con habla silenciosa (Fig. 3.9) o sílabas con habla silenciosa (Fig. 3.10). Estas señales pasan a la etapa de filtraje, donde se seleccionan las frecuencias entre 2 Hz y 13 Hz ($T_{s,v}^p$). Posteriormente, se pasa a un bloque de detección y eliminación de artefactos debidos al parpadeo ($V_{s,v}^p$). A continuación, viene un bloque de selección de intervalos en donde se piensa en la vocal o sílaba con habla silenciosa ($W_{s,v}^p$). Después sigue un bloque donde se obtiene el PSD estimado (periodograma) de la señal para las vocales con SS (ritmos δ , $\delta\theta$ y $\delta\theta\alpha$) y para sílabas (ritmos

δ) con SS ($X_{s,v}^p$). A continuación se obtienen las funciones de los datos funcionales ($\hat{g}_{s,v}(\Theta)$) y su primera derivada ($D\hat{g}_{s,v}(\Theta)$), que se hallan a partir de la posición angular de los electrodos Θ y del valor del PSD estimado. Luego se implementa el vector de características con los coeficientes de los datos funcionales y sus derivadas $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$, y después por un proceso de validación cruzada se selecciona el conjunto de datos para entrenamiento y el conjunto de datos para prueba. Para el proceso de clasificación multiclase fue seleccionada una máquina de soporte vectorial (SVM) uno contra uno (1-1). Los parámetros del SVM fueron hallados por medio de optimización con algoritmos genéticos. A continuación, se explican los respectivos bloques.

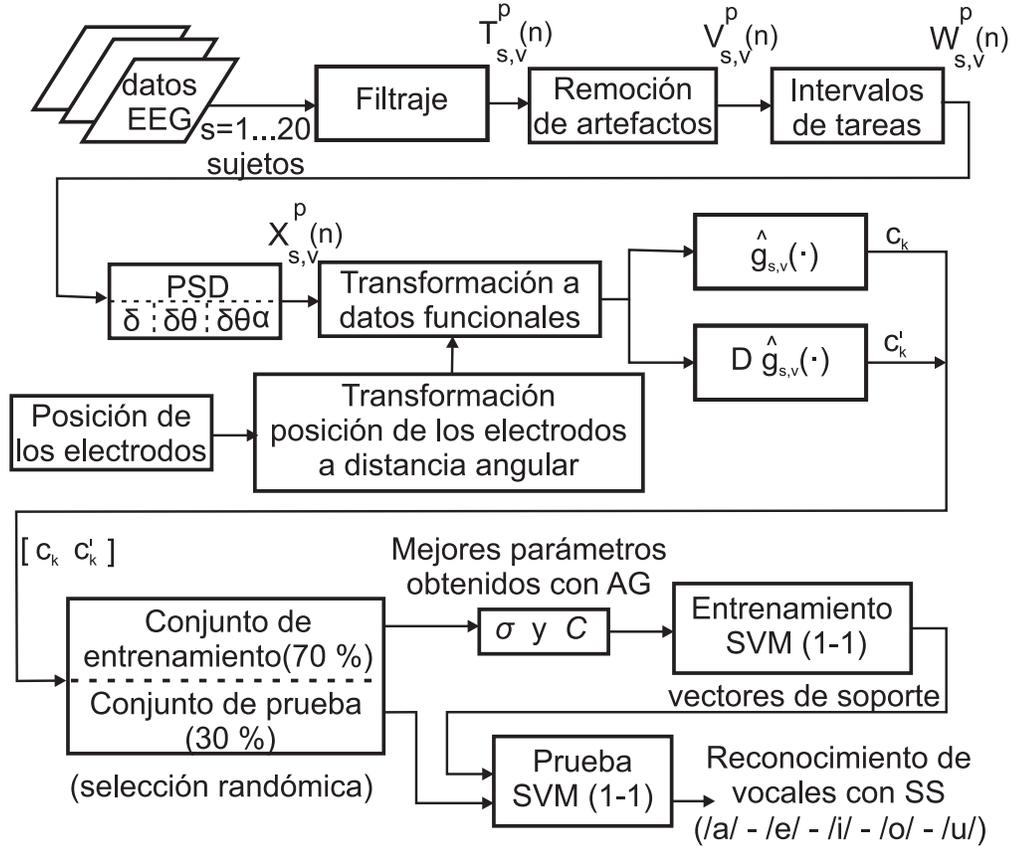


Figura 3.9: Diagrama del modelo propuesto con datos funcionales para vocales con SS

3.3.1. Filtrado

En el filtrado de señales cerebrales (Fig. 3.9 y Fig. 3.10), dos filtros elípticos con respuesta de impulso infinita (IIR) son propuestos, un filtro pasa alto (HPF) y un filtro pasa bajo (LPF) [160]. En el caso de las vocales se utiliza una frecuencia de muestreo $F_s = 500 \text{ Hz}$; el filtro pasa alto, por su parte, tiene una frecuencia de corte de 2 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 2 dB y un orden de 6. El filtro pasa bajo tiene una frecuencia de corte de 13 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 1 dB y un orden de 6. En el caso de las sílabas se utiliza una frecuencia de muestreo $F_s = 2000 \text{ Hz}$; el filtro pasa alto tiene una frecuencia de corte de 2 Hz , una banda de atenuación de 80 dB , una ganancia de paso de 2 dB y un orden de 6. El filtro pasa bajo tiene una frecuencia de

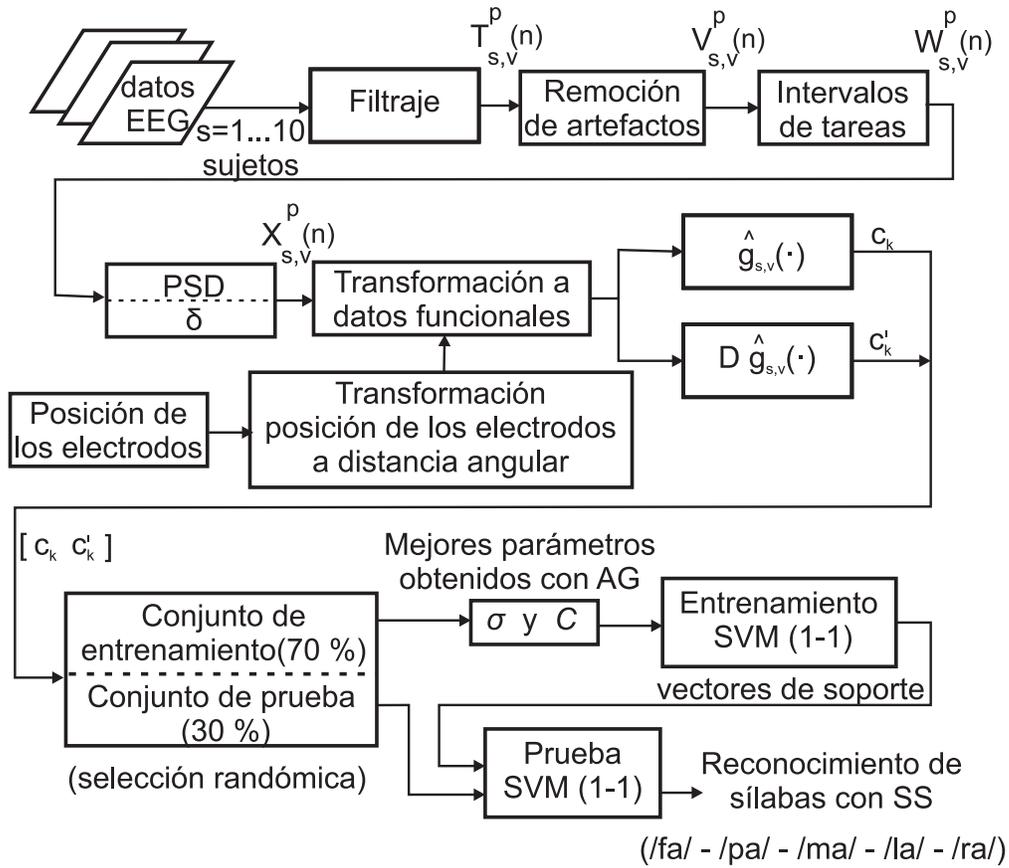


Figura 3.10: Diagrama del modelo propuesto con datos funcionales para sílabas con SS

corte de 13 Hz , una banda de atenuación de 80 dB, una ganancia de paso de 1 dB y un orden de 6.

Para el caso de vocales con SS, $T_{s,v}^p(n)$ es la señal filtrada. El subíndice s se refiere al sujeto, $s = 1, \dots, 20$, p se refiere al número de electrodos, $p = 1, \dots, 21$, y v se refiere a la vocal con habla silenciosa, $v = 1, \dots, 5$ (/a/, /e/, /i/, /o/ y /u/). La variable n denota las muestras de tiempo discreto. Los segmentos grabados proveen un total de 30000 (60 segundos) datos por vocal con habla silenciosa, por electrodo y por sujeto (Figs. 3.5).

Para el caso de las sílabas con SS, $T_{s,v}^p(n)$ es la señal filtrada. El subíndice s se refiere al sujeto, $s = 1, \dots, 10$, p se refiere al número de electrodos, $p = 1, \dots, 21$, y v se refiere a la sílaba con habla silenciosa, $v = 1, \dots, 5$ (/fa/, /pa/, /ma/, /la/ y /ra/). La variable n denota las muestras de tiempo discreto. Los segmentos grabados proveen un total de 244000 (122 segundos) datos por sílaba con habla silenciosa, por electrodo y por sujeto (Figs. 3.6).

3.3.2. Remoción artefactos de parpadeo

La remoción de artefactos por parpadeo para el método de características con datos funcionales fue realizada en dos fases (Fig. 3.9 y Fig. 3.10). La primera fase consiste en detectar el parpadeo y la segunda fase en la remoción automática del artefacto. El algoritmo para remoción de artefactos es el mismo que el descrito en la metodología basada en el vector de características con entropía de la información y la dimensión de regularización.

3.3.3. Transformación a densidad espectral de potencia

En esta sección se estima la densidad espectral de potencia, en inglés *power spectral density* (PSD), de las señales cerebrales $W_{s,v}^p(n)$. El PSD es una medida estadística que cuantifica la potencia de una señal relacionada con los componentes de frecuencia [172]. Entre las características del PSD está la posibilidad de detectar señales inmersas en ruido, y la posibilidad de utilizar señales modeladas como procesos randómicos estacionarios [173]. Si $\chi(t)$ es un proceso randómico estacionario de una vocal o sílaba con habla silenciosa (SS), entonces su función de autocorrelación ρ_{xx} es dada por

$$\rho_{xx}(\tau) = \int_{-\infty}^{\infty} \chi^*(t)\chi(t+\tau)dt, \quad (3.48)$$

$\chi^*(\cdot)$ es la conjugada de $\chi(\cdot)$, t es la variable tiempo y τ es la variable de desfase [172].

De esta forma, el teorema de Wiener-Khintchine [172] define el PSD de un proceso randómico estacionario como la transformada de Fourier de su función de autocorrelación definida como:

$$PSD(F) = \int_{-\infty}^{\infty} \rho_{xx}(\tau)e^{-j2\pi F\tau}d\tau \quad (3.49)$$

donde $\mathbf{j} = \sqrt{-1}$ y F es la variable frecuencia en tiempo continuo de $\chi(t)$. En la práctica, únicamente una realización de un proceso randómico estacionario $\chi(t)$ es habilitada, de manera que $\rho_{xx}(\tau)$ no es conocido y no es posible calcular el $PSD(F)$. En este estudio, el periodograma es usado para estimar el PSD [172] dado por:

$$X_{s,v}^p(f) = \frac{1}{A} \left| \sum_{n=0}^{A-1} W_{s,v}^p(n)e^{-j2\pi fn} \right|^2 \quad (3.50)$$

donde $X_{s,v}^p(f)$ es el periodograma, f es la frecuencia de la señal discreta, n es el tiempo discreto, A es la longitud de los datos para un sujeto s , una vocal o sílaba v y un electrodo p . Es necesario que $F_s > 2Q$, donde F_s es la frecuencia de muestreo relacionada a los datos y Q es la frecuencia más alta de las señales cerebrales con habla silenciosa. En este estudio, para vocales con habla silenciosa se utiliza $F_s = 500 \text{ Hz}$ y para sílabas con habla silenciosa se utiliza $F_s = 2000 \text{ Hz}$.

Una ventana de Hamming $\nu_H(n) = 0.54 - 0.46\cos(2\pi n/A)$ es calculada, donde n es el tiempo discreto $A = 250$ (0.5 s) para vocales con habla silenciosa y $A = 400$ (0.2 s) para sílabas con habla silenciosa. Estos valores se utilizan para calcular el periodograma de cada columna de $W_{s,v}^p(n)$. En este estudio, se selecciona para el ventaneo un solapamiento del 50%. Así, para cada columna de $W_{s,v}^p(n)$, sujeto s , vocal o sílaba v y electrodo p , el periodograma es calculado usando:

$$X_{s,v,n_1,n_2}^p(f) = \frac{1}{A} \left| \sum_{n=n_1}^{n_1+A-1} W_{s,v}^p(n)\nu_H(n)e^{-j2\pi fn} \right|^2 \quad (3.51)$$

donde $X_{s,v,n_1,n_2}^p(f)$ corresponde al periodograma entre n_1 y n_2 , donde n_1 es una fila de $W_{s,v}^p(n)$ y $n_2 = n_1 + A - 1$.

Con este método se desarrollan tres algoritmos basados en los ritmos δ , $\delta\theta$ y $\delta\theta\alpha$ para el estudio de las características de las vocales con habla silenciosa

En el caso de vocales con habla silenciosa que emplean ritmos cerebrales δ , se utilizan las frecuencias $f = 2, 3, 4 \text{ Hz}$. Los datos $X_{s,v,n_1,n_2}^p(f)$ son almacenados en una matriz $\mathbb{X}_{s,v}$, y los valores del PSD estimado desde $f = 2$ a 4 Hz son mostrados en las primeras 3 filas de $\mathbb{X}_{s,v}$. El PSD estimado de los otros intervalos es almacenado debajo de los anteriores.

$$\mathbb{X}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,0,249}^1(4) & \cdots & X_{s,v,0,249}^{21}(4) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(4) & \cdots & X_{s,v,14749,14999}^{21}(4) \end{pmatrix} \quad (3.52)$$

Siendo f_{min} la frecuencia mínima, f_{max} la frecuencia máxima y el rango de frecuencias $Range = f_{max} - f_{min} + 1$, el número de filas de \mathbb{X} es $((mm/A)Range)$, donde mm es el número de filas de \mathbb{W} y $A = 250$ es el ancho de la ventana de Hamming. En este caso, $f_{min} = 2 \text{ Hz}$, $f_{max} = 4 \text{ Hz}$ y $Range = 3$. El tamaño de $\mathbb{X}_{s,v}$ para cada sujeto y cada vocal es 180×21 .

En el caso de vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta$, se emplea el rango de frecuencias $f = 2, \dots, 8 \text{ Hz}$. Los datos $X_{s,v,n_1,n_2}^p(f)$ son almacenados en una matriz $\mathbb{X}_{s,v}$ y los valores del PSD estimado desde $f = 2$ a 8 Hz son expuestos en las primeras 7 filas de $\mathbb{X}_{s,v}$. El PSD estimado de los otros intervalos es almacenado debajo de los anteriores.

$$\mathbb{X}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,0,249}^1(8) & \cdots & X_{s,v,0,249}^{21}(8) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(8) & \cdots & X_{s,v,14749,14999}^{21}(8) \end{pmatrix} \quad (3.53)$$

Siendo f_{min} la frecuencia mínima, f_{max} la frecuencia máxima y el rango de frecuencias $Range = f_{max} - f_{min} + 1$, el número de filas de \mathbb{X} es $((mm/A)Range)$, donde mm es el número de filas de \mathbb{W} y $A = 250$ es el ancho de la ventana de Hamming. En este caso, $f_{min} = 2 \text{ Hz}$, $f_{max} = 8 \text{ Hz}$ y $Range = 7$. El tamaño de $\mathbb{X}_{s,v}$ para cada sujeto y cada vocal es 420×21 .

En el caso de vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta\alpha$, se emplea el rango de frecuencias $f = 2, \dots, 13 \text{ Hz}$. Los datos $X_{s,v,n_1,n_2}^p(f)$ son almacenados en una matriz $\mathbb{X}_{s,v}$ y los valores del PSD estimado desde $f = 2$ a 13 Hz se exponen en las primeras 12 filas de $\mathbb{X}_{s,v}$. El PSD estimado de los otros intervalos es almacenado debajo de los anteriores.

$$\mathbb{X}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,0,249}^1(13) & \cdots & X_{s,v,0,249}^{21}(13) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(13) & \cdots & X_{s,v,14749,14999}^{21}(13) \end{pmatrix} \quad (3.54)$$

Siendo f_{min} la frecuencia mínima, f_{max} la frecuencia máxima y el rango de frecuencias $Range = f_{max} - f_{min} + 1$, el número de filas de \mathbb{X} es $((mm/A)Range)$, donde mm es el número de filas de \mathbb{W} y $A = 250$ es el ancho de la ventana de Hamming. En este caso, $f_{min} = 2 Hz$, $f_{max} = 13 Hz$ y $Range = 12$. El tamaño de $\mathbb{X}_{s,v}$ para cada sujeto y cada vocal es 720×21 .

En el caso de sílabas con habla silenciosa, el ritmo cerebral seleccionado es δ para las frecuencias $f = 2, 3, 4 Hz$. Los datos $X_{s,v,n_1,n_2}^p(f)$ son almacenados en una matriz $\mathbb{X}_{s,v}$ y los valores del PSD estimado desde $f = 2$ a $4 Hz$ son expuestos en las primeras 3 filas de $\mathbb{X}_{s,v}$. El PSD estimado de los otros intervalos es almacenado debajo de los anteriores.

$$\mathbb{X}_{s,v} = \begin{pmatrix} X_{s,v,0,399}^1(2) & \cdots & X_{s,v,0,399}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,0,399}^1(4) & \cdots & X_{s,v,0,399}^{21}(4) \\ \vdots & \ddots & \vdots \\ X_{s,v,11549,11999}^1(2) & \cdots & X_{s,v,11549,11999}^{21}(2) \\ \vdots & \ddots & \vdots \\ X_{s,v,11549,11999}^1(4) & \cdots & X_{s,v,11549,11999}^{21}(4) \end{pmatrix} \quad (3.55)$$

Siendo f_{min} la frecuencia mínima, f_{max} la frecuencia máxima y, el rango de frecuencias $Range = f_{max} - f_{min} + 1$, el número de filas de \mathbb{X} es $((mm/A)Range)$, donde mm es el número de filas de \mathbb{W} y $A = 400$ es el ancho de la ventana de Hamming. En este caso, $f_{min} = 2 Hz$, $f_{max} = 4 Hz$ y $Range = 3$. El tamaño de $\mathbb{X}_{s,v}$ para cada sujeto y cada vocal es 90×21 .

Finalmente, las señales cerebrales X_{s,v,n_1,n_2}^p son consideradas como valores discretos en una matriz $\mathbb{X}_{s,v}$ y pertenecen a un proceso randómico estacionario χ [151].

3.3.4. Transformación posición de los electrodos y PSD a datos funcionales

El objetivo de esta sección es transformar el PSD estimado ($X_{s,v}^p(f)$) de las señales cerebrales, considerando la distancia angular Θ_i de los electrodos $i = 1, \dots, 21$ en curvas. En este sentido, el área de la estadística llamada datos funcionales analiza curvas o superficies, analizándolas como realizaciones en un espacio continuo de Hilbert L^2 [150].

Aplicación de los datos funcionales al habla silenciosa

El primer paso de este proceso consiste en transformar la matriz $\mathbb{X}_{s,v}$ en la matriz $\mathbb{B}_{s,v}$. Aquí se explica la transformación para vocales con habla silenciosa para los ritmos δ , $\delta\theta$ y $\delta\theta\alpha$. Posteriormente, se explica la transformación para sílabas con habla silenciosa para el ritmo δ .

En el caso de vocales con habla silenciosa para ritmos δ , se transforman los datos del PSD estimado ($\mathbb{X}_{s,v}$) organizando cada $Range = 3$ filas de $\mathbb{X}_{s,v}$ en una sola fila. Con este cambio se busca agrupar las 3 frecuencias $f = 2, 3, 4 Hz$ resultantes de cada uno de los periodogramas en una única fila. Cada uno de estos datos se almacena en $\mathbb{B}_{s,v}$, donde

$$\mathbb{B}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) & \cdots & X_{s,v,0,249}^1(4) & \cdots & X_{s,v,0,249}^{21}(4) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) & \cdots & X_{s,v,14749,14999}^1(4) & \cdots & X_{s,v,14749,14999}^{21}(4) \end{pmatrix} \quad (3.56)$$

Para vocales con habla silenciosa con ritmos δ , la matriz $\mathbb{X}_{s,v}$ con tamaño 180x21 es transformada en la matriz \mathbb{B} con tamaño $Q \times M$, donde $Range = 3$, $Q = 180/Range = 60$ y $M = (21)(Range) = 63$.

En el caso de vocales con habla silenciosa para ritmos $\delta\theta$, se transforman los datos del PSD estimado ($\mathbb{X}_{s,v}$) organizando cada $Range = 7$ filas de $\mathbb{X}_{s,v}$ en una sola fila. Con este cambio se busca agrupar las 7 frecuencias $f = 2, \dots, 8$ Hz resultantes de cada uno de los periodogramas en una única fila. Cada uno de estos datos se almacena en $\mathbb{B}_{s,v}$, donde

$$\mathbb{B}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) & \cdots & X_{s,v,0,249}^1(8) & \cdots & X_{s,v,0,249}^{21}(8) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) & \cdots & X_{s,v,14749,14999}^1(8) & \cdots & X_{s,v,14749,14999}^{21}(8) \end{pmatrix} \quad (3.57)$$

Para vocales con habla silenciosa con ritmos $\delta\theta$, la matriz $\mathbb{X}_{s,v}$ con tamaño 420x21 es transformada en la matriz \mathbb{B} con tamaño $Q \times M$, donde $Range = 12$, $Q = 420/Range = 60$ y $M = (21)(Range) = 147$.

En el caso de vocales con habla silenciosa para ritmos $\delta\theta\alpha$, se transforman los datos del PSD estimado ($\mathbb{X}_{s,v}$) organizando cada $Range = 12$ filas de $\mathbb{X}_{s,v}$ en una única una fila. Con este cambio se agrupan las 12 frecuencias $f = 2, \dots, 13$ Hz resultantes de cada uno de los periodogramas en una sola fila. Cada uno de estos datos se almacena en $\mathbb{B}_{s,v}$, donde

$$\mathbb{B}_{s,v} = \begin{pmatrix} X_{s,v,0,249}^1(2) & \cdots & X_{s,v,0,249}^{21}(2) & \cdots & X_{s,v,0,249}^1(13) & \cdots & X_{s,v,0,249}^{21}(13) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{s,v,14749,14999}^1(2) & \cdots & X_{s,v,14749,14999}^{21}(2) & \cdots & X_{s,v,14749,14999}^1(13) & \cdots & X_{s,v,14749,14999}^{21}(13) \end{pmatrix} \quad (3.58)$$

Para vocales con habla silenciosa con ritmos $\delta\theta\alpha$, la matriz $\mathbb{X}_{s,v}$ con tamaño 720x21 es transformado en la matriz \mathbb{B} con tamaño $Q \times M$, donde $Range = 12$, $Q = 720/Range = 60$ y $M = (21)(Range) = 252$.

En el caso de sílabas con habla silenciosa para ritmos δ , se transforma los datos del PSD estimado ($\mathbb{X}_{s,v}$) organizando cada $Range = 3$ filas de $\mathbb{X}_{s,v}$, a únicamente una fila. Con este cambio, se quiere agrupar las 3 frecuencias $f = 2, 3, 4$ Hz, resultantes de cada uno de los periodogramas, en únicamente una fila. Cada uno de estos datos se almacenan en $\mathbb{B}_{s,v}$ donde

$$\mathbb{B}_{s,v} = \begin{pmatrix} X_{s,v,0,399}^1(2) & \cdots & X_{s,v,0,399}^{21}(2) & \cdots & X_{s,v,0,399}^1(4) & \cdots & X_{s,v,0,399}^{21}(4) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{s,v,11599,11999}^1(2) & \cdots & X_{s,v,11599,11999}^{21}(2) & \cdots & X_{s,v,11599,11999}^1(4) & \cdots & X_{s,v,11599,11999}^{21}(4) \end{pmatrix} \quad (3.59)$$

Para sílabas con habla silenciosa con ritmos δ , la matriz $\mathbb{X}_{s,v}$ con tamaño 90x21 es transformado en la matriz \mathbb{B} con tamaño $Q \times M$, donde $Range = 3$, $Q = 90/Range = 30$ y $M = (21)(Range) = 63$.

El segundo paso de este proceso es seleccionar la función continua Θ como el dominio de la función estimada $\widehat{g}_{s,v}(\cdot)$. De esta manera, la variable seleccionada es la distancia angular de los electrodos (Eq. 2.32).

Si la posición de cada electrodo es representada por $P_i = (x_i, y_i)$ y la posición de un punto de referencia $P_r = (x_r, y_r)$, entonces la distancia angular de los electrodos es calculada con $\Theta_i = \text{atan}((y_i - y_r)/(x_i - x_r))$. Para encontrar el punto óptimo P_r se desarrolla un algoritmo genético con las siguientes características [170]: se seleccionan 64 individuos para el proceso de búsqueda; cada individuo está compuesto por un cromosoma binario (0 o 1), donde cada cromosoma está compuesto de 20 lugares (locus) (10 locus para x_r y 10 locus para y_r); la función *fitness* busca que el valor entre los ángulos de los electrodos Θ_i sea máximo y que abarquen un rango entre 0 y 360 grados. En este caso, el *fitness* está en función de la varianza entre los ángulos de los electrodos y un valor de penalización relacionado con los valores máximo y mínimo de los ángulos de los electrodos. Con los valores de *fitness* para cada individuo se seleccionan los 32 mejores y el cruce se realiza probabilísticamente entre parejas de padres, intercambiando la información de los cromosomas referidos a un locus seleccionado al azar. De esta manera se generan los individuos hijos. El valor de la probabilidad de mutación es 0.01 [171].

Para el caso de vocales con habla silenciosa, el valor encontrado por el algoritmo genético es $P_r = (30, 24.1250)$ y las respectivas distancias angulares se representan por Θ_i^* (Tabla 3.4). Para el caso de sílabas con habla silenciosa, el valor encontrado por el algoritmo genético es $P_r = (20.75, 12.75)$ y las respectivas distancias angulares se representan por Θ_i^* (Tabla 3.5).

Tabla 3.4: Valores óptimos de la distancia angular Θ_i^* para vocales con habla silenciosa, usando un algoritmo genético. E es el electrodo. Θ_i^* está en grados

i	E	Θ_i^*	i	E	Θ_i^*
1	$E16$	8.9506	12	$E6$	210.1290
2	$E15$	13.0408	13	$E1$	239.8729
3	$E14$	23.6294	14	$E2$	274.7391
4	$E21$	35.0768	15	$E7$	283.8287
5	$E20$	52.9864	16	$E3$	306.7271
6	$E13$	75.7500	17	$E4$	324.6421
7	$E19$	85.2115	18	$E5$	334.2427
8	$E18$	120.3868	19	$E8$	335.7061
9	$E17$	141.4861	20	$E9$	346.5600
10	$E12$	150.6422	21	$E10$	350.7701
11	$E11$	165.2917	—	—	—

En consecuencia, para representar adecuadamente la abscisa de $\mathbb{B}_{s,v}$ para vocales o sílabas con habla silenciosa, en términos de la variable distancia angular Θ , se llevan a cabo los siguientes pasos: primero, se toman los valores de Θ_i^* y se concatenan a la derecha con un vector inicialmente vacío $\underline{\Theta}$; después, a cada valor de Θ_i^* se le suma 360 y se concatena este valor a la derecha de $\underline{\Theta}$. Este proceso se repite hasta que la longitud del vector $\underline{\Theta}$ sea igual a M , que corresponde al número de columnas de $B_{s,v}(f)$.

En el tercer paso, se selecciona la función B-spline para estimar los datos funcionales porque las señales cerebrales de habla silenciosa son representadas como señales randómicas (Eq. 2.37). En este estudio, un orden $norden = 4$ fue seleccionado. Por lo tanto, el grado de

Tabla 3.5: Valores óptimos de la distancia angular Θ_i^* para sílabas con habla silenciosa, usando un algoritmo genético. E es el electrodo. Θ_i^* está en grados

i	E	Θ_i^*	i	E	Θ_i^*
1	$E10$	4.2974	12	$E21$	112.0419
2	$E9$	6.8013	13	$E16$	120.4833
3	$E8$	16.1134	14	$E11$	137.1476
4	$E15$	23.9932	15	$E7$	145.6197
5	$E14$	35.2382	16	$E6$	171.0983
6	$E20$	39.1810	17	$E1$	211.5690
7	$E19$	52.2942	18	$E2$	249.5612
8	$E13$	59.2942	19	$E3$	311.4237
9	$E18$	72.2926	20	$E4$	334.9256
10	$E17$	97.6745	21	$E5$	343.5746
11	$E12$	103.8610	—	—	—

los splines fue 3 (splines cúbicos).

En el cuarto paso, se hallan los valores óptimos de los coeficientes c_k (Eq. 2.44) usando estimación con mínimos cuadrados [150].

El siguiente paso es encontrar el número óptimo de bases K para estimar $\hat{g}_{s,v}$ [157]. En este caso, la ecuación 2.46 es empleada. Por medio de validación cruzada se halla el valor mínimo de K . En el capítulo de resultados utilizando datos funcionales se especifica los respectivos valores de K para vocales y sílabas con habla silenciosa.

Adicionalmente, se calcula la derivada de $\hat{g}_{s,v}(\Theta)$ con la ecuación 2.47. Los coeficientes c'_k son calculados usando mínimos cuadrados, empleando la ecuación 2.49.

El vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ es estructurado con los coeficientes c_k y c'_k (Fig. 3.9 y Fig. 3.10).

Clasificador multiclase SVM con *one against-one*

El clasificador con el método *one against-one*, o uno contra uno (1-1), toma las señales de $\mathbf{c}_{s,v}$, resultado de los coeficientes de los datos funcionales, y $\mathbf{c}'_{s,v}$, resultado de la los coeficientes de la derivada de los datos funcionales. Con estas señales se implementa el vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ por sujeto s y por vocal o sílaba v . Con los valores $\Xi_{s,v}$ se estructura Ξ_s , que serán los datos para el clasificador, donde

$$\Xi_s = \begin{pmatrix} \Xi_{s,1} \\ \Xi_{s,2} \\ \Xi_{s,3} \\ \Xi_{s,4} \\ \Xi_{s,5} \end{pmatrix} \quad (3.60)$$

Cada número $v = 1, \dots, 5$ indica las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) y las sílabas con habla silenciosa (/fa/, /pa/, /ma/, /la/, /ra/). El tamaño de Ξ_s es $5Q \times 2M$ para vocales o sílabas con habla silenciosa donde $Q \times M$ es el tamaño de la matriz $\mathbb{B}_{s,v}$.

Los datos Ξ_s se dividen en dos conjuntos usando validación cruzada como sigue: 70% de los datos (P_k) son seleccionados para entrenar el SVM (1-1); el restante 30% (T_k) es usado para evaluación (Figs. 3.9 y 3.10). Con el método (1-1) se descompone el problema en

$M(M-1)/2$ problemas binarios donde M es el número de clases (vocales o sílabas con SS) a reconocer. En este caso se busca clasificar 5 vocales o 5 sílabas con SS; por lo tanto, resulta necesario hallar 10 hiperplanos de la forma $(w^m)^T \Omega(P_k) + b^m = 0$ con máximo margen de separación, donde $m = 1, \dots, M$, w^m es el vector de parámetros de pesos, y b^m es el vector de parámetros de offset. El conjunto de datos para entrenamiento es $(P_1, y_1), \dots, (P_l, y_l)$, $\Xi_{s,k} \in \mathbb{R}^J$, donde J es el número de electrodos, l es el número de muestras, $k = 1, \dots, l$, y en $y_k \in \{1, \dots, 5\}$ cada número representa las cinco vocales con habla silenciosa o las cinco sílabas con habla silenciosa. El problema de optimización se establece como:

$$\begin{aligned}
&\text{encontrar} && w^{hq}, b^{hq}, \varepsilon^{hq} \\
&\text{minimizar} && \frac{1}{2}(w^{hq})^T w^{hq} + C \sum_{k=1}^l \varepsilon_k^{hq} \\
&\text{sujeto a} && (w^{hq})^T \Omega(P_k) + b^{hq} \geq 1 - \varepsilon_k^{hq}, \text{ if } y_k = h \\
&&& (w^{hq})^T \Omega(P_k) + b^{hq} \geq -1 + \varepsilon_k^{hq}, \text{ if } y_k = q \\
&&& \varepsilon_k^{hq} \geq 0, \quad C > 0
\end{aligned} \tag{3.61}$$

donde h y q son la h -ésima y q -ésima clase, C es el parámetro de penalidad y ε_k^{hq} las variables relajadas (*slack*).

La explicación adicional de (1-1) en esta sección es igual a la previamente desarrollada en la sección “clasificador multiclase SVM con *one against-one* de la metodología basada en vector de características con entropía de información y la dimensión de regularización.

Capítulo 4

Resultados y discusión

En este capítulo se presentan los resultados y la discusión de las dos metodologías desarrolladas para habla silenciosa utilizando señales EEG. La primera metodología desarrollada de habla silenciosa se caracteriza por utilizar la entropía de la información y la dimensión de regularización. La segunda metodología desarrollada se caracteriza por utilizar la densidad espectral de potencia (periodograma) y la información espacial de los electrodos para hacer una transformación de esas señales cerebrales en datos funcionales. Se presentan los resultados obtenidos y la discusión para vocales y sílabas con habla silenciosa.

4.1. Resultados de señales EEG para vocales con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización

4.1.1. Resultados por sujeto

La figura 4.1 muestra los resultados correspondientes al valor promedio por sujeto y por vocal de la entropía de la información ($\bar{H}_{s,v}$), para vocales con habla silenciosa. Estos valores se obtuvieron con la expresión dada por:

$$\bar{H}_{s,v} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} H_{s,v,k}^p}{l_1 l_2} \quad (4.1)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.7), $l_1 = 60$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una vocal con habla silenciosa, y p es un electrodo.

De la figura (4.1) se observa que existe una variación en los valores promedio de entropía entre sujetos y entre vocales con habla silenciosa. El rango de valores promedio de entropía está entre 110 a 400 (*bits*). Los valores promedio más altos de entropía fueron obtenidos por el sujeto 3 y los mínimos valores promedio de entropía fueron obtenidos por el sujeto 4. Teniendo en cuenta el valor promedio más alto de entropía por sujeto, se encuentra que la vocal con SS /a/ tiene el valor promedio más alto en 70 % de los sujetos, y la vocal con SS /e/ tiene el segundo más alto valor promedio de entropía en 15 % de los sujetos. En contraste,

el mínimo valor promedio de entropía por sujeto, fue obtenido por la vocal con SS /u/ en el 50 % de los casos.

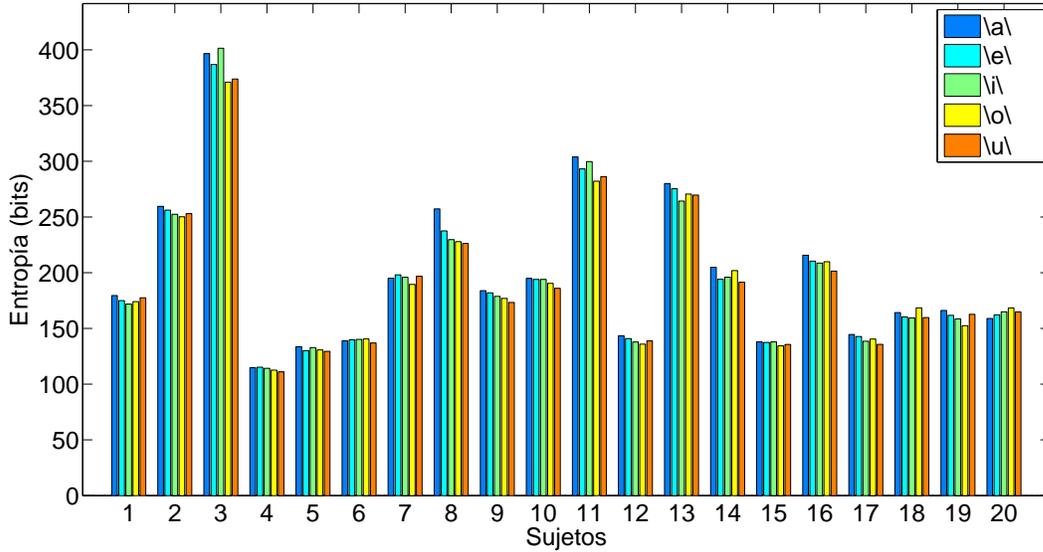


Figura 4.1: Valor promedio por sujeto y por vocal con habla silenciosa (SS) de la entropía de la información

La figura (4.2) muestra la dispersión de los datos, del valor promedio de entropía para cada vocal con SS, utilizando un diagrama de cajas o *box plots*. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.1 para cada sujeto y para cada vocal con SS (Fig. (4.1)). Los valores de la mediana para las vocales con SS (/a/, /e/, /i/, /o/, /u/) son, respectivamente, (181.70, 178.47, 175.35, 175.51 y 175.45). Los valores más altos para la mediana de la entropía de la información son para las vocales con SS /a/ y /e/. El rango intercuartil, *interquartile range* (IQR) en inglés, con la mayor dispersión es para la vocal con SS /a/ y el IQR con la menor dispersión es para la vocal con SS /u/. En el caso de la vocal con SS /a/ presenta asimetría positiva. Adicionalmente, se presentan outliers en todas las vocales con SS.

La figura 4.3 muestra los resultados correspondientes al valor promedio por sujeto y por vocal con SS de la dimensión de regularización ($\bar{\Lambda}_{s,v}$), para vocales con habla silenciosa. Estos valores se obtuvieron con la expresión dada por:

$$\bar{\Lambda}_{s,v} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{l_1 l_2} \quad (4.2)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{\Lambda}_{s,v}$ (Eq. 3.27), $l_1 = 60$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una vocal con habla silenciosa, y p es un electrodo.

De la figura (4.3) se observa que existe una variación en los valores de dimensión de regularización entre sujetos y entre vocales con habla silenciosa. El rango de valores de dimensión de regularización está entre 1.6 a 2.2. Los valores más altos de dimensión de regularización fueron obtenidos por el sujeto 14 y los mínimos valores de dimensión de regularización fueron obtenidos por el sujeto 12. Teniendo en cuenta el valor más alto de dimensión de regularización por sujeto, se encontró que la vocal con SS /a/ tiene el valor más alto en 35 % de los

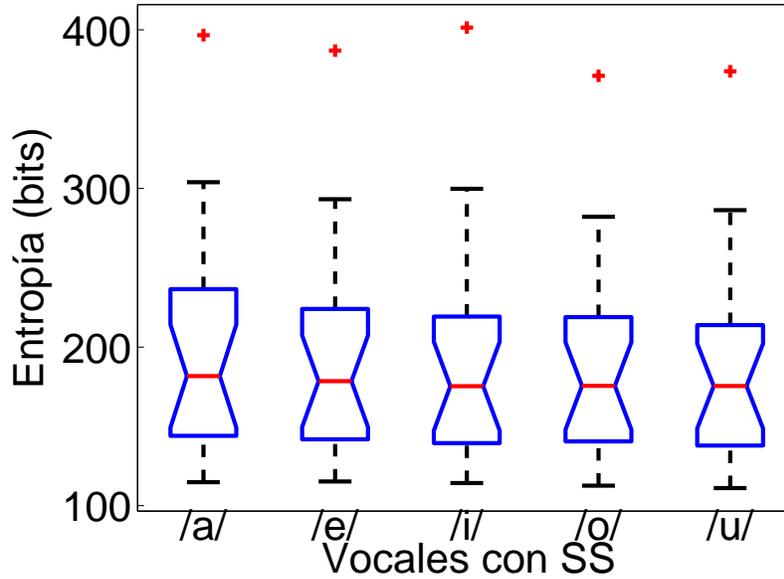


Figura 4.2: Diagrama de dispersión del valor promedio de entropía para cada vocal con habla silenciosa (SS)

sujetos, y la vocal con SS /e/ tiene el segundo valor más alto de dimensión de regularización en 30% de los sujetos. En contraste, el mínimo valor de dimensión de regularización por sujeto, fue obtenido por la vocales con SS /o/ y /u/, cada una en 30% de los casos.

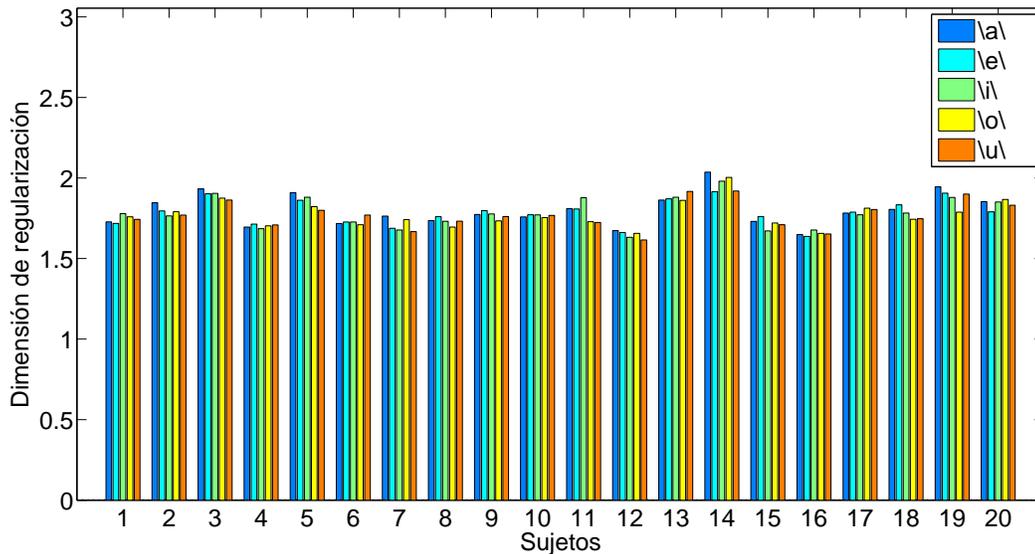


Figura 4.3: Valor promedio por sujeto y por vocal con habla silenciosa (SS) de la dimensión de regularización

La figura (4.4) muestra la dispersión de los datos, del valor promedio de la dimensión de regularización para cada vocal con SS, utilizando un diagrama de cajas o *box plots*. Los datos

para el diagrama de cajas se obtuvieron de la ecuación 4.2 para cada sujeto y para cada vocal con SS (Fig. (4.3)). Los valores de la mediana para las vocales con SS (/a/, /e/, /i/, /o/, /u/) son, respectivamente, (1.778, 1.789, 1.774, 1.749 y 1.764). El valor de la mediana es más alto para las vocales con SS /a/ y /e/ y la menor para la vocal con SS /o/. El rango intercuartil (IQR) con la mayor dispersión es para la vocal con SS /i/ y el IQR con la menor dispersión es para las vocales con SS /o/ y /u/. En el caso de las vocales con SS /a/, /i/ y /o/, se presenta asimetría positiva y para la vocal /e/ se observa asimetría negativa. Adicionalmente, se presenta un outlier en la vocal con SS /o/.

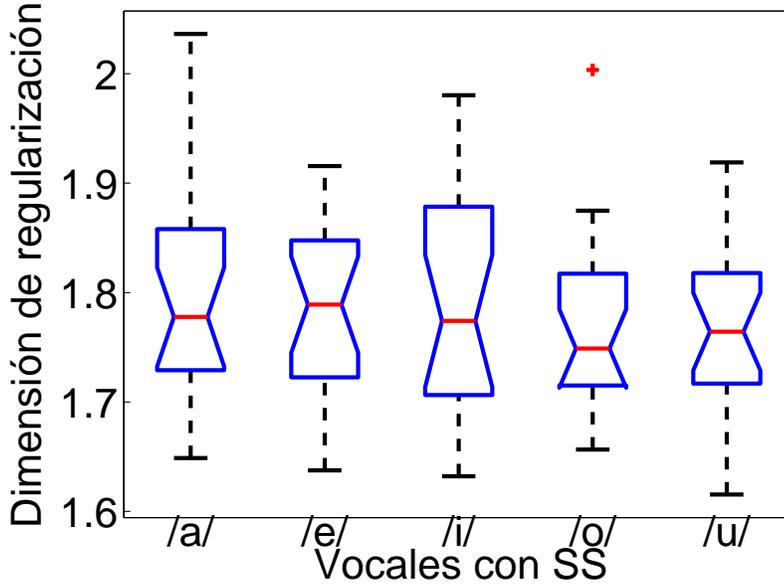


Figura 4.4: Diagrama de dispersión del valor promedio de la dimensión de regularización para cada vocal con habla silenciosa (SS)

4.1.2. Resultados por localización

Las figuras 4.5(a) a 4.5(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para la entropía de la información ($\tilde{H}_{s,v}^p$), realizando tareas de habla silenciosa con vocales. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (sujeto 14), de acuerdo a la entropía de la información (Fig. 4.10). De esta manera, el valor promedio de cada electrodo (*bits*) se obtuvo con:

$$\tilde{H}_{s,v}^p = \frac{\sum_{k=1}^{l_1} H_{s,v,k}^p}{l_1} \quad (4.3)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.7), $l_1 = 60$ es el número de filas de la matriz, k es una fila de la matriz, $s = 14$ es el sujeto seleccionado, v es una vocal con habla silenciosa, y p es un electrodo. Con los valores $\tilde{H}_{s,v}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.5(a) a 4.5(e). La distribución del color representa el valor de la entropía de la información entre 0 (azul) y 300 (rojo) (*bits*).

Los electrodos fueron posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.3). Los electrodos E1, E6, E7, E11, E12, E13, E14, E17, E18, E19 fueron ubicados sobre el área de Broca; los electrodos E2, E4, E5, E8, E9, E10, E15, E16 y E21 sobre el área de Wernicke, y el electrodo E3 sobre la corteza auditiva primaria. Estas áreas fueron seleccionadas por ser de crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las vocales con SS se presentan los siguientes resultados:

Primero, la vocal con SS /a/ (Fig. 4.5 (a)), tiene un valor promedio de posición para la entropía (*bits*) más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E4, E5, E9, E10, E15 y E16. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E11 y E12. En este caso, el área de Wernicke tiene un valor más alto de entropía que el área de Broca. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 260 (*bits*).

Segundo, la vocal con SS /e/ (Fig. 4.5 (e)) tiene un valor promedio de posición para la entropía (*bits*) más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con la más alta entropía son E4, E5, E9, E10 y E16. Para el área de Broca, el valor más alto de entropía está en E11. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 230 (*bits*).

Tercero, la vocal con SS /i/ (Fig. 4.5 (i)), tiene un valor promedio de posición para la entropía (*bits*) más alto en el área de Wernicke que en el área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E4, E5, E9, E10 y E16. Para el área de Broca, el valor más alto de entropía está en E11. De manera complementaria, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 230 (*bits*).

Cuarto, la vocal con SS /o/ (Fig. 4.5 (o)), tiene un valor promedio de posición para la entropía (*bits*) más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la son E4, E5, E9, E10, E15 y E16. Para el área de Broca, el valor más alto de entropía está en E11. En adición a lo anterior, en la corteza auditiva primaria E3 tiene un valor promedio de posición para la entropía cercano a 230 (*bits*).

Quinto, la vocal con SS /u/ (Fig. 4.5 (u)), tiene un valor promedio de posición para la entropía (*bits*) más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E4, E5, E9 y E10. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E11 y E14. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 220 (*bits*).

La figura (4.6) representa el valor promedio por electrodo de la entropía de información ($\overleftarrow{H}_{s,v}^p$), referido a cada vocal con habla silenciosa, para los 21 electrodos y los 20 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{H}_{s,v}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} H_{s,v,k}^p}{(S_T)(l_1)} \quad (4.4)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.7), $l_1 = 60$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 20$ es el número de sujetos, v es una vocal con habla silenciosa, y p es un electrodo.

En la figura (4.6), el cuadrado representa el valor promedio por electrodo de la entropía de la vocal con SS /a/, el círculo representa el valor promedio por electrodo de la entropía de la vocal con SS /e/, el rombo representa el valor promedio por electrodo de la entropía de la vocal con SS /i/, el signo (+) representa el valor promedio por electrodo de la entropía de la vocal con SS /o/ y el triángulo representa el valor promedio por electrodo de la entropía de la vocal con SS /u/. El rango de valores promedio por electrodo de entropía de la información para los 20 sujetos se encuentra entre 140 y 200 (*bits*). En el 100 % de los casos, el valor promedio por electrodo de entropía de la información para la vocal con SS /a/ es más alta en todos los electrodos que en las otras vocales con SS. La vocal con SS /u/ obtuvo el menor valor promedio por electrodo de entropía en todos los electrodos. El valor más alto de entropía lo obtuvo la vocal con SS /a/ en el electrodo 5 y el menor valor de entropía lo obtuvo la vocal con SS /u/ en el electrodo 17.

En 71 % de los electrodos el valor promedio de la vocal con SS /a/ tiende a estar separado de las otras vocales con SS. En la figura (4.6) se observa mayor promedio de entropía de la información para los electrodos E5, E9, E4 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E17, E18, E11 y E12, pertenecientes al área de Broca, presentan menor promedio de entropía. De esta manera se evidencia para el área de Wernicke un mayor promedio de entropía de la información, en tareas de habla silenciosa con vocales, en contraste a el área de Broca (ver figura 3.3).

Las figuras 4.8(a) a 4.8(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para la dimensión de regularización ($\tilde{\Lambda}_{s,v}^p$), realizando tareas de habla silenciosa con vocales. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (sujeto 14), de acuerdo a la dimensión de regularización (Fig. 4.10). De esta manera, el valor promedio de cada electrodo se obtuvo con:

$$\tilde{\Lambda}_{s,v}^p = \frac{\sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{l_1} \quad (4.5)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{\Lambda}_{s,v}$ (Eq. 3.27), $l_1 = 60$ es el número de filas de la matriz, k es una fila de la matriz, $s = 14$ es el sujeto seleccionado, v es una vocal con habla silenciosa, y p es un electrodo. Con los valores $\tilde{\Lambda}_{s,v}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.8(a) a 4.8(e). La distribución del color representa el valor de la dimensión de regularización entre 0 (azul) y 2 (rojo). Estas áreas fueron seleccionadas por ser de crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las vocales con SS se presentan los siguientes resultados:

Primero, la vocal con SS /a/ (Fig. 4.8 (a)), tiene un valor promedio de posición para la dimensión de regularización más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E4, E5, E8, E9, E10, E15 y E16. Para el área de Broca, los valores promedio de posición más altos están en E11, E12 y E18. En este caso, el área de Wernicke tiene un valor promedio de posición para la dimensión de regularización más alto que el área de Broca. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.8.

Segundo, la vocal con SS /e/ (Fig. 4.8 (b)), tiene un valor promedio de posición para la dimensión de regularización más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los valores promedio de posición más altos para la dimensión

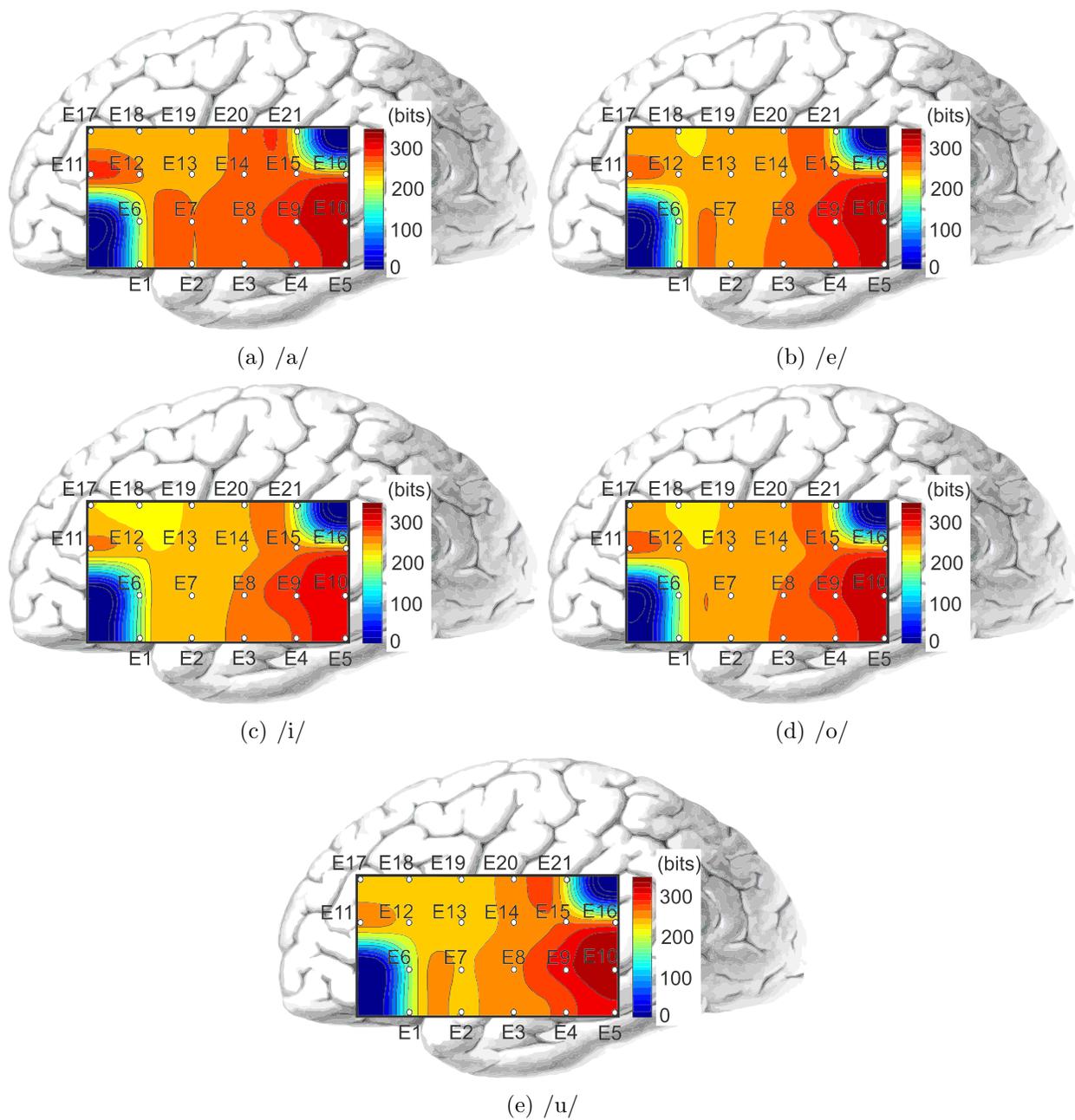


Figura 4.5: Representación espacial del valor promedio por electrodo de la entropía de la información, realizando tareas de habla silenciosa (SS) con vocales

de regularización son E4, E5, E9, E10, E15 y E16. Para el área de Broca, los valores promedio de posición más altos de dimensión de regularización están en E11 y E12. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.5.

Tercero, la vocal con SS /i/ (Fig. 4.8 (c)), tiene un valor promedio de posición para la dimensión de regularización más alto en el área de Wernicke que en el área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E4, E5, E8, E9, E10, E15 y E16. Para el área de Broca, los valores promedio de posición más altos para la dimensión

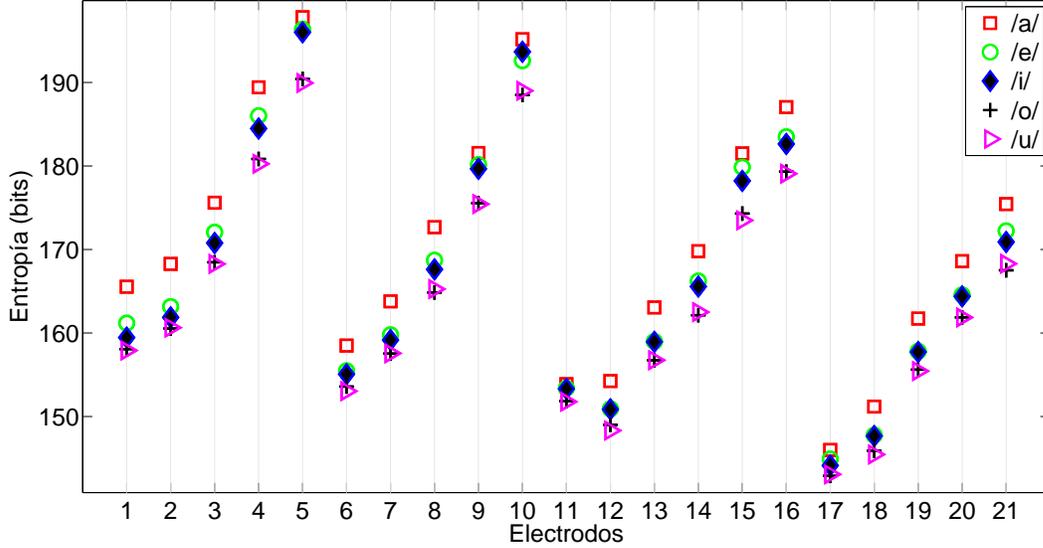


Figura 4.6: Valor promedio de entropía por electrodo en vocales con habla silenciosa (SS)

de regularización están en E11 y E12. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.8.

Cuarto, las vocales con SS /o/ y /u/ (Figs. 4.8 (d) y 4.8 (e)), tienen valores promedio de posición para la dimensión de regularización más altos en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E4, E5, E9, E10, E15 y E16. Para el área de Broca, el valor promedio de posición más alto para la dimensión de regularización está en E11. Adicionalmente, en la corteza auditiva primaria E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.6.

La figura (4.7) representa el valor promedio por electrodo de la entropía de información ($\overleftarrow{H}_{s,v}^p$), referido a cada vocal con habla silenciosa, para los 21 electrodos y los 20 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{\Lambda}_{s,v}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{(S_T)(l_1)} \quad (4.6)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{\Lambda}_{s,v}$ (Eq. 3.27), $l_1 = 60$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 20$ es el número de sujetos, v es una vocal con habla silenciosa, y p es un electrodo.

En la figura (4.7), el cuadrado representa el valor promedio por electrodo de la dimensión de regularización para la vocal con SS /a/, el círculo representa el valor promedio por electrodo de la dimensión de regularización para la vocal con SS /e/, el rombo representa el valor promedio por electrodo de la dimensión de regularización para la vocal con SS /i/, el signo (+) representa el valor promedio por electrodo de la dimensión de regularización para la vocal con SS /o/ y el triángulo representa el valor promedio por electrodo de la dimensión de regularización para la vocal con SS /u/. El rango de valores promedio por electrodo de dimensión de regularización para los 20 sujetos se encuentra entre 1.45 y 1.63. En el 48 %

de los casos, el valor promedio por electrodo de dimensión de regularización para la vocal con SS /a/ es más alta que para las otras vocales con SS. La vocal con SS /o/ obtuvo en 57% de los casos el menor valor promedio por electrodo. El valor más alto de dimensión de regularización lo obtuvo la vocal con SS /i/ en el electrodo 4 y el menor valor de dimensión de regularización lo obtuvo la vocal con SS /o/ en el electrodo 1.

En 38% de los electrodos el valor promedio por electrodo de dimensión de regularización para la vocal con SS /a/ tiende a estar separado de las otras vocales con SS. En los electrodos E4, E5, E9 y E11 se observa una mayor separación entre las vocales con SS. En la figura (4.7) se observa un mayor promedio por electrodo de dimensión de regularización para los electrodos E4, E5 y E10 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E2, E6, E7 y E11, pertenecientes mayoritariamente al área de Broca, presentan menor promedio por electrodo de dimensión de regularización. De esta manera se evidencia para el área de Wernicke un mayor promedio por electrodo de dimensión de regularización, en tareas de habla silenciosa con vocales, en contraste al área de Broca (ver figura 3.3).

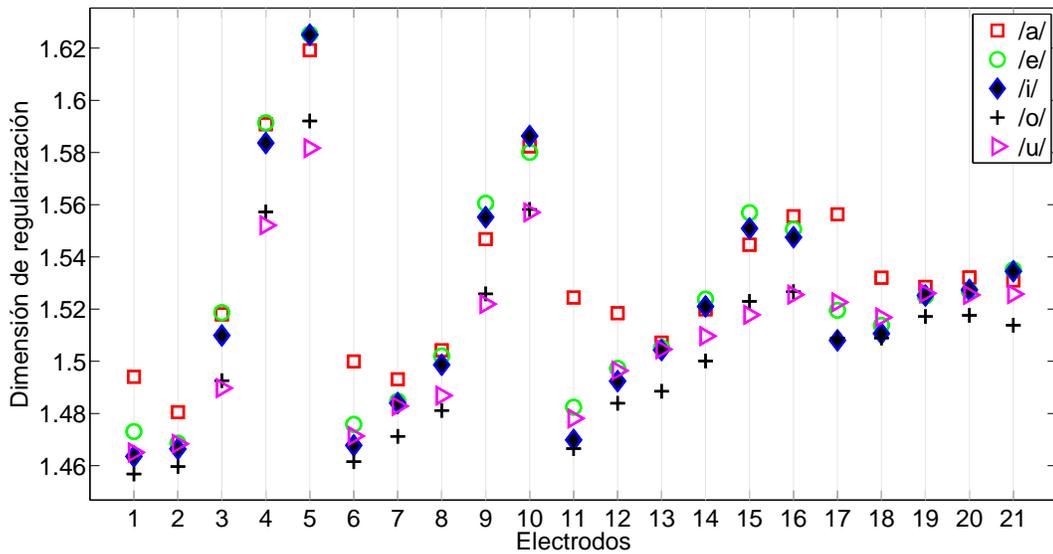


Figura 4.7: Valor promedio de dimensión de regularización por electrodo en vocales con habla silenciosa (SS)

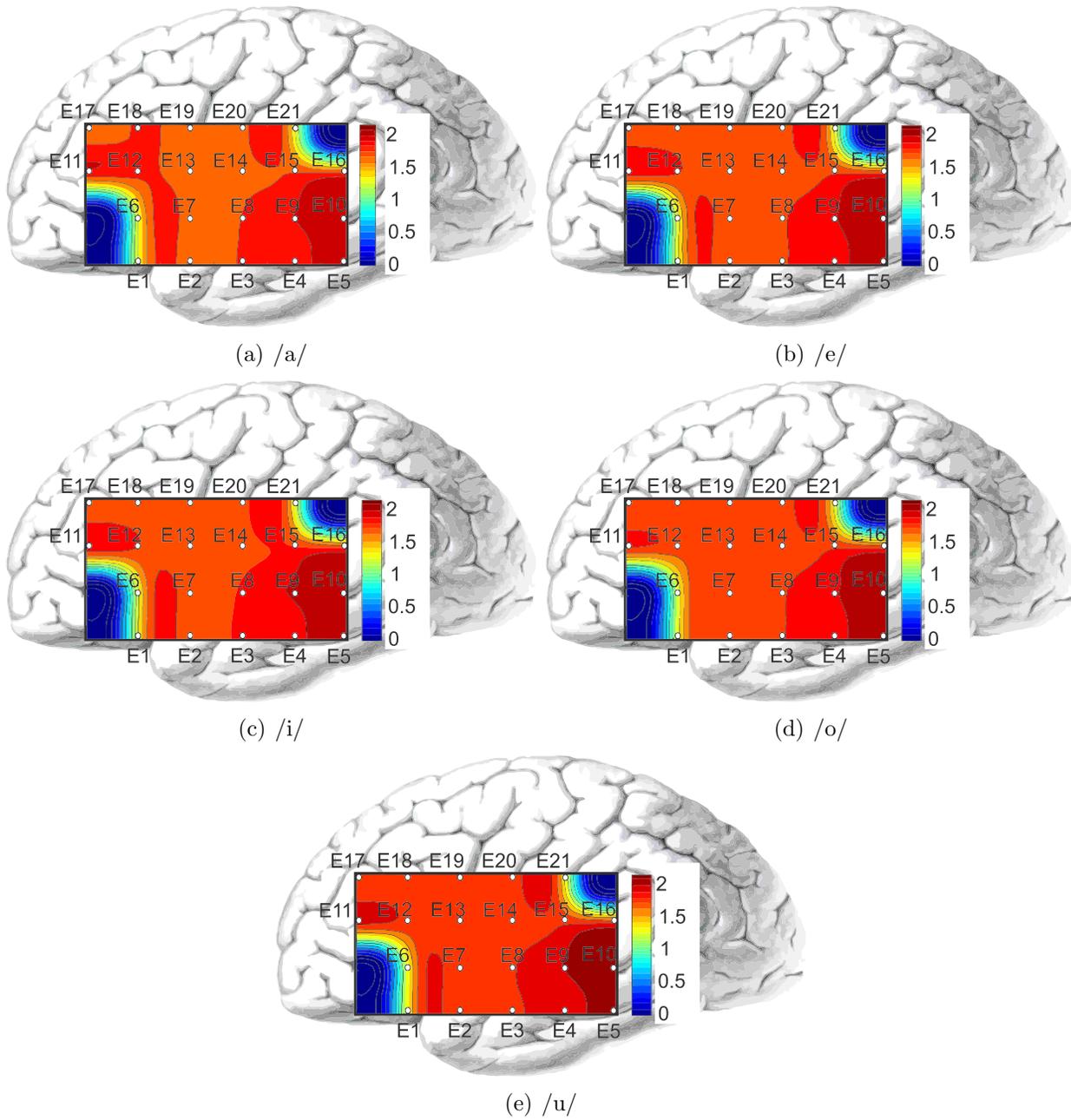


Figura 4.8: Representación espacial del valor promedio por electrodo de la dimensión de regularización en vocales con habla silenciosa (SS)

4.1.3. Resultados de la clasificación

En esta sección primero se presenta la definición de exactitud y posteriormente se continua con la presentación de los resultados de la clasificación donde se utilizan dos clasificadores multiclase SVM de tipo (1-rest) y (1-1). De esta manera, sea A una matriz de confusión de tamaño $G \times G$ donde G corresponde al número de clases a clasificar, entonces a_{gk} representa el número de objetos pertenecientes a la clase g , asignadas a la clase k . En la matriz de confusión los elementos de la diagonal de a_{gg} corresponden a los elementos correctamente clasificados. De esta manera se define la exactitud (*accuracy*) como:

$$accuracy = \frac{\sum_{g=1}^G a_{gg}}{N_T}, \quad (4.7)$$

donde N_T corresponde al número total de objetos de la matriz de confusión [174].

Los datos que toma el clasificador SVM multiclase *one against-rest* (1-rest) son el resultado de la transformación de las señales cerebrales de vocales con habla silenciosa en entropía de información $\mathbf{H}_{s,v}$ y dimensión de regularización $\mathbf{\Lambda}_{s,v}$, conformando el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$ (Fig. 3.7).

Con el método de validación cruzada, 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$ son usados para entrenamiento, y 30 % de esos datos son usados para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-rest) fue probado 10 veces para cada sujeto (tabla A.2). La exactitud obtenida con esta metodología estuvo entre 46.56 % y 85.90 %, con una media estadística de 61.41 % y una desviación estándar de 3.98 % (Fig. 4.9).

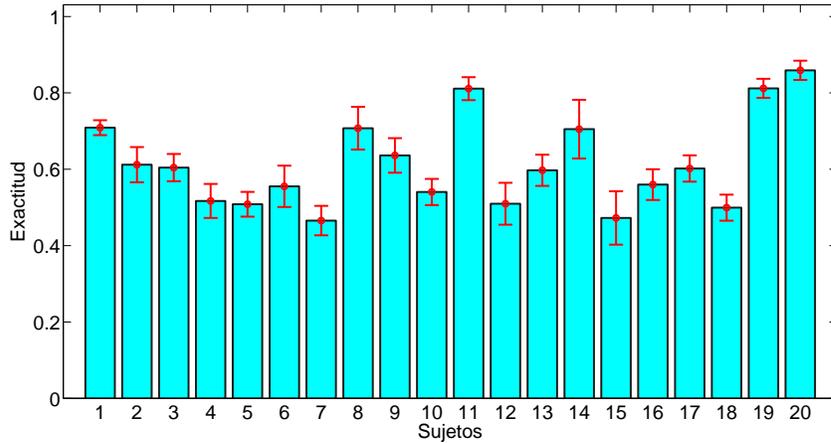


Figura 4.9: Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-rest)

Los datos que toma el clasificador SVM multiclase *one against-one* (1-1) son el resultado de la transformación de las señales cerebrales de vocales con habla silenciosa en entropía de información $\mathbf{H}_{s,v}$ y dimensión de regularización $\mathbf{\Lambda}_{s,v}$, conformando el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$ (Fig. 3.7).

Con el método de validación cruzada, 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{H}_{s,v} \ \mathbf{\Lambda}_{s,v}]$ son usados para entrenamiento, y 30 % de esos datos son usados para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto (tabla A.1). La exactitud obtenida con esta metodología estuvo entre 52.67 % y 91.43 %, con una media estadística de 69.83 % y una desviación estándar de 3.23 % (Fig. 4.10).

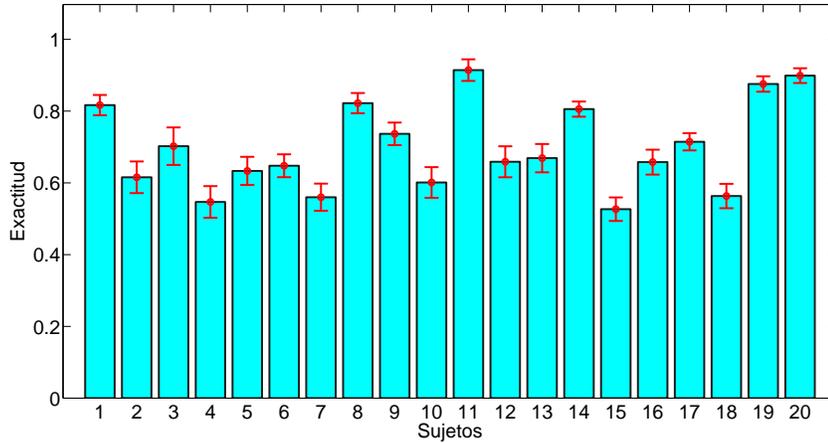


Figura 4.10: Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1)

4.2. Resultados de señales EEG para sílabas con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización

4.2.1. Resultados por sujeto

La figura 4.11 muestra los resultados correspondientes al valor promedio por sujeto y por sílaba de la entropía de la información ($\overline{H}_{s,v}$), para sílabas con habla silenciosa. Estos valores se obtuvieron con la expresión dada por:

$$\overline{H}_{s,v} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} H_{s,v,k}^p}{l_1 l_2} \quad (4.8)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.8), $l_1 = 30$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una sílaba con habla silenciosa, y p es un electrodo.

De la figura (4.11) se observa que existe una variación en los valores promedio de entropía entre sujetos y entre sílabas con habla silenciosa. El rango de valores promedio de entropía está entre 120 a 380 (*bits*). Los valores promedio más altos de entropía fueron obtenidos por el sujeto 3 y los mínimos valores promedio de entropía fueron obtenidos por el sujeto 9. Teniendo en cuenta el valor promedio más alto de entropía por sujeto, se encontró que la sílaba con SS /fa/ tiene el valor promedio más alto en el 60% de los sujetos, y la sílaba con SS /ma/ tiene el segundo valor promedio más alto de entropía en el 30% de los sujetos.

La figura (4.12) muestra la dispersión de los datos, del valor promedio de entropía para cada sílaba con SS, utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.8 para cada sujeto y para cada vocal con SS (Fig. (4.12)). Los valores de la mediana para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son, respectivamente, (176.37, 169.36, 165.30, 156.50 y 159.57). Los valores de la mediana más altos son para las sílabas con SS /fa/ y /pa/ y la menor para la sílaba con SS /la/. El rango

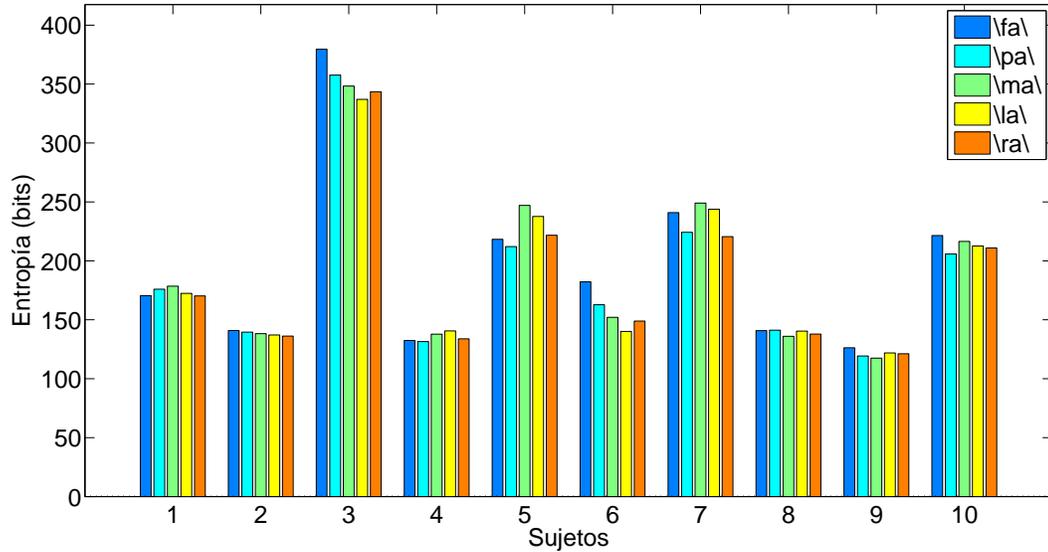


Figura 4.11: Valor promedio por sujeto y por sílaba con habla silenciosa (SS) de la entropía de la información

intercuartil (IQR) con la mayor dispersión es para la sílaba con SS /ma/ y el IQR con la menor dispersión es para la sílaba con SS /pa/. En el caso de las sílabas con SS /fa/, /pa/, /ma/, /la/ y /ra/, presentan asimetría positiva. Adicionalmente, se presentan outliers en las sílabas con SS /fa/ y /pa/.

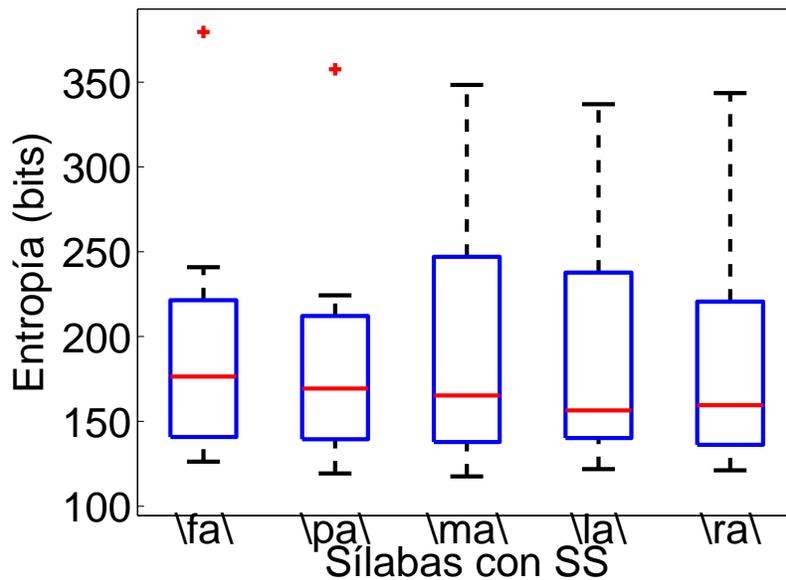


Figura 4.12: Diagrama de dispersión del valor promedio de entropía para cada sílaba con habla silenciosa (SS)

La figura 4.13 muestra los resultados correspondientes al valor promedio por sujeto y por

sílaba con SS de la dimensión de regularización ($\bar{\Lambda}_{s,v}$), para sílabas con habla silenciosa. Estos valores se obtuvieron con la expresión dada por:

$$\bar{\Lambda}_{s,v} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{l_1 l_2} \quad (4.9)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{\Lambda}_{s,v}$ (Eq. 3.28), $l_1 = 30$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una sílaba con habla silenciosa, y p es un electrodo.

De la figura (4.13) se observa que existe una variación en los valores promedio de dimensión de regularización entre sujetos y entre sílabas con habla silenciosa. El rango de valores promedio de dimensión de regularización está entre 1.3 a 1.5. Los valores promedio más altos de dimensión de regularización fueron obtenidos por el sujeto 5 y los mínimos valores promedio de dimensión de regularización fueron obtenidos por el sujeto 3. Teniendo en cuenta el valor promedio más alto de dimensión de regularización por sujeto, se encontró que la sílaba con SS /ma/ tiene el valor promedio más alto en el 50 % de los sujetos, y las sílabas con SS /fa/ y /la/ tienen el segundo valor promedio más alto de dimensión de regularización en el 40 % de los sujetos. En contraste, el mínimo valor promedio de dimensión de regularización por sujeto, fue obtenido con la sílaba con SS /ra/ en el 40 % de los casos.

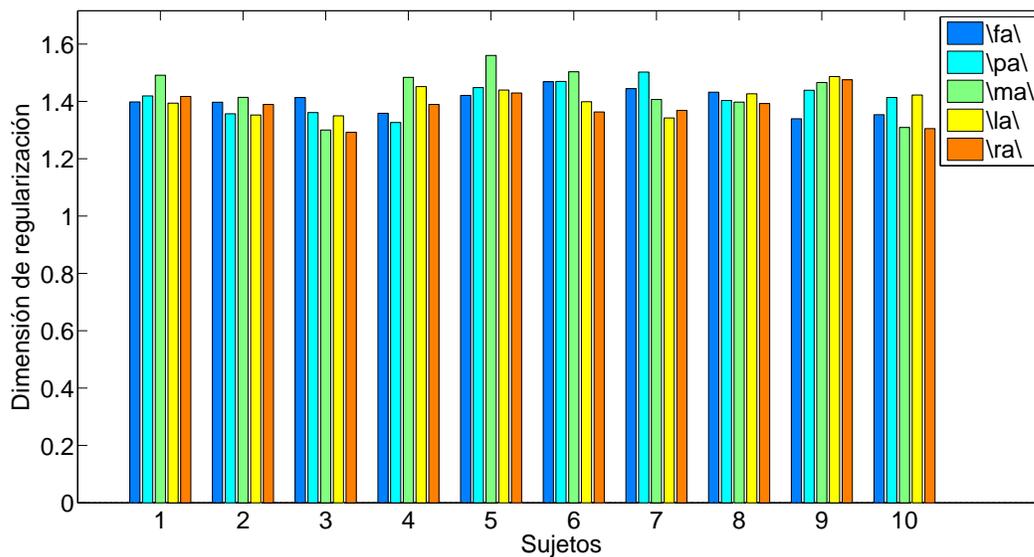


Figura 4.13: Valor promedio por sujeto y por sílaba con habla silenciosa (SS) de la dimensión de regularización

La figura (4.14) muestra la dispersión de los datos, del valor promedio de la dimensión de regularización para cada sílaba con SS, utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.9 para cada sujeto y para cada sílaba con SS (Fig. (4.13)). Los valores de la mediana para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son, respectivamente, (1.406, 1.417, 1.440, 1.411 y 1.390). El valor de la mediana es más alto para la sílaba con SS /ma/ y la menor para la sílaba con SS /ra/. El rango intercuartil (IQR) con la mayor dispersión es para la sílaba con SS /ma/ y el IQR con la

menor dispersión es para la sílaba con SS /ra/. En el caso de las sílabas con SS /fa/, /pa/, /ma/ y /la/, presentan asimetría negativa.

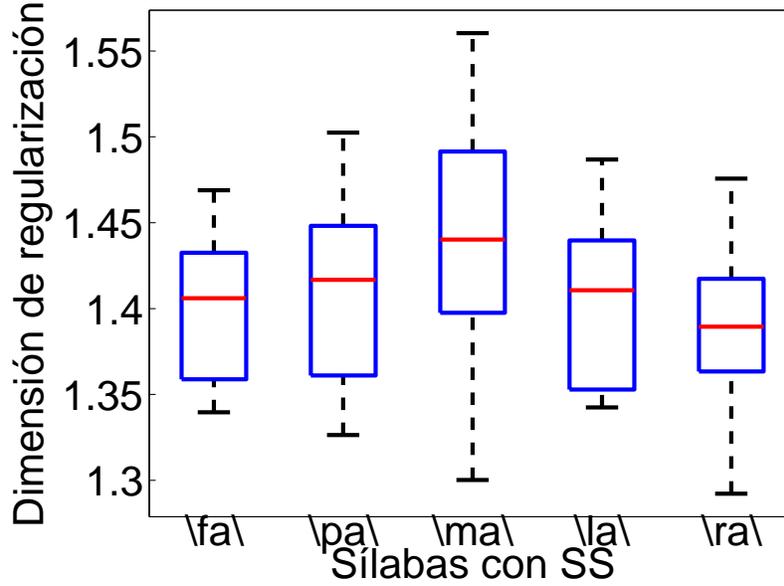


Figura 4.14: Diagrama de dispersión del valor promedio de dimensión de regularización por sílaba con habla silenciosa (SS)

4.2.2. Resultados por localización

Las figuras 4.15(a) a 4.15(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para la entropía de la información ($\tilde{H}_{s,v}^p$), realizando tareas de habla silenciosa con sílabas. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 10 sujetos de la experimentación (sujeto 7), de acuerdo a la entropía de la información (Fig. 4.20). De esta manera, el valor promedio de cada electrodo (*bits*) se obtuvo con:

$$\tilde{H}_{s,v}^p = \frac{\sum_{k=1}^{l_1} H_{s,v,k}^p}{l_1} \quad (4.10)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.8), $l_1 = 30$ es el número de filas de la matriz, k es una fila de la matriz, $s = 7$ es el sujeto seleccionado, v es una sílaba con habla silenciosa, y p es un electrodo. Con los valores $\tilde{H}_{s,v}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.15(a) a 4.15(e). La distribución del color representa el valor de la entropía de la información entre 0 (azul) y 250 (rojo) (*bits*). Los electrodos son posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.4). Los electrodos E1, E6, E7, E11, E12, E13, E16, E17, E21 fueron ubicados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E14, E15, E19 y E20 sobre el área de Wernicke, y el electrodo E3 sobre la corteza auditiva primaria. Estas áreas fueron seleccionadas porque son de crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las sílabas con SS se presenta:

Primero, las sílabas con SS /fa/ y /pa/ (Figs. 4.15 (a) y 4.15 (b)), tienen un valor promedio de posición para la entropía (*bits*) que en algunas áreas es menor para el área de Wernicke que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E4, E5, E9, E10, E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E12, E13, E16, E17 y E21. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 260 (*bits*) para la sílaba /fa/ y a 230 (*bits*) para la sílaba /pa/.

Segundo, la sílaba con SS /ma/ (Fig. 4.15 (c)), tiene un valor promedio de posición para la entropía (*bits*) que en algunas áreas es mayor para el área de Broca que para el área de Wernicke o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E9, E10, E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E12, E13, E16, E17 y E21. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 230 (*bits*).

Tercero, la sílaba con SS /la/ (Fig. 4.15 (d)), tiene un valor promedio de posición para la entropía (*bits*) que es más alto para el área de Wernicke que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E5, E9, E10, E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E12, E13, E16, E17 y E21. Adicionalmente, en la corteza auditiva primaria, E3 tiene un valor promedio de posición para la entropía cercano a 230 (*bits*).

Cuarto, la sílaba con SS /ra/ (Fig. 4.15 (e)), tiene un valor promedio de posición para la entropía (*bits*) que en algunas áreas es más alto para el área de Wernicke que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la entropía son E14 y E15. Para el área de Broca, los más altos valores promedio de posición para la entropía están en E12, E13, E16 y E21. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la entropía cercano a 220 (*bits*).

La figura (4.16) representa el valor promedio por electrodo de la entropía de información ($\overleftarrow{H}_{s,v}^p$), referida a cada sílaba con habla silenciosa, para los 21 electrodos y los 10 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{H}_{s,v}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} H_{s,v,k}^p}{(S_T)(l_1)} \quad (4.11)$$

donde $H_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{H}_{s,v}$ (Eq. 3.8), $l_1 = 30$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 10$ es el número de sujetos, v es una sílaba con habla silenciosa, y p es un electrodo.

En la figura (4.16), el cuadrado representa el valor promedio por electrodo de la entropía de la sílaba con SS /fa/, el círculo representa el valor promedio por electrodo de la entropía de la sílaba con SS /pa/, el rombo representa el valor promedio por electrodo de la entropía de la sílaba con SS /ma/, el signo (+) representa el valor promedio por electrodo de la entropía de la sílaba con SS /la/ y el triángulo representa el valor promedio por electrodo de la entropía de la sílaba con SS /ra/. El rango de valores promedio por electrodo de entropía de la información para los 10 sujetos se encuentra entre 160 y 195 (*bits*). En 86 % de los casos la entropía de la sílaba con SS /fa/ es más alta en los electrodos que las otras sílabas con

SS. La sílaba con SS /ra/ obtuvo el menor valor de entropía en todos los electrodos. El valor más alto de entropía lo obtuvo la sílaba con SS /fa/ en el electrodo 5 y el menor valor de entropía lo obtuvo la sílaba con SS /ra/ en el electrodo 11.

En la figura (4.16) los valores promedio de las silabas con SS tienden a estar separados en cada uno de los electrodos. En la figura (4.16) se observa mayor promedio de entropía de la información para los electrodos E5, E9, E10, E14, E15 y E20 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E6, E11, E16 y E21, pertenecientes al área de Broca, presentan menor promedio de entropía. De esta manera se evidencia para el área de Wernicke un mayor promedio de entropía de la información, en tareas de habla silenciosa con vocales, en contraste a el área de Broca (ver figura 3.4).

Las figuras 4.17(a) a 4.17(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para la dimensión de regularización ($\tilde{\Lambda}_{s,v}^p$), realizando tareas de habla silenciosa con sílabas. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 10 sujetos de la experimentación (sujeto 7), de acuerdo a la dimensión de regularización (Fig. 4.20). De esta manera, el valor promedio de cada electrodo se obtuvo con:

$$\tilde{\Lambda}_{s,v}^p = \frac{\sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{l_1} \quad (4.12)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbf{\Lambda}_{s,v}$ (Eq. 3.28), $l_1 = 30$ es el número de filas de la matriz, k es una fila de la matriz, $s = 7$ es el sujeto seleccionado, v es una sílaba con habla silenciosa, y p es un electrodo. Con los valores $\tilde{\Lambda}_{s,v}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.17(a) a 4.17(e). La distribución del color representa el valor de la dimensión de regularización entre 0 (azul) y 1.6 (rojo). Estas áreas fueron seleccionadas por ser de crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las sílabas con SS se presentan los siguientes resultados:

Primero, la sílaba con SS /fa/ (Fig. 4.17 (a)), tiene un valor promedio de posición para la dimensión de regularización que en algunas áreas es menor para el área de Wernicke que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición para la dimensión de regularización están en E11, E12, E13, E17 y E21. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.4.

Segundo, la sílaba con SS /pa/ (Fig. 4.17 (b)), tiene un valor promedio de posición para la dimensión de regularización que en algunas áreas es mayor para el área de Broca que para el área de Wernicke. En el área de Wernicke los electrodos con los valores promedio de posición más altos para la dimensión de regularización son E14, E15, E19 y E20. Para el área de Broca, los valores más altos de dimensión de regularización están en E1, E6, E7, E11, E12, E13, E16, E17 y E21. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.4.

Tercero, la sílaba con SS /ma/ (Fig. 4.17 (c)), tiene un valor promedio de posición para la dimensión de regularización que en algunas áreas es mayor para el área de Broca que para el área de Wernicke o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la de dimensión de regularización son E14, E15, E19 y E20. Para el área de Broca, los valores promedio de posición más altos de dimensión de regularización están en E12, E13, E16, E17 y E21. Complementariamente, en el

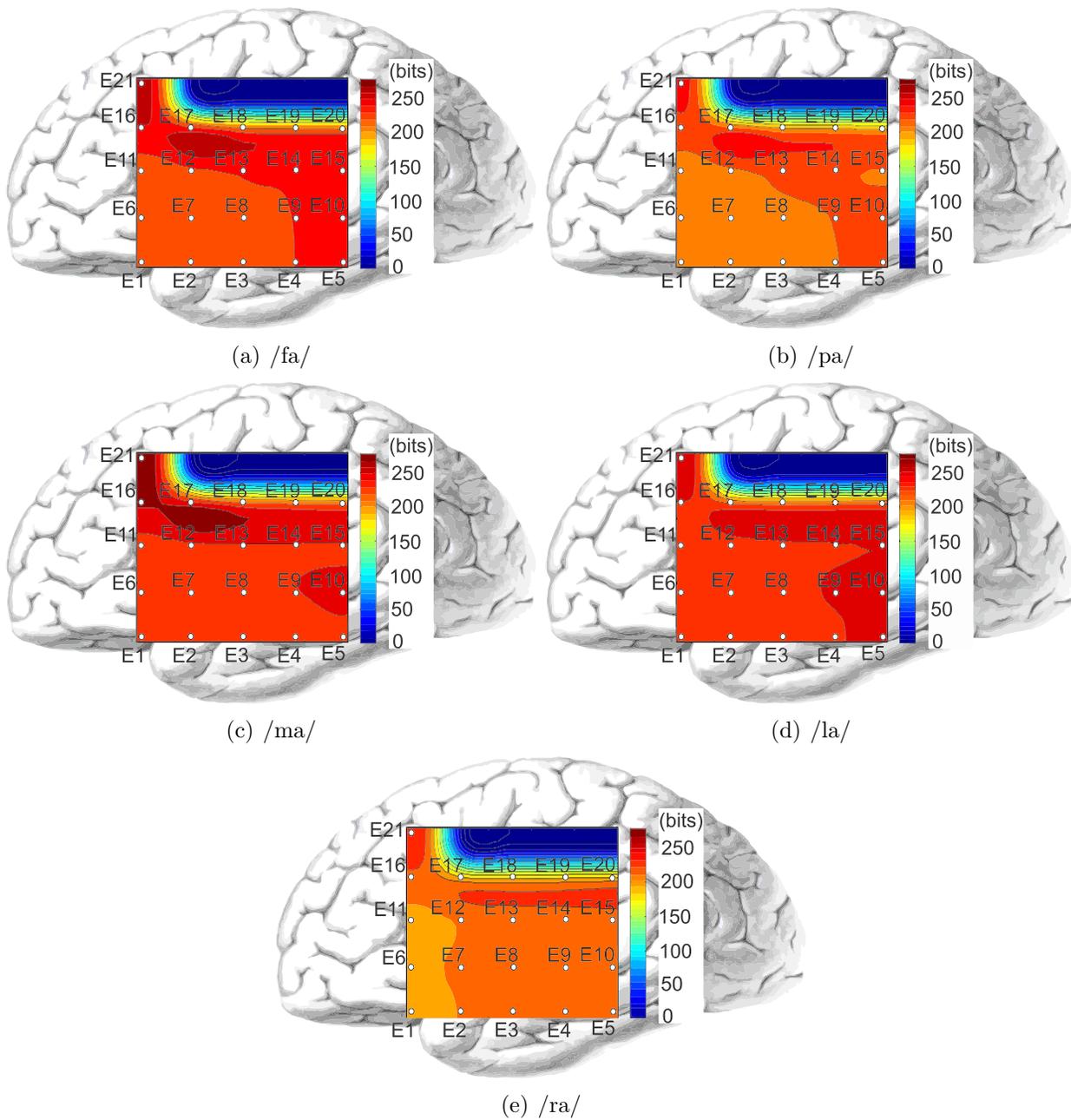


Figura 4.15: Representación espacial del valor promedio por electrodo de la entropía para sílabas con habla silenciosa (SS)

área auditiva primaria el electrodo E3 tiene un valor promedio de posición para la dimensión de regularización cercano a 1.2.

Cuarto, la sílaba con SS /la/ (Fig. 4.17 (d)), tiene un valor promedio de posición para la dimensión de regularización que en algunas áreas es mayor para el área de Broca que para el área de Wernicke o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E14 y E19. Para el área de Broca, los más altos valores promedio de posición para la dimensión de regularización están en E16 y E21. Además, en la corteza auditiva primaria E3 tiene un valor de dimensión de regularización cercano a 1.2.

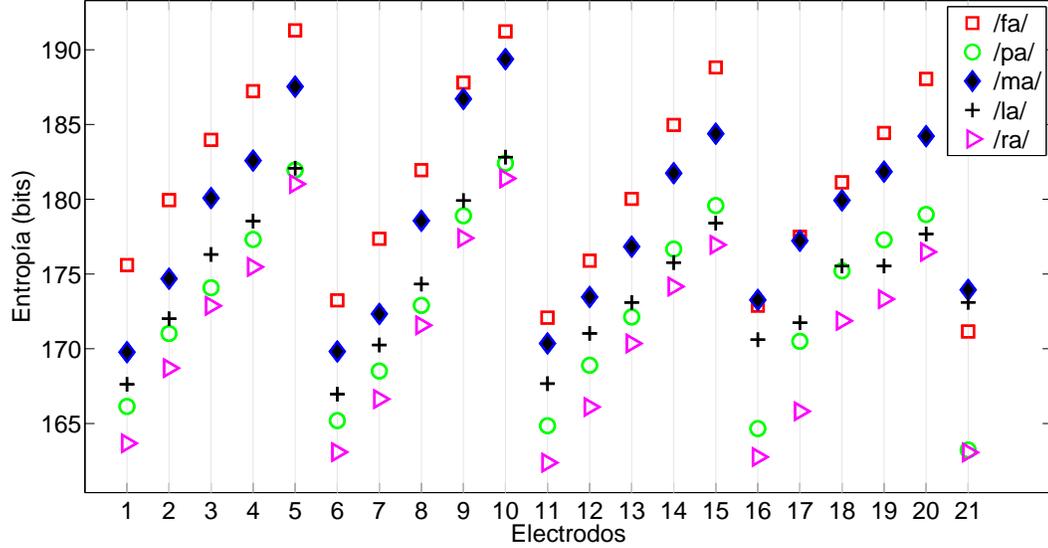


Figura 4.16: Valor promedio de entropía por electrodo para sílabas con habla silenciosa (SS)

Quinto, la sílaba con SS /ra/ (Fig. 4.17 (e)), tiene un valor promedio de posición para la dimensión de regularización que en algunas áreas es mayor para el área de Broca o para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición para la dimensión de regularización son E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición para la dimensión de regularización están en E16 y E21. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor de dimensión de regularización cercano a 1.2.

La figura (4.18) representa el valor promedio por electrodo de la entropía de información ($\overleftarrow{H}_{s,v}^p$), referido a cada sílaba con habla silenciosa, para los 21 electrodos y los 10 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{\Lambda}_{s,v}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} \Lambda_{s,v,k}^p}{(S_T)(l_1)} \quad (4.13)$$

donde $\Lambda_{s,v,k}^p$ corresponde a cada elemento de la matriz $\Lambda_{s,v}$ (Eq. 3.28), $l_1 = 60$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 10$ es el número de sujetos, v es una sílaba con habla silenciosa, y p es un electrodo.

En la figura (4.18), el cuadrado representa la dimensión de regularización por electrodo de la sílaba con SS /fa/, el círculo representa la dimensión de regularización por electrodo de la sílaba con SS /pa/, el rombo representa la dimensión de regularización por electrodo de la sílaba con SS /ma/, el signo (+) representa la dimensión de regularización por electrodo de la sílaba con SS /la/ y el triángulo representa la dimensión de regularización por electrodo de la sílaba con SS /ra/. El rango de valores de dimensión de regularización se encuentra entre 1.22 y 1.34. En el 100% de los casos la dimensión de regularización de la sílaba /ma/ con SS es más alta que para las otras sílabas con SS. La sílaba con SS /ra/ obtuvo el menor valor de dimensión de regularización en el 95% de los casos. El valor más alto de dimensión de regularización lo obtuvo la sílaba con SS /ma/ en el electrodo 16 y el menor valor de dimensión de regularización lo obtuvo la sílaba con SS /ra/ en el electrodo 5.

En la figura (4.18), los valores promedio de las sílabas con SS tienden a estar separados en cada uno de los electrodos. En los electrodos E5, E10 y E15 se observa una mayor separación entre las sílabas con SS. En la figura (4.18) se observa un menor promedio por electrodo de dimensión de regularización para los electrodos E2, E4, E5, E15 y E20 pertenecientes al área de Wernicke. En contraste, los electrodos E11, E16, E17 y E19, pertenecientes mayoritariamente al área de Broca, presentan mayor promedio por electrodo de dimensión de regularización. De esta manera se evidencia para el área de Wernicke un menor promedio por electrodo de dimensión de regularización, en tareas de habla silenciosa con sílabas, en contraste al área de Broca (ver figura 3.4).

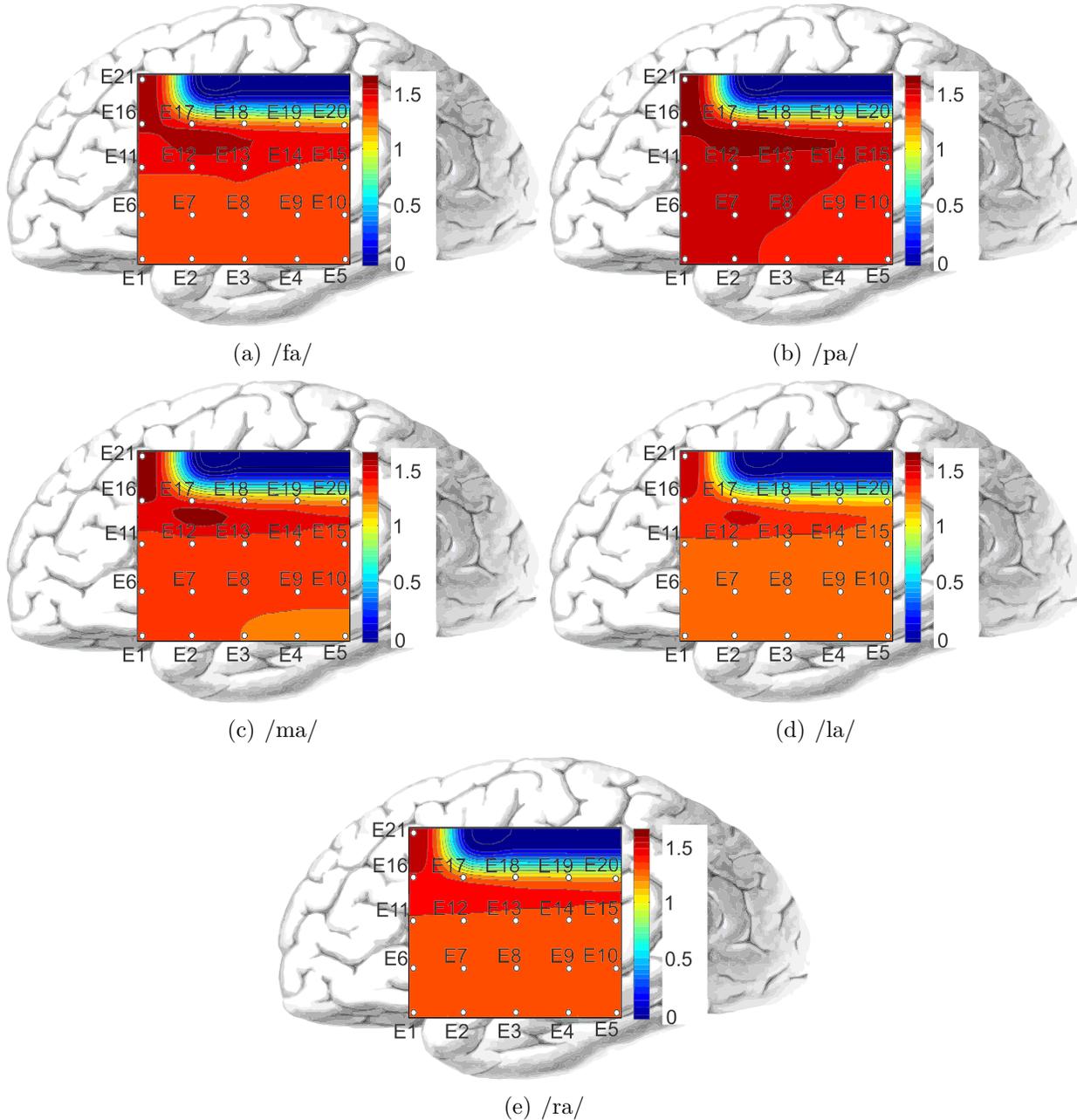


Figura 4.17: Representación espacial del valor promedio por electrodo de la dimensión de regularización para sílabas con habla silenciosa (SS)

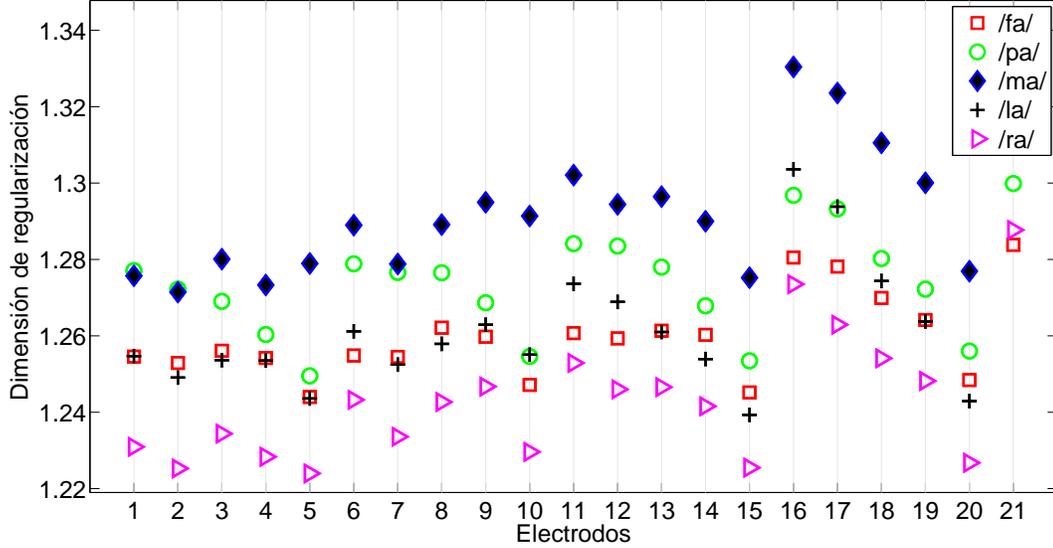


Figura 4.18: Valor promedio de dimensión de regularización por electrodo para sílabas con habla silenciosa (SS)

4.2.3. Resultados de la clasificación

En esta sección se presenta de los resultados de la clasificación de sílabas con habla silenciosa donde se utilizan dos clasificadores multiclase SVM. Primero se explica los resultados de la clasificación con (1-rest) y posteriormente se explica los resultados de la clasificación con (1-1). Con el primer clasificador, los datos que toma el clasificador SVM multiclase *one against-rest* (1-rest) son el resultado de la transformación de las señales cerebrales de sílabas con habla silenciosa en entropía de información $\mathbf{H}_{s,v}$ y dimensión de regularización $\mathbf{\Lambda}_{s,v}$, conformando el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \mathbf{\Lambda}_{s,v}]$ (Fig. 3.8).

Con el método de validación cruzada, 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{H}_{s,v} \mathbf{\Lambda}_{s,v}]$ son usados para entrenamiento, y 30 % de esos datos son usados para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-rest) fue probado 10 veces para cada sujeto (tabla A.7). La exactitud obtenida con esta metodología estuvo entre 47.97 % y 74.68 %, con una media estadística de 58.57 % y una desviación estándar de 4.92 % (Fig. 4.19).

Con el segundo clasificador, los datos que toma el clasificador SVM multiclase *one against-one* (1-1) son el resultado de la transformación de las señales cerebrales de sílabas con habla silenciosa en entropía de información $\mathbf{H}_{s,v}$ y dimensión de regularización $\mathbf{\Lambda}_{s,v}$, conformando el vector de características $\Xi_{s,v} = [\mathbf{H}_{s,v} \mathbf{\Lambda}_{s,v}]$ (Fig. 3.8).

Con el método de validación cruzada, 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{H}_{s,v} \mathbf{\Lambda}_{s,v}]$ son usados para entrenamiento, y 30 % de esos datos son usados para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto (tabla A.6). La exactitud obtenida con esta metodologías estuvo entre 55.11 % y 84.22 % con una media estadística de 66.89 % y una desviación estándar de 4.03 % (Fig. 4.20).

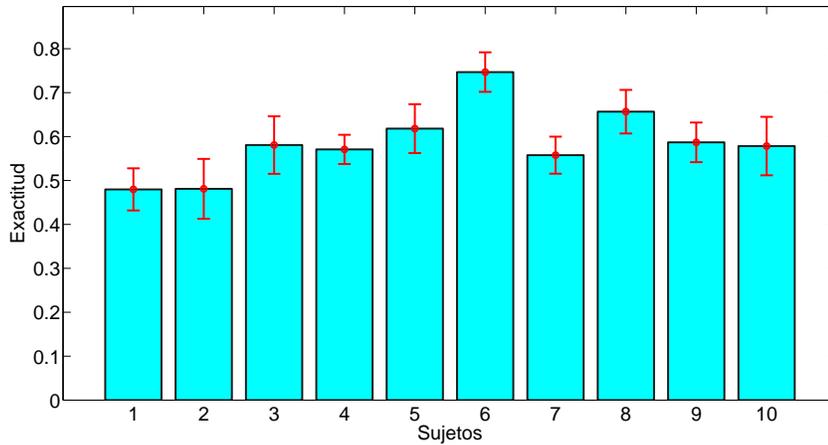


Figura 4.19: Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-rest)

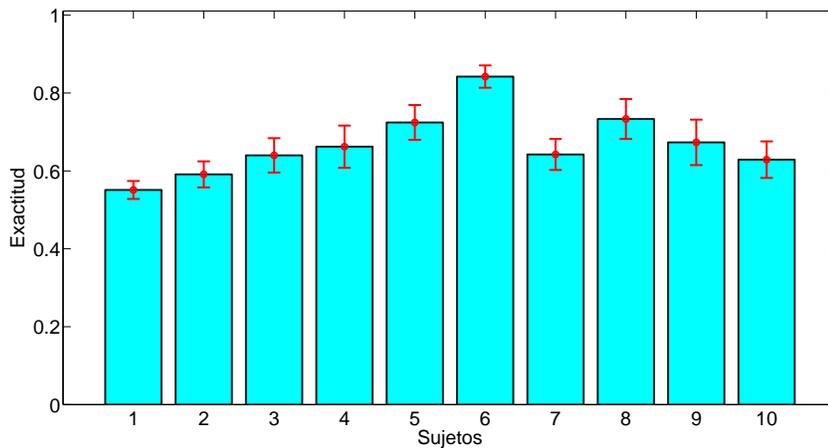


Figura 4.20: Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-1)

4.3. Comparación entre los resultados de la metodología basada en vector de características con entropía de la información y dimensión de regularización para vocales y sílabas con habla silenciosa

Al comparar los diagramas de dispersión relacionados con entropía de la información de la figura (4.2) con la figura (4.12) se encuentra que:

Los valores de las medianas para las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) son respectivamente (181.70, 178.47, 175.35, 175.51, 175.45) (*bits*) y una media de estos valores es 177.30 (*bits*). Los valores de las medianas para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son respectivamente (176.35, 169.36, 165.30, 156.50, 159.57) (*bits*) y una media de estos valores es 165.42 (*bits*). En este caso, se puede evidenciar que el cerebro tiende a generar un mayor valor de entropía de la información para las vocales con SS, en comparación con las sílabas con SS. También, se encuentra que existe una mayor variación en el rango intercuartil

(IQR) de las sílabas con SS, en comparación a las vocales con SS. El mayor IQR lo presenta la vocal con SS /a/ y la sílaba con SS /ma/. Adicionalmente, se observa una mayor asimetría positiva para las sílabas con SS, en comparación a las vocales con SS.

Al contrastar los diagramas de dispersión relacionados con la dimensión de regularización de la figura (4.4) con la figura (4.14) se observa que:

Los valores de las medianas para las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) son respectivamente (1.778, 1.789, 1.774, 1.749, 1.764) y una media de estos valores es 1.771. Los valores de las medianas para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son respectivamente (1.406, 1.417, 1.440, 1.411, 1.390) y una media de estos valores es 1.413. En este caso, se puede evidenciar que el cerebro tiende a generar un mayor valor de dimensión de regularización para las vocales con SS, en comparación con las sílabas con SS. Adicionalmente, se encuentra que el mayor valor de la mediana para las vocales con SS es /e/ y para las sílabas con SS es /ma/. De otra manera, el mayor IQR lo presenta la vocal con SS /a/ y la sílaba con SS /ma/. Complementariamente, se observa una mayor asimetría positiva para las sílabas con SS, en comparación a las vocales con SS.

Al confrontar los diagramas de entropía de la información por electrodo de la figuras (4.6) y (4.16) se encuentra que:

Para vocales con SS el rango de valores promedio por electrodo de entropía de la información se encuentra entre 140 y 200 (*bits*); mientras para sílabas con SS está entre 160 y 195 (*bits*). En el 100 % de los casos, el valor promedio por electrodo de entropía de la información para la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /u/ en todos los casos es la más baja por electrodo; mientras en 86 % de los casos la entropía de la sílaba con SS /fa/ es la más alta por electrodo y la sílaba con SS /ra/ en todos los casos es la más baja por electrodo.

En 71 % el valor promedio por electrodo de entropía de la información para la vocal con SS /a/ tiende a estar separado, en cada electrodo, de las otras vocales con SS; mientras los valores promedio por electrodo de entropía de la información para las sílabas con SS tienden a estar separados entre sí.

Para las vocales con SS, se observa mayor promedio de entropía de la información para los electrodos E5, E9, E4 y E16 pertenecientes al área de Wernicke; en contraste, los electrodos E17, E18, E11 y E12, pertenecientes al área de Broca, presentan menor promedio de entropía. En el caso de las sílabas con SS, se observa mayor promedio de entropía de la información para los electrodos E5, E9, E10, E14, E15 y E20 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E6, E11, E16 y E21, pertenecientes al área de Broca, presentan menor promedio de entropía. De esta manera se evidencia para el área de Wernicke un mayor promedio de entropía de la información, en tareas de habla silenciosa con vocales y sílabas, en contraste al área de Broca.

Al relacionar los diagramas de dimensión de regularización por electrodo de la figuras (4.7) y (4.18) se encuentra que:

Para vocales con SS el rango de valores promedio por electrodo de dimensión de regularización se encuentra entre 1.45 y 1.63; mientras para sílabas con SS está entre 1.22 y 1.34.

En el 48 % de los casos, el valor promedio por electrodo de dimensión de regularización para la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /o/ en 57 % de los casos es la más baja por electrodo; mientras en 100 % de los casos la dimensión de regularización

de la sílaba con SS /ma/ es la más alta por electrodo y la sílaba con SS /ra/ en 95 % de los casos es la más baja por electrodo.

En 38 % de los electrodos, el valor promedio por electrodo de dimensión de regularización de la vocal con SS /a/ tiende a estar separado, en cada electrodo, de las otras vocales con SS; mientras los valores promedio por electrodo de dimensión de regularización de las sílabas con SS tienden a estar separados en cada uno de los electrodos.

Para las vocales con SS, se observa mayor promedio por electrodo de dimensión de regularización para los electrodos E4, E5 y E10 pertenecientes al área de Wernicke; en contraste, los electrodos E1, E2, E6, E7 y E11 pertenecientes mayoritariamente al área de Broca, presentan menor promedio por electrodo de dimensión de regularización. En el caso de las sílabas con SS, se observa mayor promedio de dimensión de regularización para los electrodos E2, E4, E5, E15 y E20 pertenecientes al área de Wernicke. En contraste, los electrodos E11, E16, E17 y E19 pertenecientes al área de Broca, presentan menor promedio de dimensión de regularización. De esta manera se evidencia para el área de Wernicke un mayor promedio de dimensión de regularización, en tareas de habla silenciosa con vocales y sílabas, en contraste al área de Broca.

4.4. Discusión sobre señales EEG para vocales y sílabas con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización

4.4.1. Determinación del método de clasificación

A continuación se hace una discusión de los resultados obtenidos con la metodología basada en vector de características con entropía de la información para la clasificación de las vocales y sílabas con habla silenciosa.

En el caso de vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización se seleccionaron dos algoritmos de clasificación multiclase (1-rest) y (1-1). Para SVM con (1-rest) se alcanzó en la clasificación una exactitud entre 46.56 % y 85.90 % con una media estadística de 61.41 % y una desviación estándar de 3.98 % (Fig. 4.9). Para SVM con (1-1) se alcanzó en la clasificación una exactitud entre 52.67 % y 91.43 % con una media estadística de 69.83 % y una desviación estándar de 3.23 % (Fig. 4.10).

Para el caso de sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización, se seleccionaron dos algoritmos de clasificación multiclase (1-rest) y (1-1). Para SVM con (1-rest) se alcanzó en la clasificación una exactitud entre 47.97 % y 74.68 % con una media estadística de 58.57 % y una desviación estándar de 4.92 % (Fig. 4.19). Para SVM con (1-1) se alcanzó en la clasificación una exactitud entre 55.11 % y 84.22 % con una media estadística de 66.89 % y una desviación estándar de 4.03 % (Fig. 4.20).

Asimismo, se encontró que utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización, para cinco vocales y cinco sílabas con habla silenciosa el clasificador multiclase SVM (1-1) presenta mayor exactitud

que el clasificador multiclase SVM (1-*rest*). Estos resultados están en concordancia con Hsu et al. [167] y Pal et al. [175], donde se especifican las mejores prestaciones del clasificador (1-1), en comparación que con el (1-*rest*).

Contrastando los resultados de la metodología desarrollada con otros trabajos de habla silenciosa en cuanto a exactitud (*accuracy*), se encuentra que:

En el caso de vocales con habla silenciosa la metodología desarrollada en esta tesis clasifica 5 vocales en contraste con: Da Salla et al. [29] que logró de 68 % a 78 % de exactitud, Yoshimura et al. [122] con 49.9 %, Riaz et al. [123] con 75 %, clasificando todos ellos únicamente 2 vocales con habla silenciosa. El trabajo desarrollado en esta tesis tiene mayor exactitud que aquel de Kamala et al. [125], que alcanza una exactitud de 44 % donde se clasifican 5 vocales de habla inglesa con EEG. Igualmente, el trabajo desarrollado es comparable a estudios de reconocimiento para 5 vocales japonesas con EEG presentados por Matsumoto et al. [126] con un 85 % de exactitud, Hori et al. [127] con 57 % a 87 % y Matsumoto et al. [30] con un 57 %-87 %.

En el caso de sílabas con habla silenciosa, la metodología desarrollada clasifica 5 sílabas con habla silenciosa, en contraste con las siguientes estudios que clasifican exclusivamente 2 sílabas: D'Zmura [128] que logra una exactitud promedio de 75 %, Brigham et al. [31] con una exactitud de 56 % a 88 %, Deng et al. [32] con una 54 % a 73 %.

En el caso de caracteres y palabras con habla silenciosa, la metodología desarrollada en esta tesis es: mejor en función del número de clases reconocidas que Wang et al. [129], con una exactitud de 79 % a 88 %, y Wang et al. [130] con 73 % a 95 %, que sólo reconocen dos caracteres chinos; también mejor en función del número de clases reconocidas que Salamanca et al. [33], con una exactitud de 47 % a 67 %, y en Kazi et al. [34], con 27 % a 91 %, donde se reconocen dos palabras en inglés; y con resultados superiores a Porbadnigk et al. [131], García et al. [132], con una exactitud promedio de 45.5 % y 42 %, respectivamente, donde se reconocen 5 palabras en inglés y 5 en español.

En contraste con el estudio desarrollado en esta tesis, otros estudios sobre reconocimiento de vocales con habla silenciosa han utilizado algoritmos como: SVM para dos clases [29]; combinación de los algoritmos SVM con kernel gaussiano (SVM-G), máquinas de vectores relevantes con kernel gaussiano (RVM-G) y máquinas de vectores relevantes lineales (RVM-L) [127]; el algoritmo de regresión logístico (SLR-VAR) [122]; Naive Bayes [137]; modelos ocultos de Markov (HMM), con el algoritmo de los k vecinos (k-NN) [123]; y una red neuronal artificial *Backpropagation* [125]. En el caso de reconocimiento de sílabas con habla silenciosa, se ha utilizado el algoritmo de producto interno generado por la envoltura de filtros [128]; el vecino más cercano (k-NN) [31]; el clasificador Bayesiano y el análisis discriminante lineal (LDA) [32], entre otros. Para el reconocimiento de palabras con habla silenciosa se han utilizado clasificadores basados en HMM [134], entre otros. El clasificador utilizado en este estudio es un SVM multiclase (1-*rest*) y (1-1), donde el clasificador (1-1) presentó mejor exactitud para clasificar sílabas y vocales con habla silenciosa. Una de las claves para la óptima prestación de estos algoritmos fue la selección de los parámetros σ y C encontrados con algoritmos genéticos.

4.4.2. Desempeño de vector de características en función de la metodología desarrollada

Otra de las claves para el buen desempeño de los algoritmos desarrollados en función de la exactitud, es la selección del vector de características con la entropía de información y la

dimensión de regularización. En las figuras (4.5(a) a 4.5(e)) y (4.8(a) a 4.8(e)), se muestra la representación espacial, sobre el área de Wernicke y Broca, en términos de entropía de información para vocales (sujeto 14) y sílabas (sujeto 7) con habla silenciosa. También, en las figuras (4.15(a) a 4.15(e)) y (4.17(a) a 4.17(e)) se muestra la representación espacial en términos de dimensión de regularización para vocales (sujeto 14) y sílabas (sujeto 7) con habla silenciosa.

En contraste, otras investigaciones de vocales con habla silenciosa han utilizado los siguientes vectores de características: descriptores estadísticos [125]; transformada rápida de Fourier (FFT) [28]; patrones espaciales comunes (CSP) [29],[126], [127], [30]; periodogramas [124] y coeficientes cepstrales en las frecuencias de Mel (MFCC) [123]. En el caso de sílabas con habla silenciosa se han utilizado: CSP [129],[130]; análisis de componentes independientes (ICA) con modelos de autorregresión (AR) [31]; espectro de Hilbert [32] y envoltura generada por el espacio de Hilbert [128]. En el caso de palabras con habla silenciosa se ha utilizado: descriptores estadísticos [125]; FFT [133]; modelos ocultos de Markov (HMM) [134], [131] y transformadas de wavelet discretas (DWT) [132].

4.4.3. Criterios para la selección de la metodología desarrollada de habla silenciosa

A continuación se explican algunas de las razones por las cuales fueron seleccionados la ubicación de los electrodos y los componentes de la metodología desarrollada.

La ubicación de los electrodos estuvo supeditada a los modelos neurológicos del lenguaje denominados: Litchtheim, Geshwind y modelo de Hickok y Poeppel. En el modelo de Litchtheim el centro auditivo está relacionado con la comprensión del habla y el centro conceptual con el almacenamiento del significado de componentes del lenguaje (área de Wernicke); adicionalmente, el centro motor está relacionado con la producción del habla (áreas de Broca) y estas áreas con el hemisferio izquierdo (Fig. 2.9). El modelo de Geshwind relaciona tres centros en el hemisferio izquierdo (Fig. 2.10): el centro de producción del lenguaje (área de Broca), el centro de comprensión del lenguaje y el centro conceptual, ubicado en la circunvolución supramarginal y circunvolución angular, esta últimas ubicadas a su vez en el área de Wernicke [49]. En el modelo de Hickok y Poeppel hay múltiples áreas del cerebro (bloques) que trabajan juntas para el procesamiento del lenguaje [54] entre ellas se encuentran los bloques de: análisis espectral, red fonológica, análisis espectral, interfaz léxica, red combinatorial y red conceptual [56].

Por consiguiente, para vocales con habla silenciosa, los electrodos E1, E6, E7, E11, E12, E13, E14, E17, E18, E19 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E15, E16 y E21 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.3). Para sílabas con habla silenciosa, los electrodos E1, E6, E7, E11, E12, E13, E16, E17, E21 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E14, E15, E19 y E20 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.4).

Para seleccionar el número de sujetos en la experimentación de vocales y sílabas con habla silenciosa, se tuvo como referencia la cantidad de sujetos en otras experimentaciones relacionadas con habla silenciosa. Para vocales con habla silenciosa se ha utilizado la siguiente cantidad de sujetos: 3 en [29], 4 en [126], 5 en [127], 8 en [28] y 13 en [125]. Para sílabas con habla silenciosa se ha utilizado la siguiente cantidad de sujetos: 4 en [128] y 7 en [31]. Para caracteres chinos y palabras con habla silenciosa se utilizó la siguiente cantidad de sujetos: 2

en [34], 7 en [33], 8 en [130] y 21 en [131]. En esta tesis se buscó que la cantidad de sujetos estuviera por encima de la media de esos trabajos, de manera que para vocales con habla silenciosa se contó con la participación de 20 sujetos y para sílabas con habla silenciosa se contó con 10 sujetos.

Como uno de los requerimientos, en la experimentación para vocales y sílabas con habla silenciosa se les pidió a los sujetos permanecer con los ojos cerrados durante todo el tiempo. Lo anterior, por un lado, con el fin de disminuir el número de artefactos debidos al parpadeo; y por el otro, en relación con el modelo cognitivo de Benedet [52], con el fin de disminuir la cantidad de entradas de información no deseadas para el proceso de lenguaje por parte del sujeto. Cuando el sujeto cierra los ojos no tiene entrada de señales visuales relacionadas a objetos del entorno. De esta forma la información del sistema semántico se relaciona a procesos internos del cerebro, excluyendo la entrada de léxico grafémico.

Otro factor a tener en cuenta es la frecuencia de muestreo. La frecuencia de muestreo de vocales con habla silenciosa fue de 500 Hz , y un parámetro para la selección estuvo determinado por la frecuencia de muestreo mínima del equipo de adquisición de señales cerebrales, amplificador NicoletOne™ V32 (Viasys Healthcare, Inc. Conshohocken, Pennsylvania). Para el caso de las sílabas con habla silenciosa, se buscó la máxima frecuencia de muestreo del equipo de adquisición, 2000 Hz , teniendo en cuenta que la duración promedio de las sílabas para la experimentación es 200 ms o 400 muestras [176]. Tanto para sílabas como para vocales, se tuvo en cuenta que cumplieran con el criterio de Nyquist [160]. En este sentido, la metodología basada en entropía de información y dimensión de regularización la máxima frecuencia utilizada fue 50 Hz .

El parpadeo es uno de los artefactos que más afecta el desempeño de un sistema de procesamiento de señales cerebrales aplicados a BCI. En esta tesis se desarrolló un algoritmo que, primero, detecta el parpadeo (Eq. 3.1) y posteriormente lo elimina. El algoritmo desarrollado para eliminar el parpadeo está basado en wavelets y tiene tres etapas: (1) descomposición de la señal observable, (2) umbralización, y (3) reconstrucción de la señal estimada [162]. Entre las principales ventajas que presenta este algoritmo están: el hecho de que los wavelets son procesos aditivos y de esta manera se puede representar las señales cerebrales como una sumatoria de la señal a extraer, más ruido (artefactos) que deseamos eliminar; el tipo de artefacto a detectar se puede particularizar seleccionando el wavelet apropiado; los artefactos se pueden modelar como ruido blanco; y el tiempo computacional lo hace factible para ser utilizado on-line, entre otras. Las ventajas anteriores lo hacen más robusto, en comparación con métodos para eliminar el parpadeo tales como: *second order blind identification* (SOBI) [177], *second order cyclostationary statistics* (SOCS) [178], análisis de correlación canónica (CCA) [179], y *Automatic Wavelet Independent Component Analysis* (AWICA) [180], entre otros.

Las características de la metodología con entropía de la información y dimensión de la regularización se seleccionaron de acuerdo con los siguientes criterios: Shanon define la entropía de la información como el promedio de información por símbolo emitido, midiendo la incertidumbre de una fuente de información en unidades como el *bit* [140]. De esta manera, si el cerebro es un sistema de información, entonces se puede analizar como un generador de fuentes de información de tipo aleatorio [138]. Adicionalmente, se puede considerar al cerebro como un sistema de alta complejidad y en particular al lenguaje como uno de los procesos cognitivos más complejos que realiza el ser humano [44]; en este sentido, sería ideal poder medir esta complejidad. Existen distintas maneras para medir la complejidad: el exponente de Hurst, el exponente de Lyapunov, las órbitas caóticas y la dimensión fractal, entre otras

[181], [145]. En esta tesis, la dimensión de regularización fue seleccionada como una medida de la dimensión fractal en razón a las siguientes ventajas: el kernel de regularización puede ser escogido, las versiones suavizadas son adaptativas por construcción, el parámetro de suavamiento puede ser variado en muy pequeños pasos y presenta tolerancia a niveles de ruidos [149].

En la metodología de vector de características con entropía de la información y dimensión de la regularización, la selección de los algoritmos de clasificación basados en SVM estuvo supeditada al análisis comparativo de otros autores sobre clasificadores multiclase de señales EEG aplicadas a BCI. Lotte et al. [182], hacen un estudio comparativo entre los siguientes clasificadores aplicados a señales EEG: análisis discriminante lineal (LDA), SVM, redes neuronales, modelos ocultos de Markov (HMM), clasificadores Bayesianos, el vecino más cercano (kNN) y combinación de clasificadores como Boosting, entre otros. Lotte et al. presentan un cuadro comparativo para la selección del óptimo clasificador en términos de: linealidad-no linealidad, generativo-discriminativo, estático-dinámico, estable-inestable y robustez a la alta dimensión. Para el caso asincrónico y sincrónico, el clasificador SVM presta un alto desempeño, en comparación con los otros clasificadores. Lee et al. [183] concluyen que para señales EEG de alta dimensión el clasificador SVM es más preciso que otros clasificadores como análisis discriminante lineal.

4.5. Resultados de señales EEG para vocales con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales

En este caso, se presentan los resultados de vocales con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales para ritmos δ , $\delta\theta$ y $\delta\theta\alpha$.

4.5.1. Resultados por sujeto

La figura 4.21 muestra los resultados correspondientes al valor promedio por sujeto y por vocal de la estimación del PSD ($\bar{X}_{s,v,\delta}$), para vocales con habla silenciosa utilizando ritmos cerebrales δ con frecuencias entre 2, \dots , 4 Hz. Estos valores se obtuvieron con la expresión dada por:

$$\bar{X}_{s,v,\delta} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1 l_2} \quad (4.14)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.52), $l_1 = 180$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una vocal con habla silenciosa, y p es un electrodo.

De la figura (4.21) se observa que existe una variación en los valores promedio del PSD estimado entre sujetos y entre vocales con habla silenciosa. El rango de valores promedio estuvo entre 0.07 a 0.9 ($\mu V^2/Hz$). Los valores promedio más altos de PSD estimado fueron obtenidos por el sujeto 16 y los mínimos valores promedio de PSD estimado fueron obtenidos por el sujeto 5. Teniendo en cuenta el valor promedio más alto de PSD estimado por sujeto, se encuentra que la vocal con SS /a/ tiene el valor promedio más alto en 40 % de los sujetos, y

la vocal con SS /o/ tienen el segundo valor promedio más alto de PSD estimado con 25 % de los sujetos. En contraste, el mínimo valor promedio de PSD estimado por sujeto, fue obtenido por la vocal con SS /u/ en 45 % de los casos.

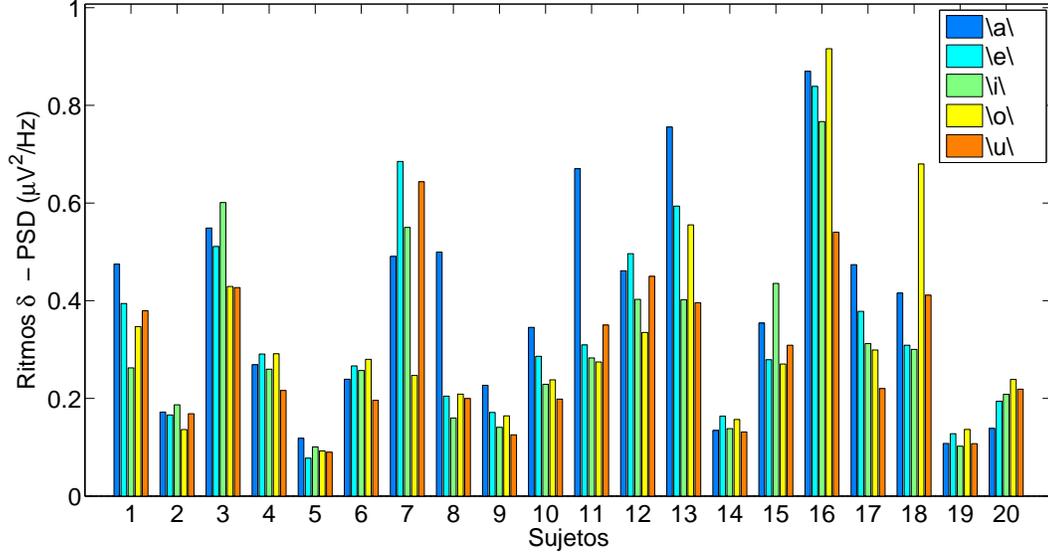


Figura 4.21: Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos δ

La figura (4.22) muestra la dispersión de los datos, del valor promedio del PSD estimado para cada vocal con SS con ritmos cerebrales δ , utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.14 para cada sujeto y para cada vocal con SS (Fig. (4.21)). Los valores promedio para las vocales con SS (/a/, /e/, /i/, /o/, /u/) son, respectivamente, (0.385, 0.289, 0.261, 0.273, 0.220) ($\mu V^2/Hz$). El valor de la mediana es más alto para las vocales con SS /a/ y /e/ y la menor para la vocal con SS /u/. El rango intercuartil (IQR), con la mayor dispersión es para la vocal con SS /a/ y el IQR con la menor dispersión es para la vocal con SS /o/. En el caso de las vocales con SS /e/, /i/ y /u/ se presenta asimetría positiva, y para las vocales con SS /a/ y /o/ se presenta asimetría negativa. Adicionalmente, se presentan outliers en las vocales con SS /i/ y /o/.

La figura 4.23 muestra los resultados correspondientes al valor promedio por sujeto y por vocal de la estimación del PSD ($\bar{X}_{s,v,\delta\theta}$), para vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta$ con frecuencias entre 2, ..., 8 Hz. Estos valores se obtuvieron con la expresión dada por:

$$\bar{X}_{s,v,\delta\theta} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1 l_2} \quad (4.15)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.53), $l_1 = 420$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una vocal con habla silenciosa, y p es un electrodo.

De la figura (4.23) se observa que existe una variación en los valores promedio del PSD estimado entre sujetos y entre vocales con habla silenciosa. El rango de valores promedio estuvo entre 0.06 a 0.62 ($\mu V^2/Hz$). Los valores promedio más altos de PSD estimado fueron

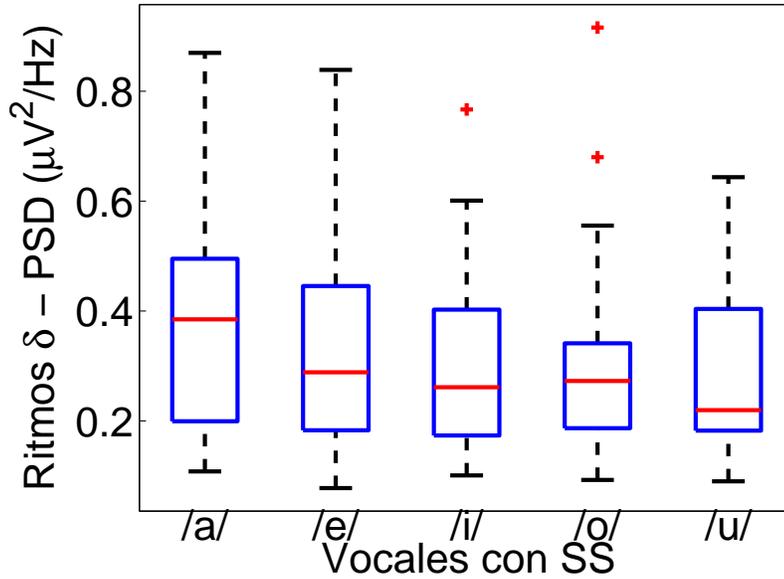


Figura 4.22: Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos δ

obtenidos por el sujeto 16 y los mínimos valores promedio de PSD estimado fueron obtenidos por el sujeto 5. Teniendo en cuenta el valor promedio más alto de PSD estimado por sujeto, se encontró que la vocal con SS /a/ tiene el valor promedio más alto en 50% de los sujetos, y las vocales con SS /e/ y /o/ tienen, respectivamente, el segundo valor promedio más alto de PSD estimado, con 25% de los sujetos cada una. En contraste, el mínimo valor promedio de PSD estimado por sujeto, fue obtenido por la vocal con SS /u/ en 40% de los casos.

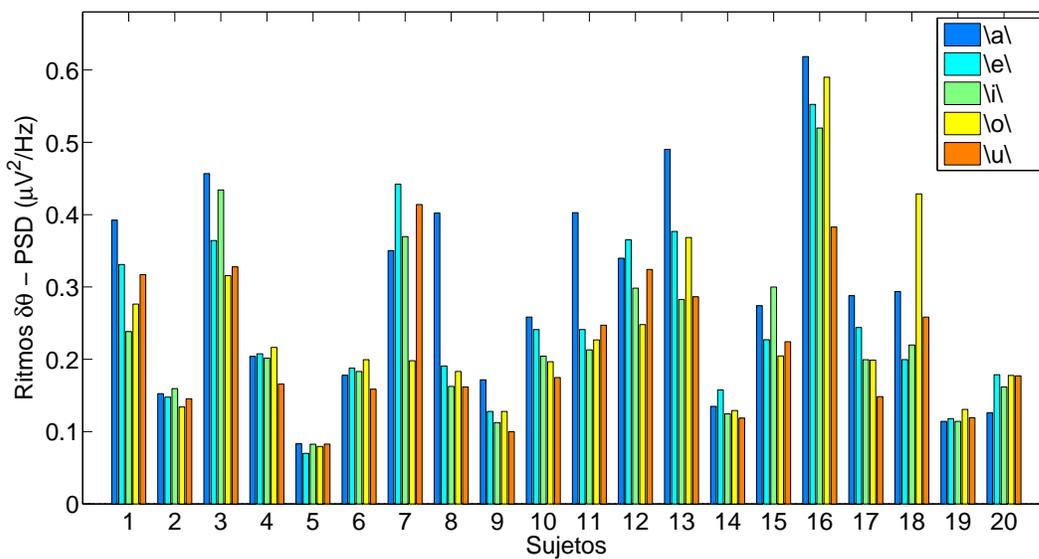


Figura 4.23: Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos $\delta\theta$

La figura (4.24) muestra la dispersión de los datos, del valor promedio del PSD estimado para cada vocal con SS con ritmos cerebrales $\delta\theta$, utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.15 para cada sujeto y para cada vocal con SS (Fig. (4.23)). Los valores promedio para las vocales con SS (/a/, /e/, /i/, /o/, /u/) son, respectivamente, (0.281, 0.217, 0.203, 0.199, 0.176) ($\mu V^2/Hz$). El valor de la mediana es más alto para las vocales con SS /a/ y /e/ y la menor para la vocal con SS /u/. El rango intercuartil (IQR), con la mayor dispersión es para la vocal con SS /a/ y el IQR con la menor dispersión es para la vocal con SS /o/. En el caso de las vocales con SS /e/, /i/, /o/ y /u/ presentan asimetría positiva. Adicionalmente, existen outliers en las vocales con SS /i/ y /o/.

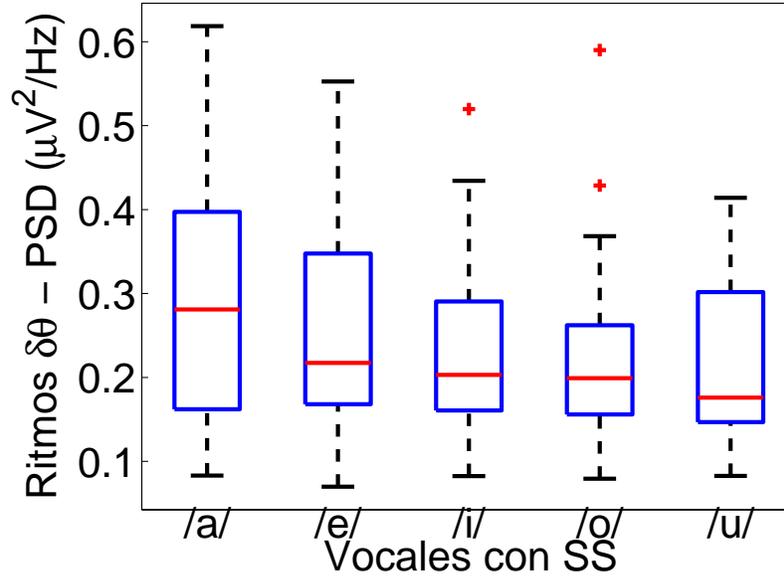


Figura 4.24: Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos $\delta\theta$

La figura 4.25 muestra los resultados correspondientes al valor promedio por sujeto y por vocal de la estimación del PSD ($\bar{X}_{s,v,\delta\theta\alpha}$), para vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta\alpha$ con frecuencias entre 2, \dots , 13 Hz. Estos valores se obtuvieron con la expresión dada por:

$$\bar{X}_{s,v,\delta\theta\alpha} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1 l_2} \quad (4.16)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.54), $l_1 = 720$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una vocal con habla silenciosa, y p es un electrodo.

De la figura (4.25) se observa que existe una variación en los valores promedio del PSD estimado entre sujetos y entre vocales con habla silenciosa. El rango de valores promedio estuvo entre 0.08 a 0.55 ($\mu V^2/Hz$). Los valores promedio más altos de PSD estimado fueron obtenidos por el sujeto 16 y los mínimos valores promedio de PSD estimado fueron obtenidos por el sujeto 5. Teniendo en cuenta el valor más alto de PSD estimado por sujeto, se encontró que la vocal con SS /a/ tiene el valor más alto en el 40% de los sujetos. En contraste,

el mínimo valor de PSD estimado por sujeto, fue obtenido por la vocal con SS /u/ en el 35 % de los casos.

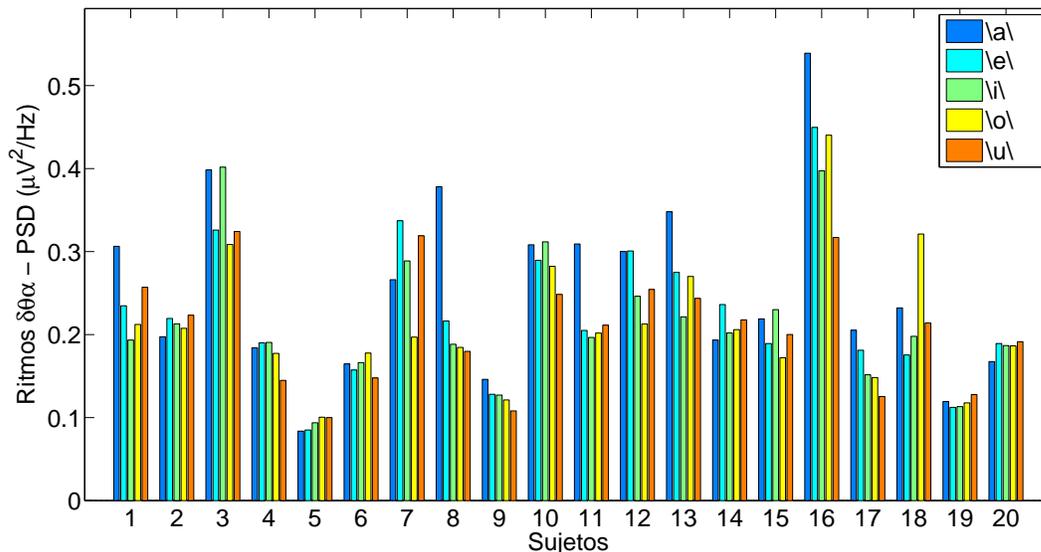


Figura 4.25: Valor promedio por sujeto y por vocal con habla silenciosa (SS) del PSD estimado - ritmos $\delta\theta\alpha$

La figura (4.26) muestra la dispersión de los datos, del valor promedio del PSD estimado para cada vocal con SS con ritmos cerebrales $\delta\theta\alpha$, utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.16 para cada sujeto y para cada vocal con SS (Fig. (4.23)). Los valores promedio para las vocales con SS (/a/, /e/, /i/, /o/, /u/) son, respectivamente, (0.225, 0.211, 0.197, 0.199, 0.213) ($\mu V^2/Hz$). El valor de la mediana es más alto para las vocales con SS /a/ y /u/ y la menor para la vocal con SS /i/. El rango intercuartil (IQR), con la mayor dispersión es para la vocal con SS /a/ y el IQR con la menor dispersión es para la vocal con SS /i/. En el caso de las vocales con SS /a/, /e/, /i/ y /o/ presentan asimetría positiva. Al contrario, la vocal con SS /u/ presenta asimetría negativa. Adicionalmente, se presentan outliers en las vocales con SS /a/, /e/, /i/ y /o/.

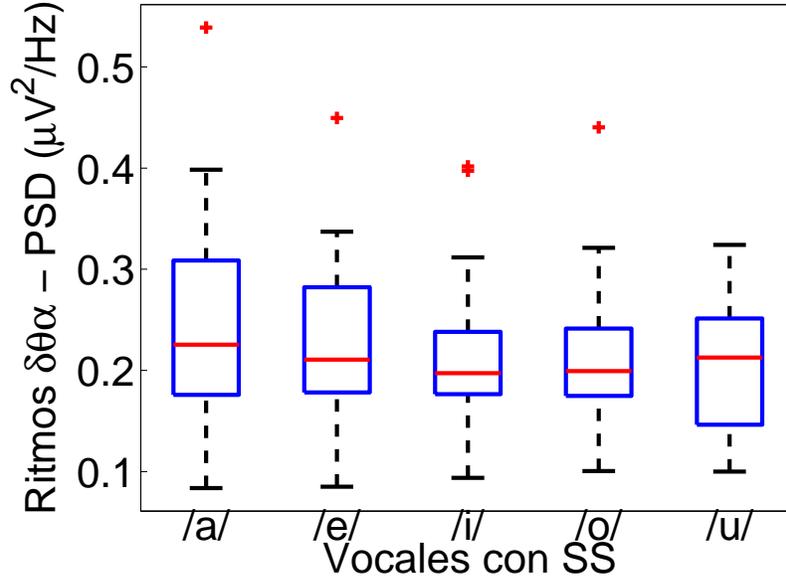


Figura 4.26: Diagrama de dispersión del valor promedio de PSD estimado por vocal con habla silenciosa (SS) - ritmos $\delta\theta\alpha$

4.5.2. Resultados por localización

Las figuras 4.27(a) a 4.27(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para el PSD estimado ($\tilde{X}_{s,v,\delta}^p$), realizando tareas de habla silenciosa con vocales. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (sujeto 1), de acuerdo al PSD estimado (Fig. 4.42). De esta manera, el valor promedio de cada electrodo ($\mu V^2/Hz$) se obtuvo con:

$$\tilde{X}_{s,v,\delta}^p = \frac{\sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1} \quad (4.17)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.52), $l_1 = 180$ es el número de filas de la matriz, k es una fila de la matriz, $s = 1$ es el sujeto seleccionado, v es una vocal con habla silenciosa, y p es un electrodo. Con los valores $\tilde{H}_{s,v,\delta}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.27(a) a 4.27(e). La distribución del color representa el valor del PSD estimado entre 0 (azul) y 0.7 (rojo) ($\mu V^2/Hz$). Los electrodos son posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.3). Los electrodos E1, E6, E7, E11, E12, E13, E14, E17, E18, E19 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E15, E16 y E21 sobre el área de Wernicke, y el electrodo E3 sobre la corteza auditiva primaria. Estas áreas fueron seleccionadas por su crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las vocales con SS se presenta:

Primero, la vocal con SS /a/ (Fig. 4.27 (a)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que

en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E15 y E16. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E11 y E13. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.55 (\mu V^2/Hz)$.

Segundo, la vocal con SS /e/ (Fig. 4.27 (b)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con el más alto PSD estimado son E4, E5, E9 y E10. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E13. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.5 (\mu V^2/Hz)$.

Tercero, la vocal con SS /i/ (Fig. 4.27 (c)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en el área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9 y E10. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E11. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.2 (\mu V^2/Hz)$.

Cuarto, la vocal con SS /o/ (Fig. 4.27 (d)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10 y E16. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7 y E13. Además, en la corteza auditiva primaria E3 tiene un valor promedio de posición del PSD estimado cercano a $0.4 (\mu V^2/Hz)$.

Quinto, la vocal con SS /u/ (Fig. 4.27 (e)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E15 y E16. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E11. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.45 (\mu V^2/Hz)$.

La figura (4.28) representa el valor promedio por electrodo del PSD estimado ($\overleftarrow{X}_{s,v,\delta}^p$), referido a cada vocal con habla silenciosa (ritmos δ), para los 21 electrodos y los 20 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{X}_{s,v,\delta}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} X_{s,v,k}^p}{(S_T)(l_1)} \quad (4.18)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.52), $l_1 = 180$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 20$ es el número de sujetos, v es una vocal con habla silenciosa, y p es un electrodo.

En la figura (4.28), el cuadrado representa el PSD estimado por electrodo de la vocal con SS /a/, el círculo representa la PSD estimado por electrodo de la vocal con SS /e/, el rombo representa el PSD estimado por electrodo de la vocal con SS /i/, el signo (+) representa el PSD estimado por electrodo de la vocal con SS /o/ y el triángulo representa el PSD estimado por electrodo de la vocal con SS /u/. El rango de valores de PSD estimado se encuentra entre

0.10 y 0.60 ($\mu V^2/Hz$) para los ritmos cerebrales δ . En el 95 % de los casos el PSD estimado de la vocal con SS /a/ es más alto en la mayoría de electrodos que en otras vocales con SS. La vocal con SS /u/ obtuvo el menor valor de PSD estimado en el 81 % de los electrodos. El valor más alto de PSD estimado lo obtuvo la vocal con SS /a/ en el electrodo 2 y el menor valor de PSD estimado lo obtuvo la vocal con SS /u/ en el electrodo 18.

En 76 % de los electrodos el valor promedio de la vocal con SS /a/ tiende a estar separado de las otras vocales con SS y en 48 % de los electrodos el valor promedio de la vocal con SS /e/ tiende a estar separado de las otras vocales con SS. En la figura (4.28) se observa mayor promedio de PSD estimado para los electrodos E2, E5, E10 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el área de Wernicke un mayor promedio de PSD estimado para ritmos δ , en tareas de habla silenciosa con vocales, en contraste a el área de Broca (ver figura 3.3).

Las figuras 4.29(a) a 4.29(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para el PSD estimado ($\tilde{X}_{s,v,\delta\theta}^p$), realizando tareas de habla silenciosa con vocales. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (sujeto 1), de acuerdo al PSD estimado (Fig. 4.42). De esta manera, el valor promedio de cada electrodo ($\mu V^2/Hz$) se obtuvo con:

$$\tilde{X}_{s,v,\delta\theta}^p = \frac{\sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1} \quad (4.19)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.53), $l_1 = 420$ es el número de filas de la matriz, k es una fila de la matriz, $s = 1$ es el sujeto seleccionado, v es una vocal con habla silenciosa, y p es un electrodo. Con los valores $\tilde{H}_{s,v,\delta\theta}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.29(a) a 4.29(e). La distribución del color representa el valor del PSD estimado entre 0 (azul) y 0.7 (rojo) ($\mu V^2/Hz$). Los electrodos son posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.3). Estas áreas fueron seleccionadas por su crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las vocales con SS se presenta:

Primero, la vocal con SS /a/ (Fig. 4.29 (a)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta$ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E15 y E16. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7, E11 y E13. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.55 ($\mu V^2/Hz$).

Segundo, la vocal con SS /e/ (Fig. 4.29 (b)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) en los ritmos cerebrales $\delta\theta$ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con más alto promedio de posición del PSD estimado son E4, E5, E9 y E10. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7 y E13. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.5 ($\mu V^2/Hz$).

Tercero, la vocal con SS /i/ (Fig. 4.29 (c)) tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) en los ritmos cerebrales $\delta\theta$ más alto en el área de Wernicke que en el

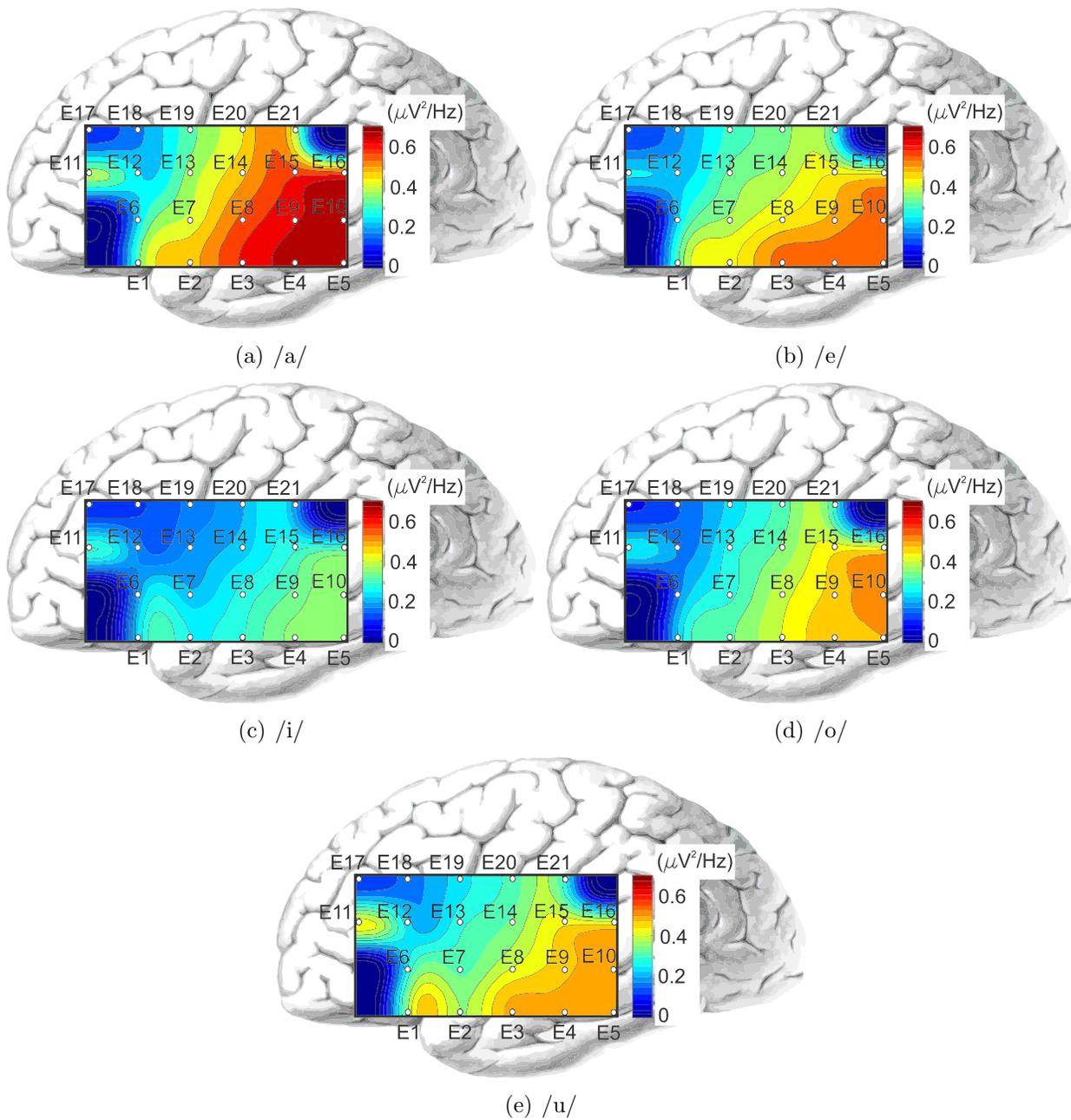


Figura 4.27: Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos δ

área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E5 y E10. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7 y E11. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.3 (\mu V^2/Hz)$.

Cuarto, las vocales con SS /o/ y /u/ (Figs. 4.29 (d) y 4.29 (e)), tienen valores promedio de posición del PSD estimado ($\mu V^2/Hz$) en los ritmos cerebrales $\delta\theta$ más altos en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10 y E16. Para el área de

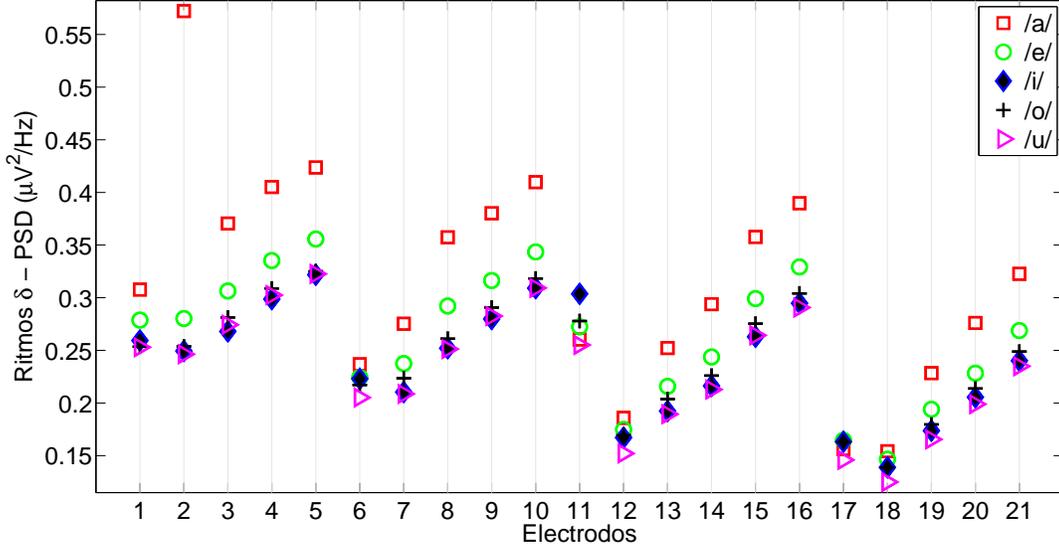


Figura 4.28: Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos δ

Broca, el valor promedio de posición más alto de PSD estimado está en E7 en el caso de la vocal con SS/o/ ,y en E1 y E11 en el caso de la vocal con SS /u/. Además, en la corteza auditiva primaria E3 tiene un valor promedio de posición del PSD estimado cercano a 0.3 y 0.5 respectivamente ($\mu V^2/Hz$).

La figura (4.30) representa el valor promedio por electrodo del PSD estimado ($\overleftarrow{X}_{s,v,\delta\theta}^p$), referido a cada vocal con habla silenciosa (ritmos $\delta\theta$), para los 21 electrodos y los 20 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{X}_{s,v,\delta\theta}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} X_{s,v,k}^p}{(S_T)(l_1)} \quad (4.20)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.53), $l_1 = 420$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 20$ es el número de sujetos, v es una vocal con habla silenciosa, y p es un electrodo.

En la figura (4.30), el rango de valores promedio por electrodo del PSD estimado se encuentra entre 0.10 y 0.35 ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta$. En el 90 % de los casos el PSD estimado de la vocal con SS /a/ es más alta en la mayoría de electrodos que para en las otras vocales con SS. La vocal con SS /u/ obtuvo el menor valor de PSD estimado en 100 % de los electrodos. El valor promedio por electrodo más alto de PSD estimado lo obtuvo la vocal con SS /a/ en el electrodo 2 y el menor valor promedio por electrodo del PSD estimado lo obtuvo la vocal con SS /u/ en el electrodo 18.

En 76 % de los electrodos el valor promedio de la vocal con SS /a/ tiende a estar separado de las otras vocales con SS y en 48 % de los electrodos el valor promedio de la vocal con SS /e/ tiende a estar separado de las otras vocales con SS. En la figura (4.30) se observa mayor promedio de PSD estimado para los electrodos E2, E4, E5, E9, E10, E15 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el

área de Wernicke un mayor promedio de PSD estimado para ritmos $\delta\theta$, en tareas de habla silenciosa con vocales, en contraste a el área de Broca (ver figura 3.3).

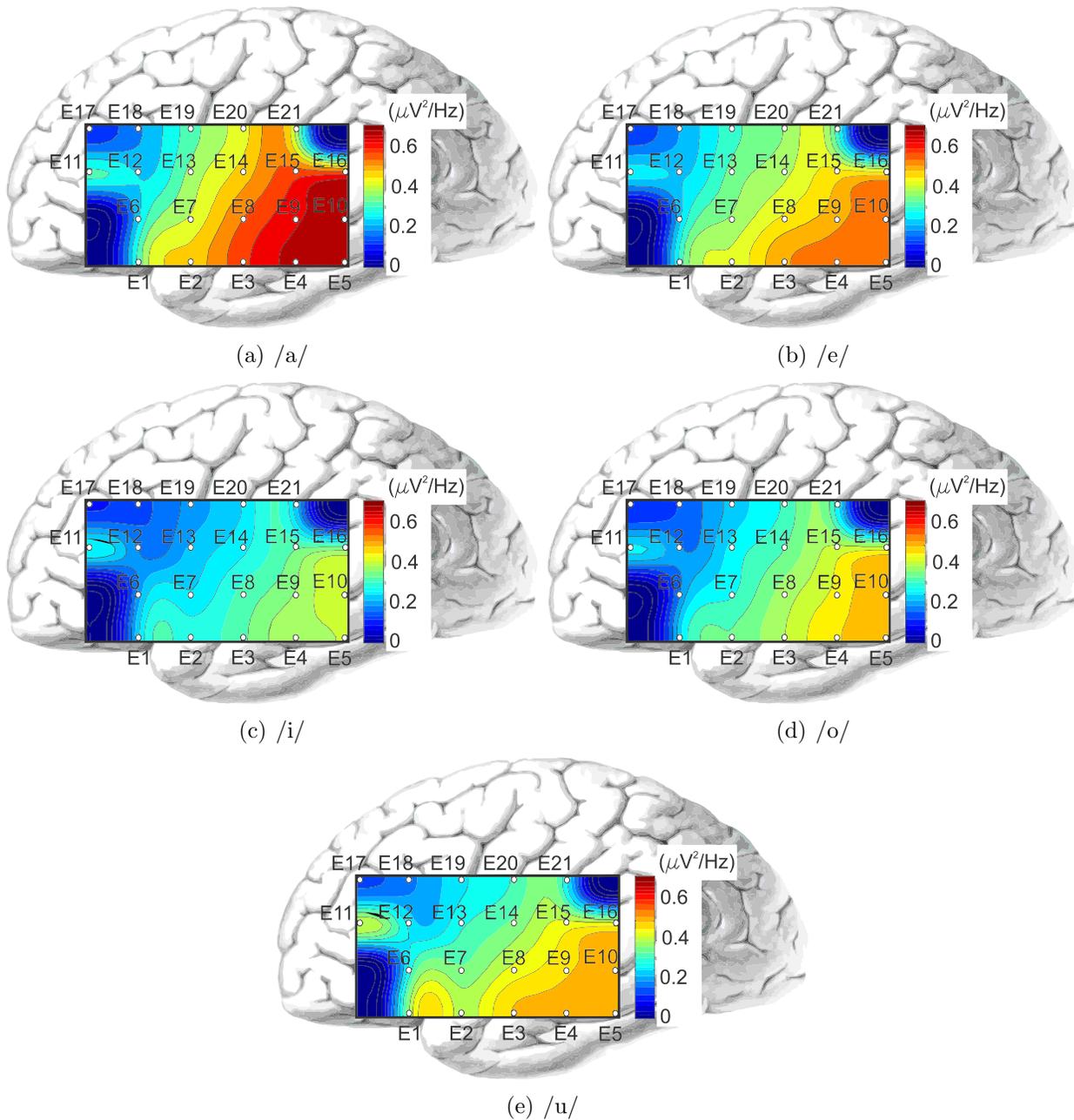


Figura 4.29: Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos $\delta\theta$

Las figuras 4.31(a) a 4.31(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para el PSD estimado ($\tilde{X}_{s,v,\delta\theta\alpha}^p$), realizando tareas de habla silenciosa con vocales. En este caso, se seleccionó un individuo calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (sujeto 1), de acuerdo al PSD estimado (Fig. 4.42). De esta manera, el valor promedio de cada electrodo ($\mu V^2/Hz$) se obtuvo con:

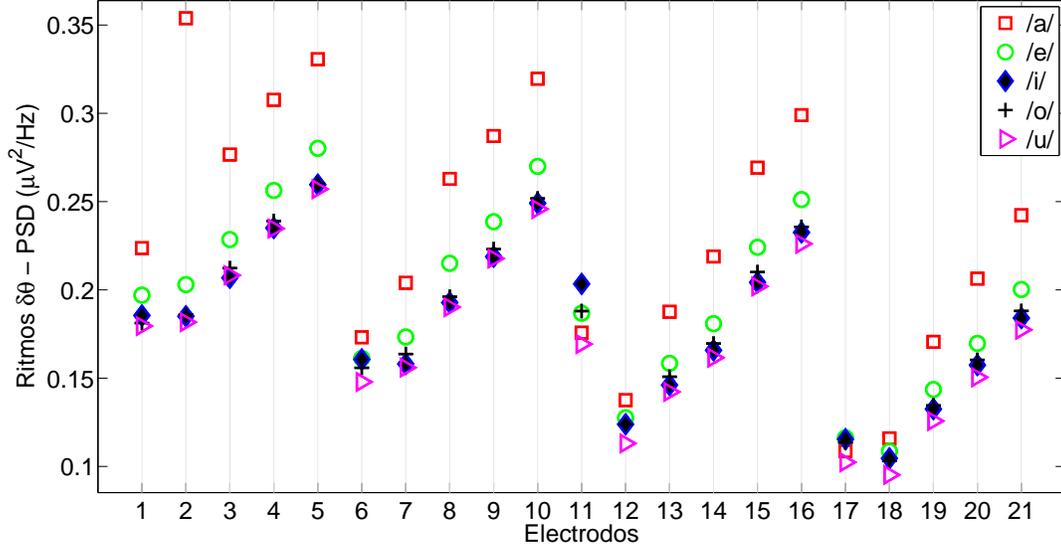


Figura 4.30: Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos $\delta\theta$

$$\tilde{X}_{s,v,\delta\theta\alpha}^p = \frac{\sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1} \quad (4.21)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.54), $l_1 = 720$ es el número de filas de la matriz, k es una fila de la matriz, $s = 1$ es el sujeto seleccionado, v es una vocal con habla silenciosa, y p es un electrodo. Con los valores $\tilde{X}_{s,v,\delta\theta\alpha}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.31(a) a 4.31(e). La distribución del color representa el valor del PSD estimado entre 0 (azul) y 0.5 (rojo) ($\mu V^2/Hz$). Los electrodos son posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.3). Estas áreas fueron seleccionadas por su crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las vocales con SS se presenta:

Primero, la vocal con SS /a/ (Fig. 4.31 (a)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta\alpha$ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E15 y E16. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E7. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.35 ($\mu V^2/Hz$).

Segundo, la vocal con SS /e/ (Fig. 4.31 (b)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta\alpha$ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con más alto promedio de posición del PSD estimado son E4, E5, E9 y E10. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E7. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.3 ($\mu V^2/Hz$).

Tercero, la vocal con SS /i/ (Fig. 4.31 (c)), tiene un valor promedio de posición del PSD

estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta\alpha$ más alto para el área de Wernicke que para el área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E5 y E10. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7 y E11. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.25 (\mu V^2/Hz)$.

Cuarto, la vocal con SS /o/ (Fig. 4.31 (d)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta\alpha$ más alto para el área de Wernicke que para las otras áreas. En el área de Wernicke el electrodo con el más alto valor promedio de posición del PSD estimado es E10. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E7. Además, en la corteza auditiva primaria E3 tiene un valor promedio de posición del PSD estimado cercano a $0.28 (\mu V^2/Hz)$.

Quinto, la vocal con SS /u/ (Fig. 4.31 (e)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales $\delta\theta\alpha$ más alto para el área de Wernicke que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9 y E10. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E7 y E11. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a $0.32 (\mu V^2/Hz)$.

La figura (4.32) representa el valor promedio por electrodo del PSD estimado ($\overleftarrow{X}_{s,v,\delta\theta\alpha}^p$), referido a cada vocal con habla silenciosa (ritmos $\delta\theta\alpha$), para los 21 electrodos y los 20 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{X}_{s,v,\delta\theta\alpha}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} X_{s,v,k}^p}{(S_T)(l_1)} \quad (4.22)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.54), $l_1 = 720$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 20$ es el número de sujetos, v es una vocal con habla silenciosa, y p es un electrodo.

En la figura (4.32), el rango de valores promedio por electrodo del PSD estimado se encuentra entre 0.07 y $0.35 (\mu V^2/Hz)$ para los ritmos cerebrales $\delta\theta\alpha$. En el 90 % de los casos el PSD estimado de la vocal con SS /a/ es más alta en la mayoría de electrodos que para en las otras vocales con SS. La vocal con SS /u/ obtuvo el menor valor de PSD estimado en 95 % de los electrodos. El valor promedio por electrodo más alto de PSD estimado lo obtuvo la vocal con SS /a/ en el electrodo 5 y el menor valor promedio por electrodo del PSD estimado lo obtuvo la vocal con SS /u/ en el electrodo 17.

En 76 % de los electrodos el valor promedio de la vocal con SS /a/ tiende a estar separado de las otras vocales con SS. En la figura (4.32) se observa mayor promedio de PSD estimado para los electrodos E4, E5, E10 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el área de Wernicke un mayor promedio de PSD estimado para ritmos $\delta\theta\alpha$, en tareas de habla silenciosa con vocales, en contraste a el área de Broca (ver figura 3.3).

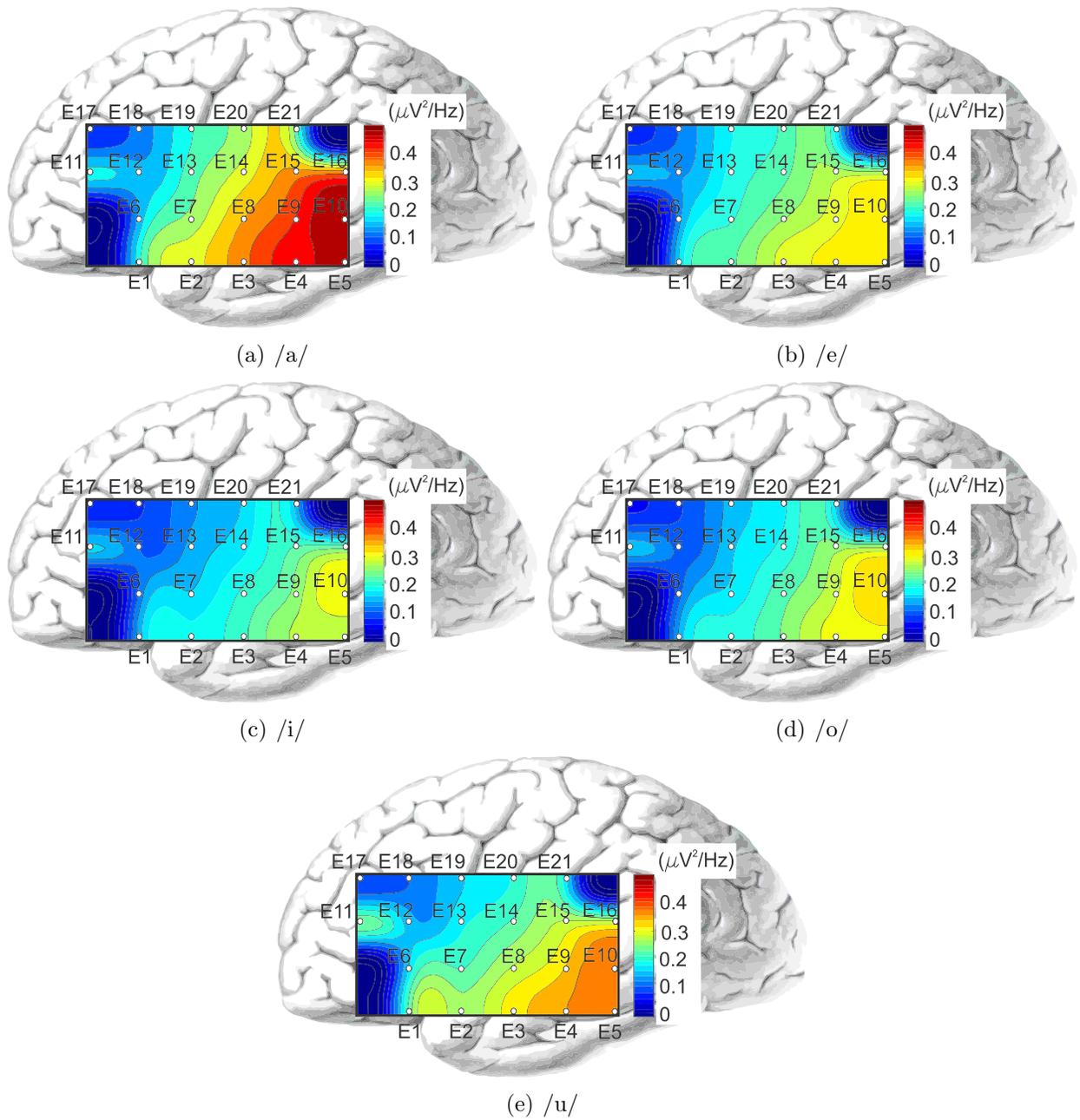


Figura 4.31: Representación espacial del valor promedio por electrodo del PSD estimado para vocales con habla silenciosa (SS) - ritmos $\delta\theta\alpha$

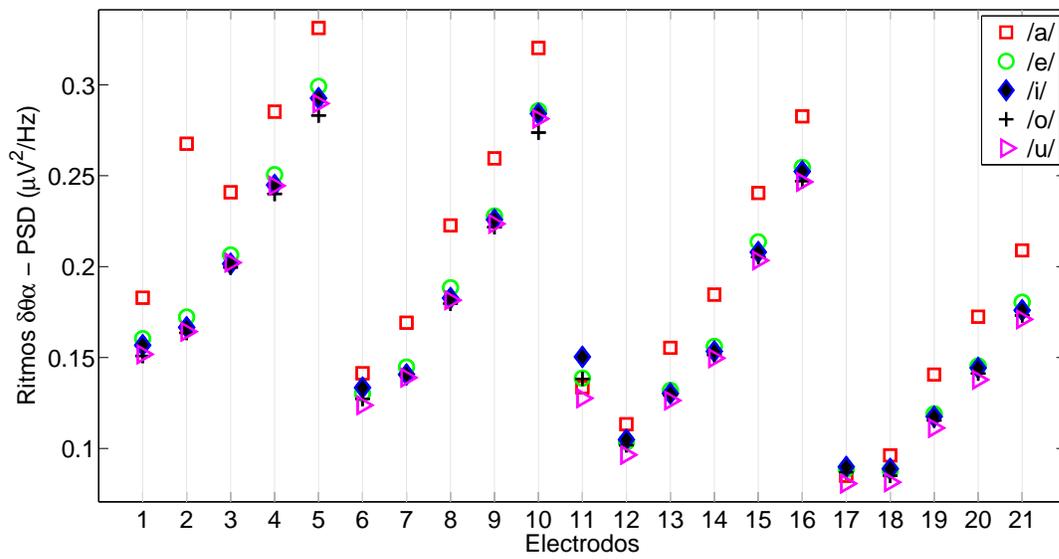


Figura 4.32: Valor promedio de PSD estimado por electrodo para vocales con habla silenciosa (SS) - ritmos $\delta\theta\alpha$

4.5.3. Resultados utilizando datos funcionales

La tabla (4.1) muestra el valor mínimo obtenido del error cuadrático medio (MSE) aplicado a los ritmos cerebrales δ , $\delta\theta$ y $\delta\theta\alpha$ cuando los sujetos piensan vocales con SS. Cada dato en la tabla (4.1) se obtuvo con la ecuación (2.46) variando el número de bases y los ritmos cerebrales para cada sujeto. Para seleccionar el número de bases óptimo K por ritmo cerebral para los 20 sujetos, se utiliza el valor de la mediana. Para el grupo δ el número de bases óptimo es $K = 10$, para el grupo $\delta\theta$ el número de bases óptimo es $K = 11$ y para el grupo $\delta\theta\alpha$ el número de bases óptimo es $K = 12$.

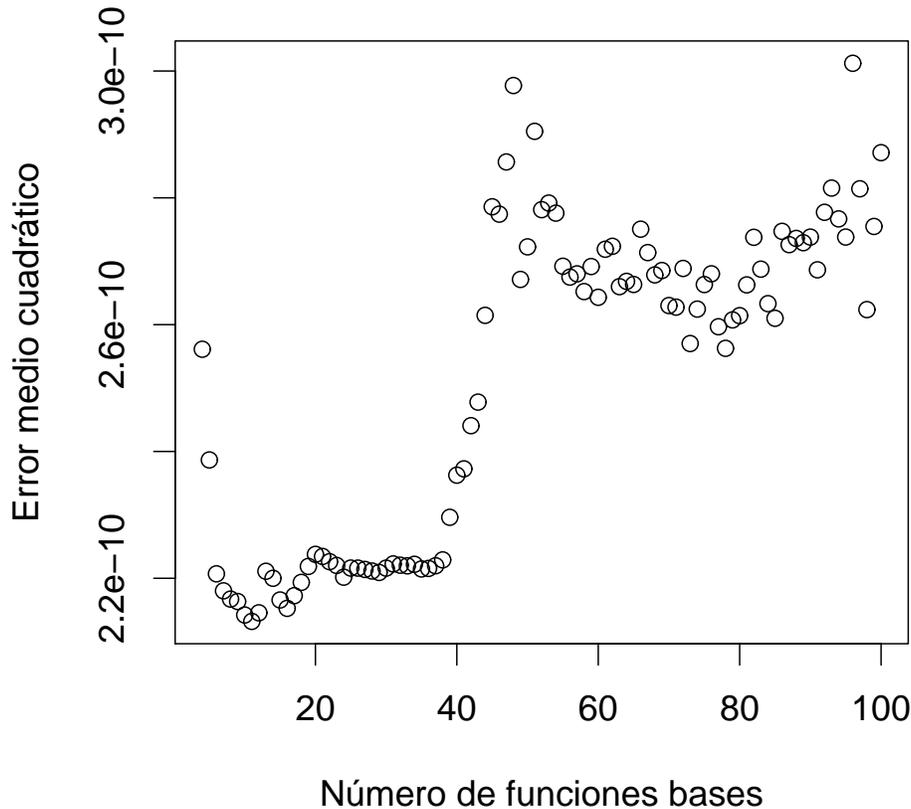


Figura 4.33: Error medio cuadrático (MSE) relacionado al número de bases para vocales con habla silenciosa (SS), utilizando ritmos $\delta\theta\alpha$

Para el sujeto 1 se muestra la variación del número de bases con los ritmos cerebrales $\delta\theta\alpha$ (ver gráfica 4.33), en donde el valor óptimo (MSE) es 11.

Las figuras 4.34 (a) a (f) y 4.35 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v,\delta}(\cdot)$ y sus respectivas derivadas $D\hat{g}_{s,v,\delta}(\cdot)$ en donde el dominio corresponde a la la distancia angular entre electrodos, para tareas de habla silenciosa con vocales (ritmos δ). En este caso, se seleccionó un individuo (sujeto 1) calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (Fig. 4.42). De esta manera, las curvas funcionales se obtuvieron con:

Tabla 4.1: MSE variando el número de bases para vocales con habla silenciosa (SS) - ritmos cerebrales δ , $\delta\theta$ y $\delta\theta\alpha$

Sujeto	δ	$\delta\theta$	$\delta\theta\alpha$
1	6	11	11
2	19	17	38
3	6	7	9
4	14	11	18
5	13	11	36
6	13	13	12
7	14	9	10
8	6	11	35
9	10	10	37
10	6	11	38
11	100	7	7
12	10	9	9
13	6	7	10
14	14	11	36
15	6	11	11
16	6	11	11
17	6	11	11
18	6	10	10
19	13	18	38
20	14	11	18
mediana	10	11	11.5

$$\widehat{g}_{s,v,\delta}(\Theta) = \sum_{k=1}^K c_k \aleph_k^m(\Theta) \quad (4.23)$$

donde $\widehat{g}_{s,v,\delta}(\cdot)$ son las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 10$ es el número de bases (ver tabla 4.1), k es la referencia a las curvas funcionales estimadas, c_k son los coeficientes, $\aleph_k^m(\cdot)$ corresponde a los B-splines con grado $m = 3$.

De manera similar, las derivadas de las curvas funcionales se obtuvieron con:

$$D\widehat{g}_{s,v,\delta}(\Theta) = \sum_{k=1}^K c'_k D\aleph_k^m(\Theta) \quad (4.24)$$

donde $D\widehat{g}_{s,v,\delta}(\cdot)$ corresponde a la primera derivada de las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 10$ es el número de bases (ver tabla 4.1), k es la referencia a la primera derivada de las curvas funcionales estimadas, c'_k son los coeficientes de la primera derivada, $D\aleph_k^m(\cdot)$ corresponde a la primera derivada de los B-splines con grado $m = 3$. La media de las curvas funcionales y la media de sus respectivas derivadas se calculan basados en la ecuación (Eq. 2.50).

Los datos funcionales (Figs. 4.34 y 4.35) son curvas abiertas con una dimensión topológica igual a 1. La medias de las funciones $\widehat{g}_{s,v,\delta}(\cdot)$ tienen un máximo global (entre 400 a 500 grados) y un mínimo global (cerca a 900 grados). Los puntos globales de las medias tienen un valor de cero en sus respectivas derivadas $D\widehat{g}_{s,v,\delta}(\cdot)$ (figures 4.34 (b, d, f) y 4.35 (b, d)). Hay ciertas diferencias entre las curvas referidas a la media estadística de las vocales con SS. Entre ellas se encuentran:

El máximo global tiene un valor aproximado de 0.7 ($\mu V^2/Hz$) para la media de la función vocal con SS /a/ y un valor aproximado de 0.35 ($\mu V^2/Hz$) para la media de la función vocal con SS /i/. El valor del máximo global para las medias de las otras funciones (/e/, /o/ y /u/) se encuentra entre los dos rangos anteriores.

De acuerdo con el número de puntos de inflexión, (en inglés, *inflection points* (IP)), el número de IPs varía entre las medias de las curvas de las funciones vocales con SS antes del máximo global. En este caso, se encuentra que el número de IPs claramente demarcados son: 0 IP para la media de la función vocal con SS /o/, 1 IP para la media de las funciones vocales con SS /a/, /i/ y /o/, y 2 IPs para la media de las funciones vocales con SS /e/ y /u/. El número de IPs claramente demarcados entre el máximo y mínimo global para las medias de las curvas funciones vocales son: 2 IPs para la media de la función vocal con SS /e/, y 1 IP para las medias de las funciones vocales con SS /a/, /i/, /o/ y /u/.

Las figuras 4.36 (a) a (f) y 4.37 (a) a (d) muestran las curvas funcionales $\widehat{g}_{s,v,\delta\theta}(\cdot)$ y sus respectivas derivadas $D\widehat{g}_{s,v,\delta\theta}(\cdot)$ en donde el dominio corresponde a la la distancia angular entre electrodos, para tareas de habla silenciosa con vocales (ritmos $\delta\theta$). En este caso, se seleccionó un individuo (sujeto 1) calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (Fig. 4.42). De esta manera, las curvas funcionales se obtuvieron con:

$$\widehat{g}_{s,v,\delta\theta}(\Theta) = \sum_{k=1}^K c_k \aleph_k^m(\Theta) \quad (4.25)$$

donde $\widehat{g}_{s,v,\delta\theta}(\cdot)$ son las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 11$ es el número de bases (ver tabla 4.1), k es la referencia a las curvas funcionales

estimadas, c_k son los coeficientes, $\aleph_k^m(\cdot)$ corresponde a los B-splines con grado $m = 3$.

De manera similar, las derivadas de las curvas funcionales se obtuvieron con:

$$D\widehat{g}_{s,v,\delta\theta}(\Theta) = \sum_{k=1}^K c'_k D\aleph_k^m(\Theta) \quad (4.26)$$

donde $D\widehat{g}_{s,v,\delta\theta}(\cdot)$ corresponde a la primera derivada de las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 11$ es el número de bases (ver tabla 4.1), k es la referencia a la primera derivada de las curvas funcionales estimadas, c'_k son los coeficientes de la primera derivada, $D\aleph_k^m(\cdot)$ corresponde a la primera derivada de los B-splines con grado $m = 3$. La media de las curvas funcionales y la media de sus respectivas derivadas se calculan basados en la ecuación (Eq. 2.50).

Las medias de las funciones $\widehat{g}_{s,v}(\cdot)$ tienen un máximo global (cerca a 1000 grados) y un mínimo global (cerca a 2150 grados). Los puntos globales tienen un valor de cero en las respectivas medias de las derivadas $D\widehat{g}_{s,v}(\cdot)$ (figuras 4.36 (b, d, f) y 4.37 (b, d)). Hay diferencias entre curvas de habla silenciosa relacionadas con los valores de la curva media tales como:

El máximo global, con un valor cercano a $0.55 (\mu V^2/Hz)$ para la media de la función vocal con SS /a/ y cercano a $0.35 (\mu V^2/Hz)$ para la media de la función vocal con SS /i/; y el valor del máximo global para las otras medias de las funciones (/e/, /o/ y /u/) que se encuentran entre 0.35 a $0.5 (\mu V^2/Hz)$. De acuerdo al número de puntos de inflexión, el número de IPs varía entre las medias de las curvas de las funciones vocales con SS, antes del máximo global. En este caso, se encuentra que el número de IPs claramente demarcados son: 1 IP para las medias de las funciones vocales con SS /a/, /i/ y /o/, y 2 IPs para las medias de las funciones vocales con SS /e/ y /u/. El número de IPs claramente demarcados entre el máximo y mínimo globales para las medias de las curvas funciones vocales son: 1 IP para las medias de las funciones vocales con SS /a/, /i/ y /o/, y 2 IPs para la media de la función vocal con SS /e/.

Las figuras 4.38 (a) a (f) y 4.39 (a) a (d) muestran las curvas funcionales $\widehat{g}_{s,v,\delta\theta\alpha}(\cdot)$ y sus respectivas derivadas $D\widehat{g}_{s,v,\delta\theta\alpha}(\cdot)$ en donde el dominio corresponde a la la distancia angular entre electrodos, para tareas de habla silenciosa con vocales (ritmos $\delta\theta\alpha$). En este caso, se seleccionó un individuo (sujeto 1) calculando la media estadística de los valores de exactitud para los 20 sujetos de la experimentación (Fig. 4.42). De esta manera, las curvas funcionales se obtuvieron con:

$$\widehat{g}_{s,v,\delta\theta\alpha}(\Theta) = \sum_{k=1}^K c_k \aleph_k^m(\Theta) \quad (4.27)$$

donde $\widehat{g}_{s,v,\delta\theta\alpha}(\cdot)$ son las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 12$ es el número de bases (ver tabla 4.1), k es la referencia a las curvas funcionales estimadas, c_k son los coeficientes, $\aleph_k^m(\cdot)$ corresponde a los B-splines con grado $m = 3$.

De manera similar, las derivadas de las curvas funcionales se obtuvieron con:

$$D\widehat{g}_{s,v,\delta\theta\alpha}(\Theta) = \sum_{k=1}^K c'_k D\aleph_k^m(\Theta) \quad (4.28)$$

donde $D\widehat{g}_{s,v,\delta\theta\alpha}(\cdot)$ corresponde a la primera derivada de las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 12$ es el número de bases (ver tabla 4.1), k es la

referencia a la primera derivada de las curvas funcionales estimadas, c'_k son los coeficientes de la primera derivada, $D\mathbb{N}_k^m(\cdot)$ corresponde a la primera derivada de las B-splines con grado $m = 3$. La media de las curvas funcionales y la media de sus respectivas derivadas se calculan basados en la ecuación (Eq. 2.50).

Los datos funcionales (Figs. 4.38 y 4.39) son curvas abiertas con una dimensión topológica igual a 1. Las medias de las funciones $\widehat{g}_{s,v}(\cdot)$ tienen un máximo global (cerca a 1900 grados), un mínimo global (cerca a 3500 grados) y un máximo local cerca a 4100 grados. Los puntos locales y globales tienen un valor de cero en las respectivas medias de las derivadas $D\widehat{g}_{s,v}(\cdot)$ (figures 4.38 (b, d, f) y 4.39 (b, d)). Hay algunas diferencias entre las medias de las curvas de habla silenciosa relacionadas a los valores de la curva media o el número de puntos de inflexión.

El máximo global tiene un valor mayor a $0.45 (\mu V^2/Hz)$ para la media de la función vocal con SS /a/ y menor que $0.3 (\mu V^2/Hz)$ para la media de la función vocal con SS /i/. El valor del máximo global para las medias de las otras funciones (/e/, /o/ y /u/) se encuentra entre los dos rangos anteriores. De acuerdo al número de puntos de inflexión, el número de IPs varía entre las medias de las curvas de las funciones vocales con SS antes del máximo global. En este caso, se encuentra que el número de IPs claramente demarcados son: 1 IP para la media de la función vocal con SS /a/, 2 IPs para las medias de las funciones vocales con SS /e/, /i/ y /o/ y 3 IPs para la media de la función vocal con SS /u/. El número de IPs claramente demarcados entre 1900 a 3500 grados para las medias de las curvas de las funciones vocales son: 1 IP para la media de la función vocal con SS /i/, y 3 IPs para las medias de las funciones vocales con SS /a/, /e/, /o/ y /u/.

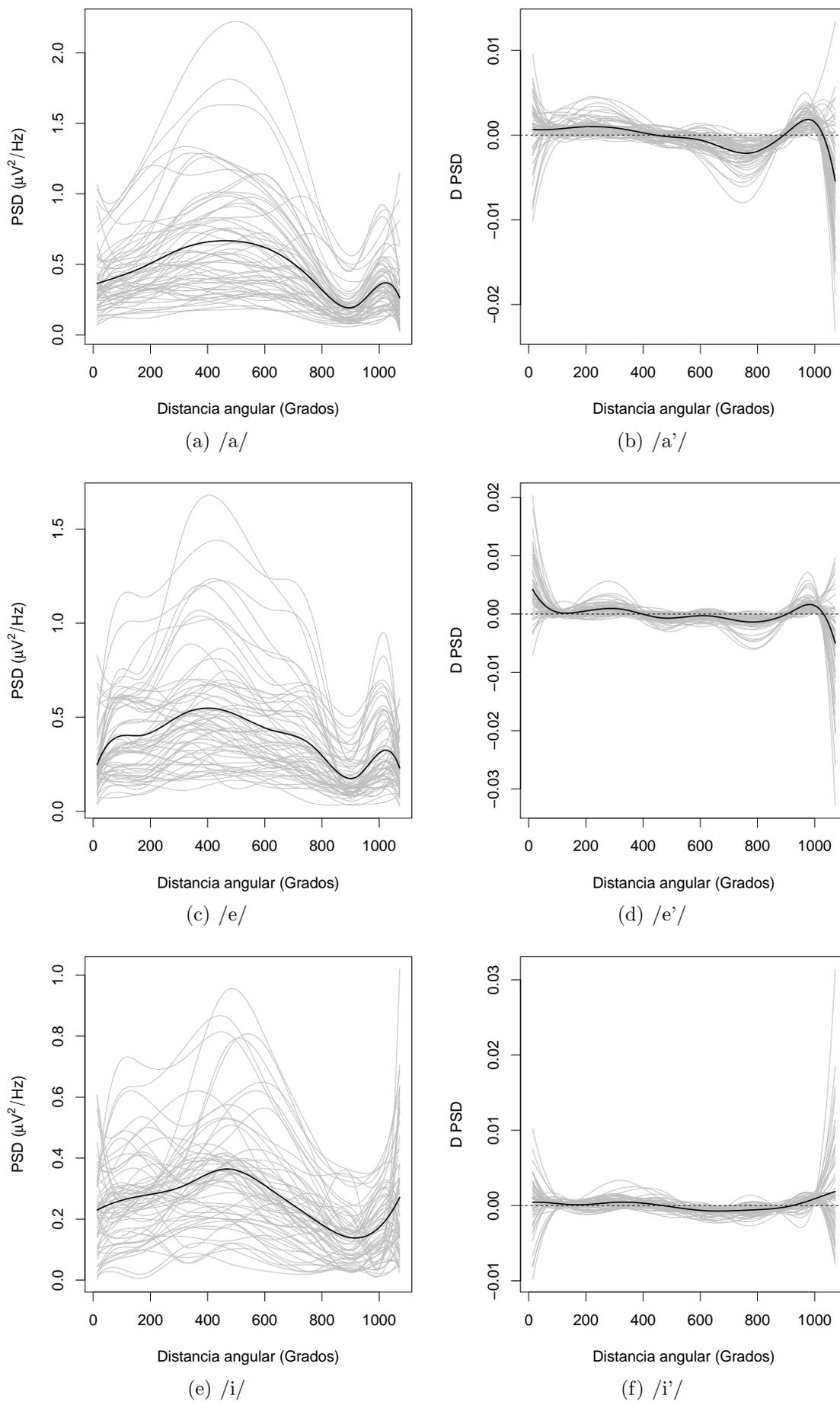


Figura 4.34: Datos funcionales con ritmos δ . La curva en color negro corresponde a la media.

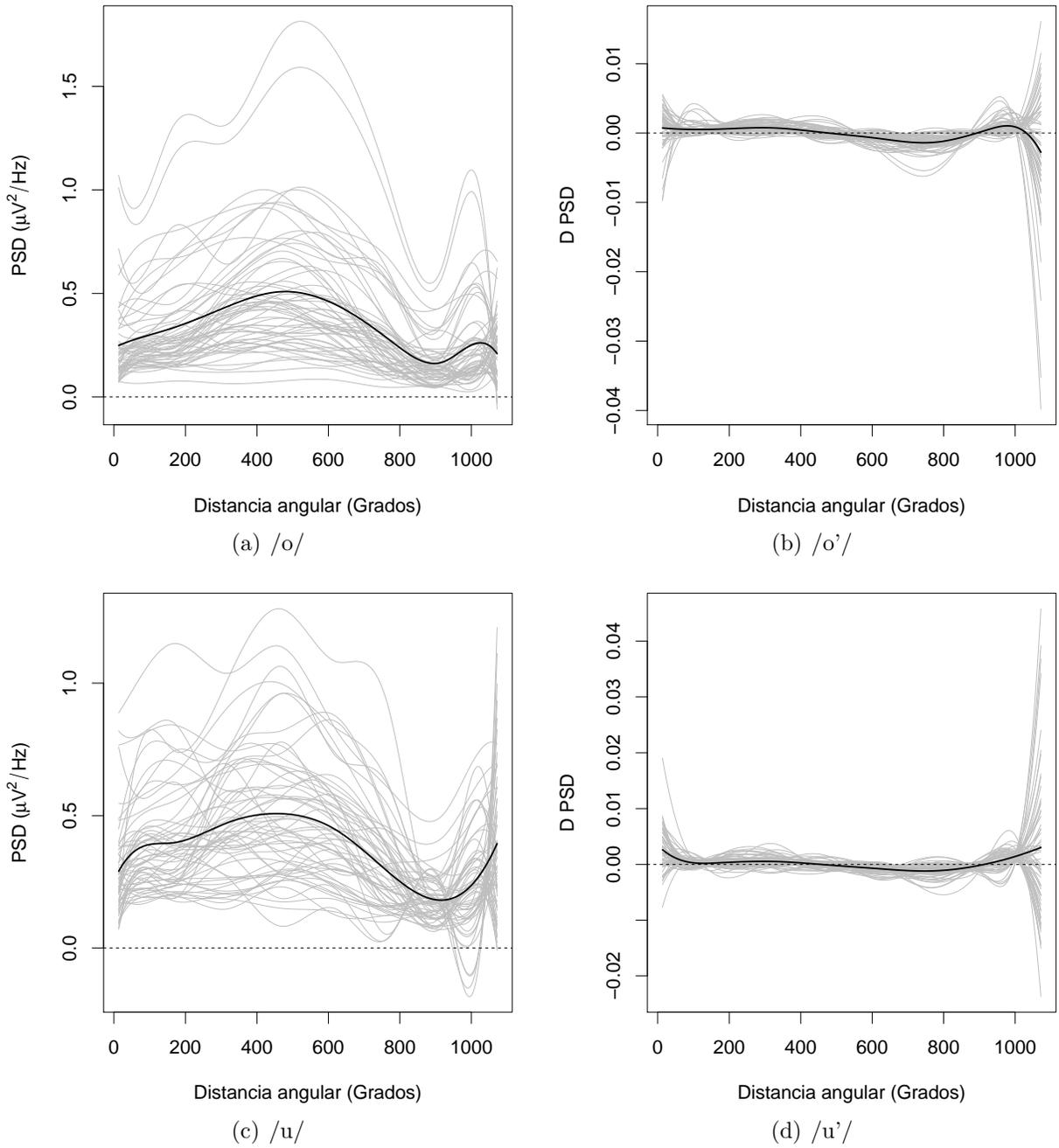


Figura 4.35: Datos funcionales con ritmos δ . La curva en color negro corresponde a la media.

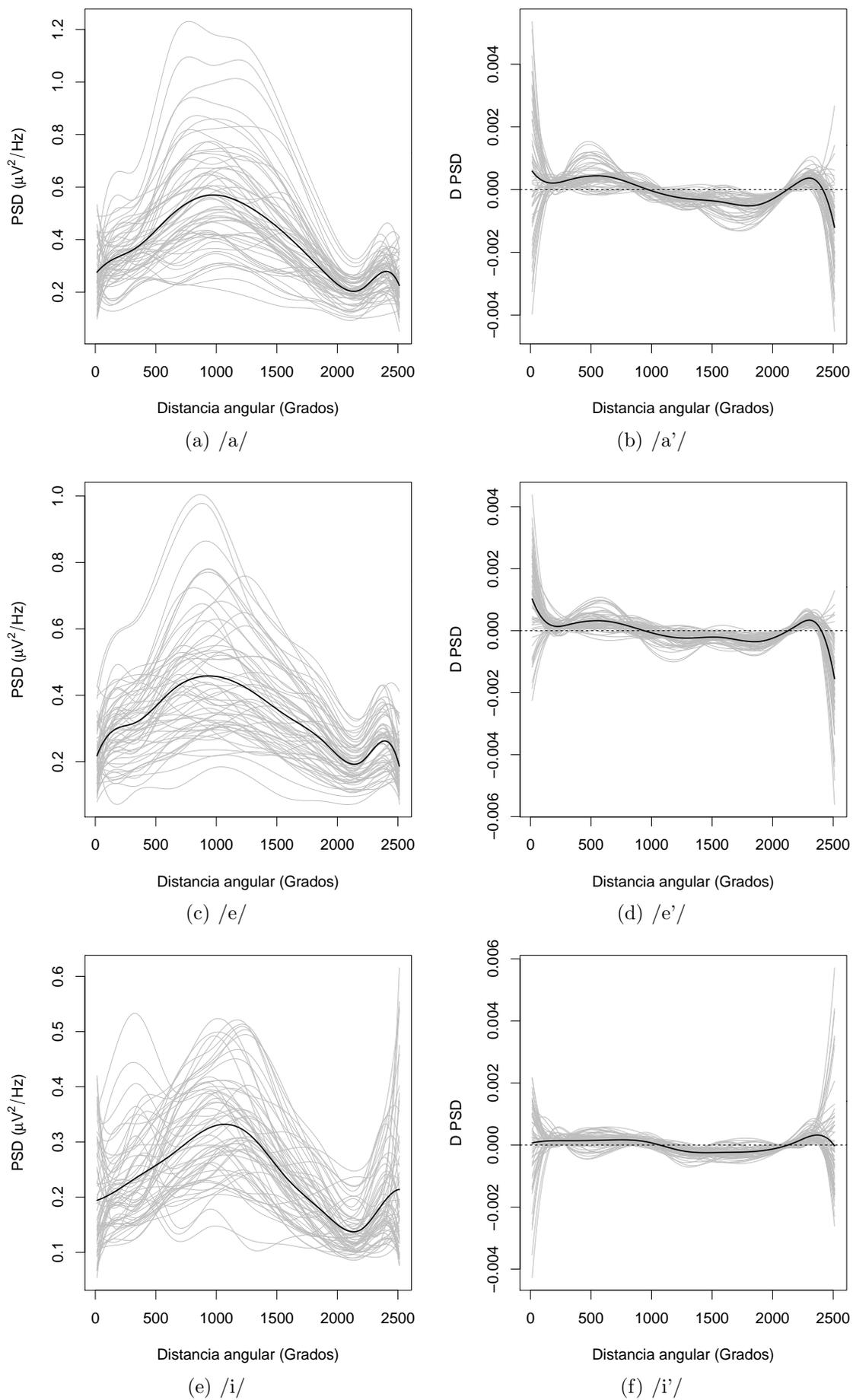
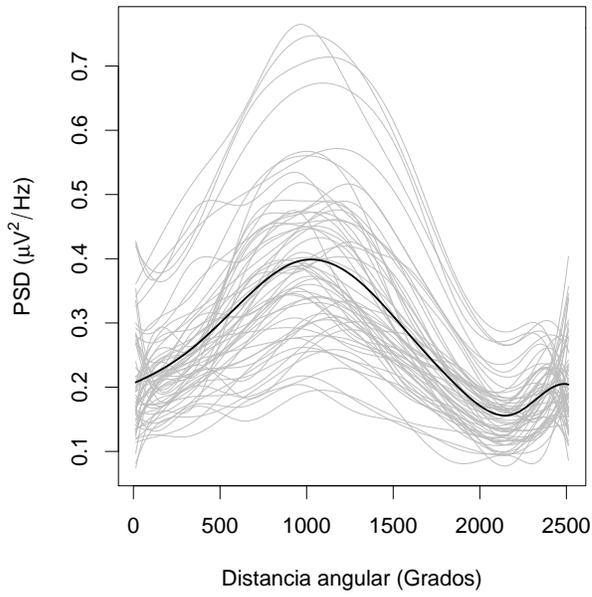
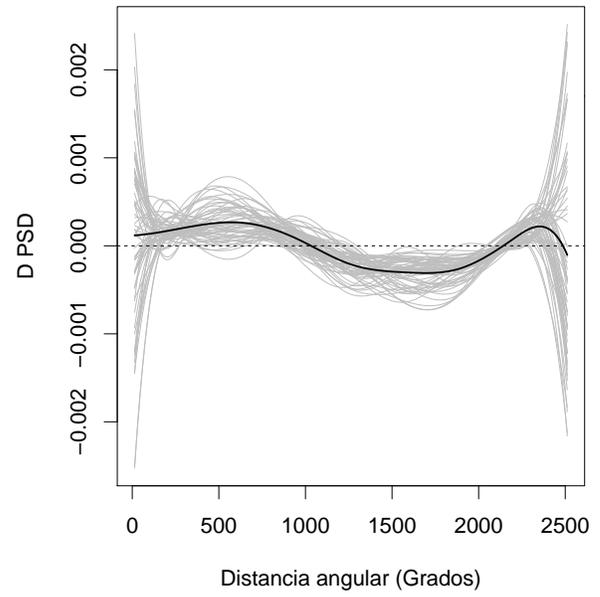


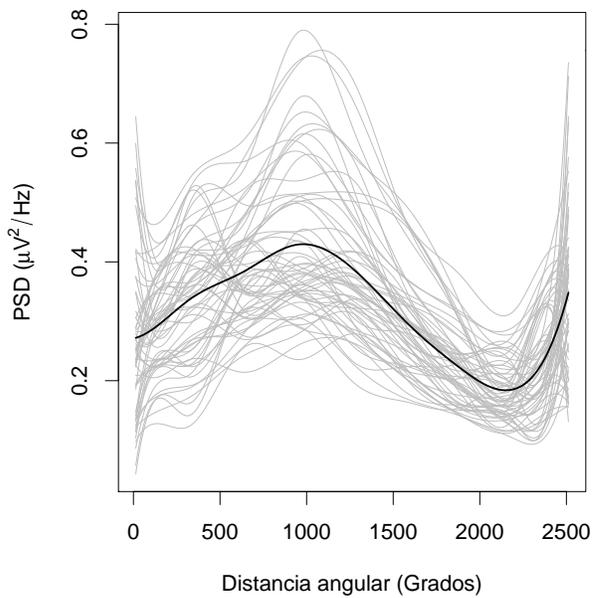
Figura 4.36: Datos funcionales con ritmos $\delta\theta$. La curva en color negro corresponde a la media.



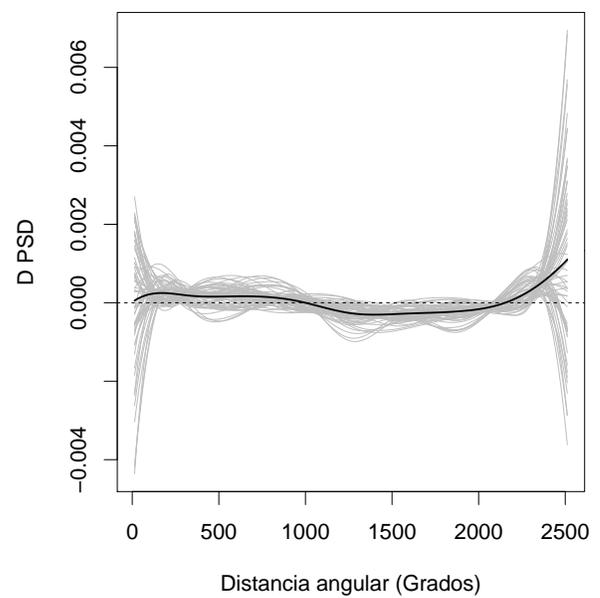
(a) /o/



(b) /o'/

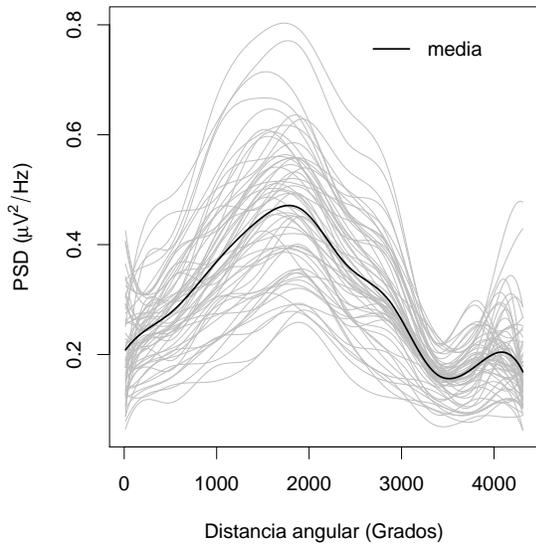


(c) /u/

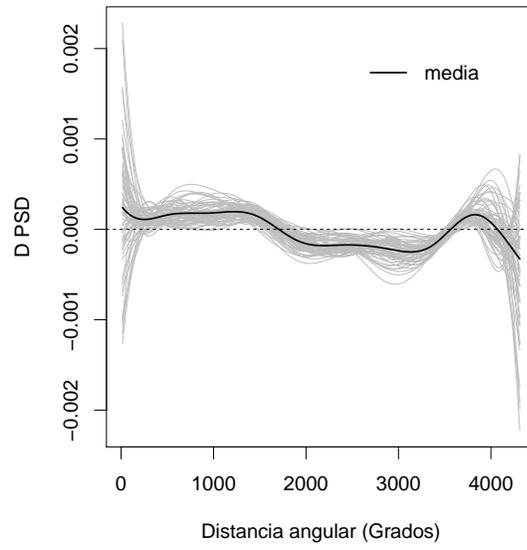


(d) /u'/

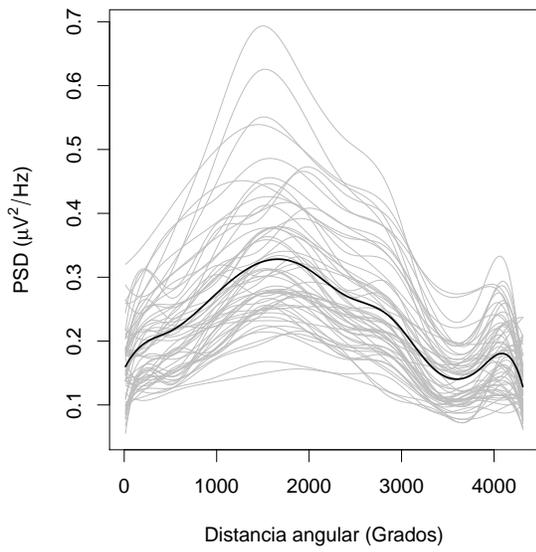
Figura 4.37: Datos funcionales con ritmos $\delta\theta$. La curva en color negro corresponde a la media.



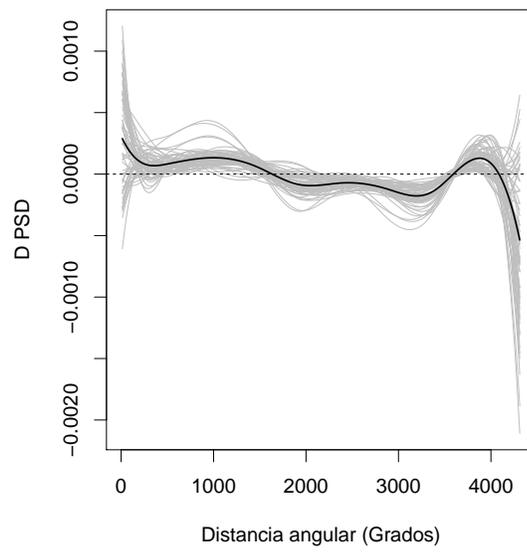
(a) /a/



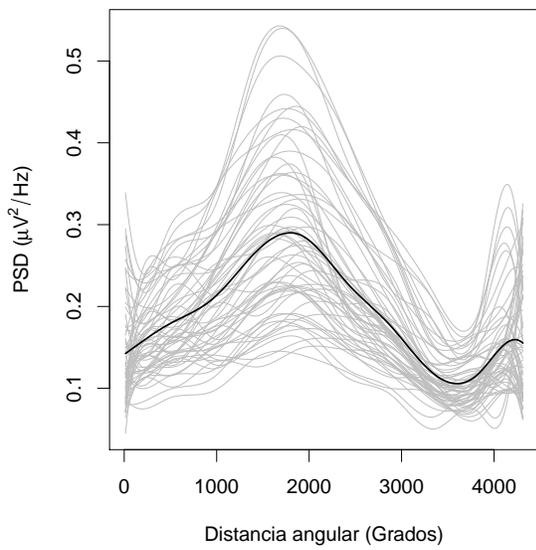
(b) /a'/



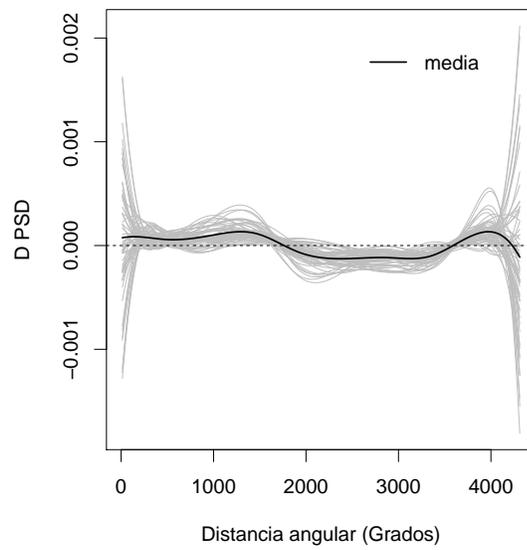
(c) /e/



(d) /e'/



(e) /i/



(f) /i'/

Figura 4.38: Datos funcionales con ritmos $\delta\theta\alpha$. La curva en color negro corresponde a la media.

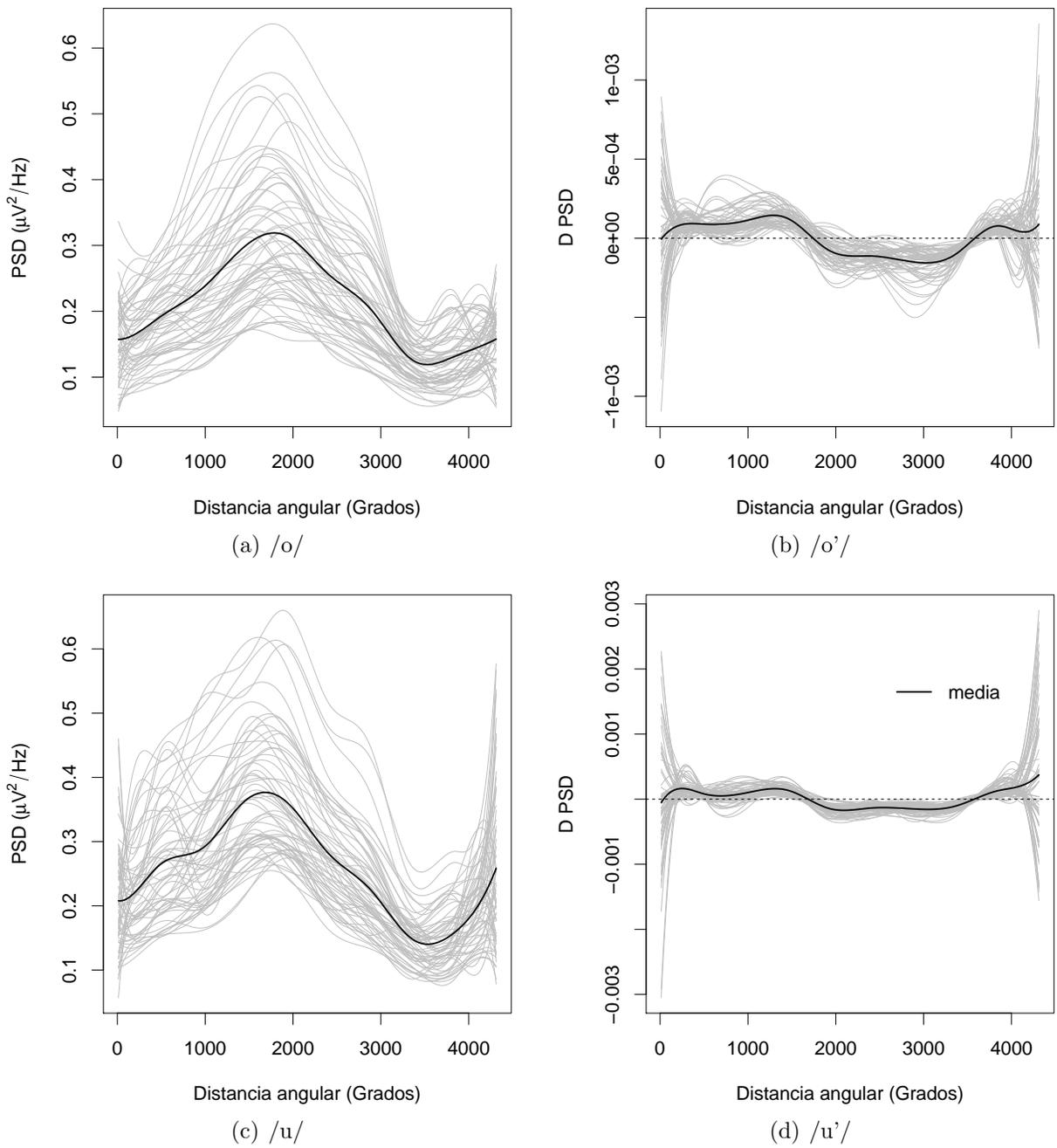


Figura 4.39: Datos funcionales con ritmos $\delta\theta\alpha$. La curva en color negro corresponde a la media.

4.5.4. Resultados de la clasificación

Los datos para los ritmos cerebrales δ que toma el clasificador SVM multiclase *one against-one* (1-1) son el resultado del vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$. Estos coeficientes corresponden al valor de los datos funcionales $\hat{g}(\cdot)$ y sus respectivas derivadas $D\hat{g}(\cdot)$ de las vocales con habla silenciosa utilizando ritmos cerebrales δ (Fig. 3.9).

Con el método de validación cruzada, el 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ son usados para entrenamiento, y el 30 % para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto (tabla A.3). La exactitud obtenida con esta metodologías estuvo entre 47.33 % y 68.56 %, con una media estadística de 57.49 % y una desviación estándar de 4.56 % (Fig. 4.40).

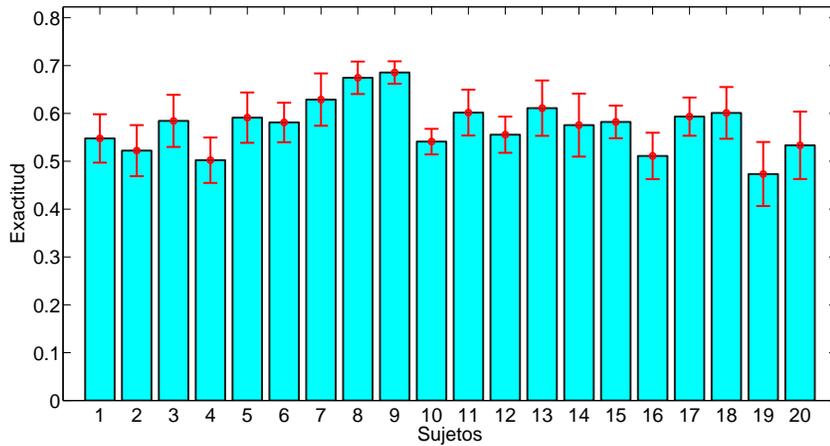


Figura 4.40: Clasificación de vocales con habla silenciosa utilizando (SS) SVM (1-1) - ritmos δ

Los datos para los ritmos cerebrales $\delta\theta$ que toma el clasificador SVM multiclase (1-1) son el resultado del vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$. Estos coeficientes corresponden al valor de los datos funcionales $\hat{g}(\cdot)$ y sus respectivas derivadas $D\hat{g}(\cdot)$ de las vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta$ (Fig. 3.9).

Con el método de validación cruzada, el 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ son usados para entrenamiento, y el 30 % para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto (tabla A.4). La exactitud obtenida con esta metodologías estuvo entre 53.56 % y 78.78 %, con una media estadística de 69.37 % y una desviación estándar de 4.08 % (Fig. 4.41).

Los datos para los ritmos cerebrales $\delta\theta\alpha$ que toma el clasificador SVM multiclase *one against-one* (1-1) son el resultado del vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$. Estos coeficientes corresponden al valor de los datos funcionales $\hat{g}(\cdot)$ y sus respectivas derivadas $D\hat{g}(\cdot)$ de las vocales con habla silenciosa utilizando ritmos cerebrales $\delta\theta\alpha$ (Fig. 3.9).

Con el método de validación cruzada, el 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ son usados para entrenamiento, y el 30 % para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto

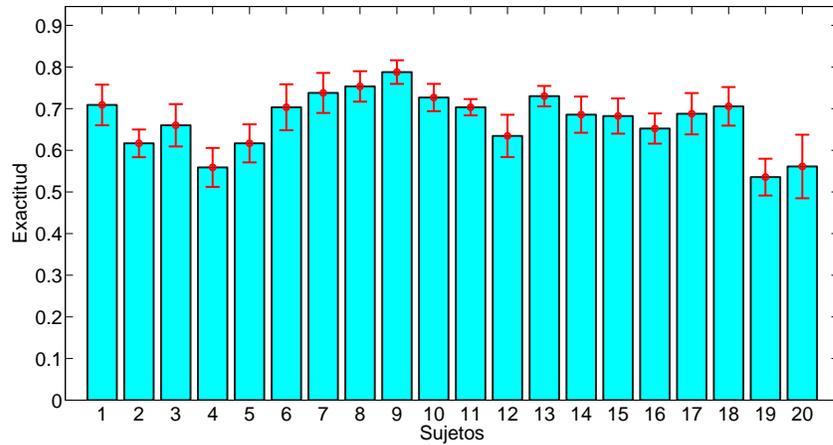


Figura 4.41: Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1)- ritmos $\delta\theta$

(tabla A.5). La exactitud obtenida con esta metodologías estuvo entre 65.22% y 78.0% con una media estadística de 71.92% y una desviación estándar de 4.56% (Fig. 4.42).

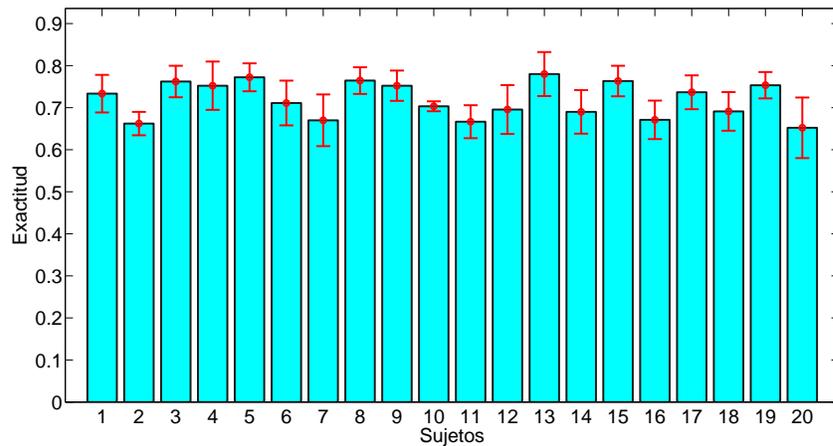


Figura 4.42: Clasificación de vocales con habla silenciosa (SS) utilizando SVM (1-1) - ritmos $\delta\theta\alpha$

4.6. Resultados de señales EEG para sílabas con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales

4.6.1. Resultados por sujeto

La figura 4.43 muestra los resultados correspondientes al valor promedio por sujeto y por sílaba de la estimación del PSD ($\bar{X}_{s,v,\delta}$), para sílabas con habla silenciosa utilizando ritmos

cerebrales δ con frecuencias entre $2, \dots, 4$ Hz. Estos valores se obtuvieron con la expresión dada por:

$$\bar{X}_{s,v,\delta} = \frac{\sum_{p=1}^{l_2} \sum_{k=1}^{l_1} X_{s,v,k}^p}{l_1 l_2} \quad (4.29)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.55), $l_1 = 90$ es el número de filas de la matriz, $l_2 = 21$ es el número de columnas de la matriz, k es una fila de la matriz, s es un sujeto, v es una sílaba con habla silenciosa, y p es un electrodo.

De la figura (4.43) se observa que existe una variación en los valores promedio del PSD estimado entre sujetos y entre sílabas con habla silenciosa. El rango de valores promedio estuvo entre 0.1 a 3.4 ($\mu V^2/Hz$). Los valores promedio más altos de PSD estimado fueron obtenidos por el sujeto 7 y los mínimos valores promedio de PSD estimado fueron obtenidos por el sujeto 8. Teniendo en cuenta el valor promedio más alto de PSD estimado por sujeto, la sílaba con SS /la/ tiene el valor promedio más alto en el 50% de los sujetos, y la sílaba con SS /fa/ tiene, respectivamente, el segundo valor promedio más alto de PSD estimado, con 30% de los sujetos cada una. En contraste, el mínimo valor promedio de PSD estimado por sujeto, fue obtenido por la sílaba con SS /pa/ en el 60% de los casos.

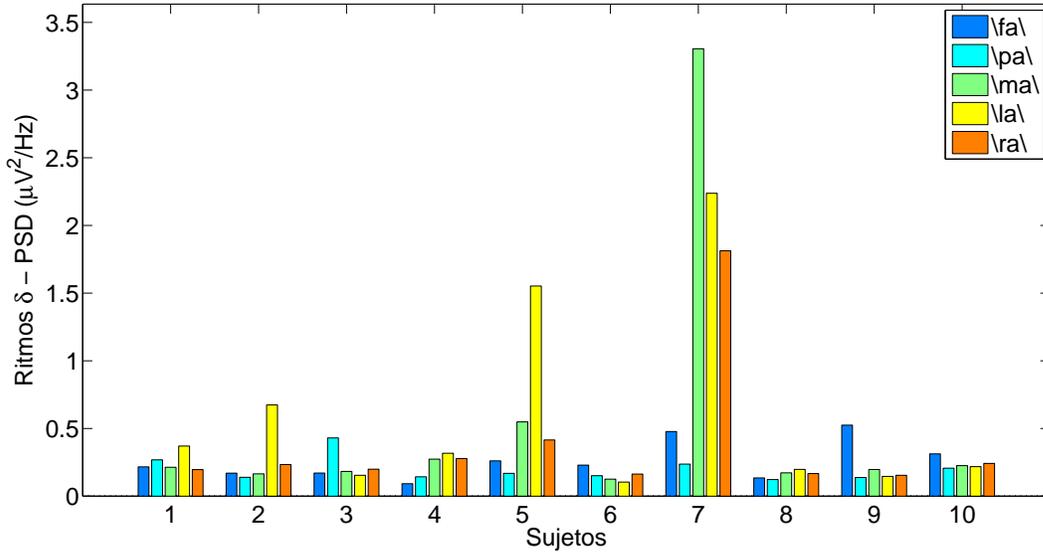


Figura 4.43: Valor promedio del PSD estimado por sujeto y por sílaba con habla silenciosa (SS) - ritmos δ

La figura (4.44) muestra la dispersión de los datos, del valor promedio del PSD estimado para cada sílaba con SS con ritmos cerebrales δ , utilizando un diagrama de cajas. Los datos para el diagrama de cajas se obtuvieron de la ecuación 4.29 para cada sujeto y para cada sílaba con SS (Fig. (4.43)). Los valores promedio para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son, respectivamente, (0.223, 0.160, 0.205, 0.267, 0.216) ($\mu V^2/Hz$). El valor de la mediana es más alto para la sílaba con SS /la/ y menor para la sílaba con SS /pa/. El rango intercuartil (IQR) con la mayor dispersión es para la sílaba con SS /la/, y el IQR con la menor dispersión es para la sílaba SS /pa/. En el caso de las sílabas con SS /fa/, /pa/, /ma/

y /la/ presentan asimetría positiva. Adicionalmente, se presentan outliers en las sílabas con SS /pa/, /ma/, /la/ y /ra/.

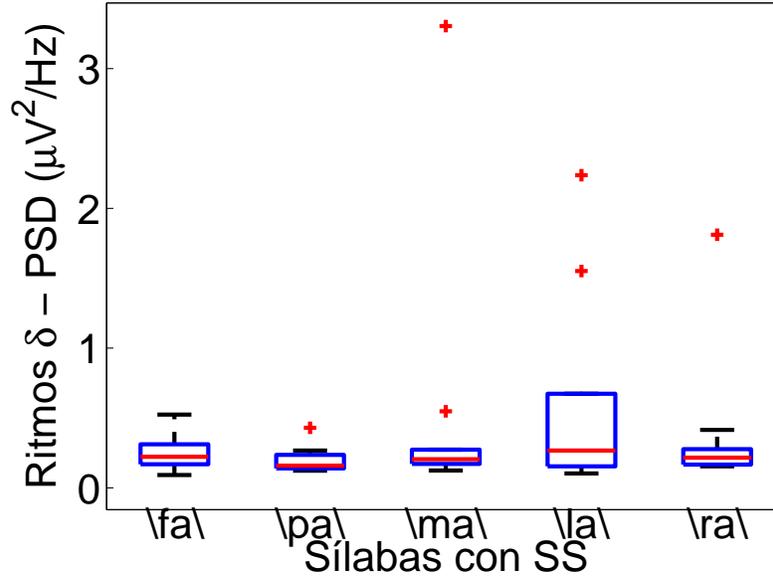


Figura 4.44: Diagrama de dispersión del valor promedio de PSD estimado por sílaba con habla silenciosa (SS) - ritmos δ

4.6.2. Resultados por localización

Las figuras 4.46(a) a 4.46(e) muestran la representación espacial de los 21 electrodos (E1 hasta E21) en términos del valor promedio de posición para el PSD estimado ($\tilde{X}_{s,v,\delta}^p$), realizando tareas de habla silenciosa con sílabas. En este caso, se seleccionó el valor promedio de los 10 individuos. De esta manera, el valor promedio de cada electrodo ($\mu V^2/Hz$) se obtiene con:

$$\tilde{X}_{s,v,\delta}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} X_{s,v,k}^p}{(S_T)(l_1)} \quad (4.30)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.55), $l_1 = 90$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 10$ es el número de sujetos, v es una sílaba con habla silenciosa, y p es un electrodo. Con los valores $\tilde{H}_{s,v,\delta}^p$ se realizó una interpolación cúbica cuyos resultados se muestran en las figuras 4.46(a) a 4.46(e). La distribución del color representa el valor del PSD estimado entre 0 (azul) y 0.65 (rojo) ($\mu V^2/Hz$). Los electrodos son posicionados sobre las áreas de Broca, Wernicke y la corteza auditiva primaria (ver fig. 3.4). Los electrodos E1, E6, E7, E11, E12, E13, E16, E17 y E21 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E14, E15, E19 y E20 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria. Estas áreas fueron seleccionadas por su crucial importancia en las funciones del lenguaje. De esta forma, para cada una de las sílabas con SS se presenta:

Primero, la sílaba con SS /fa/ (Fig. 4.46 (a)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10 y E15. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E1, E7 y E13. En la corteza auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.2 ($\mu V^2/Hz$).

Segundo, la sílaba con SS /pa/ (Fig. 4.46 (b)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el electrodo E18. En el área de Wernicke los electrodos con más alto promedio de posición del PSD estimado son E5 E9 y E19. Para el área de Broca, el valor promedio de posición más alto de PSD estimado está en E13. Adicionalmente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.1 ($\mu V^2/Hz$).

Tercero, la sílaba con SS /ma/ (Fig. 4.46 (c)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto para el área de Wernicke que en el área de Broca o el área de la corteza auditiva primaria. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E1, E7, E12 y E13. Complementariamente, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.52 ($\mu V^2/Hz$).

Cuarto, la sílaba con SS /la/ (Fig. 4.46 (d)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto en el área de Wernicke que en las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9, E10, E14, E15, E19 y E20. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E1, E7, E12 y E13. Además, en la corteza auditiva primaria E3 tiene un valor promedio de posición del PSD estimado cercano a 0.6 ($\mu V^2/Hz$).

Quinto, la sílaba con SS /ra/ (Fig. 4.46 (e)), tiene un valor promedio de posición del PSD estimado ($\mu V^2/Hz$) para los ritmos cerebrales δ más alto para el electrodo E18 que para las otras áreas. En el área de Wernicke los electrodos con los más altos valores promedio de posición del PSD estimado son E4, E5, E9 y E10. Para el área de Broca, los más altos valores promedio de posición del PSD estimado están en E1, E7 y E13. En este caso, en el área auditiva primaria el electrodo E3 tiene un valor promedio de posición del PSD estimado cercano a 0.35 ($\mu V^2/Hz$).

La figura (4.45) representa el valor promedio por electrodo del PSD estimado ($\overleftarrow{X}_{s,v,\delta}^p$), referido a cada sílaba con habla silenciosa (ritmos δ), para los 21 electrodos y los 10 sujetos de la experimentación. De esta manera, los valores se calcularon utilizando:

$$\overleftarrow{X}_{s,v,\delta}^p = \frac{\sum_{s=1}^{S_T} \sum_{k=1}^{l_1} X_{s,v,k}^p}{(S_T)(l_1)} \quad (4.31)$$

donde $X_{s,v,k}^p$ corresponde a cada elemento de la matriz $\mathbb{X}_{s,v}$ (Eq. 3.55), $l_1 = 90$ es el número de filas de la matriz, k es una fila de la matriz, s es un sujeto, $S_T = 10$ es el número de sujetos, v es una sílaba con habla silenciosa, y p es un electrodo.

En la figura (4.45), el cuadrado representa el PSD estimado por electrodo de la sílaba con SS /fa/, el círculo representa la PSD estimado por electrodo de la sílaba con SS /pa/,

el rombo representa el PSD estimado por electrodo de la sílaba con SS /ma/, el signo (+) representa el PSD estimado por electrodo de la sílaba con SS /la/ y el triángulo representa el PSD estimado por electrodo de la sílaba con SS /ra/. El rango de valores de PSD estimado se encuentra entre 0.10 y 0.70 ($\mu V^2/Hz$) para los ritmos cerebrales δ . En el 95 % de los casos el PSD estimado de la sílaba con SS /la/ es más alta en la mayoría de electrodos que en las otras sílabas con SS. La sílaba con SS /pa/ obtuvo el menor valor de PSD estimado en el 90 % de los electrodos. El valor más alto de PSD estimado lo obtuvo la sílaba con SS /la/ en el electrodo 5 y el menor valor de PSD estimado lo obtuvo la sílaba con SS /pa/ en el electrodo 16.

En la figura (4.45), se observa que el valor promedio por electrodo de las sílabas con SS, tiende a estar separados de las otras sílabas con SS. En la figura (4.45) se observa mayor promedio de PSD estimado para los electrodos E4, E5, E9, E10 y E15 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E6, E7, E9, E11, E12, E16, E17 y E21, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el área de Wernicke un mayor promedio de PSD estimado para ritmos δ , en tareas de habla silenciosa con sílabas, en contraste a el área de Broca (ver figura 3.4).

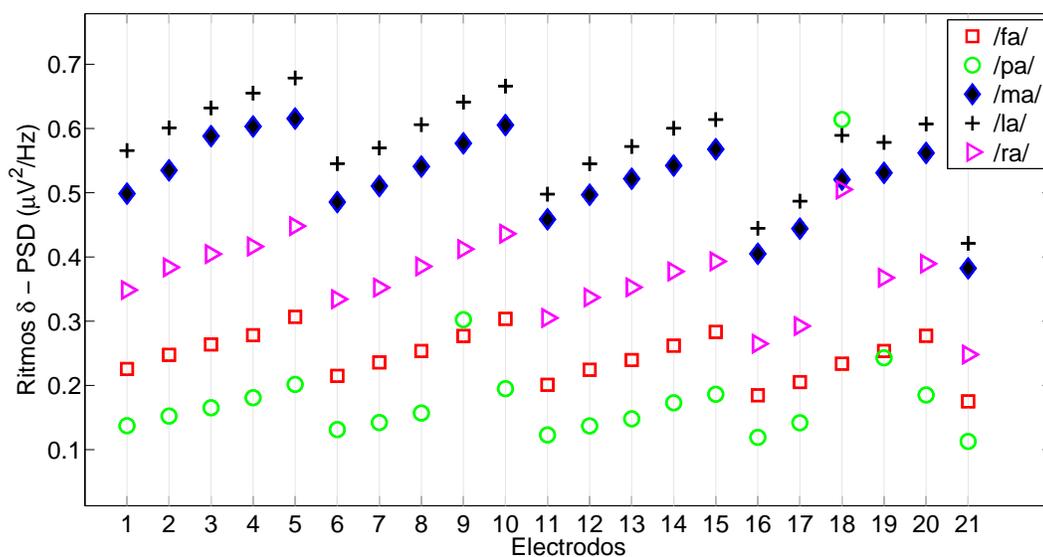


Figura 4.45: Valor promedio de PSD estimado por electrodo para sílabas con habla silenciosa (SS) - ritmos δ

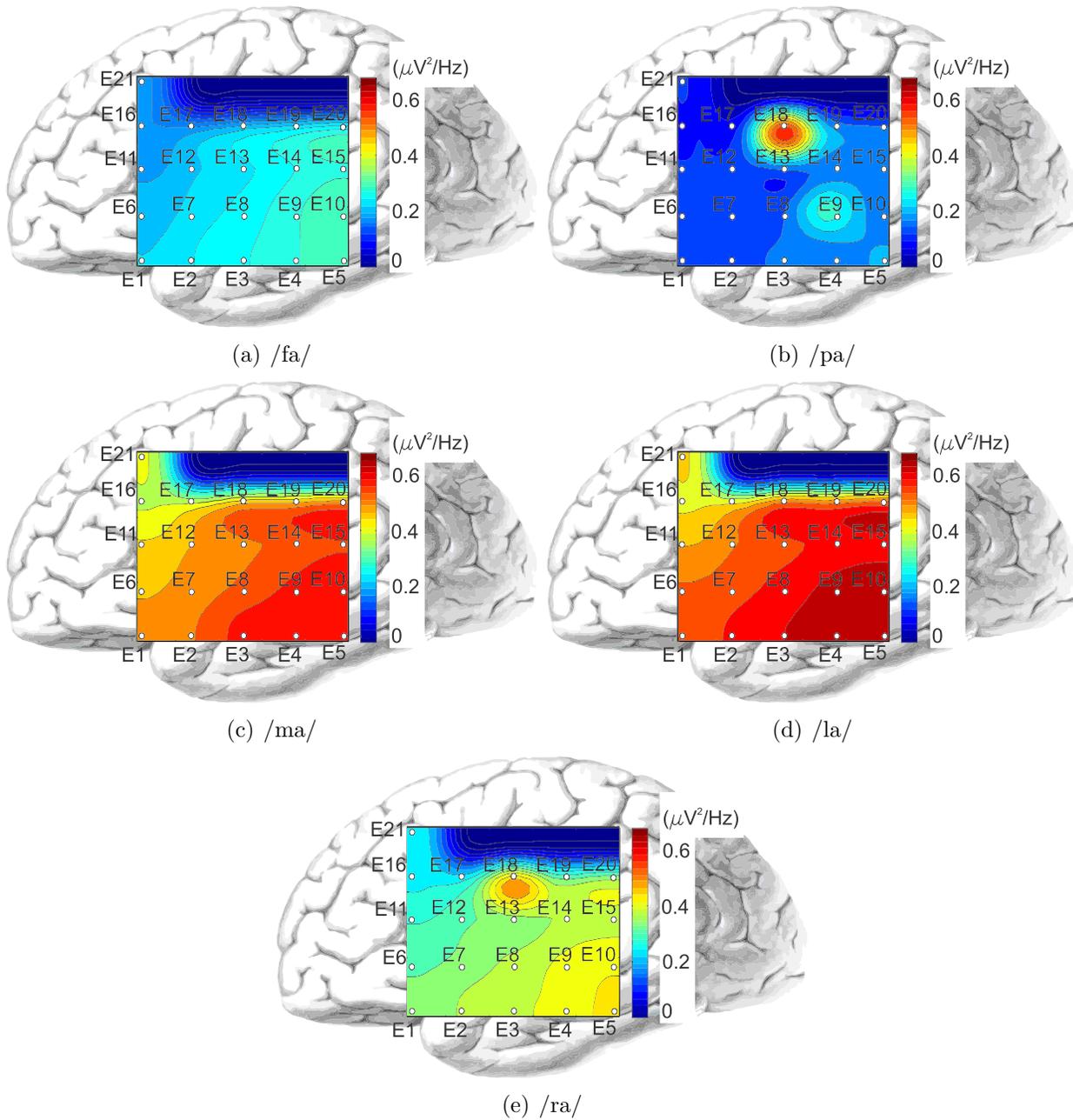


Figura 4.46: Representación espacial del valor promedio por electrodo del PSD estimado para sílabas con habla silenciosa (SS) - ritmos δ

4.6.3. Resultados utilizando datos funcionales

En la tabla (4.2) se muestra el valor mínimo obtenido del error cuadrático medio (MSE) aplicado a los ritmos cerebrales δ cuando los sujetos piensan en sílabas con SS. Cada dato en la tabla (4.2) se obtiene con la ecuación (2.46) variando el número de bases y los ritmos cerebrales para cada sujeto. Para seleccionar el número de bases óptimo K por ritmo cerebral para los 10 sujetos, se utiliza el valor de la mediana. Para el grupo δ el número de bases óptimo es $K = 11$,

Las figuras 4.47 (a) a (f) y 4.48 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v,\delta}(\cdot)$ y sus

Tabla 4.2: MSE variando el número de bases para sílabas con habla silenciosa (SS) - ritmos cerebrales δ

Sujeto	δ
1	12
2	5
3	7
4	16
5	10
6	9
7	12
8	15
9	7
10	20
mediana	11

respectivas derivadas $D\hat{g}_{s,v,\delta}(\cdot)$ en donde el dominio corresponde a la la distancia angular entre electrodos, para tareas de habla silenciosa con sílabas (ritmos δ). En este caso, se selecciona un individuo (sujeto 6) calculando la media estadística de los valores de exactitud para los 10 sujetos de la experimentación (Fig. 4.49). De esta manera, las curvas funcionales se obtienen con:

$$\hat{g}_{s,v,\delta}(\Theta) = \sum_{k=1}^K c_k \aleph_k^m(\Theta) \quad (4.32)$$

donde $\hat{g}_{s,v,\delta}(\cdot)$ son las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 11$ es el número de bases (ver tabla 4.2), k es la referencia a las curvas funcionales estimadas, c_k son los coeficientes, $\aleph_k^m(\cdot)$ corresponde a los B-splines con grado $m = 3$.

De manera similar, las derivadas de las curvas funcionales se obtienen con:

$$D\hat{g}_{s,v,\delta}(\Theta) = \sum_{k=1}^K c'_k D\aleph_k^m(\Theta) \quad (4.33)$$

donde $D\hat{g}_{s,v,\delta}(\cdot)$ corresponde a la primera derivada de las curvas estimadas, Θ es la distancia angular entre los electrodos, $K = 11$ es el número de bases (ver tabla 4.2), k es la referencia a la primera derivada de las curvas funcionales estimadas, c'_k son los coeficientes de la primera derivada, $D\aleph_k^m(\cdot)$ corresponde a la primera derivada de los B-splines con grado $m = 3$. La media de las curvas funcionales y la media de sus respectivas derivadas se calculan basados en la ecuación (Eq. 2.50).

Los datos funcionales (Figs. 4.47 y 4.48) son curvas abiertas con una dimensión topológica igual a 1. Las medias de las funciones $\hat{g}_{s,v,\delta}(\cdot)$ tienen un máximo global (entre 240 a 700 grados) y un mínimo global (entre 80 a 900 grados). Los puntos globales tienen un valor de cero en sus respectivas medias de las derivadas $D\hat{g}_{s,v,\delta}(\cdot)$ (figures 4.47 (b, d, f) y 4.48 (b, d)). Hay ciertas diferencias entre las curvas referidas a la media estadística de las sílabas con SS. Entre ellas se encuentran:

El máximo global, dada la media de la función para la sílaba con SS /fa/ es un valor cercano a 0.3 ($\mu V^2/Hz$) a 250 grados; la media de la función para la sílaba con SS /pa/

es $0.2 (\mu V^2/Hz)$ a 250 grados; la media de la función para la sílaba con SS /ma/ es $0.25 (\mu V^2/Hz)$ a 190 grados; la media de la función para la sílaba con SS /la/ es $0.15 (\mu V^2/Hz)$ a 240 grados; y la media de la función para la sílaba con SS /ra/ es $0.4 (\mu V^2/Hz)$ a 700 grados.

La media del mínimo global, para la sílaba con SS /fa/ es un valor cercano a $0.16 (\mu V^2/Hz)$ a 80 grados; la media de la función para la sílaba con SS /pa/ es $0.12 (\mu V^2/Hz)$ a 80 grados; la media de la función para la sílaba con SS /ma/ es $0.05 (\mu V^2/Hz)$ a 80 grados; la media de la función para la sílaba con SS /la/ es $0.09 (\mu V^2/Hz)$ a 900 grados; y la media de la función para la sílaba con SS /ra/ es $0.1 (\mu V^2/Hz)$ a 80 grados.

De acuerdo al número de puntos de inflexión (IP) después del respectivo máximo global, se encuentra: para la media de la función de la sílaba con SS /fa/ son 7 IPs; para la media de la función de sílaba con SS /pa/ son 5 IPs; para la media de la función de la sílaba con SS /ma/ son 6 IPs; para la media de la función de sílaba con SS /la/ son 4 IPs; y para la media de la función de sílaba con SS /ra/ son 3 IPs.

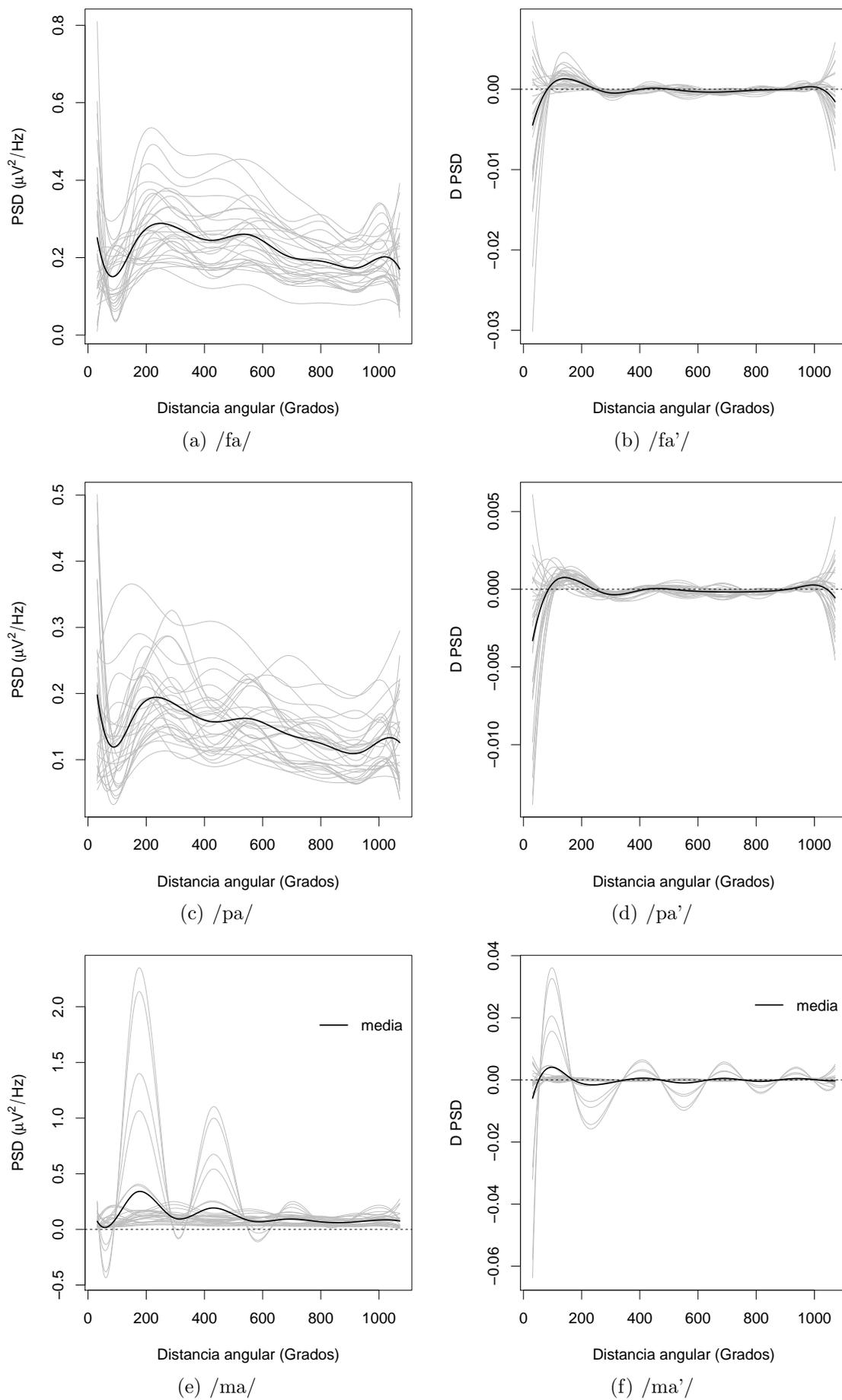


Figura 4.47: Datos funcionales para sílabas con ritmos δ . La curva en color negro corresponde a la media

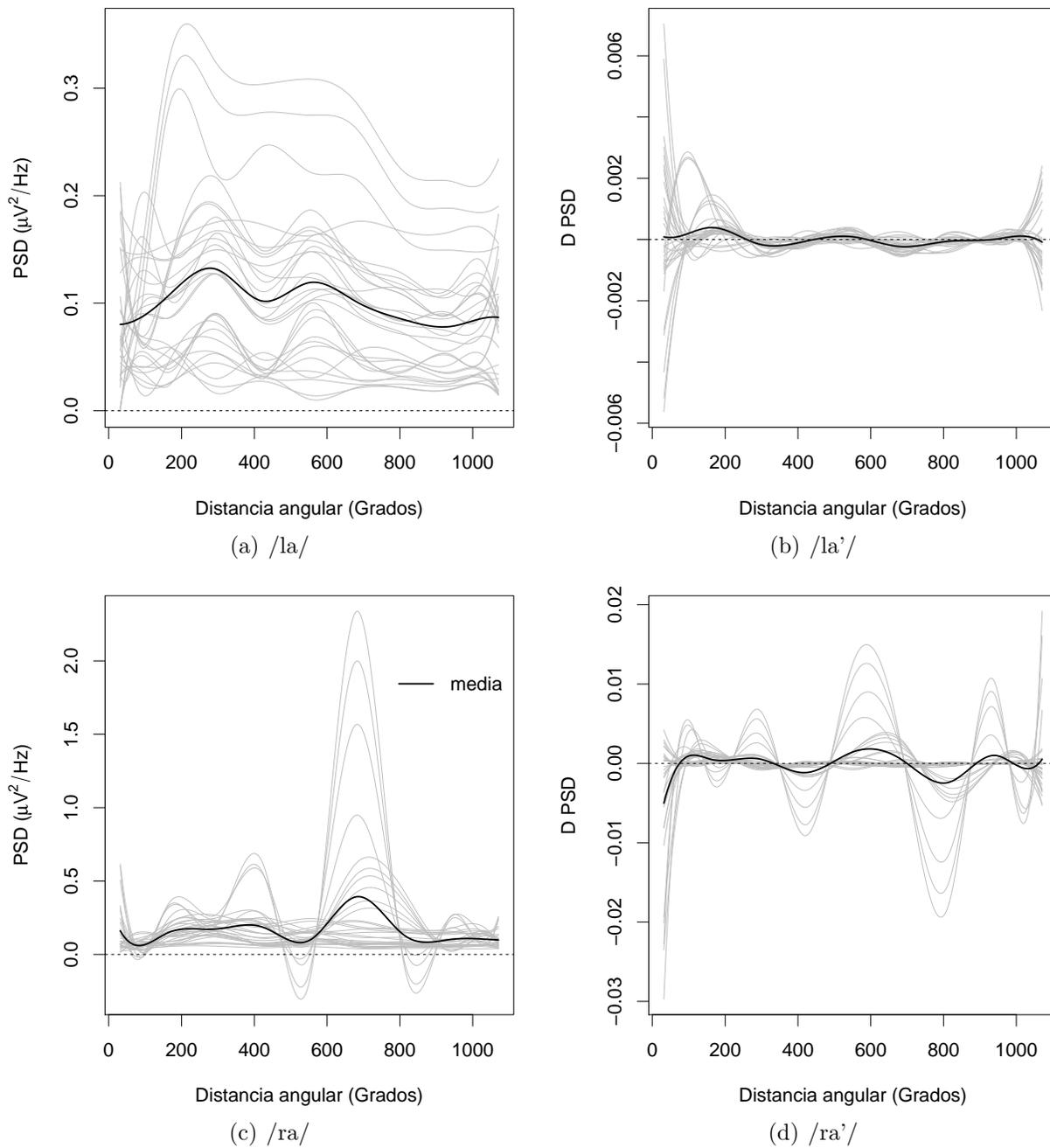


Figura 4.48: Datos funcionales para sílabas con ritmos δ . La curva en color negro corresponde a la media.

4.6.4. Resultados de la clasificación

Los datos para los ritmos cerebrales δ que toma el clasificador SVM multiclase one-against-one (1-1) son el resultado del vector de características $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$. Estos coeficientes son los datos funcionales $\hat{f}(\cdot)$ y sus respectivas derivadas $D\hat{f}(\cdot)$ de las sílabas con habla silenciosa utilizando ritmos cerebrales δ (Fig. 3.10).

Con el método de validación cruzada, el 70 % de los datos relacionados con $\Xi_{s,v} = [\mathbf{c}_{s,v} \ \mathbf{c}'_{s,v}]$ son usados para entrenamiento, y el 30 % para evaluación. Un algoritmo genético es aplicado con los datos de entrenamiento para encontrar los parámetros σ y C (ver algoritmo 1). El conjunto de datos de prueba con SVM multiclase (1-1) fue probado 10 veces para cada sujeto (tabla A.8). La exactitud obtenida con esta metodologías estuvo entre 60.67 % y 72.67 %, con una media estadística de 67.13 % y una desviación estándar de 6.61 % (Fig. 4.49).

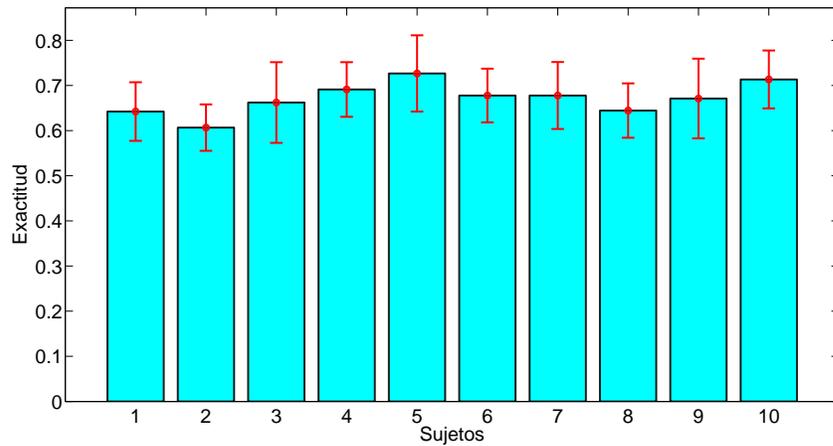


Figura 4.49: Clasificación de sílabas con habla silenciosa (SS) utilizando SVM (1-1) - ritmos δ

4.7. Comparación entre los resultados de la metodología basada en vector de características con datos funcionales para vocales y sílabas con habla silenciosa

Al comparar los diagramas de dispersión relacionados con el PSD estimado de las figuras (4.22), (4.24) y (4.26) se encuentra que:

Los valores de las medianas para las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) son respectivamente (0.385, 0.289, 0.261, 0.273) ($\mu V^2/Hz$) para los ritmos δ , (0.281, 0.217, 0.203, 0.199, 0.176) ($\mu V^2/Hz$) para los ritmos $\delta\theta$, y (0.225, 0.211, 0.197, 0.199, 0.213) ($\mu V^2/Hz$) para los ritmos $\delta\theta\alpha$. De los valores anteriores se observa que el PSD estimado de las medianas más altas son para los ritmos cerebrales δ .

Para los ritmos cerebrales δ , $\delta\theta$ y $\delta\theta\alpha$ en términos de PSD estimado, el mayor IQR lo presenta la vocal con SS /a/ y el menor IQR lo presenta la vocal con SS /o/.

Para las vocales con SS /e/ e /i/ utilizando los ritmos δ , $\delta\theta$ y $\delta\theta\alpha$, existe asimetría positiva.

La vocal con SS /u/ presenta asimetría positiva para los ritmos δ y $\delta\theta$. La vocal con SS /o/ presenta asimetría positiva para los ritmos $\delta\theta$ y $\delta\theta\alpha$. La vocal con SS /a/ presenta asimetría positiva para los ritmos $\delta\theta\alpha$. Las vocales con SS /a/ y /o/ presentan asimetría negativa para los ritmos δ y $\delta\theta$. La vocal con SS /u/ presenta asimetría negativa para los ritmos $\delta\theta\alpha$.

Al contrastar los diagramas de PSD estimado por electrodo de la figuras (4.28), (4.30) y (4.32) se encuentra que:

Para vocales con SS el rango de valores promedio por electrodo de PSD estimado se encuentra entre 0.10 hasta 0.60 ($\mu V^2/Hz$) para ritmos δ , 0.10 hasta 0.35 ($\mu V^2/Hz$) para ritmos $\delta\theta$ y 0.07 hasta 0.35 ($\mu V^2/Hz$) para ritmos $\delta\theta\alpha$.

Para los ritmos δ , en el 95 % de los casos el PSD estimado de la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /u/ en 81 % de los casos es la más baja por electrodo. Para los ritmos $\delta\theta$, en el 90 % de los casos el PSD estimado de la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /u/ en 100 % de los casos es la más baja por electrodo. Para los ritmos $\delta\theta\alpha$, en el 90 % de los casos el PSD estimado de la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /u/ en 95 % de los casos es la más baja por electrodo. En 71 % de los electrodos, el valor promedio por electrodo de PSD estimado para la vocal con SS /a/ tiende a estar separado, en cada electrodo, de las otras vocales con SS; mientras los valores promedio por electrodo de PSD estimado para las sílabas con SS tienden a estar separados en cada uno de los electrodos.

Para las vocales con SS, se observa mayor promedio de PSD estimado para los electrodos E5, E9, E4 y E16 pertenecientes al área de Wernicke; en contraste, los electrodos E17, E18, E11 y E12, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. En el caso de las sílabas con SS, se observa mayor promedio de PSD estimado para los electrodos E5, E9, E10, E14, E15 y E20 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E6, E11, E16 y E21, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el área de Wernicke un mayor promedio de PSD estimado, en tareas de habla silenciosa con vocales y sílabas, en contraste al área de Broca. Para los ritmos δ , $\delta\theta$ y $\delta\theta\alpha$, en 76 % de los electrodos el valor promedio de la vocal con SS /a/ tiende a estar separado de las otras vocales con SS. Para los ritmos δ y $\delta\theta$, en 48 % de los electrodos el valor promedio de la vocal con SS /e/ tiende a estar separado de las otras vocales con SS.

Para los ritmos δ , se observa mayor promedio de PSD estimado para los electrodos E2, E5, E10 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. Para los ritmos $\delta\theta$, se observa mayor promedio de PSD estimado para los electrodos E2, E4, E5, E9, E10, E15 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. Para los ritmos $\delta\theta\alpha$, se observa mayor promedio de PSD estimado para los electrodos E4, E5, E10 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado.

De esta manera se evidencia para ritmos δ , $\delta\theta$ y $\delta\theta\alpha$, en tareas de habla silenciosa con vocales, que el área de Wernicke presenta un mayor promedio de PSD estimado, en comparación al área de Broca.

Al relacionar los diagramas de dispersión relacionados con PSD estimado de la figura (4.22) con la figura (4.44), ambos con ritmos δ , se encuentra que:

Los valores de las medianas para las vocales con habla silenciosa (/a/, /e/, /i/, /o/, /u/) son

respectivamente (0.385, 0.289, 0.261, 0.273, 0.220) ($\mu V^2/Hz$) y una media de estos valores es 0.286 ($\mu V^2/Hz$). Los valores de las medianas para las sílabas con SS (/fa/, /pa/, /ma/, /la/, /ra/) son respectivamente (0.223, 0.160, 0.205, 0.267, 0.216) ($\mu V^2/Hz$) y una media de estos valores es 0.214 ($\mu V^2/Hz$). En este caso, se puede evidenciar que el cerebro tiende a generar un mayor valor de PSD estimado utilizando ritmos δ para las vocales con SS, en comparación con las sílabas con SS.

También, se encuentra en términos de PSD estimado que el mayor IQR lo presenta la vocal con SS /a/ y la sílaba con SS /la/. En contraste, el menor IQR lo presenta la vocal con SS /o/ y la sílaba con SS /pa/

Adicionalmente, se observa asimetría positiva para las vocales con SS /e/, /i/ y /u/ , y asimetría negativa para las vocales con SS /a/ y /o/, utilizando PSD estimado con ritmos δ . En el caso de las sílabas con SS /fa/, /pa/, /ma/ y /la/ presentan asimetría positiva.

Al confrontar los diagramas de PSD estimado por electrodo de la figuras (4.28) y (4.45) se encuentra que:

Para vocales con SS (ritmos δ), el rango de valores promedio por electrodo de PSD estimado se encuentra entre 0.10 y 0.60 ($\mu V^2/Hz$); mientras para sílabas con SS (ritmos δ) está entre 0.10 y 0.70 ($\mu V^2/Hz$).

En el 95 % de los casos (ritmos δ), el valor promedio por electrodo de PSD estimado para la vocal con SS /a/ es la más alta por electrodo y la vocal con SS /u/ en 81 % de los casos es la más baja por electrodo; mientras en 95 % de los casos (ritmos δ) la entropía de la sílaba con SS /la/ es la más alta por electrodo y la sílaba con SS /pa/ en 90 % de los casos es la más baja por electrodo.

En 76 % de los electrodos el valor promedio de la vocal con SS /a/ (ritmos δ) tiende a estar separado de las otras vocales con SS y en 48 % de los electrodos el valor promedio de la vocal con SS /e/ tiende a estar separado de las otras vocales con SS; mientras los valores promedio por electrodo de PSD estimado para las silabas con SS (ritmos δ) tienden a estar separados en cada uno de los electrodos.

Para las vocales con SS (ritmos δ), se observa mayor promedio de PSD estimado para los electrodos E2, E5, E10 y E16 pertenecientes al área de Wernicke. En contraste, los electrodos E12, E17 y E18, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. En el caso de las sílabas con SS (ritmos δ), se observa mayor promedio de PSD estimado para los electrodos E4, E5, E9, E10 y E15 pertenecientes al área de Wernicke. En contraste, los electrodos E1, E6, E7, E9, E11, E12, E16, E17 y E21, pertenecientes al área de Broca, presentan menor promedio de PSD estimado. De esta manera se evidencia para el área de Wernicke un mayor promedio de PSD estimado para ritmos δ , en tareas de habla silenciosa con vocales y sílabas, en contraste a el área de Broca.

Al comparar la estimación de las curvas funcionales (ritmos δ) entre vocales con SS (Figs. 4.34 y 4.35) y sílabas con SS (Figs. 4.47 y 4.48) se encuentra:

Para vocales con SS (ritmos δ), la media de las funciones $\widehat{g}_{s,v,\delta}(\cdot)$ tienen un máximo global (entre 400 a 500 grados) y un mínimo global (cerca a 900 grados) . Los puntos globales de las medias tienen un valor de cero en sus respectivas derivadas $D\widehat{g}_{s,v,\delta}(\cdot)$. Para sílabas con SS (ritmos δ), la media de las funciones $\widehat{g}_{s,v,\delta}(\cdot)$ tienen un máximo global (entre 240 a 700 grados) y un mínimo global (entre 80 a 900 grados). Los puntos globales tienen un valor de cero en sus respectivas medias de las derivadas $D\widehat{g}_{s,v,\delta}(\cdot)$.

El máximo global para vocales con SS (ritmos δ), tiene un valor aproximado de 0.7 ($\mu V^2/Hz$) para la media de la función vocal con SS /a/ y un valor aproximado de 0.35 ($\mu V^2/Hz$) para la media de la función vocal con SS /i/. El valor del máximo global para las medias de las

otras funciones (/e/, /o/ y /u/) se encuentra entre los dos valores anteriores. El máximo global para sílabas con SS (ritmos δ), dada la media de la función para la sílaba con SS /fa/ es un valor cercano a 0.3 ($\mu V^2/Hz$) a 250 grados; la media de la función para la sílaba con SS /pa/ es 0.2 ($\mu V^2/Hz$) a 250 grados; la media de la función para la sílaba con SS /ma/ es 0.25 ($\mu V^2/Hz$) a 190 grados; la media de la función para la sílaba con SS /la/ es 0.15 ($\mu V^2/Hz$) a 240 grados; y la media de la función para la sílaba con SS /ra/ es 0.4 ($\mu V^2/Hz$) a 700 grados.

El número de puntos de inflexión (IPs) para vocales con SS (ritmos δ), entre el máximo y mínimo global para las medias de las curvas funciones vocales son: 2 IPs para la media de la función vocal con SS /e/, y 1 IP para las medias de las funciones vocales con SS /a/, /i/, /o/ y /u/. Para sílabas con SS (ritmos δ), el número de IPs después del respectivo máximo global es para la media de la función de la sílaba con SS: /fa/, 7 IPs; /pa/, 5 IPs; /ma/, 6 IPs; /la/, 4 IPs; y /ra/, 3 IPs.

4.8. Discusión sobre señales EEG para vocales y sílabas con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales

4.8.1. Determinación del método de clasificación

A continuación se hace una discusión de los resultados obtenidos con la metodología basada en vector de características con datos funcionales para la clasificación de las vocales y sílabas con habla silenciosa.

En el caso de vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales se seleccionó el algoritmo SVM de clasificación multiclase (1-1), ya que en la metodología basada en vector de características con entropía de la información y dimensión de regularización presentó mayor exactitud, tanto para vocales como para sílabas con habla silenciosa. A continuación se describen los valores de clasificación alcanzados con la metodología basada en vector de características con datos funcionales. Para el caso de vocales con habla silenciosa utilizando ritmos δ se alcanzó en la clasificación una exactitud entre 47.33 % y 68.56 % con una media estadística de 57.49 % y una desviación estándar de 4.56 % (Fig. 4.40). Para ritmos $\delta\theta$ se alcanzó en la clasificación una exactitud entre 53.56 % y 78.78 % con una media estadística de 69.37 % y una desviación estándar de 4.08 % (Fig. 4.41). Para ritmos $\delta\theta\alpha$ se alcanzó en la clasificación una exactitud entre 65.22 % y 78.0 % con una media estadística de 71.92 % y una desviación estándar de 4.56 % (Fig. 4.42). En el caso de vocales con habla silenciosa utilizando la metodologías de vector de características con datos funcionales, los ritmos $\delta\theta\alpha$ presentan mayor exactitud que los ritmos $\delta\theta$ y los ritmos δ .

Para el caso de sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales se seleccionó el algoritmo de clasificación multiclase (1-1). Con este clasificador se alcanzó una exactitud entre 60.67 % y 72.67 % con una media estadística de 67.13 % y una desviación estándar de 6.61 % (Fig. 4.49).

Contrastando los resultados de esta tesis con otros trabajos de habla silenciosa en cuanto a exactitud (*accuracy*), se encuentra que:

En el caso de vocales con habla silenciosa la metodología desarrollada en esta tesis clasifica 5 vocales en contraste con: Da Salla et al. [29] de 68 % a 78 %, Yoshimura et al. [122] con 49.9 %, Riaz et al. [123] con 75 %, donde clasifican todos ellos únicamente 2 vocales con habla silenciosa. El trabajo desarrollado en esta tesis tiene mayor exactitud que aquel de Kamala et al. [125], que alcanza una exactitud de 44 % donde se clasifican 5 vocales de habla inglesa con EEG. Igualmente, el trabajo desarrollado es comparable a estudios de reconocimiento para 5 vocales japonesas con EEG presentados por Matsumoto et al. [126] con un 85 % de exactitud, Hori et al. [127] con 57 % a 87 % y Matsumoto et al. [30] con un 57 %-87 %.

En el caso de sílabas con habla silenciosa, la metodología desarrollada clasifica 5 sílabas con habla silenciosa en contraste con las siguientes estudios que clasifican exclusivamente 2 sílabas: D'Zmura [128] que logra una exactitud promedio de 75 %, Brigham et al. [31] con una exactitud de 56 % a 88 %, Deng et al. [32] con una 54 % a 73 %.

En el caso de caracteres y palabras con habla silenciosa, la metodología desarrollada en esta tesis es: mejore en función del número de clases reconocidas que Wang et al. [129], con una exactitud de 79 % a 88 %, y Wang et al. [130] con 73 % a 95 %, que sólo reconocen dos caracteres chinos; también mejor en función del número de clases reconocidas que Salamanca et al. [33], con una exactitud de 47 % a 67 %, y en Kazi et al. [34], con 27 % a 91 %, donde se reconocen dos palabras en inglés; y con resultados superiores a Porbadnigk et al. [131], García et al. [132], con una exactitud promedio de 45.5 % y 42 %, respectivamente, donde se reconocen 5 palabras en inglés y 5 en español.

En contraste con el estudio desarrollado en esta tesis, otros estudios sobre reconocimiento de vocales con habla silenciosa han utilizado algoritmos como: SVM para dos clases [29]; combinación de los algoritmos SVM con kernel gaussiano (SVM-G), máquinas de vectores relevantes con kernel gaussiano (RVM-G) y máquinas de vectores relevantes lineales (RVM-L) [127]; el algoritmo de regresión logístico (SLR-VAR) [122]; Naive Bayes [137]; modelos ocultos de Markov (HMM), con el algoritmo de los k vecinos (k-NN) [123]; y una red neuronal artificial *Backpropagation* [125]. En el caso de reconocimiento de sílabas con habla silenciosa, se ha utilizado el algoritmo de producto interno generado por la envoltura de filtros [128]; el vecino más cercano (k-NN) [31]; el clasificador Bayesiano y el análisis discriminante lineal (LDA) [32], entre otros. Para el reconocimiento de palabras con habla silenciosa se han utilizado clasificadores basados en HMM [134], entre otros. El clasificador utilizado con esta metodología es un SVM multiclase (1-1). Una de las claves para la óptima prestación de estos algoritmos fue la selección de los parámetros σ y C encontrados con algoritmos genéticos.

4.8.2. Desempeño del vector de características en función de la metodología desarrollada

Otra de las claves para el buen desempeño de los algoritmos desarrollados en función de la exactitud, es la selección del vector de características con datos funcionales y sus respectivas derivadas. Las figuras 4.34 (a) a (f) y 4.35 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v}(\cdot)$ que fueron obtenidas calculando la Eq. (2.37), y sus respectivas derivadas $D\hat{g}_{s,v}(\cdot)$ usando Eq. (2.47) del sujeto 1 con los ritmos cerebrales δ . Las figuras 4.36 (a) a (f) y 4.37 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v}(\cdot)$ que fueron obtenidas calculando la Eq. (2.37), y sus respectivas derivadas $D\hat{g}_{s,v}(\cdot)$ usando Eq. (2.47) del sujeto 1 con los ritmos cerebrales $\delta\theta$. Las figuras 4.38 (a) a (f) y 4.39 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v}(\cdot)$ que fueron

obtenidas calculando la Eq. (2.37), y sus respectivas derivadas $D\hat{g}_{s,v}(\cdot)$ usando Eq. (2.47) del sujeto 1 con los ritmos cerebrales $\delta\theta\alpha$. Para el caso de sílabas con habla silenciosa, las figuras 4.47 (a) a (f) y 4.48 (a) a (d) muestran las curvas funcionales $\hat{g}_{s,v}(\cdot)$ que fueron obtenidas calculando la Eq. (2.37), y sus respectivas derivadas $D\hat{g}_{s,v}(\cdot)$ usando Eq. (2.47) del sujeto 6 con los ritmos cerebrales δ .

En comparación, otras investigaciones de vocales con habla silenciosa han utilizado los siguientes vectores de características: descriptores estadísticos [125]; transformada rápida de Fourier (FFT) [28]; patrones espaciales comunes (CSP) [29],[126], [127], [30]; periodogramas [124] y coeficientes cepstrales en las frecuencias de Mel (MFCC) [123]. En el caso de sílabas con habla silenciosa se han utilizado: CSP [129],[130]; análisis de componentes independientes (ICA) con modelos de autorregresión (AR) [31]; espectro de Hilbert [32] y envoltura generada por el espacio de Hilbert [128]. En el caso de palabras con habla silenciosa se ha utilizado: descriptores estadísticos [125]; FFT [133]; modelos ocultos de Markov (HMM) [134], [131] y transformadas de wavelet discretas (DWT) [132].

4.8.3. Criterio para la selección de la metodología desarrollada de habla silenciosa

A continuación se explican algunas de las razones por las cuales fueron seleccionados la ubicación de los electrodos y los componentes del modelo desarrollado.

La ubicación de los electrodos estuvo supeditada a los modelos neurológicos del lenguaje denominados: Litchtheim, Geshwind y modelo de Hickok y Poeppel. En el modelo de Litchtheim el centro auditivo está relacionado con la comprensión del habla y el centro conceptual con el almacenamiento del significado de componentes del lenguaje (área de Wernicke); adicionalmente, el centro motor está relacionado con la producción del habla (áreas de Broca) y estas áreas con el hemisferio izquierdo (Fig. 2.9). El modelo de Geshwind relaciona tres centros en el hemisferio izquierdo (Fig. 2.10): el centro de producción del lenguaje (área de Broca), el centro de comprensión del lenguaje y el centro conceptual, ubicado en la circunvolución supramarginal y circunvolución angular, esta últimas ubicadas a su vez en el área de Wernicke [49].

En consecuencia, para vocales con habla silenciosa, los electrodos E1, E6, E7, E11, E12, E13, E14, E17, E18, E19 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E15, E16 y E21 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.3). Para sílabas con habla silenciosa, los electrodos E1, E6, E7, E11, E12, E13, E16, E17, E21 fueron localizados sobre el área de Broca, los electrodos E2, E4, E5, E8, E9, E10, E14, E15, E19 y E20 sobre el área de Wernicke y el electrodo E3 sobre la corteza auditiva primaria (Fig. 3.4).

Para determinar el número de sujetos en la experimentación de vocales y sílabas con habla silenciosa, se tuvo presente la cantidad de sujetos en otras experimentaciones relacionadas con habla silenciosa. Para vocales con habla silenciosa se ha utilizado la siguiente cantidad de sujetos: 3 en [29], 4 en [126], 5 en [127], 8 en [28] y 13 en [125]. Para sílabas con habla silenciosa se ha utilizado la siguiente cantidad de sujetos: 4 en [128] y 7 en [31]. Para caracteres chinos y palabras con habla silenciosa se utilizó la siguiente cantidad de sujetos: 2 en [34], 7 en [33], 8 en [130] y 21 en [131]. En esta tesis se buscó que la cantidad de sujetos estuviera por encima de la media de esos trabajos, de manera que para vocales con habla silenciosa se contó con la participación de 20 sujetos y para sílabas con habla silenciosa se contó con 10 sujetos.

Durante la experimentación, para vocales y sílabas con habla silenciosa, se les pidió a los sujetos permanecer con los ojos cerrados durante todo el tiempo. Esta solicitud tenía como propósito disminuir el número de artefactos debidos al parpadeo; y por el otro, en relación con el modelo cognitivo de Benedet [52], se buscaba disminuir la cantidad de entradas de información no deseadas para el proceso de lenguaje por parte del sujeto. Cuando el sujeto cierra los ojos no tiene entrada de señales visuales relacionadas a objetos del entorno. De esta forma la información del sistema semántico se relaciona a procesos internos del cerebro, excluyendo la entrada de léxico gráfico.

Otro factor a tener en cuenta es la frecuencia de muestreo. La frecuencia de muestreo de vocales con habla silenciosa fue de 500 Hz , y un parámetro para la selección estuvo determinado por la frecuencia de muestreo mínima del equipo de adquisición de señales cerebrales, amplificador NicoletOne™ V32 (Viasys Healthcare, Inc. Conshohocken, Pennsylvania). Para el caso de las sílabas con habla silenciosa, se buscó la máxima frecuencia de muestreo del equipo de adquisición, 2000 Hz , teniendo en cuenta que la duración promedio de las sílabas para la experimentación es 200 ms o 400 muestras [176]. Tanto para sílabas como para vocales, se tuvo en cuenta que cumplieran con el criterio de Nyquist [160]. En este sentido, en la metodología basada en datos funcionales la máxima frecuencia utilizada fue 13 Hz .

El parpadeo es uno de los artefactos que más afecta el desempeño de un sistema de procesamiento de señales cerebrales aplicados a BCI. En esta tesis se desarrolló un algoritmo que, primero, detecta el parpadeo (Eq. 3.1) y posteriormente lo elimina. El algoritmo desarrollado para eliminar el parpadeo está basado en wavelets y tiene tres etapas: (1) descomposición de la señal observable, (2) umbralización, y (3) reconstrucción de la señal estimada [162]. Entre las principales ventajas que presenta este algoritmo están: el hecho de que los wavelets son procesos aditivos y de esta manera se puede representar las señales cerebrales como una sumatoria de la señal a extraer, más ruido (artefactos) que deseamos eliminar; el tipo de artefacto a detectar se puede particularizar seleccionando el wavelet apropiado; los artefactos se pueden modelar como ruido blanco; y el tiempo computacional lo hace factible para ser utilizado on-line, entre otras. Las ventajas anteriores lo hacen más robusto, en comparación con métodos para eliminar el parpadeo tales como: *second order blind identification* (SOBI) [177], *second order cyclostationary statistics* (SOCS) [178], análisis de correlación canónica (CCA) [179], y *Automatic Wavelet Independent Component Analysis* (AWICA) [180], entre otros.

Las características de la metodología con datos funcionales se definieron de acuerdo con los siguientes criterios: Las señales de sílabas o vocales de habla silenciosa se transforman en estimaciones de PSD (periodogramas). El PSD es una medida estadística que cuantifica la potencia de una señal relacionada a los componentes de frecuencia [172]. Entre las características del PSD está la posibilidad de detectar señales inmersas en ruido, y la posibilidad de utilizar señales modeladas como procesos randómicos estacionarios [173]; en este sentido, el PSD estimado permite, en principio, separar, en cierto grado, las señales de habla silenciosa del ruido (artefactos). Otra ventaja es que permite seleccionar los ritmos cerebrales que se desean trabajar. Para el análisis de vocales con habla silenciosa se utilizaron los ritmos δ , $\delta\theta$ y $\delta\theta\alpha$, y en el caso de sílabas con habla silenciosa se utilizaron los ritmos δ .

En contraste al trabajo desarrollado en esta tesis, otros estudios de EEG con habla silenciosa han utilizado para reconocimiento de vocales con habla silenciosa ritmos: θ , α y β [137]; δ , θ y α [124]; α , β y γ [123]; δ , θ , α , β y γ [29], [122], [30]; y δ , θ , α , β , γ y γ alta [126], [127]. En el caso de reconocimiento de sílabas con habla silenciosa se han utilizado ritmos: δ , θ y α [32] y, θ , α y β [128], [31], [129], [130]. Finalmente, para reconocimiento de palabras

con habla silenciosa se han utilizado ritmos: β [34]; θ , α y β [132] y, δ , θ , α , β y γ [133], [134], [33].

Con los valores del PSD estimado, se transformó esta información en curvas funcionales. el área de la estadística llamada datos funcionales analiza curvas o superficies, describiéndolas como realizaciones en un espacio continuo de Hilbert L^2 [150]. Para calcular una estimación de los datos funcionales, un conjunto de funciones base independientes entre sí fue utilizado [155]. Seleccionando estas funciones base es posible aproximar funciones arbitrarias, calculando una combinación lineal de las respectivas bases [150]. Existen diferentes tipos de funciones base aplicables a datos funcionales [155]: exponenciales, polinomiales, poligonales, wavelets, Fourier y B-splines, entre otras. Las funciones base más utilizadas son las bases de Fourier, aplicadas a datos periódicos, y bases B-spline, aplicadas a datos no periódicos [150]. Como las señales cerebrales con habla silenciosa son representadas como una señal randómica, entonces se selecciona las bases B-spline [150], [155]. De esta manera, los datos funcionales estimados con las funciones B-spline son representados en términos de los coeficientes, el número de bases y las bases B-splines (Eq. 2.37). Adicionalmente, los datos funcionales permiten utilizar la información espacial de los electrodos como el dominio de las curvas funcionales generando un sistema más robusto en cuanto a integración de información. Otra ventaja que posee los datos funcionales es la utilidad de expresar las curvas como espacios continuos de Hilbert en L^2 ; de esta manera se puede calcular transformaciones matemáticas de las funciones como la primera o segunda derivada. En este estudio, para conformar el vector de características de vocales y sílabas con habla silenciosa, se utilizaron los coeficientes de las curvas funcionales y los coeficientes de la derivada de las curvas funcionales; las derivadas se utilizaron con el ánimo de caracterizar las razones de cambio en las curvas funcionales.

En la metodología de vector de características con datos funcionales, la selección del algoritmo de clasificación basado en SVM estuvo supeditada al análisis comparativo de otros autores sobre clasificadores multiclase de señales EEG aplicadas a BCI. Lotte et al. [182], hacen un estudio comparativo entre los siguientes clasificadores aplicados a señales EEG: análisis discriminante lineal (LDA), SVM, redes neuronales, modelos ocultos de Markov (HMM), clasificadores Bayesianos, el vecino más cercano (kNN) y combinación de clasificadores como Boosting, entre otros. Lotte et al. presentan un cuadro comparativo para la selección del óptimo clasificador en términos de: linealidad-no linealidad, generativo-discriminativo, estático-dinámico, estable-inestable y robustez a la alta dimensión. Para el caso asincrónico y sincrónico, el clasificador SVM presta un alto desempeño, en comparación con los otros clasificadores. Lee et al. [183] concluyen que para señales EEG de alta dimensión el clasificador SVM es más preciso que otros clasificadores como análisis discriminante lineal.

4.8.4. Viabilidad de las metodologías propuestas de habla silenciosa aplicadas a BCIs

En esta parte se busca explicar cómo se podrían aplicar las metodologías desarrolladas a dispositivos (*spellers*, movimiento de mouse, sillas de ruedas y prótesis, entre otros), contrastándolos con otras metodologías de BCI.

En BCI se utilizan los *spellers* para el deletreo de caracteres utilizando señales cerebrales. En este caso, el estímulo que se le presenta al sujeto es información visual a través de una pantalla. En ella se muestra una matriz de 6x6 que contiene las letras del alfabeto, los números del 1 al 9 y un símbolo de espacio que sirve de separación entre las posibles palabras de una oración. En la pantalla, los caracteres se prenden y apagan en forma aleatoria. En su base

de datos se almacena la fila y columna de ese caracter y el tiempo en que se generó el evento respectivo. Cuando la persona fija su atención en el caracter que quiere codificar y este se enciende, se genera en el sujeto un potencial evocado. Esta información junto con la fila y columna que se iluminó sirve para determinar el caracter en que el sujeto estaba pensando en ese momento [76], [77]. Señales como P300 y SSVEP se han utilizado para esta tarea [19], pero aún poseen grandes limitaciones como el requerir de un estímulo externo (matriz de caracteres en una pantalla de computador) y el estar completamente concentrado esperando el evento del mundo externo; además, tiene el inconveniente de generar fatiga después de un tiempo determinado. Como la información se genera caracter a caracter, lo hace costoso en el tiempo para transmitir información relativa a oraciones. En contraste, la utilización de las metodologías desarrolladas en esta tesis tienen la factibilidad de proyectar *spellers* con nuevas ventajas. Una primera etapa con las metodologías desarrolladas sería entrenar al sistema para reconocer los caracteres de un idioma, y posteriormente la persona sólo debería pensar en el respectivo caracter. Esto lo podría hacer sin requerir de una pantalla de computador o del esfuerzo cognitivo de atención. Como la persona no debería esperar a que en la pantalla de un computador se encendiera el caracter, esto incrementaría el tiempo de respuesta.

El control de sillas de ruedas actualmente se hace utilizando señales P300, SSVEP e imagería motora. En el caso de P300 o SSVEP, el sujeto debe estar mirando la pantalla de un computador y de acuerdo a la dirección donde desea dirigirse, se debe concentrar en una parte de la pantalla de un computador que parpadea a una determinada frecuencia [20], [16], [23]. De acuerdo a esa frecuencia, el BCI interpreta la dirección en que el sujeto desea desplazarse. En el caso de la imagería motora, el sujeto debe concentrarse en imaginar una actividad relacionada con la dirección a donde desea desplazarse. Los métodos anteriores requieren de la existencia de un estímulo externo y de que el sujeto tenga un alto nivel de concentración; a su vez, tienen el inconveniente de generar fatiga después de un tiempo determinado. En contraste, en esta tesis la persona podría utilizar las vocales o sílabas con habla silenciosa para realizar movimientos con la silla de ruedas: adelante, atrás, girar a la izquierda, girar a la derecha y detenerse. Esto lo podría hacer sin requerir de un estímulo externo y con un mínimo nivel de concentración.

Pfurtscheller et al. han controlado prótesis de brazo utilizando señales EEG con SSVEP [16]. En este método se han colocado 4 leds en diferentes partes de la prótesis. Los movimientos que puede realizar son: abrir la mano (led a 8 Hz), cerrar la mano (led a 13 Hz), girar a la izquierda la muñeca (led a 6 Hz) y girar a la derecha la muñeca (led a 7 Hz). En este caso, el movimiento de una articulación está supeditado al led que observe con detenimiento. En contraste, las metodologías desarrolladas en esta tesis le permitiría a una persona amputada de miembro superior controlar una prótesis, utilizando vocales o sílabas con habla silenciosa. Por consiguiente, podría realizar movimientos como: abrir pinza, cerrar pinza, girar muñeca a la izquierda, girar muñeca a la derecha y detenerse. Esto lo podría hacer sin requerir de un estímulo externo y con un mínimo nivel de concentración. Si se llegara a requerir de prótesis con más grados de libertad, sólo sería necesario entrenar una mayor cantidad de unidades de lenguaje.

En términos generales, las metodologías desarrolladas de habla silenciosa tienen una gran cantidad de ventajas como las siguientes: no ser invasivos; no requerir de procesos de entrenamiento dispendiosos como en el caso de la imagería motora [22]; no requerir de un proceso de atención riguroso como en SSVEP [23], P300 [14] o imagería motora [24]; no requerir de un estímulo externo como en el caso de SSVEP [25] o P300 [26]; y no requerir de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imagería

motora [27]. Adicionalmente, utilizan señales cerebrales que están relacionadas con el área de lenguaje y tienen la posibilidad de utilizar la innumerable cantidad de palabras (léxico) de un lenguaje.

4.8.5. Proyección de mejoras a las metodologías desarrolladas de habla silenciosa

Entre las mejoras a desarrollar, en esta tesis, se encuentran: el hecho que el trabajo es desarrollado off-line y para aplicaciones de BCIs on-line, sería necesario tener presente el tiempo computacional de los algoritmos empleados. También, se debe tener presente que las metodologías desarrolladas fueron realizadas en un laboratorio con condiciones controladas. Por consiguiente, para controlar dispositivos con BCI se requiere de un estudio en ambientes externos a un laboratorio.

Capítulo 5

Conclusiones

El lenguaje es un proceso cognitivo donde interactúan distintas partes del cerebro. Diferentes modelos se han desarrollado para tratar de entender este proceso. Broca y Wernicke desarrollaron los primeros modelos neurológicos del lenguaje. Broca relacionó el área en la circunvolución frontal inferior del hemisferio izquierdo [47] con la producción del lenguaje y Wernicke relacionó el área la circunvolución temporal superior en el hemisferio izquierdo con los procesos de comprensión del lenguaje [48]. El modelo neurológico de Geschwind retoma los trabajos de Wernicke y Litchtheim adicionando el centro conceptual en la circunvolución supramarginal y circunvolución angular, ambas pertenecientes a la zona posterior e inferior del lóbulo parietal en el hemisferio izquierdo. [49].

Durante las últimas dos décadas se han desarrollado avanzados modelos de lenguaje basados en neuroimágenes. El modelo de comprensión oral de Hickok y Poeppel establece el proceso neurológico del lenguaje como un conjunto de redes especializadas que se interconectan. Este modelo tiene una entrada de información que puede ser auditiva (corteza auditiva primaria) o visual en los lóbulos occipitales. En el caso de la información auditiva existe una red para el análisis espectrotemporal de las señales del lenguaje, una red fonológica donde se reconocen los fonemas de un lenguaje, una interfaz léxica donde se le asigna una representación conceptual a los fonemas que han sido reconocidos en la red fonológica, una red combinatorial donde se conforman sistemas complejos del lenguaje con carácter sintáctico, una interfaz sensoriomotora donde se conforma la representación motora de los fonemas y una red articular localizada en la circunvolución precentral para que sea producida la señal de voz por el aparato fonador [54].

Uno de los primeros estudios sobre habla silenciosa con EEG se reportó en 1994 con la vocal /a/ [28]. Durante la primera década de este siglo se reportaron trabajos de habla silenciosa utilizando EEG con dos vocales o sílabas: en el 2009 se reportó la clasificación de las vocales silenciosas /a/ y /u/ [29] y en el 2009 y 2010 se reportó la clasificación de las sílabas silenciosas /ba/ y /ku/ [128], [31], [32].

Un primer estudio sobre habla silenciosa con EEG donde se reporta la clasificación de cinco palabras fue realizado en el 2009 [131]. En esta última década, estudios que han utilizando EEG con habla silenciosa han reportado la clasificación de: dos caracteres silenciosos chinos "zuo" y "yi" [129], [130] y dos palabras silenciosas [33], [34]. En el 2014 el autor de esta tesis clasificó la clase de vocales silenciosas abiertas-medias, en contraste con la clase de vocales silenciosas cerradas [124]. Los trabajos más recientes han logrado clasificar hasta cinco vocales en idioma inglés [125] y cinco en idioma japonés [126],[127], [30]. En esta tesis, por medio de señales EEG se logró clasificar cinco vocales con habla silenciosa y cinco sílabas con habla

silenciosa. La primera metodología estuvo basada en un vector de características con entropía de la información y dimensión de regularización; y la segunda, en un vector de características con datos funcionales.

5.1. Metodología basada en vector de características con entropía de la información y dimensión de regularización

El algoritmo denominado metodología basada en vector de características con entropía de la información y dimensión de regularización, fue desarrollado en esta investigación para el reconocimiento de señales EEG con habla silenciosa (vocales y sílabas). El algoritmo toma las señales cerebrales del área del lenguaje, centradas en las áreas de Broca y Wernicke, y posteriormente realiza un proceso de filtraje, luego lleva a cabo una detección y eliminación de artefactos. De las señales cerebrales resultantes se extraen características de entropía de la información y características de dimensión de regularización. La entropía de información mide la incertidumbre de las señales de habla silenciosa [138] y la dimensión de regularización mide la complejidad de las señales de habla silenciosa en términos de la dimensión fractal [149]. La implementación numérica de este algoritmo se verificó con 20 sujetos, clasificando cinco vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa, y se verificó con 10 sujetos, clasificando cinco sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa. La clasificación de estas señales de habla silenciosa fue desarrollada utilizando los clasificadores multiclase SVM del tipo uno contra el resto (1-*rest*) y uno contra uno (1-1).

De acuerdo a los resultados de esta tesis, la metodología basada en vector de características con entropía de la información y dimensión de regularización está en concordancia con los modelos de Litchtheim, Geshwind y el modelo de Hickok y Poeppel para vocales y sílabas con habla silenciosa. De los resultados de la respectiva metodología se encuentra que, para vocales y sílabas con habla silenciosa, existe una mayor activación del centro de comprensión del lenguaje y el centro conceptual (área de Wernicke), en relación al área de Broca [49].

En el caso de vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización se seleccionaron dos algoritmos de clasificación multiclase (1-*rest*) y (1-1). Para SVM con (1-*rest*) se alcanzó en la clasificación una exactitud entre 46.56 % y 85.90 % con una media estadística de 61.41 % y una desviación estándar de 3.98 %. Para SVM con (1-1) se alcanzó en la clasificación una exactitud entre 52.67 % y 91.43 % con una media estadística de 69.83 % y una desviación estándar de 3.23 %. Para el caso de sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización, se seleccionaron dos algoritmos de clasificación multiclase (1-*rest*) y (1-1). Para SVM con (1-*rest*) se alcanzó en la clasificación una exactitud entre 47.97 % y 74.68 % con una media estadística de 58.57 % y una desviación estándar de 4.92 %. Para SVM con (1-1) se alcanzó en la clasificación una exactitud entre 55.11 % y 84.22 % con una media estadística de 66.89 % y una desviación estándar de 4.03 %. Asimismo, se encontró que utilizando la metodología basada en vector de características con entropía de la información y dimensión de regularización, para cinco vocales y cinco sílabas con habla silenciosa, el clasificador multiclase SVM (1-1) presenta mayor exactitud que el clasificador multiclase SVM (1-*rest*). Estos resultados están en concordancia con Hsu et al.

[167] y Pal et al. [175], donde se especifican las mejores prestaciones del clasificador (1-1) en comparación a (1-rest).

En términos de exactitud y para el caso de vocales con habla silenciosa, los resultados de este algoritmo son comparables, en el número de vocales clasificadas, a otras investigaciones reportadas de habla silenciosa con señales EEG de cinco vocales como la de Kamala et al. [125] en idioma inglés, y las de Matsumoto [126],[127], [30] en idioma japonés. Para el caso de sílabas, la exactitud de este algoritmo con habla silenciosa es superior, en el número de sílabas clasificadas, a otras investigaciones reportadas de habla silenciosa con señales EEG donde se han trabajado hasta dos sílabas como la de D’Zmura et al. [128], Brigham et al. [31] y Deng et al.[32].

En contraste a otros métodos de BCI con señales EEG, la metodología basada en vector de características con entropía de la información y dimensión de regularización presenta las siguientes ventajas: no es invasivo; no requiere de procesos de entrenamiento dispendiosos como en el caso de la imaginería motora [22]; no requiere de un proceso de atención riguroso como en SSVEP [23] , P300 [14] o imaginería motora [24]; no requiere de un estímulo externo como en el caso de SSVEP [25] o P300 [26]; y no requiere de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imaginería motora [27]. Como otra ventaja, utiliza señales cerebrales que están relacionadas con el área de lenguaje y tienen la posibilidad de utilizar la innumerable cantidad de palabras (léxico) de un lenguaje.

De esta manera, se responde a la pregunta de investigación y se cumple con los objetivos de la tesis, demostrando la siguiente hipótesis:

La identificación de señales cerebrales EEG a través del desarrollo de un algoritmo de clasificación basado en vector de características con entropía y dimensión de regularización permiten el reconocimiento del habla silenciosa.

5.2. Metodología basada en vector de características con datos funcionales

El algoritmo denominado metodología basada en vector de características con datos funcionales fue desarrollado en esta investigación para el reconocimiento de señales EEG con habla silenciosa (vocales y sílabas). El algoritmo toma las señales cerebrales del área del lenguaje, centradas en las áreas de Broca y Wernicke, y posteriormente realiza un proceso de filtraje; luego realiza una detección y eliminación de artefactos. De las señales cerebrales resultantes se obtiene el PSD estimado (periodograma) para las vocales con habla silenciosa (ritmos δ , $\delta\theta$ y $\delta\theta\alpha$) y para sílabas (ritmos δ) [172]. Con los datos de PSD estimado y la distancia angular de los electrodos se obtienen las funciones de los datos funcionales y su primera derivada [150].

La implementación numérica de este algoritmo se verificó clasificando cinco vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa y clasificando cinco sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa en 20 y 10 sujetos, respectivamente. La clasificación de estas señales de habla silenciosa fue desarrollada utilizando el clasificador multiclase SVM del tipo uno contra uno (1-1).

De acuerdo a los modelos neurológicos del lenguaje, en el modelo de Litchtheim el centro auditivo está relacionado con la comprensión del habla, y el centro conceptual con el almacenamiento del significado de componentes del lenguaje (área de Wernicke); adicionalmente, el centro motor está relacionado con la producción del habla (áreas de Broca) y estas áreas

con el hemisferio izquierdo (Fig. 2.9). En el modelo de Hickok y Poeppel se relaciona varios componentes con las áreas de Wernicke y Broca como los bloques de: red fonológica, análisis espectrotemporal, interfaz léxica y red combinacional [54]. Por consiguiente, en los resultados de esta tesis, se encuentra que la metodología basada en vector de características con datos funcionales está en concordancia con los modelos de Litchtheim, Geshwind y el modelo de Hickok y Poeppel para vocales y sílabas con habla silenciosa. En el caso de vocales y sílabas con habla silenciosa, una mayor activación del centro de comprensión del lenguaje y el centro conceptual (área de Wernicke) se presenta, en relación con el área de Broca [49].

Para vocales (/a/, /e/, /i/, /o/, /u/) con habla silenciosa utilizando la metodología basada en vector de características con datos funcionales se seleccionó el algoritmo SVM de clasificación multiclase (1-1), ya que la metodología basada en vector de características con entropía de la información y dimensión de regularización presentó mayor exactitud, tanto para vocales como para sílabas con habla silenciosa. A continuación se describen los valores de clasificación alcanzados con la metodología basada en vector de características con datos funcionales. Primero, se utilizó vocales con habla silenciosa para ritmos δ , donde se alcanzó en la clasificación una exactitud entre 47.33 % y 68.56 % con una media estadística de 57.49 % y una desviación estándar de 4.56 %. Después, se utilizó ritmos $\delta\theta$ donde se alcanzó en la clasificación una exactitud entre 53.56 % y 78.78 % con una media estadística de 69.37 % y una desviación estándar de 4.08 %. Finalmente, para ritmos $\delta\theta\alpha$ se alcanzó en la clasificación una exactitud entre 65.22 % y 78.0 % con una media estadística de 71.92 % y una desviación estándar de 4.56 %. En el caso de vocales con habla silenciosa utilizando la metodología de vector de características con datos funcionales, los ritmos $\delta\theta\alpha$ presentan mayor exactitud que los ritmos $\delta\theta$ y los ritmos δ . Para el caso de sílabas (/fa/, /pa/, /ma/, /la/, /ra/) con habla silenciosa (ritmos δ) utilizando la metodología basada en vector de características con datos funcionales se seleccionó el algoritmo de clasificación multiclase (1-1). Con este clasificador se alcanzó una exactitud entre 60.67 % y 72.67 % con una media estadística de 67.13 % y una desviación estándar de 6.61 %. De los resultados de la clasificación se concluye que la exactitud alcanzada, para vocales y sílabas con habla silenciosa, utilizando la metodología basada en vector de características con datos funcionales, es más alta que con la metodología basada en vector de características con entropía de la información y dimensión de regularización.

La exactitud de la metodología basada en vector de características con datos funcionales, es comparable en el número de vocales clasificadas, a otras investigaciones reportadas de habla silenciosa con señales EEG de cinco vocales en inglés [125] y cinco vocales en japonés [126], [127], [30]. Matsumoto et al. [127], [30] clasificaron cinco vocales japonesas con una exactitud entre 57 % a 87 % utilizando un rango de frecuencia entre 1 a 125 Hz . En contraste, la metodología desarrollada basada en vector de características con datos funcionales alcanzó una exactitud comparable, utilizando solamente un rango de frecuencias entre 2 a 13 Hz (ritmos $\delta\theta\alpha$). Para el caso de sílabas utilizando la metodología basada en vector de características con datos funcionales, la exactitud de este algoritmo con habla silenciosa es superior, en el número de sílabas clasificadas, a otras investigaciones reportadas de habla silenciosa con señales EEG donde se han clasificado hasta dos sílabas con habla silenciosa como los trabajos desarrollados por D'Zmura et al. [128], Brigham et al. [31] y Deng et al. [32].

En términos generales, la metodología basada en vector de características con datos funcionales tienen grandes ventajas similares a la metodología basada en vector de características con entropía de la información y dimensión de regularización como: no requiere de procesos de entrenamiento dispendiosos como en el caso de la imagería motora; no requiere de un proceso de atención riguroso como en SSVEP , P300 o imagería motora; no requiere de

un estímulo externo como en el caso de SSVEP o P300; y no requiere de tareas cognitivas que generen fatiga muscular o cognitiva como en el caso de la imaginación motora. De otra manera, la metodología desarrollada tiene la potencialidad de reconocer más de cinco vocales y de cinco sílabas con habla silenciosa. En contraste, otros métodos de BCI han alcanzado un limitado número de patrones a clasificar. En el caso de P300 reconoce, en general, un solo potencial evocado [77], [78], [79]. Con el método de atención selectiva SSVEP la tendencia es a clasificar tres patrones [84] y cuatro patrones [16], [15], [23]. En recientes trabajos de SSSVEP, se ha logrado clasificar hasta 8 patrones [85]. Para el caso de imaginación motora se ha logrado clasificar hasta cuatro patrones [98], [94], [87] y en un reciente estudio se logró clasificar cinco patrones [100]. Para mejorar la respuesta de las BCIs se ha tratado de realizar combinaciones de los métodos de atención selectiva e imaginación motora como: P300 y SSVEP [20]; P300 e imaginación motora [21] y SSVEP con imaginación motora [18], pero el número de patrones a clasificar sigue siendo limitado. En consecuencia, los métodos desarrollados en esta tesis tienen la potencialidad de superar esta limitación y se convierten en una opción importante para el control de dispositivos con BCI.

De esta manera se responde a la pregunta de investigación y se alcanzan los objetivos de la tesis, demostrando la siguiente hipótesis:

La identificación de señales cerebrales EEG a través del desarrollo de un algoritmo de clasificación basado en vector de características con datos funcionales permiten el reconocimiento del habla silenciosa.

Considerando las metodologías desarrolladas, esta tesis presenta dos novedosos sistemas de procesamiento de señales basados en electroencefalografía, con la habilidad de clasificar vocales y sílabas con habla silenciosa. Una de las metodologías se basa en las características de la entropía de la información con la dimensión de regularización, y el otro en las características de datos funcionales en el espacio de Hilbert L^2 , utilizando los datos de la posición de los electrodos. Las ventajas expuestas de las metodologías desarrolladas con habla silenciosa utilizando señales EEG, la convierten en una opción importante en el campo de BCIs, para ser aplicada como señales de control en dispositivos como: exoesqueletos aplicados en recuperar la movilidad de personas con esclerosis lateral amiotrófica (ALS); sillas de ruedas aplicados en personas con paraplejia; dispositivos protésicos en personas con amputación de miembro superior o miembro inferior; entre otros.

5.3. Recomendaciones y proyecciones

Los dos algoritmos desarrollados (metodología basada en características de entropía de la información y dimensión de regularización, y metodología basada en la transformación de la densidad espectral de potencia a datos funcionales), constituyen promisorias metodologías para el desarrollo de interfaces cerebro-computador (BCIs) utilizando habla silenciosa con señales EEG.

En esta investigación se clasificaron cinco vocales con habla silenciosa y cinco sílabas con habla silenciosa. Las características léxicas de un idioma posibilitan la implementación de una gran cantidad de palabras para ser aplicadas a BCIs. En este sentido, las posibilidades son ilimitadas. Para futuros trabajos de BCIs con habla silenciosa, se sugiere la clasificación de otras sílabas, palabras y oraciones.

En este estudio, la experimentación se desarrolló en condiciones controladas de luminosidad y ruido. Para que las metodologías desarrolladas se puedan aplicar en ambientes reales

se requiere que sean robustos a condiciones de ruido, luminosidad, artefactos generados por el movimiento del cuerpo, etc.

El estudio realizado en esta tesis fue off-line y para aplicarse como un sistema BCI, se requiere la implementación on-line. De esta manera, los algoritmos desarrollados tienen que prever los tiempos de respuesta para el control de dispositivos.

Una proyección futura de este trabajo es la aplicación de las señales de habla silenciosa con EEG para ser utilizadas en sistemas de BCIs, como señales para el control de: prótesis de miembro superior, prótesis de miembro inferior, robots, etc. Igualmente, se podría utilizar, con ayuda de un exoesqueleto, en personas con esclerosis lateral múltiple.

En trabajos venideros será posible plantear sistemas inteligentes basados en el reconocimiento automático de características y la posibilidad de comunicar varias personas para el desarrollo de trabajo cooperativo y colaborativo.

En un futuro cercano, es posible que dicho sistema BCI pueda comunicar diferentes agentes naturales o artificiales. En el caso de agentes naturales, podría ser aplicado en sistemas de comunicación como los celulares, o en ambientes de alto ruido donde a las personas se les dificulta comunicarse en voz alta.

5.4. Publicación de resultados y difusión

C. Sarmiento y A. Tovar. Control de prótesis de múltiples grados de libertad mediante potenciales evocados visuales modulados. Primer Congreso Internacional sobre Tecnologías Avanzadas de Mecatrónica, Diseño y Manufactura AMDM 2012. Universidad Tecnológica de Pereira. 2012.

C. Sarmiento. Interface Cerebro Computador para el control de dispositivos. II Congreso Internacional de Ingeniería Mecatrónica y Automatización. Universidad de la Salle - Bogotá. 2013.

C. Sarmiento, A. Tovar y C. Cortés. Interface Cerebro Computador (BCI) para la identificación del estado de pensamiento de las vocales a través de señales electroencefalográficas (EEG). VII Seminario internacional de Ingeniería biomédica. Universidad de los Andes - Bogotá. Memorias ISSN 2322-7702. 2014.

C. Sarmiento, A. Tovar, C. Cortés, J. Bacca, P. Lorenzana y J. Arcos. Brain computer interface (BCI) with EEG signals for automatic vowel recognition based on articulation mode. In Biosignals and Robotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC). 5th ISSNIP-IEEE. Universidad Federal de Bahía. Brasil. DOI: 10.1109 / BRC.2014.6880997. IEEE CONFERENCE PUBLICATIONS. pp. 712-717, IEEE, 2014.

C. Sarmiento, A. Tovar, C. Cortés, J. Bacca y P. Lorenzana. Pensar en las vocales activaría prótesis. UN Periódico. Publicación de la Universidad Nacional de Colombia. N°188 - mayo 2015 - Pag. 15. ISSN 1657-0987.

C. Sarmiento. Brain-computer interface (BCI). Reunión mensual de la Asociación Colombiana de Neurología (ACN), en el área de la neurofisiología. Transmisión a todas las sedes de la Asociación en el país por vía electrónica. Noviembre 27 del 2015.

Carlos Sarmiento, Andrés Tovar, Carlos Cortés, Jan Bacca, Pablo Lorenzana y Rubén Guevara. Classification of Electroencephalography Silent Speech Vowels Using Functional Data for Brain-Computer Interfaces. Artículo en elaboración para presentarse en: IEEE Transactions on Audio, Speech, and Language Processing.

Apéndice A

Tablas de la experimentación

Tabla A.1: Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para vocales con SS - SVM (1-1)

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.8222	0.8333	0.8667	0.7778	0.8444	0.7889	0.8000	0.8333	0.7889	0.8111	0.8167	0.0269
2	0.5889	0.5444	0.6556	0.6667	0.5556	0.6333	0.6222	0.6667	0.6333	0.5889	0.6156	0.0419
3	0.6667	0.6667	0.7000	0.8222	0.7444	0.6333	0.7222	0.6778	0.6889	0.7000	0.7022	0.0496
4	0.5333	0.5222	0.5560	0.4444	0.5444	0.5333	0.6000	0.5889	0.5667	0.5778	0.5467	0.0418
5	0.5778	0.6000	0.6889	0.6778	0.6444	0.5889	0.6000	0.6444	0.6667	0.6444	0.6333	0.0372
6	0.5889	0.6444	0.6333	0.6778	0.6889	0.6778	0.6667	0.6111	0.6556	0.6333	0.6478	0.0303
7	0.5111	0.5222	0.5444	0.5333	0.5556	0.6111	0.5556	0.6111	0.6111	0.5444	0.5600	0.0359
8	0.8222	0.8667	0.7778	0.8111	0.8444	0.8111	0.8556	0.7889	0.8333	0.8111	0.8222	0.0268
9	0.7111	0.7889	0.7222	0.7556	0.7222	0.7111	0.7111	0.7444	0.7889	0.7111	0.7367	0.0298
10	0.5889	0.5889	0.5778	0.6000	0.6778	0.6000	0.5667	0.5778	0.6778	0.5556	0.6011	0.0405
11	0.9222	0.9222	0.8778	0.8556	0.9222	0.9333	0.9000	0.9100	0.9556	0.9444	0.9143	0.0286
12	0.6000	0.6778	0.6111	0.6444	0.6222	0.7222	0.6444	0.6889	0.6556	0.7222	0.6589	0.0410
13	0.6333	0.6778	0.7000	0.6667	0.6444	0.7111	0.7444	0.6333	0.6222	0.6556	0.6689	0.0375
14	0.7667	0.8000	0.7889	0.8111	0.7889	0.8222	0.8000	0.8333	0.8111	0.8333	0.8056	0.0200
15	0.5000	0.4889	0.5222	0.5444	0.5000	0.5556	0.5667	0.5556	0.5556	0.4778	0.5267	0.0311
16	0.6778	0.6556	0.6444	0.6667	0.6111	0.7000	0.6667	0.6778	0.5889	0.6889	0.6578	0.0329
17	0.7222	0.6778	0.7000	0.7667	0.7000	0.7111	0.7111	0.7333	0.7222	0.7000	0.7144	0.0228
18	0.5444	0.6000	0.5889	0.5778	0.5556	0.5000	0.5444	0.5333	0.6111	0.5778	0.5633	0.0322
19	0.8444	0.8889	0.9111	0.8889	0.8889	0.8889	0.8667	0.8667	0.8667	0.8444	0.8756	0.0204
20	0.8889	0.8889	0.8667	0.8889	0.9333	0.9111	0.9000	0.8778	0.9222	0.9111	0.8989	0.0195

Tabla A.2: Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para vocales con SS - SVM (1-rest)

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.6972	0.7080	0.7232	0.6990	0.6667	0.7087	0.7358	0.7282	0.7182	0.7034	0.7088	0.0185
2	0.6506	0.5619	0.6071	0.6563	0.5288	0.5977	0.6277	0.5843	0.6220	0.6825	0.6119	0.0438
3	0.6182	0.5688	0.6465	0.6282	0.6442	0.6111	0.5476	0.6000	0.6237	0.5547	0.6043	0.0339
4	0.5000	0.5062	0.5223	0.5488	0.5625	0.4435	0.6023	0.5057	0.4949	0.4839	0.5170	0.0423
5	0.5439	0.5310	0.5225	0.5347	0.4537	0.5145	0.5135	0.5321	0.4643	0.4730	0.5083	0.0308
6	0.6716	0.5625	0.5495	0.5143	0.4857	0.5443	0.5317	0.5545	0.6197	0.5185	0.5552	0.0513
7	0.5204	0.4207	0.4393	0.4674	0.4412	0.4545	0.4455	0.5043	0.4336	0.5287	0.4656	0.0365
8	0.7363	0.7051	0.8158	0.7356	0.6783	0.7320	0.6606	0.7284	0.6724	0.6102	0.7075	0.0530
9	0.6264	0.6818	0.6593	0.6381	0.6273	0.6489	0.5565	0.5630	0.6813	0.6782	0.6361	0.0430
10	0.6104	0.5154	0.5377	0.5313	0.4911	0.5085	0.5657	0.5455	0.5300	0.5688	0.5404	0.0326
11	0.7810	0.7905	0.8077	0.7699	0.8019	0.8269	0.8387	0.7925	0.8351	0.8667	0.8111	0.0285
12	0.4892	0.4747	0.5887	0.5039	0.6180	0.4495	0.4737	0.5377	0.4874	0.4744	0.5097	0.0521
13	0.5340	0.6721	0.5483	0.5833	0.6000	0.6076	0.5897	0.6196	0.6400	0.5784	0.5973	0.0388
14	0.7170	0.6822	0.7553	0.6387	0.5341	0.8049	0.7609	0.6961	0.7609	0.7000	0.7050	0.0729
15	0.4365	0.4561	0.5222	0.4000	0.5341	0.3386	0.5172	0.5000	0.4486	0.5714	0.4725	0.0664
6	0.6588	0.5271	0.5784	0.5347	0.5789	0.5583	0.5593	0.5500	0.5283	0.5238	0.5598	0.0382
17	0.5676	0.6098	0.5935	0.6212	0.6699	0.5870	0.5833	0.6410	0.5556	0.5909	0.6020	0.0327
18	0.4946	0.5070	0.4773	0.4659	0.4845	0.4463	0.5000	0.5181	0.5397	0.5606	0.4994	0.0323
19	0.8462	0.8351	0.7980	0.8081	0.8384	0.8220	0.7818	0.7855	0.7788	0.8247	0.8119	0.0236
20	0.8247	0.8652	0.8804	0.8941	0.8515	0.8571	0.8247	0.8333	0.8764	0.8830	0.8590	0.0238

Tabla A.3: Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos δ

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.5222	0.6333	0.5556	0.5444	0.5444	0.4778	0.5667	0.5667	0.6000	0.4667	0.5478	0.0479
2	0.5778	0.4667	0.4778	0.5667	0.5111	0.5000	0.6333	0.5000	0.5111	0.4778	0.5222	0.0507
3	0.6333	0.5666	0.5222	0.6667	0.6111	0.5000	0.6444	0.5889	0.5444	0.5667	0.5844	0.0517
4	0.5222	0.5000	0.4778	0.5333	0.5778	0.5000	0.4222	0.4333	0.5333	0.5222	0.5022	0.0449
5	0.5444	0.6000	0.5778	0.6667	0.5222	0.5444	0.6000	0.6222	0.6778	0.5556	0.5911	0.0499
6	0.5778	0.6000	0.6000	0.5667	0.5444	0.5444	0.6000	0.5111	0.6556	0.6111	0.5811	0.0392
7	0.6222	0.6778	0.6444	0.6111	0.6556	0.7000	0.6889	0.5667	0.5333	0.5889	0.6289	0.0519
8	0.7000	0.6556	0.6222	0.6778	0.7111	0.6889	0.6556	0.6222	0.7111	0.7000	0.6745	0.0322
9	0.7222	0.6556	0.6556	0.7111	0.6889	0.6889	0.6778	0.7111	0.6778	0.6667	0.6856	0.0222
10	0.6000	0.5333	0.5556	0.5222	0.5444	0.5444	0.5222	0.5333	0.5000	0.5556	0.5411	0.0254
11	0.5333	0.6222	0.5667	0.6000	0.5333	0.5889	0.6556	0.6556	0.6000	0.6622	0.6018	0.0454
12	0.5667	0.6333	0.5444	0.5444	0.5222	0.5556	0.5444	0.6000	0.5444	0.5000	0.5555	0.0358
13	0.5111	0.6000	0.5778	0.6444	0.6889	0.6111	0.5333	0.6444	0.6222	0.6778	0.6111	0.0549
14	0.5111	0.5889	0.6222	0.6000	0.6444	0.5000	0.5222	0.5778	0.6889	0.5000	0.5756	0.0624
15	0.5889	0.6556	0.5444	0.6000	0.5889	0.5778	0.5333	0.5889	0.5556	0.5889	0.5822	0.0323
16	0.6000	0.4667	0.5222	0.5222	0.5333	0.4222	0.4667	0.5222	0.5333	0.5222	0.5111	0.0461
17	0.6000	0.6556	0.5333	0.5556	0.6000	0.6444	0.5667	0.5556	0.6111	0.6111	0.5933	0.0379
18	0.6222	0.5222	0.6222	0.5889	0.6333	0.5889	0.7111	0.5444	0.6222	0.5556	0.6011	0.0513
19	0.4111	0.4444	0.4667	0.5333	0.5333	0.4111	0.3556	0.4889	0.5333	0.5556	0.4733	0.0635
20	0.4111	0.6333	0.5778	0.5111	0.5889	0.5667	0.4222	0.5556	0.5444	0.5222	0.5333	0.0669

Tabla A.4: Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos $\delta\theta$

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.6667	0.6778	0.7444	0.6333	0.8000	0.7444	0.7333	0.6778	0.6889	0.7222	0.7089	0.0463
2	0.5778	0.6000	0.6111	0.6556	0.6111	0.6333	0.6778	0.6333	0.5889	0.5778	0.6167	0.0315
3	0.6667	0.6333	0.6778	0.5667	0.6778	0.6556	0.7111	0.5889	0.7222	0.7000	0.6600	0.0482
4	0.6000	0.5111	0.6222	0.6222	0.5778	0.5444	0.5444	0.5444	0.5444	0.4778	0.5589	0.0445
5	0.5889	0.6222	0.6000	0.5333	0.7111	0.6111	0.6000	0.6111	0.6444	0.6444	0.6167	0.0434
6	0.7444	0.7000	0.6556	0.7556	0.7889	0.7333	0.6778	0.6222	0.6333	0.7222	0.7033	0.0524
7	0.7889	0.8000	0.6444	0.7667	0.7222	0.7111	0.7000	0.7111	0.7667	0.7667	0.7378	0.0456
8	0.7889	0.7000	0.7778	0.8111	0.7222	0.7444	0.7333	0.7667	0.7111	0.7778	0.7533	0.0347
9	0.7556	0.7556	0.8000	0.8444	0.7889	0.7889	0.8000	0.7778	0.7556	0.8111	0.7878	0.0269
10	0.6556	0.7444	0.7333	0.7556	0.7222	0.7556	0.7000	0.7111	0.7667	0.7222	0.7267	0.0311
11	0.7111	0.7000	0.7000	0.6667	0.6889	0.7111	0.7000	0.7111	0.7444	0.7000	0.7033	0.0186
12	0.6000	0.6000	0.5667	0.6222	0.7444	0.5889	0.6556	0.6556	0.6667	0.6444	0.6345	0.0483
13	0.7222	0.7444	0.7333	0.7667	0.6889	0.7444	0.7556	0.7000	0.7111	0.7333	0.7300	0.0233
14	0.6667	0.7111	0.7000	0.6889	0.7556	0.6222	0.7333	0.6333	0.7000	0.6444	0.6856	0.0413
15	0.7333	0.6444	0.6889	0.7333	0.7111	0.6444	0.7333	0.6556	0.6444	0.6333	0.6822	0.0401
16	0.7000	0.6222	0.6556	0.6778	0.6667	0.6222	0.7111	0.6111	0.6111	0.6444	0.6522	0.0345
17	0.6444	0.6444	0.7556	0.7333	0.6889	0.7667	0.6556	0.6889	0.6222	0.6778	0.6878	0.0470
18	0.6556	0.7222	0.7444	0.7444	0.6667	0.7222	0.7556	0.7444	0.6222	0.6778	0.7056	0.0439
19	0.4778	0.5667	0.6000	0.5222	0.5111	0.5222	0.5222	0.5556	0.4778	0.6000	0.5356	0.0418
20	0.6111	0.4556	0.6556	0.5667	0.5111	0.5111	0.6667	0.5667	0.4556	0.6111	0.5611	0.0724

Tabla A.5: Metodología basada en vector de características con datos funcionales - Exactitud para vocales con SS - ritmos $\delta\theta\alpha$

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.7444	0.7556	0.6889	0.7889	0.7444	0.6667	0.7889	0.7111	0.7667	0.6778	0.7333	0.0425
2	0.6667	0.6222	0.6667	0.6889	0.6889	0.6333	0.7000	0.6222	0.6667	0.6667	0.6622	0.0264
3	0.7444	0.7667	0.7778	0.7556	0.7889	0.7333	0.8111	0.7778	0.6778	0.7889	0.7622	0.0356
4	0.7111	0.6444	0.7667	0.8444	0.7889	0.7111	0.7556	0.7667	0.8111	0.7222	0.7522	0.0545
5	0.8222	0.7778	0.7556	0.7778	0.7556	0.7222	0.8000	0.7222	0.8000	0.7889	0.7722	0.0315
6	0.7444	0.7000	0.6556	0.8222	0.7333	0.7222	0.6667	0.7444	0.6556	0.6667	0.7111	0.0504
7	0.7556	0.6667	0.6778	0.5333	0.6444	0.7000	0.6556	0.7000	0.7333	0.6333	0.6700	0.0584
8	0.7778	0.7667	0.8222	0.7000	0.7778	0.7778	0.7778	0.7556	0.7444	0.7444	0.7645	0.0301
9	0.7222	0.7222	0.8000	0.7556	0.7444	0.7222	0.7889	0.7667	0.7000	0.8000	0.7522	0.0341
10	0.7000	0.7000	0.6778	0.7111	0.7111	0.7000	0.7222	0.7000	0.7111	0.7000	0.7033	0.0112
11	0.6333	0.6444	0.6333	0.6889	0.6778	0.6444	0.6222	0.7111	0.7444	0.6667	0.6667	0.0372
12	0.7444	0.6778	0.6667	0.8222	0.6667	0.7111	0.6778	0.7222	0.6111	0.6556	0.6956	0.0551
13	0.7667	0.7778	0.7889	0.8000	0.8000	0.7333	0.6667	0.7778	0.8556	0.8333	0.7800	0.0496
14	0.6778	0.6556	0.7444	0.7444	0.7333	0.7444	0.6111	0.6444	0.6333	0.7111	0.6900	0.0490
15	0.7667	0.8222	0.7111	0.8000	0.7111	0.7667	0.7667	0.7667	0.7889	0.7333	0.7633	0.0344
16	0.7111	0.7444	0.5889	0.6667	0.6333	0.6667	0.7222	0.6778	0.6444	0.6556	0.6711	0.0434
17	0.7333	0.7333	0.8222	0.7222	0.7333	0.6889	0.7000	0.7889	0.7111	0.7333	0.7367	0.0382
18	0.6667	0.6556	0.7444	0.7667	0.6444	0.6778	0.6778	0.6556	0.6667	0.7556	0.6911	0.0435
19	0.7333	0.7778	0.7444	0.7889	0.8000	0.7444	0.7556	0.7556	0.6889	0.7444	0.7533	0.0297
20	0.6333	0.6556	0.6000	0.7889	0.6889	0.6444	0.7222	0.6111	0.6556	0.5222	0.6522	0.0683

Tabla A.6: Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para sílabas con SS - SVM (1-1)

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.5778	0.5111	0.5556	0.5333	0.5556	0.5333	0.5778	0.5778	0.5333	0.5556	0.5511	0.0218
2	0.6000	0.5778	0.5333	0.5778	0.6000	0.6222	0.5556	0.6444	0.6222	0.5778	0.5911	0.0317
3	0.6000	0.6444	0.6000	0.6444	0.5778	0.6667	0.7111	0.6667	0.6000	0.6889	0.6400	0.0419
4	0.6000	0.7333	0.6444	0.6222	0.6889	0.6667	0.7111	0.7333	0.6444	0.5778	0.6622	0.0514
5	0.7556	0.7111	0.8222	0.7111	0.6889	0.7333	0.7556	0.6667	0.6889	0.7111	0.7245	0.0424
6	0.8667	0.8000	0.8444	0.8000	0.8667	0.8222	0.8667	0.8667	0.8667	0.8222	0.8422	0.0271
7	0.6000	0.6889	0.6222	0.6000	0.6667	0.6444	0.6000	0.6222	0.7111	0.6667	0.6422	0.0378
8	0.6889	0.7778	0.7556	0.8000	0.6444	0.7333	0.6889	0.8000	0.7333	0.7111	0.7333	0.0487
9	0.6222	0.6889	0.7556	0.6000	0.7333	0.6222	0.6667	0.6667	0.7556	0.6222	0.6733	0.0554
10	0.7111	0.5778	0.6444	0.6444	0.5778	0.5778	0.6889	0.6000	0.6444	0.6222	0.6289	0.0445

Tabla A.7: Metodología basada en vector de características con entropía de la información y la dimensión de regularización - Exactitud para sílabas con SS - SVM (1-rest)

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.3947	0.4490	0.4500	0.5610	0.5135	0.4878	0.4906	0.4769	0.4444	0.5294	0.4797	0.0455
2	0.4348	0.4651	0.4407	0.4478	0.4308	0.4242	0.4918	0.5854	0.4677	0.6216	0.4810	0.0647
3	0.6038	0.6250	0.5217	0.5532	0.6944	0.5962	0.5397	0.5000	0.5143	0.6591	0.5807	0.0622
4	0.5846	0.5938	0.6182	0.5517	0.6071	0.5373	0.5077	0.5574	0.5789	0.5714	0.5708	0.0316
5	0.5686	0.6000	0.6122	0.6250	0.6000	0.6800	0.5167	0.6000	0.6885	0.6906	0.6182	0.0528
6	0.7885	0.6964	0.6842	0.8000	0.7544	0.7736	0.8000	0.7547	0.6923	0.7241	0.7468	0.0426
7	0.5625	0.6207	0.5085	0.5313	0.5217	0.5797	0.5902	0.5224	0.6170	0.5224	0.5576	0.0399
8	0.6667	0.7115	0.5857	0.5536	0.6981	0.6538	0.6905	0.6667	0.6739	0.6667	0.6567	0.0470
9	0.5800	0.5254	0.6000	0.6750	0.6111	0.6038	0.5424	0.5857	0.6129	0.5323	0.5869	0.0427
10	0.5400	0.6458	0.6000	0.4545	0.6410	0.5556	0.5556	0.5636	0.5424	0.6857	0.5784	0.0631

Tabla A.8: Metodología basada en vector de características con datos funcionales - Exactitud para sílabas con SS - ritmos δ

Subject	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10	mean	std
1	0.6444	0.7111	0.5778	0.7111	0.6667	0.6889	0.6889	0.5111	0.6222	0.6000	0.6422	0.0616
2	0.6444	0.5556	0.5778	0.6000	0.5333	0.6667	0.5556	0.6667	0.6667	0.6000	0.6067	0.0487
3	0.7111	0.5778	0.6444	0.6667	0.7556	0.6000	0.5556	0.5556	0.7778	0.7778	0.6622	0.0848
4	0.6222	0.6000	0.6000	0.6889	0.7333	0.7333	0.7333	0.7556	0.7333	0.7111	0.6911	0.0575
5	0.6667	0.8222	0.8000	0.7333	0.6667	0.6000	0.8000	0.8444	0.6667	0.6667	0.7267	0.0801
6	0.6889	0.7111	0.6889	0.6444	0.6889	0.5556	0.6889	0.7111	0.6222	0.7778	0.6778	0.0564
7	0.6667	0.5556	0.7778	0.7111	0.6444	0.6222	0.8000	0.6667	0.6222	0.7111	0.6778	0.0704
8	0.6444	0.6222	0.5778	0.7556	0.6444	0.6667	0.6444	0.5333	0.6667	0.6889	0.6444	0.0571
9	0.7111	0.7778	0.7778	0.6889	0.7333	0.5778	0.6889	0.6667	0.5333	0.5556	0.6711	0.0836
10	0.6444	0.7778	0.6889	0.6444	0.6889	0.6222	0.7556	0.7333	0.7778	0.8000	0.7133	0.0608

Bibliografía

- [1] R. S. Snell, *Clinical neuroanatomy*. Lippincott Williams & Wilkins, 2010.
- [2] H. Mitsumoto, *Amyotrophic lateral sclerosis: a guide for patients and families*. Demos Medical Publishing, 2009.
- [3] H. Sontheimer, *Diseases of the Nervous System*. Elsevier Science, 2015.
- [4] D. S. Younger, *Motor disorders*. Lippincott Williams & Wilkins, 2005.
- [5] W. H. Organization, *Neurological disorders: public health challenges*. World Health Organization, 2006.
- [6] B. J. May, *Amputations and prosthetics: a case study approach*. FA Davis Company, 1996.
- [7] T. R. Dillingham, L. E. Pezzin, and E. J. MacKenzie, “Limb amputation and limb deficiency: epidemiology and recent trends in the united states,” *Southern medical journal*, vol. 95, no. 8, pp. 875–884, 2002.
- [8] P. F. Adams, G. E. Hendershot, and M. A. Marano, “Current estimates from the national health interview survey, 1996.,” *Vital and Health Statistics. Series 10, Data from the National Health Survey*, no. 200, pp. 1–203, 1999.
- [9] A. Muzumdar, *Powered Upper Limb Prostheses: Control, Implementation and Clinical Application; 11 Tables*. Springer Science & Business Media, 2004.
- [10] D. G. Shurr, J. W. Michael, and T. M. Cook, *Prosthetics and orthotics*. Prentice Hall, 2002.
- [11] G. R. Muller-Putz, R. Rupp, and G. Pfurtscheller, “Graz brain-computer interface: Control of neuroprostheses for the upper extremity,” in *Applied Sciences on Biomedical and Communication Technologies, 2008. ISABEL’08. First International Symposium on*, pp. 1–2, IEEE, 2008.
- [12] C. Sarmiento, “Prosthetic hand for amputees hand and wrist,” Master’s thesis, National University of Colombia, Bogotá, 2007.
- [13] H. Lee and D. J. Roberson, “A systemic view of an upper extremity prosthesis,” in *System of Systems Engineering, 2007. SoSE’07. IEEE International Conference on*, pp. 1–6, IEEE, 2007.

- [14] B. Rebsamen, E. Burdet, C. Guan, H. Zhang, C. L. Teo, Q. Zeng, M. Ang, and C. Laugier, “A brain-controlled wheelchair based on p300 and path guidance,” in *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on*, pp. 1101–1106, IEEE, 2006.
- [15] C. Mandel, T. Luth, T. Laue, T. Rofer, A. Graser, and B. Krieg-Bruckner, “Navigating a smart wheelchair with a brain-computer interface interpreting steady-state visual evoked potentials,” in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pp. 1118–1125, IEEE, 2009.
- [16] G. R. Muller-Putz and G. Pfurtscheller, “Control of an electrical prosthesis with an ssvep-based bci,” *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 1, pp. 361–364, 2008.
- [17] M. Severens, M. Perusquia-Hernandez, B. Nienhuis, J. Farquhar, and J. Duysens, “Using actual and imagined walking related desynchronization features in a bci,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 23, no. 5, pp. 877–886, 2015.
- [18] F. Duan, D. Lin, W. Li, and Z. Zhang, “Design of a multimodal eeg-based hybrid bci system with visual servo module,” *Autonomous Mental Development, IEEE Transactions on*, vol. 7, no. 4, pp. 332–341, 2015.
- [19] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, “A speedy hybrid bci spelling approach combining p300 and ssvep,” 2014.
- [20] Y. Li, J. Pan, F. Wang, and Z. Yu, “A hybrid bci system combining p300 and ssvep and its application to wheelchair control,” *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 11, pp. 3156–3166, 2013.
- [21] R. Zhang, Y. Li, Y. Yan, H. Zhang, S. Wu, T. Yu, and Z. Gu, “Control of a wheelchair in an indoor environment based on a brain-computer interface and automated navigation,” 2015.
- [22] N. Sharma, V. M. Pomeroy, and J.-C. Baron, “Motor imagery a backdoor to the motor system after stroke?,” *Stroke*, vol. 37, no. 7, pp. 1941–1952, 2006.
- [23] N. Chumerin, N. V. Manyakov, M. van Vliet, A. Robben, A. Combaz, and M. Van Hulle, “Steady-state visual evoked potential-based computer gaming on a consumer-grade eeg device,” *Computational Intelligence and AI in Games, IEEE Transactions on*, vol. 5, no. 2, pp. 100–110, 2013.
- [24] B. Graimann, B. Z. Allison, and G. Pfurtscheller, *Brain-computer interfaces: Revolutionizing human-computer interaction*. Springer, 2010.
- [25] D. S. Tan and A. Nijholt, eds., *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*. London: Springer Verlag, July 2010.
- [26] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2008.
- [27] D. Drubach, E. Benarroch, and F. Mateen, “Imaginación: definición, utilidad y neurobiología,” *Rev Neurol*, vol. 45, no. 6, pp. 353–358, 2007.

- [28] N. Fujimaki, F. Takeuchi, T. Kobayashi, S. Kuriki, and S. Hasuo, “Event-related potentials in silent speech,” *Brain topography*, vol. 6, no. 4, pp. 259–267, 1994.
- [29] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, “Single-trial classification of vowel speech imagery using common spatial patterns,” *Neural Networks*, vol. 22, no. 9, pp. 1334–1339, 2009.
- [30] M. Matsumoto, “Brain computer interface using silent speech for speech assistive device,” 2014.
- [31] K. Brigham and B. V. Kumar, “Imagined speech classification with eeg signals for silent communication: a preliminary investigation into synthetic telepathy,” in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, pp. 1–4, IEEE, 2010.
- [32] S. Deng, R. Srinivasan, T. Lappas, and M. D’Zmura, “Eeg classification of imagined syllable rhythm using hilbert spectrum methods,” *Journal of neural engineering*, vol. 7, no. 4, p. 046006, 2010.
- [33] M. Salama, L. ElSherif, H. Lashin, and T. Gamal, “Recognition of unspoken words using eeg signals,” in *COGNITIVE 2014, The Sixth International Conference on Advanced Cognitive Technologies and Applications*, pp. 51–55, 2014.
- [34] Y. Kazi and S. Sudhir, “Imagined speech eeg signal processing for brain computer interface,” *International Journal of Application or Innovation in Engineering and Management (IJAIEM)*, vol. 3, no. 7, 2014.
- [35] D. L. Cheney and R. M. Seyfarth, “Why animals don’t have language,” *Tanner lectures on human values*, vol. 19, pp. 173–210, 1998.
- [36] M. Rouse, *Neuroanatomy for Speech Language Pathology and Audiology*. Jones & Bartlett Publishers, 2015.
- [37] Y. Grodzinsky, L. P. Shapiro, and D. Swinney, *Language and the brain: Representation and processing*. Academic Press, 2000.
- [38] A. Longstaff, *Neuroscience*. Bios instant notes, Taylor & Francis, 2005.
- [39] *Neuroscience, author=Purves,Dale and Augustine, George and Fitzpatrick, David, year=2004*.
- [40] R. S. Snell, *Neuroanatomía clínica*. Ed. Médica Panamericana, 2007.
- [41] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience*, vol. 2. Lippincott Williams & Wilkins, 2007.
- [42] I. Singh, *Textbook of human neuroanatomy*. JAYPEE BROTHERS PUBLISHERS, 2006.
- [43] S. Standring, *Gray’s anatomy: the anatomical basis of clinical practice*. Elsevier Health Sciences, 2015.

- [44] M. W. Passer and R. E. Smith, “Psychology: The science of mind and behavior,” *New York*, 2009.
- [45] M. Petrides, *Neuroanatomy of language regions of the human brain*. Academic Press, 2013.
- [46] A. A. Sleeper, *Speech and language*. Infobase Publishing, 2006.
- [47] F. Cuetos, “Neurociencia del lenguaje,” *Bases neurológicas e implicaciones clínicas*. Madrid: Médica Panamericana, 2012.
- [48] H. Schnelle, *Language in the Brain*. Cambridge University Press, 2010.
- [49] N. Geschwind, *Disconnexion syndromes in animals and man*. Springer, 1974.
- [50] J. Friedenberg and G. Silverman, *Cognitive science: An introduction to the study of mind*. Sage, 2011.
- [51] D. Poeppel and G. Hickok, “Towards a new functional anatomy of language,” *Cognition*, vol. 92, no. 1, pp. 1–12, 2004.
- [52] M. J. Benedet, *Acercamiento neurolingüístico a las alteraciones del lenguaje: Fundamento teórico de la neurolingüística: Procesamiento normal del lenguaje*. Eos, 2006.
- [53] E. E. Smith and S. M. Kosslyn, *Cognitive Psychology: Pearson New International Edition: Mind and Brain*. Pearson Higher Ed, 2013.
- [54] G. Hickok and D. Poeppel, “The cortical organization of speech processing,” *Nature Reviews Neuroscience*, vol. 8, no. 5, pp. 393–402, 2007.
- [55] D. Poeppel, “Pure word deafness and the bilateral processing of the speech code,” *Cognitive Science*, vol. 25, no. 5, pp. 679–693, 2001.
- [56] G. Hickok and D. Poeppel, “Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language,” *Cognition*, vol. 92, no. 1, pp. 67–99, 2004.
- [57] P. Indefrey and A. Cutler, “Prelexical and lexical processing in listening,” in *The cognitive neurosciences III.*, pp. 759–774, MIT Press, 2004.
- [58] H.-J. Hwang, S. Kim, S. Choi, and C.-H. Im, “Eeg-based brain-computer interfaces: A thorough literature survey,” *International Journal of Human-Computer Interaction*, vol. 29, no. 12, pp. 814–826, 2013.
- [59] B. Stemmer and H. A. Whitaker, *Handbook of the Neuroscience of Language*. Academic Press, 2008.
- [60] C. Mulert and L. Lemieux, *EEG-fMRI: physiological basis, technique, and applications*. Springer Science & Business Media, 2009.
- [61] M. F. Bear, B. W. Connors, and M. A. Paradiso, *NEUROCIENCIA La exploración*. 2008.

- [62] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
- [63] B. J. Fisch and R. Spehlmann, *Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG*. Elsevier Health Sciences, 1999.
- [64] J. M. Stern, *Atlas of EEG patterns*. Lippincott Williams & Wilkins, 2005.
- [65] E. E. Fetz, "Real-time control of a robotic arm by neuronal ensembles," *Nature neuroscience*, vol. 2, pp. 583–584, 1999.
- [66] D. M. Taylor, S. I. H. Tillery, and A. B. Schwartz, "Direct cortical control of 3d neuroprosthetic devices," *Science*, vol. 296, no. 5574, pp. 1829–1832, 2002.
- [67] G. Santhanam, S. I. Ryu, M. Y. Byron, A. Afshar, and K. V. Shenoy, "A high-performance brain–computer interface," *nature*, vol. 442, no. 7099, pp. 195–198, 2006.
- [68] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, "Brain-machine interface: Instant neural control of a movement signal," *Nature*, vol. 416, no. 6877, pp. 141–142, 2002.
- [69] J. K. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex," *Nature neuroscience*, vol. 2, no. 7, pp. 664–670, 1999.
- [70] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. Nicolelis, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, no. 6810, pp. 361–365, 2000.
- [71] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner, "Neuronal population coding of movement direction," *Science*, vol. 233, no. 4771, pp. 1416–1419, 1986.
- [72] A. P. Georgopoulos, G. Pellizzer, A. V. Poliakov, and M. H. Schieber, "Neural coding of finger and wrist movements," *Journal of Computational Neuroscience*, vol. 6, no. 3, pp. 279–288, 1999.
- [73] A. V. Poliakov and M. H. Schieber, "Limited functional grouping of neurons in the motor cortex hand area during individuated finger movements: a cluster analysis," *Journal of Neurophysiology*, vol. 82, no. 6, pp. 3488–3505, 1999.
- [74] M. I. Posner and S. E. Petersen, "The attention system of the human brain," tech. rep., DTIC Document, 1989.
- [75] J. V. Ruiloba, "Introducción a la psicopatología y la psiquiatría," 2000.
- [76] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: assessing the speed of a p300-based brain-computer interface," *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 2, pp. 174–179, 2000.
- [77] P.-J. Kindermans, H. Verschore, and B. Schrauwen, "A unified probabilistic approach to improve spelling in an event-related potential-based brain–computer interface," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 10, pp. 2696–2705, 2013.

- [78] W. Speier, C. Arnold, J. Lu, A. Deshpande, and N. Pouratian, “Integrating language information with a hidden markov model to improve communication rate in the p300 speller,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 22, no. 3, pp. 678–684, 2014.
- [79] T. Zeyl, E. Yin, M. Keightley, and T. Chau, “Adding real-time bayesian ranks to error-related potential scores improves error detection and auto-correction in ap300 speller,” 2015.
- [80] M. E. Thurlings, J. B. Van Erp, A.-M. Brouwer, and P. Werkhoven, “Controlling a tactile erp-bci in a dual task,” *Computational Intelligence and AI in Games, IEEE Transactions on*, vol. 5, no. 2, pp. 129–140, 2013.
- [81] M. Spuler, A. Walter, W. Rosenstiel, and M. Bogdan, “Spatial filtering based on canonical correlation analysis for classification of evoked or event-related potentials in eeg data,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 22, no. 6, pp. 1097–1103, 2014.
- [82] J. Long, Y. Li, T. Yu, and Z. Gu, “Target selection with hybrid feature for bci-based 2-d cursor control,” *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 1, pp. 132–140, 2012.
- [83] Y. Xu and Y. Nakajima, “A two-level predictive event-related potential-based brain-computer interface,” *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 10, pp. 2839–2847, 2013.
- [84] P.-L. Lee, H.-C. Chang, T.-Y. Hsieh, H.-T. Deng, and C.-W. Sun, “A brain-wave-actuated small robot car using ensemble empirical mode decomposition-based approach,” *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 42, no. 5, pp. 1053–1064, 2012.
- [85] K.-K. Shyu, Y.-J. Chiu, P.-L. Lee, J.-M. Liang, and S.-H. Peng, “Adaptive ssvep-based bci system with frequency and pulse duty-cycle stimuli tuning design,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 21, no. 5, pp. 697–703, 2013.
- [86] C.-H. Han, H.-J. Hwang, and C.-H. Im, “Classification of visual stimuli with different spatial patterns for single-frequency, multi-class ssvep bci,” *Electronics Letters*, vol. 49, no. 22, pp. 1374–1376, 2013.
- [87] T. Matsunaga, Y. Katayama, T. Hayami, and K. Iramina, “Measurements of the mu/beta erd and gamma ers during the imagination of body parts movement,” in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 4130–4133, IEEE, 2008.
- [88] A. R. Murguialday, V. Aggarwal, A. Chatterjee, Y. Cho, R. Rasmussen, B. O’Rourke, S. Acharya, and N. V. Thakor, “Brain-computer interface for a prosthetic hand using local machine control and haptic feedback,” in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pp. 609–613, IEEE, 2007.

- [89] J. R. Millan, F. Renkens, J. Mouriño, and W. Gerstner, “Noninvasive brain-actuated control of a mobile robot by human eeg,” *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1026–1033, 2004.
- [90] M. Polak and A. Kostov, “Training setup for control of neural prosthesis using brain-computer interface,” in *[Engineering in Medicine and Biology, 1999. 21st Annual Conference and the 1999 Annual Fall Meeting of the Biomedical Engineering Society] BMES/EMBS Conference, 1999. Proceedings of the First Joint*, vol. 1, pp. 446–vol, IEEE, 1999.
- [91] A. Erfanian and M. Gerivany, “Eeg signals can be used to detect the voluntary hand movements by using an enhanced resource-allocating neural network,” in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 1, pp. 721–724, IEEE, 2001.
- [92] C. W. Chen, M. S. Ju, and C.-C. Lin, “Real time identification of μ wave with wavelet neural networks,” in *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pp. 218–220, IEEE, 2003.
- [93] C.-T. Lin, L.-W. Ko, J.-C. Chiou, J.-R. Duann, R.-S. Huang, S.-F. Liang, T.-W. Chiu, and T.-P. Jung, “Noninvasive neural prostheses using mobile and wireless eeg,” *Proceedings of the IEEE*, vol. 96, no. 7, pp. 1167–1183, 2008.
- [94] E. Curran, P. Sykacek, M. Stokes, S. J. Roberts, W. Penny, I. Johnsrude, and A. M. Owen, “Cognitive tasks for driving a brain-computer interfacing system: a pilot study,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 12, no. 1, pp. 48–54, 2004.
- [95] W. Jia, X. Zhao, H. Liu, X. Gao, S. Gao, and F. Yang, “Classification of single trial eeg during motor imagery based on erd,” in *Engineering in Medicine and Biology Society, 2004. IEMBS’04. 26th Annual International Conference of the IEEE*, vol. 1, pp. 5–8, IEEE, 2004.
- [96] A. Akrami, S. Solhjoo, A. Motie-Nasrabadi, and M.-R. Hashemi-Golpayegani, “Eeg-based mental task classification: linear and nonlinear classification of movement imagery,” *proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual*, pp. 1–4, 2005.
- [97] L. Wang, G. Xu, J. Wang, S. Yang, and W. Yan, “Feature extraction of mental task in bci based on the method of approximate entropy,” in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 1941–1944, IEEE, 2007.
- [98] V. Morash, O. Bai, S. Furlani, P. Lin, and M. Hallett, “Prediction of multiple movement intentions from cnv signal for multi-dimensional bci,” in *Complex Medical Engineering, 2007. CME 2007. IEEE/ICME International Conference on*, pp. 1946–1949, IEEE, 2007.
- [99] M. Li, L. Yang, Y. Zhang, and Y. Luo, “Discrimination of movement imagery eeg based on ar and svm,” in *Automation and Logistics (ICAL), 2011 IEEE International Conference on*, pp. 351–355, IEEE, 2011.

- [100] L. F. Nicolas-Alonso, R. Corralejo, J. Gomez-Pilar, D. Alvarez, and R. Hornero, "Adaptive stacked generalization for multiclass motor imagery-based brain computer interfaces," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 23, no. 4, pp. 702–712, 2015.
- [101] A. S. Aghaei, M. S. Mahanta, and K. N. Plataniotis, "Separable common spatio-spectral patterns for motor imagery bci systems," *Biomedical Engineering, IEEE Transactions on*, vol. 63, no. 1, pp. 15–29, 2016.
- [102] H. Baali, A. Khorshidtalab, M. Mesbah, and M. J. Salami, "A transform-based feature extraction approach for motor imagery tasks classification," *Translational Engineering in Health and Medicine, IEEE Journal of*, vol. 3, pp. 1–8, 2015.
- [103] B. Denby, T. Schultz, K. Honda, T. Hueber, J. M. Gilbert, and J. S. Brumberg, "Silent speech interfaces," *Speech Communication*, vol. 52, no. 4, pp. 270–287, 2010.
- [104] M. Wester and T. Schultz, "Unspoken speech-speech recognition based on electroencephalography," *Master's thesis, Universität Karlsruhe (TH), Karlsruhe, Germany*, 2006.
- [105] J. Hummel, M. Figl, W. Birkfellner, M. R. Bax, R. Shahidi, C. Maurer Jr, and H. Bergmann, "Evaluation of a new electromagnetic tracking system using a standardized assessment protocol," *Physics in medicine and biology*, vol. 51, no. 10, p. N205, 2006.
- [106] M. Fagan, S. Ell, J. Gilbert, E. Sarrazin, and P. Chapman, "Development of a (silent) speech recognition system for patients following laryngectomy," *Medical engineering & physics*, vol. 30, no. 4, pp. 419–425, 2008.
- [107] S. Levinson, *Mathematical models for speech technology*. John Wiley & Sons, 2005.
- [108] T. Hueber, G. Chollet, B. Denby, M. Stone, and L. Zouari, "Ouisper: corpus based synthesis driven by articulatory data," in *International Congress of Phonetic Sciences, Saarbrücken, Germany*, pp. 2193–2196, 2007.
- [109] B. Denby, Y. Oussar, G. Dreyfus, and M. Stone, "Prospects for a silent speech interface using ultrasound imaging.," in *ICASSP (1)*, pp. 365–368, 2006.
- [110] T. Hueber, G. Aversano, G. Chollet, B. Denby, G. Dreyfus, Y. Oussar, P. Roussel, and M. Stone, "Eigentongue feature extraction for an ultrasound-based silent speech interface," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, pp. I–1245, IEEE, 2007.
- [111] T. Hueber, E.-L. Benaroya, G. Chollet, B. Denby, G. Dreyfus, and M. Stone, "Development of a silent speech interface driven by ultrasound and optical images of the tongue and lips," *Speech Communication*, vol. 52, no. 4, pp. 288–300, 2010.
- [112] R. Merletti and P. A. Parker, *Electromyography: physiology, engineering, and non-invasive applications*, vol. 11. John Wiley & Sons, 2004.
- [113] G. Kamen and D. Gabriel, *Essentials of electromyography*. Human Kinetics, 2010.

- [114] N. Sugie and K. Tsunoda, “A speech prosthesis employing a speech synthesizer-vowel discrimination from perioral muscle activities and vowel production,” *Biomedical Engineering, IEEE Transactions on*, no. 7, pp. 485–490, 1985.
- [115] M. Morse, Y. Gopalan, and M. Wright, “Speech recognition using myoelectric signals with neural networks,” in *Engineering in Medicine and Biology Society, 1991. Vol. 13: 1991., Proceedings of the Annual International Conference of the IEEE*, pp. 1877–1878, IEEE, 1991.
- [116] C. Jorgensen, D. D. Lee, and S. Agabont, “Sub auditory speech recognition based on emg signals,” in *Neural Networks, 2003. Proceedings of the International Joint Conference on*, vol. 4, pp. 3128–3133, IEEE, 2003.
- [117] L. Maier-Hein, F. Metze, T. Schultz, and A. Waibel, “Session independent non-audible speech recognition using surface electromyography,” in *Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on*, pp. 331–336, IEEE, 2005.
- [118] T. Schultz and M. Wand, “Modeling coarticulation in emg-based continuous speech recognition,” *Speech Communication*, vol. 52, no. 4, pp. 341–353, 2010.
- [119] X. Pei, J. Hill, and G. Schalk, “Silent communication: toward using brain signals,” *Pulse, IEEE*, vol. 3, no. 1, pp. 43–46, 2012.
- [120] X. Pei, D. L. Barbour, E. C. Leuthardt, and G. Schalk, “Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans,” *Journal of neural engineering*, vol. 8, no. 4, p. 046028, 2011.
- [121] X. Pei, E. C. Leuthardt, C. M. Gaona, P. Brunner, J. R. Wolpaw, and G. Schalk, “Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition,” *Neuroimage*, vol. 54, no. 4, pp. 2960–2972, 2011.
- [122] N. Yoshimura, A. Satsuma, C. S. DaSalla, T. Hanakawa, M.-a. Sato, and Y. Koike, “Usability of eeg cortical currents in classification of vowel speech imagery,” in *Virtual Rehabilitation (ICVR), 2011 International Conference on*, pp. 1–2, IEEE, 2011.
- [123] A. Riaz, S. Akhtar, S. Iftikhar, A. A. Khan, and A. Salman, “Inter comparison of classification techniques for vowel speech imagery using eeg sensors,” in *Systems and Informatics (ICSAI), 2014 2nd International Conference on*, pp. 712–717, IEEE, 2014.
- [124] L. Sarmiento, P. Lorenzana, C. Cortes, W. Arcos, J. Bacca, and A. Tovar, “Brain computer interface (bci) with eeg signals for automatic vowel recognition based on articulation mode,” in *Biosignals and Birobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), 5th ISSNIP-IEEE*, pp. 1–4, IEEE, 2014.
- [125] R. Kamalakkannan, R. Rajkumar, R. Madan, W. Arcos, and D. Shenbaga, “Imagined speech classification using eeg,” *Advances in Biomedical Science and Engineering*, vol. 1, no. 2, pp. 20–32, 2014.
- [126] M. Matsumoto and J. Hori, “Classification of silent speech using adaptive collection,” in *Computational Intelligence in Rehabilitation and Assistive Technologies (CIRAT), 2013 IEEE Symposium on*, pp. 5–12, IEEE, 2013.

- [127] M. Matsumoto and J. Hori, “Classification of silent speech using support vector machine and relevance vector machine,” *Applied Soft Computing*, vol. 20, pp. 95–102, 2014.
- [128] M. D’Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, “Toward eeg sensing of imagined speech,” in *Human-Computer Interaction. New Trends*, pp. 40–48, Springer, 2009.
- [129] L. Wang, X. Zhang, and Y. Zhang, “Extending motor imagery by speech imagery for brain-computer interface,” in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pp. 7056–7059, IEEE, 2013.
- [130] L. Wang, X. Zhang, X. Zhong, and Y. Zhang, “Analysis and classification of speech imagery eeg for bci,” *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 901–908, 2013.
- [131] A. Porbadnigk, M. Wester, and T. S. Jan-p Calliess, “Eeg-based speech recognition impact of temporal effects,” 2009.
- [132] A. A. T. García, C. A. R. García, and L. V. Pineda, “Toward a silent speech interface based on unspoken speech.,” in *BIOSIGNALS*, pp. 370–373, 2012.
- [133] P. Suppes, Z.-L. Lu, and B. Han, “Brain wave recognition of words,” *Proceedings of the National Academy of Sciences*, vol. 94, no. 26, pp. 14965–14969, 1997.
- [134] A. Porbadnigk and T. Schultz, “Eeg-based speech recognition: Impact of experimental design on performance,” *Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe (TH), Karlsruhe, Germany*, 2008.
- [135] P. Suppes, B. Han, and Z.-L. Lu, “Brain-wave recognition of sentences,” *Proceedings of the National Academy of Sciences*, vol. 95, no. 26, pp. 15861–15866, 1998.
- [136] S. Deng, R. Srinivasan, and M. D’Zmura, “Cortical signatures of heard and imagined speech envelopes,” tech. rep., DTIC Document, 2013.
- [137] X. Chia, J. B. Hagedorna, D. Schoonovera, and M. D’Zmuraa, “Eeg-based discrimination of imagined speech phonemes,” *International Journal of Bioelectromagnetism*, vol. 13, no. 4, pp. 201–206, 2011.
- [138] M. Borda, *Fundamentals in information theory and coding*. Springer Science & Business Media, 2011.
- [139] M. Burgin, *Theory of information: fundamentality, diversity and unification*, vol. 1. World Scientific, 2010.
- [140] R. M. Gray, *Entropy and information theory*. Springer Science & Business Media, 2011.
- [141] B. A. Mathew, “Entropy of electroencephalogram (eeg) signals changes with sleep state,” 2006.
- [142] A. Shiriyayev, *Selected works of AN Kolmogorov*. Kluwer, 1993.

- [143] K. Falconer, *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.
- [144] G. Rubiano, “Iteración y fractales (con mathematica),” 2009.
- [145] H. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*. Springer New York, 2013.
- [146] G. Arenas and S. Sabogal, “Una introducción a la geometría fractal,” *Ediciones Universidad Industrial de Santander, Bucaramanga*, 2011.
- [147] M. Al-Akaidi, *Fractal speech processing*. Cambridge university press, 2004.
- [148] F. Roueff and J. L. Véhel, “A regularization approach to fractional dimension estimation,” *Proc. of Fractals*, vol. 98, 1998.
- [149] Z. Feng, M. J. Zuo, and F. Chu, “Application of regularization dimension to gear damage assessment,” *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1081–1098, 2010.
- [150] B. Silverman and J. Ramsay, *Functional Data Analysis*. Springer, 2005.
- [151] F. Ferraty and P. Vieu, *Nonparametric functional data analysis: theory and practice*. Springer Science & Business Media, 2006.
- [152] L. Díaz and M. Morales, “Análisis estadístico de datos multivariados,” *Bogotá: Universidad Nacional de Colombia*, 2012.
- [153] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*. Springer Science & Business Media, 2013.
- [154] O. Christensen, *Functions, spaces, and expansions: mathematical tools in physics and engineering*. Springer Science & Business Media, 2010.
- [155] S. Graves, G. Hooker, and J. Ramsay, “Functional data analysis with r and matlab,” 2009.
- [156] H. Prautzsch, W. Boehm, and M. Paluszny, *Bézier and B-spline techniques*. Springer Science & Business Media, 2013.
- [157] T. Hsing and R. Eubank, *Theoretical foundations of functional data analysis, with an introduction to linear operators*. John Wiley & Sons, 2015.
- [158] L. Horváth and P. Kokoszka, *Inference for functional data with applications*, vol. 200. Springer Science & Business Media, 2012.
- [159] R. R. A. Española, *Ortografía de la lengua española*. Espasa, 2010.
- [160] V. Ingle and J. Proakis, *Digital signal processing using MATLAB*. Cengage Learning, 2011.
- [161] S. Tong and N. V. Thakor, *Quantitative EEG analysis methods and clinical applications*. Artech House, 2009.

- [162] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelets and their Applications*. John Wiley & Sons, 2013.
- [163] R. Cohen, “Signal denoising using wavelets,” *Project Report, Department of Electrical Engineering Technion, Israel Institute of Technology, Haifa*, 2012.
- [164] A. Antoniadis, J. Bigot, and T. Sapatinas, “Wavelet estimators in nonparametric regression: a comparative simulation study,” *Journal of Statistical Software*, vol. 6, pp. pp–1, 2001.
- [165] C. Ciro Martínez Bencardino, *Estadística y muestreo (incluye CD)*. Serie textos universitarios: Area matemáticas, Ecoe Ediciones, 2005.
- [166] P. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Pearson Addison Wesley, 2006.
- [167] C.-W. Hsu and C.-J. Lin, “A comparison of methods for multiclass support vector machines,” *Neural Networks, IEEE Transactions on*, vol. 13, no. 2, pp. 415–425, 2002.
- [168] Z. Gu, Z. Yu, Z. Shen, and Y. Li, “An online semi-supervised brain–computer interface,” *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 9, pp. 2614–2623, 2013.
- [169] D.-G. Chen, H.-Y. Wang, and E. C. Tsang, “Generalized mercer theorem and its application to feature space related to indefinite kernels,” in *Machine Learning and Cybernetics, 2008 International Conference on*, vol. 2, pp. 774–777, IEEE, 2008.
- [170] A. A. Ghanbari, A. Broum, H. Navidi, and A. Ahmadi, “Brain computer interface with genetic algorithm,” *International Journal of Information*, vol. 2, no. 1, 2012.
- [171] S. Sivanandam and S. Deepa, *Introduction to genetic algorithms*. Springer Science & Business Media, 2007.
- [172] J. Proakis and D. Manolakis, *Digital Signal Processing: principles algorithms and applications*. Prentice Hall, 2006.
- [173] Y. Li and K. M. Wong, “Riemannian distances for signal classification by power spectral density,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 4, pp. 655–669, 2013.
- [174] D.-W. Sun, *Infrared spectroscopy for food quality analysis and control*. Academic Press, 2009.
- [175] M. Pal, “Multiclass approaches for support vector machine based land cover classification,” *arXiv preprint arXiv:0802.2411*, 2008.
- [176] M. Almeida, *Contribuciones al estudio de la lingüística hispánica: homenaje al profesor Ramón Trujillo*, vol. 2. Editorial Montesinos, 1997.
- [177] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, “A blind source separation technique using second-order statistics,” *Signal Processing, IEEE Transactions on*, vol. 45, no. 2, pp. 434–444, 1997.

- [178] K. Abed-Meraim, Y. Xiang, J. H. Manton, and Y. Hua, “Blind source separation using second-order cyclostationary statistics,” *IEEE Transactions on Signal Processing*, vol. 49, no. 4, 2001.
- [179] Y. Tran, R. A. Thiraisingham, A. Craig, and H. Nguyen, “Evaluating the efficacy of an automated procedure for eeg artifact removal,” in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 376–379, IEEE, 2009.
- [180] N. Mammone, F. L. Foresta, and F. C. Morabito, “Automatic artifact rejection from multichannel scalp eeg by wavelet ica,” *Sensors Journal, IEEE*, vol. 12, no. 3, pp. 533–542, 2012.
- [181] R. C. Hilborn, *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. Oxford University Press on Demand, 2000.
- [182] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for eeg-based brain–computer interfaces,” *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.
- [183] F. Lee, R. Scherer, R. Leeb, C. Neuper, H. Bischof, and G. Pfurtscheller, “A comparative analysis of multi-class eeg classification for brain computer interface,” in *Proceedings of the 10th Computer Vision Winter Workshop*, pp. 195–204, 2005.