



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**Maestría en Ciencias (Neurobiología)
Instituto de Neurobiología**

“Aplicación de la electroencefalografía para el estudio de diferencias individuales cuantitativas en la percepción de estímulos acústicos.”

TESIS QUE PARA OPTAR POR EL GRADO DE: MAESTRO EN CIENCIAS

PRESENTA:

Ricardo Alonso Echavarría Solana

TUTORA PRINCIPAL:

**Dra. M. Florencia Assaneo
Instituto de Neurobiología**

Juriquilla, Querétaro, a 09 de Agosto del 2023



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas

Tesis Digitales

Restricciones de uso

DERECHOS RESERVADOS ©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

El presente proyecto de maestría se llevó a cabo en el laboratorio de Percepción y Producción de Habla (C-14) del Instituto de Neurobiología de la Universidad Nacional Autónoma de México, en Juriquilla, Querétaro, con la dirección de la Dra. M. Florencia Assaneo, titular del laboratorio.

Además, fue realizado con la guía del comité tutorial integrado por:

- Dr. Victor De Lafuente Flores (INB, UNAM).
- Dra. Fabiola Reveca Gómez Velázquez (CUCBA, UG).

El presente proyecto de maestría fue financiado por el programa PAPIIT-2021 IA200223 y por la beca otorgada por el CONACYT (CVU: 1135904).

Agradecimientos

A la doctora Flor, por su conocimiento y diligencia, pero también por su paciencia y humanidad, y por mostrarnos en el laboratorio cómo hacer las cosas con pasión.

A todos los integrantes pasados y presentes del Hablab, por su compañerismo y su apoyo a lo largo de esta maestría. Les agradezco por haber compartido tiempo y espacio en un laboratorio tan único, empezando por Cecy, quien me animó a dar mi primer paso para entrar al laboratorio y por la amistad que me dejó en adelante, a Fer por su presteza para orientarnos a todos en muchas situaciones y muy a su manera ser el hermano mayor, a Ireri, Moramay y Montserrat por todas las ideas, las risas y momentos muy preciados dentro y fuera del laboratorio, y sobre todo a Lily por ser mi compañera y mejor amiga en el delito de cursar esta maestría y en mil cosas más, y sobre todo por hacerme sentir en casa.

A mis compañeros de generación, por todos los momentos de desvelo, estudio, diversión y por las amistades que surgieron, y que espero duren muchísimo. A Oscar Fer, July, Chuy, César, Darel y todos los integrantes de esta generación, las tkm a todas y todos. Y en especial a mis roomies favoritos, Waldo, Andrés y Sara, porque han sido mi familia sustituta por 2 años, y no cabe duda de que juntos hemos crecido y aprendido mucho.

A mis tres hermanas por ser, cada una a su manera, la luz de mi vida, mis ejemplos a seguir, confidentes, y compañeras de vida desde siempre y para siempre. Kishi por ser la persona increíble, graciosa y ocurrente que es, Abril por ser la hermana mayor y por tener el don de llevar la fiesta a todas partes donde va, Piti por su personalidad única y su inteligencia ¡Las amo!

Y por sobre todas las cosas, a mi mamá y mi papá, porque cualquier logro mío es su logro y es gracias a todo el amor, todas las batallas y las lágrimas. Por que tuve la suerte de tenerlos guiándome, alentándome, siendo pacientes, volviendo a intentar. Los amo como no tienen idea.

ÍNDICE

Introducción	1
Antecedentes	4
Hipótesis	6
Objetivos	6
Materiales y Métodos	7
Participantes	7
Estímulos	7
Prueba de Sincronización Espontánea del Habla ...	8
Paradigma Oddball Musical	9
Adquisición de data Electroencefalográfica	10
Análisis de datos	11
Preprocesamiento	11
Cálculo de ERPs	11
Cálculo del componente MMN	11
Aprendizaje Automático	12
Representación de Topografías	13
Análisis de Distancias Euclidianas	13
Análisis Estadísticos	13
Resultados	14
Discusión	20
Conclusiones	22
Referencias	23
Anexos	27

Introducción

Sincronización auditivo-motora del habla

Varias habilidades y experiencias de nuestro día a día, como el habla, la música, el baile y algunas actividades atléticas, implican la capacidad de coordinación temporal entre la ejecución motora de una conducta y un estímulo. A este fenómeno se le conoce como sincronización sensoriomotora, y es observable en muy pocas especies animales, incluyendo el ser humano (Repp, 2005). Particularmente, la capacidad de sincronización a eventos auditivos, conocida como sincronización auditivo-motora, es posible en el humano para un amplio rango de estímulos, ritmos y efectores (ver, por ejemplo, Repp, 2013). Se sabe que la capacidad de sincronización sensoriomotora en el humano es más eficiente en exactitud y consistencia para la modalidad auditiva que para la táctil y la visual (Whitton & Jiang, 2023), y se ha observado que los tiempos de reacción ante estímulos auditivos son menores que para estímulos en otras modalidades, lo que refleja tiempos más rápidos de conducción a la corteza motora y de procesamiento en la corteza auditiva (Shelton & Kumar, 2010); estas observaciones apuntan a la naturaleza temporal de los estímulos auditivos con los que el ser humano interactúa, y que son de importancia sobre todo en el contexto de la percepción y producción del habla y la música.

Históricamente, el estudio de la sincronización auditivo-motora se ha abordado desde enfoques conductuales y neuro-funcionales. Descubrimientos en esta última línea han permitido describir correlatos neuronales para la sincronización auditivo-motora, sugiriendo la existencia de mecanismos propios encargados de establecerla y adaptarla a un entorno cambiante. Resulta interesante mencionar, por ejemplo, que en el ser humano los estímulos auditivos rítmicos como el habla (Doelling, 2014), o la música (Doelling & Poppel, 2015) inducen actividad cerebral sincrónica en la corteza auditiva, lo que parece ser crítico para la percepción auditiva de los propios estímulos. A su vez, se ha demostrado que la actividad de la corteza motora durante la sincronización auditivo-motora es sincrónica a la del estímulo auditivo percibido (Thaut, 2015), lo que implica que las áreas motoras y auditivas corticales operan de forma coordinada durante este tipo de conductas.

Por otro lado, la sincronización auditivo-motora ha sido extensamente estudiada mediante tareas conductuales de *finger tapping*, que en su forma clásica consisten en el registro de los golpeteos de los dedos contra una superficie por parte del participante frente a la presentación de un estímulo auditivo sencillo, como un metrónomo (Repp, 2005). Sin embargo, este tipo de aproximaciones, aunque muy valiosas a la hora de explorar y caracterizar la sincronización auditivo-motora, no han permitido describir apropiadamente la sincronización auditivo-motora del habla

(SAMH), que hace uso de estímulos y efectores más complejos. De forma interesante, se ha sugerido que la capacidad de sincronización durante el habla podría depender del acoplamiento entre las cortezas auditiva y motora mediante un mecanismo de oscilación neural (Assaneo et al., 2018), lo que abre la posibilidad de que la precisión y consistencia de la SAMH en la población general estén sujetas a variación individual en función de diferencias en estos u otros mecanismos neurales. En este contexto, se ha desarrollado un experimento conductual sencillo que hace posible cuantificar la capacidad de sincronización espontánea del habla de un participante frente a un tren de sílabas de ritmo conocido (Assaneo et al., 2019). Este experimento ha descubierto que, efectivamente, la SAMH no es homogénea en la población general, y que estas diferencias cualitativas en la capacidad de sincronización vienen acompañadas de diferencias estructurales, funcionales y cognitivas (Assaneo et al., 2019). Particularmente, se ha observado que la mayor capacidad para la SAMH correlaciona fuertemente con un mayor volumen total del fascículo arqueado izquierdo (LAF), un tracto de materia blanca que conecta a áreas temporales y frontales de forma directa o indirecta mediante tres segmentos: (i) un segmento largo de conexión directa temporal-frontal, (ii) un segmento posterior, de conexión indirecta temporal-parietal, y (iii) un segmento anterior, de conexión indirecta parietal-frontal (Bernard et al., 2019). Además de estas observaciones estructurales, Assaneo et al. (2019) y otros estudios han desenmascarado diferencias funcionales y cognitivas importantes entre estos grupos con distinta precisión en la sincronización auditivo-motora del habla (Assaneo et al., 2020; Kern et al., 2021; Assaneo et al., 2022; Orpella et al., 2022).

La observación de que la capacidad para la SAMH es más compleja que lo previamente pensado, y que su variación se asocia a correlatos estructurales, funcionales y cognitivos, invita a explorar desde otro enfoque lo que se conoce sobre la capacidad humana para percibir y producir ritmos. Particularmente, resulta necesario cuestionar cómo esta variabilidad en la capacidad de sincronización específica al habla podría interactuar o asociarse con fenómenos más basales o pre-atencionales de la percepción del ritmo, como lo es la detección automática de cambios en patrones rítmicos, un fenómeno cognitivo muy investigado y cuya variación correlaciona con condiciones clínicas como las dislexias, pero que hasta ahora no han sido abordados considerando dicha variabilidad individual.

El potencial de disparidad

En el contexto de la electroencefalografía (EEG), el potencial de disparidad o *mismatch negativity* (MMN), es un componente de los potenciales relacionados a eventos (ERP) ampliamente descrito en la literatura (ver, por ejemplo, Näätänen et

al., 2007). Concretamente, el componente MMN es una señal cerebral de origen frontocentral y voltaje negativo, que se produce durante la detección de alteraciones en estímulos acústicos regulares presentados durante un paradigma *oddball* (Näätänen et al., 1978; Wacongne et al., 2012). En su forma clásica, dicho paradigma consiste en la presentación secuencial y aleatorizada de un estímulo estándar con alta probabilidad de ocurrencia frente a otro estímulo atípico (el *oddball*) de menor probabilidad. En tareas *oddball* auditivas de tipo pasivo (es decir, en las que el sujeto no requiere realizar ninguna acción consciente) (Kropotov, 2009), el MMN emerge como una respuesta cerebral pre-atencional provocada por cambios en características de primer orden como la frecuencia e intensidad del sonido (Schröger, 1998), al igual que por cambios en características más complejas, como alteraciones de frecuencia en secuencias de tonos descendentes (Tervaniemi et al., 1994) y violaciones gramaticales (Pulvermüller & Shtyrov, 2003). Además, cambios temporales inesperados (por ejemplo, retrasar la presentación de un tono en una secuencia periódica) también suscitan la respuesta MMN (Ford & Hillyard, 1981; Näätänen, et al., 1993; Kisley, et al., 2004). Actualmente, se considera que los sustratos neurales generadores del componente MMN se encuentran en tres regiones corticales: (i) las regiones auditivas primaria y secundaria del giro temporal superior, (ii) los giros frontales inferior y medio junto con áreas prefrontales, y (iii) regiones parietales inferiores (Vaquero et al., 2021).

Además de su papel como un marcador electrofisiológico de importancia para las ciencias cognitivas en el estudio de procesos de discriminación auditiva (Näätänen et al., 1978), el componente MMN cuenta con aplicaciones clínicas prometedoras. Estudios en niños han encontrado que diferencias en la amplitud y latencia del componente MMN en paradigmas *oddball* de variación de frecuencia (fMMN) se asocian al futuro desarrollo de dislexia en grupos de riesgo (Lohvansuu et al., 2021). También, el componente MMN frente a variaciones de duración (dMMN) está propuesto como un biomarcador robusto para la esquizofrenia, que puede reflejar la progresión de la patología e incluso ser un predictor de la psicosis (Nagai et al., 2013), y la amplitud de la señal es un buen predictor del pronóstico de recuperación en pacientes en coma.

La generación de la señal MMN también ha sido explorada mediante la presentación de estímulos musicales, y se ha demostrado que este componente se produce como consecuencia de la violación de reglas musicales implícitas como alteraciones de timbre, melodía, clave o ritmo, y que la intensidad de la respuesta (en amplitud) correlaciona con el grado de musicalidad de los participantes, con los músicos mostrando respuestas significativamente mayores (ver, por ejemplo, Putkinen et al., 2014; Vuust et al., 2016). De manera relevante para este trabajo, se ha demostrado también que la amplitud de la señal MMN frente a modulaciones de ritmo en estímulos musicales correlaciona positivamente

con el volumen del segmento largo del fascículo arqueado izquierdo en participantes sin educación musical (Vaquero et al., 2021), lo que permite asociar la variación en esta respuesta neurofisiológica a diferencias estructurales presentes en la población general.

Antecedentes.

Prueba de sincronización espontánea del habla (SSS-Test).

La prueba de sincronización espontánea del habla (SSS-Test) es un protocolo conductual que permite cuantificar el grado de alineamiento entre la tasa silábica producida por un participante y la frecuencia de un tren rítmico de sílabas percibido (Assaneo et al., 2019). Durante esta tarea, el participante recibe la instrucción de producir la sílaba “tah” de forma continua mientras es presentado concurrentemente, por medio de audífonos, con un tren rítmico de sílabas. La señal acústica producida por el participante es registrada, y las envolventes de dicha señal y del estímulo son generadas para calcular el valor de acoplamiento de fase (PLV) promedio entre ambos. Esta medida cuantifica en qué grado un participante modifica su ritmo silábico para acoplarlo al ritmo percibido.

Actualmente, existen dos versiones distintas de la prueba: (i) La *versión fija implícita*, en la que el estímulo externo tiene una tasa fija de 4.5 Hz, y en la que el participante no es instruido de forma explícita para acoplar su ritmo de producción; (ii) La *versión explícita acelerada*, en la que el estímulo silábico varía en su frecuencia de 4.3 a 4.7 Hz en incrementos de 0.1 Hz cada 10 segundos, y en la que el sujeto recibe la instrucción de sincronizarse (Lizcano-Cortés et al., 2022).

Estatus de sincronización auditivo-motora del habla: Sincronizadores altos y bajos. La aplicación del SSS-Test a un grupo de participantes en el estudio original de Assaneo et al. (2019) permitió descubrir una distribución bimodal en las mediciones de sincronización, distinguiéndose sujetos con una alta capacidad de sincronizar su producción a un ritmo externo (los sincronizadores altos, o *highs*), y sujetos que no presentan esta capacidad (los sincronizadores bajos, o *lows*). Estas diferencias cualitativas en la capacidad de sincronización permanecieron estables en el tiempo y permitieron categorizar a los sujetos en dos grupos distintos. Las características demográficas como edad y sexo no fueron significativamente distintas entre ambos grupos con excepción de la musicalidad, que correlacionó positivamente con la capacidad de sincronización (Assaneo et al., 2019). Aunado a esto, se describieron diferencias fundamentales entre ambos grupos en su desempeño en tareas cognitivas y en características cerebrales estructurales y funcionales. Concretamente, los sujetos de alta sincronía muestran un mejor desempeño en una tarea de aprendizaje de palabras (Assaneo et al.,

2019). Además, el uso de técnicas de resonancia magnética pesada con tensor de difusión (DW-MRI) permitió descubrir diferencias estructurales entre ambos grupos de sincronización: los sujetos de alta sincronía presentan mayor volumen en el fascículo arqueado izquierda (principal vía de materia blanca conectando regiones auditivas y frontales) (Assaneo et al., 2019). Por último, el estatus de sincronía también arrojó diferencias neuro-funcionales medidas mediante un experimento de magnetoencefalografía (MEG), en el que los participantes escucharon de forma pasiva un estímulo rítmico de sílabas mientras se realizaba un registro electrofisiológico. Este experimento demostró que la precisión de la sincronización auditivo-motora del habla se asocia a un aumento en la sincronía entre el estímulo y la actividad neural en los giros frontales inferior y medio izquierdos, correspondientes a las áreas de Brodmann 44, 45, 46 y 9 izquierdas (Assaneo et al., 2019).

A estas observaciones iniciales, se han sumado evidencias de otros estudios que han continuado la exploración de las diferencias cognitivas entre sincronizadores altos y bajos. Por ejemplo, se ha descubierto que estos grupos difieren en sus estrategias a la hora de distinguir estímulos silábicos en condiciones de ruido, con los sincronizadores de alta mostrando mejoras perceptuales facilitadas por la producción de sílabas a un ritmo preferido, sugiriendo que el acoplamiento auditivo-motor necesario para la percepción es oscilatorio en su naturaleza (Assaneo et al., 2020). Así mismo, se han observado diferencias perceptuales en la sensibilidad temporal entre sincronizadores de alta y baja, con el desempeño de los primeros cubriendo un rango de frecuencias relativamente más amplio (Kern et al., 2021) dentro del rango theta. También, se han descrito diferencias en el desempeño en tareas de aprendizaje estadístico, donde los sincronizadores de alta muestran ventajas asociadas al reclutamiento de una vía frontoparietal adicional en condiciones de escucha pasiva, y que es inhibida sólo para este grupo por supresión articulatoria (Orpella et al., 2022). Por otro lado, el estatus de sincronía también se ha asociado a diferencias en la severidad auto-reportada de tartamudeo, con los sincronizadores de baja percibiéndose a sí mismos con un tartamudeo más severo (Assaneo et al., 2022).

Paradigma oddball musical, el potencial de disparidad y el fascículo arqueado.

Para explorar la capacidad de detección automática de cambios en estímulos musicales, han sido desarrollados paradigmas *oddball* musicales variados en múltiples características. Particularmente, Putkinen et al. (2014) han desarrollado un paradigma *oddball* musical consistente en la presentación secuencial de melodías que presentan uno de seis posibles tipos de cambios ocasionales en características musicales de bajo nivel (desentonación, error de ritmo, cambio de timbre) o alto nivel (modulación de ritmo, modulación de melodía, cambio de clave)

en una de cuatro posibles notas cada vez. La aplicación de este paradigma musical en el estudio original encontró respuestas MMN significativas para todos los tipos de cambios, y, junto con otros estudios previos, ha permitido demostrar que el componente MMN frente a cambios en características musicales difiere entre músicos y no músicos, con los primeros mostrando una mayor respuesta (Vuust et al., 2012; Putkinen et al., 2014). Recientemente, mediante la aplicación de este mismo paradigma con un grupo de sujetos no músicos, Vaquero et al. (2021) encontraron que un mayor volumen en el segmento largo del LAF se asocia a componentes MMN de mayor amplitud sólo durante cambios en la modulación del ritmo.

Estas observaciones son sugerentes, y plantean la posibilidad de que la capacidad de detección automática de cambios en patrones rítmicos, cuantificada mediante la amplitud de la respuesta MMN, interactúe de forma relevante con otros fenómenos conductuales que dependen de la percepción de ritmos, como la sincronización auditivo-motora del habla, una propiedad que también se correlaciona con la conectividad estructural del fascículo arqueado izquierdo. En el presente trabajo, exploramos estas posibles diferencias en la amplitud de la respuesta MMN en función del grupo de sincronización auditivo-motora del habla presentes en la población general mediante el registro electroencefalográfico durante un paradigma oddball musical.

Hipótesis.

Los sincronizadores altos tendrán diferencias en la distinción de estímulos atípicos y estándares respecto de los sincronizadores bajos frente a cambios inesperados en la modulación del ritmo, pero no en los cambios de melodía, presentados mediante un paradigma oddball musical.

Objetivos.

Objetivo general

Determinar si existen diferencias individuales en la detección automática de cambios en patrones rítmicos en la población general en función del grupo de sincronización auditivo-motora del habla.

Objetivos específicos

1. Rediseño del paradigma oddball musical introducido por Putkinen y colaboradores (Putkinen et al., 2014) para evaluar la respuesta cerebral a cambios en la modulación del ritmo y cambios de melodía en un estímulo

musical, para establecer si existen diferencias en las respuestas de MMN entre sujetos de alta y baja sincronía.

2. Uso de estrategias clásicas (Análisis de ERPs) y de aprendizaje automático para la prueba de hipótesis experimental y el hallazgo de posibles diferencias entre grupos de sincronización auditivo-motora del habla.

Materiales y Métodos.

Participantes.

Todos los procedimientos experimentales se llevaron a cabo en dos grupos de sujetos: Sincronizadores altos ($n=19$, 12 hombres, edad promedio: 25.9 ± 3.75 años), y Sincronizadores bajos ($n=20$, 10 hombres, edad promedio: 27.21 ± 4.45 años). Los participantes fueron seleccionados a partir de una base de datos existente en el laboratorio, o fueron reclutados. Se excluyó a los voluntarios que reportaran entrenamiento musical formal mayor a 1 año. Los participantes firmaron un consentimiento informado y recibieron una gratificación económica por su participación. Todos los procedimientos fueron evaluados y aprobados por el comité de ética del Instituto de Neurobiología de la Universidad Nacional Autónoma de México (protocolo 096.H).

Estímulos.

Los estímulos fueron adaptados del paradigma musical descrito en Putkinen et al. (2014), y consistieron en melodías monofónicas con un compás de tres tiempos ($\frac{3}{4}$) de duración a un ritmo de 90 pulsaciones por minuto, sintetizados con la biblioteca PyPiano 0.1.0.

Cada melodía, que denominaremos también compás, inicia con una tríada de Do mayor invertida (E4-G4-C5) de 300 ms de duración, seguida de 50 ms de silencio. Enseguida, se presentan de forma consecutiva cuatro notas en la porción intermedia del compás, o intertonos, de los cuales dos son intertonos de corta duración (125 ms, $\frac{1}{4}$ de tiempo de compás), y los otros dos son de larga duración (300 ms, $\frac{1}{2}$ de tiempo de compás). Cada uno de estos intertonos puede ser un Re (D4), Mi (E4), Fa (F4) o Sol (G4) central, y cada intertono es seguido por un silencio de 50 ms. El compás concluye con un Do central (C4) de 575 ms de duración (aproximadamente 1 tiempo de compás) seguido de 125 ms de silencio, sumando una duración total de 2.1 segundos por compás.

La modulación de ritmo fue provocada al intercambiar la duración de dos intertonos contiguos de diferente duración, respecto del compás inmediato anterior (**Fig. 1**, Odd R). Por su parte, la modulación de melodía fue provocada al variar aleatoriamente la nota de uno de los intertonos largos por alguna otra de las notas permitidas, respecto del compás anterior (**Fig. 1**, Odd M). Sólo un tipo de cambio se realizó a la vez sobre un mismo compás, y el nuevo compás se repitió al menos dos veces, y en promedio tres veces para ambos casos, antes de la próxima modificación.

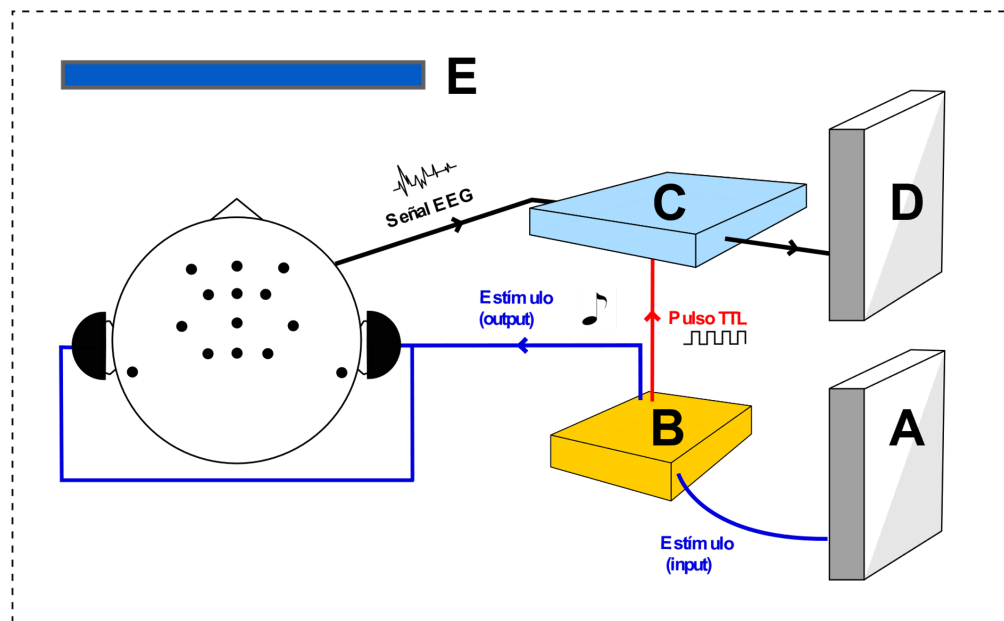
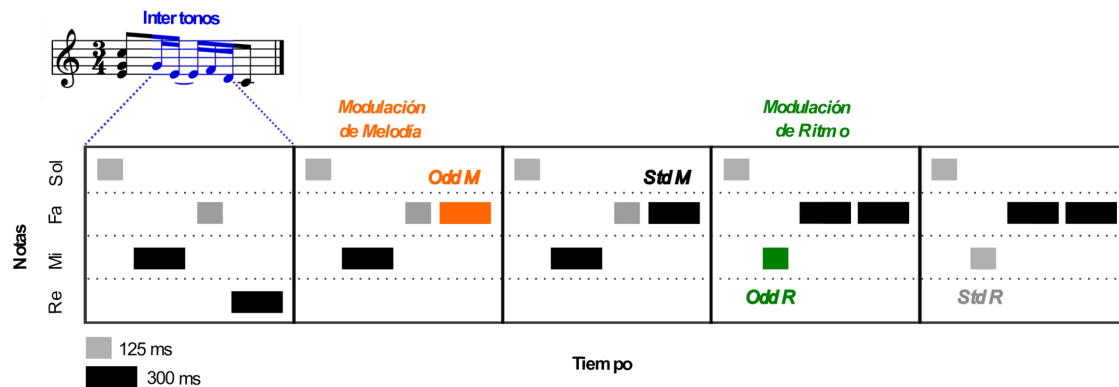


Figura 1; Diseño Experimental. Arriba: Se muestra la estructura y variación de los intertonos dentro de cada compás, que siempre incluyen a dos intertonos largos (300 ms, barras negras), y dos intertonos cortos (125 ms, cuadros grises). Cada uno de estos intertonos puede ser la nota de Re (D4), Mi (E4), Fa (F4) o Sol (G4) central. Las modulaciones de ritmo en un compás se llevan a cabo al intercambiar la duración entre dos intertonos contiguos de distinta duración (Odd R) sin alterar sus notas. Las modulaciones de melodía, por su parte, se llevan a cabo al cambiar aleatoriamente la nota de uno de los intertonos largos (Re → Fa, Odd M), por cualquiera de las otras notas permitidas. Sólo un tipo de cambio se realiza a la vez,

y el nuevo compás se repite inmediatamente al menos dos veces. *Abajo:* El sujeto recibe un tren de estímulos acústicos de presentación binaural generados mediante el software Psychopy (A, Equipo de Cómputo 1), y temporalmente acoplados a pulsos TTL o triggers mediante el uso del QUAD StimTracker (B), con la finalidad de marcar el inicio de cada compás durante la adquisición. A su vez, el sujeto es presentado con una película para desviar la atención de los estímulos acústicos a través de una pantalla ubicada a 30 cm del participante (E). La señal entrante de 14 electrodos (F3, Fz, F4, FCz, FC1, FC2, C3, Cz, C4, CP1, CPz, CP2, TP9 y TP10) es amplificada a través del equipo de EEG Neuroscan Neuvo de 64 canales (C), y la data es registrada en un equipo de cómputo (D, Equipo de Cómputo 2) mediante el software Curry 8.0.2.

Prueba de Sincronización Espontánea del Habla.

Tanto para los sujetos reclutados de la base de datos como para los caracterizados previamente a la aplicación del experimento, se llevó a cabo la versión acelerada explícita del SSS-Test, de acuerdo con lo descrito por Lizcano-Cortés et al. (2022). El participante recibió la instrucción de susurrar la sílaba “tah” de forma continua y concurrente con un estímulo de sílabas de ritmo acelerado durante un minuto. Previamente, el participante fue entrenado para realizar correctamente la tarea, bajo supervisión del experimentador. El experimento se corrió en dos bloques, con una sesión previa de entrenamiento cada uno. Las señales acústicas producidas fueron registradas, e inspeccionadas visualmente mediante el software Praat. Para cada bloque, las envolventes del estímulo y de la respuesta del participante fueron calculadas tomando el valor absoluto de la transformada de Hilbert de las señales acústicas para calcular el valor de acoplamiento de fase (PLV) mediante la ecuación:

$$PLV = \frac{1}{T} \left| \sum_{t=1}^T e^{i(\theta_1(t) - \theta_2(t))} \right| \quad (1)$$

Donde t es el tiempo discretizado, T es el número total de puntos temporales, y θ_1 y θ_2 son las fases de las dos envolventes; para más detalle, ver Assaneo et al. (2019).

El PLV de ambos bloques fue promediado para la obtención del PLV promedio final. Basándonos en el punto de cruce de la distribución ajustada a dos gaussianas descrito para la población mexicana por Gómez-Varela (2023), se eligió un valor de corte de $PLV = 0.62$, con el cual el sujeto fue clasificado en el grupo de baja ($PLV < 0.62$) o alta sincronía ($PLV > 0.62$). Tanto el experimento como el análisis se realizó con código personalizado en Matlab.

Paradigma oddball musical.

El experimento principal consistió en la presentación secuencial de 1387 compases distribuidos en cuatro bloques de aproximadamente 12.1 minutos de duración cada uno, con una duración aproximada total de 1 hora por experimento.

Los estímulos fueron presentados biauralmente mediante audífonos especializados ER1 (*Etymotic*) con auriculares de espuma insertables de un solo uso. A lo largo del experimento, se presentaron un total de 200 compases con modulaciones aleatorias de ritmo, y 200 compases con modulaciones aleatorias de melodía. Como regla general del experimento, solo un tipo de cambio fue inducido en un mismo compás, y éste nuevo compás fue repetido como mínimo una vez antes de la próxima modificación. El estándar para un compás oddball dado fue seleccionado buscando a través del experimento la última repetición de ese mismo compás.

Para desviar la atención de los estímulos auditivos, el sujeto fue presentado con una de cuatro películas de cine mudo sin intertítulos, en una pantalla a 30 cm de distancia del participante (Fig. 1E). Las películas fueron asignadas a cada bloque de forma aleatoria, e iniciaron de forma aleatoria en un rango de los primeros 0 a 10 segundos, con el fin de aleatorizar la fase entre el estímulo acústico y el video. Al final de cada bloque, se aplicó un cuestionario de cuatro reactivos sobre los eventos ocurridos en la película, funcionando como un intermediario para evaluar la atención del sujeto. Los sujetos con un porcentaje de aprobación menor al 70% fueron excluidos del análisis. El experimento tuvo una duración total aproximada de 50 minutos. El audio y el video reproducido en cada bloque fueron presentados mediante el software Psychopy v2021.2.3 en el equipo de cómputo iMac (**Fig. 1A**).

Adquisición de datos electrofisiológicos por electroencefalografía (EEG).

La adquisición de registros electroencefalográficos durante la aplicación del paradigma oddball musical se llevó a cabo mediante un amplificador Neuroscan Neuvo de 64 canales (**Fig. 1C**). Se registraron señales de un subgrupo de canales formado por F3, Fz, F4, FCz, FC1, FC2, C3, Cz, C4, CP1, CPz y CP2, que son los que muestran respuestas significativas de tipo MMN en Vaquero et al. (2021), más los electrodos TP9 y TP10 para la re-referencia *offline*. La tierra y la referencia *online* fueron colocadas en el entrecejo y en el canto ocular externo derecho del participante, respectivamente. Los datos fueron colectados a una frecuencia de muestreo de 500 Hz, por medio del software Curry 8.0.2 (Compumedics Neuroscan), en un equipo de cómputo HP 205 G3 AiO (**Fig. 1D**, Equipo de cómputo 2). El onset de cada compás fue registrado por el amplificador junto con los datos continuos de voltaje a través de un pulso TTL o trigger de 5 volts acoplado temporalmente al estímulo, producido mediante el equipo marcador de eventos *QUAD Stim Tracker* (Cedrus co.; **Fig. 1B**). Todo el protocolo se llevó a cabo en una habitación eléctrica y acústicamente aislada.

Análisis de Datos

Preprocesamiento de señales de EEG.

La señal continua de EEG fue preprocesada mediante el paquete MNE-Python (Gramfort et al., 2013) aplicando un filtro pasabandas tipo butterworth de 2 a 30 Hz, basado en transformada rápida de fourier (FFT). Los artefactos de parpadeo y electrocardiograma fueron removidos mediante análisis de componentes independientes (ICA), tras lo cual los datos fueron re-referenciados a la media de los dos mastoideos (TP9 y TP10).

Cálculo de Potenciales Relacionados a Eventos (ERP).

A partir del registro de EEG continuo, se establecieron épocas de 500 ms en torno al onset del intertono estándar u oddball, incluyendo los 100 ms previos a la presentación del intertono. Todos los análisis subsecuentes se llevaron a cabo con los datos dentro de estas ventanas temporales, descartando épocas que contengan amplitudes pico a pico por encima de los 80 μ V. El cálculo de las épocas incluyó corrección de línea base utilizando la actividad registrada durante un período de 100 ms previos al onset del estímulo, para substraer ruido no relacionado a la condición experimental.

Cálculo del componente MMN.

El potencial de disparidad MMN se calculó en cada sujeto para cada tipo de variación y cada canal a partir de la onda de diferencia obtenida al sustraer el ERP de las épocas estándar promedio al ERP de las épocas atípicas (*oddball*) promedio registradas de 100 ms antes del onset del intertono hasta 400 ms tras la presentación del estímulo. Para visualizar las diferencias a partir de los promedios individuales, también se calculan los promedios grupales de las ondas de diferencia para las dos condiciones de sincronización. Para identificar la señal MMN en cada grupo de sincronía, se corrió una prueba pareada de permutación de clusters contra cero utilizando las ondas de diferencia individuales, lo que permitió determinar la ventana temporal en que el voltaje de la señal es significativamente distinto de cero para cada grupo de sincronía (**Fig. 2b**, izquierda).

A partir de las señales MMN identificadas, se calcularon la latencia de pico y la amplitud media para cada individuo por canal (**Anexos**, Tabla 1) y para la señal media a través de canales (**Fig. 2b**, derecha). La latencia de pico de la señal MMN

para la modulación de ritmo se calculó en cada sujeto como el tiempo transcurrido en milisegundos desde la presentación del estímulo y el pico más negativo en una ventana de tiempo de 200 a 350 ms tras la presentación del estímulo, que se consideró la amplitud media de pico.

Aprendizaje Automático.

Como análisis alternativo y complementario a los potenciales relacionados a eventos, se entrenó un algoritmo categorizador de aprendizaje automático con el fin de comparar dos sets de datos por sujeto: las épocas estándar y *oddball* de todos los canales, para cada tipo de modulación y para cada grupo de sincronía. El algoritmo ajustó un modelo de regresión logística en ventanas de media móvil de 20 ms, llevándose a cabo una validación cruzada a 5 iteraciones. El desempeño de este categorizador fue evaluado en cada punto temporal graficado mediante los puntajes de área bajo la curva (AUC) para la curva de la característica operativa del receptor (ROC) obtenidos. Los puntajes de AUC de cada sujeto fueron promediados para cada tipo de variación, para construir promedios grupales. Los puntajes AUC promedios de cada grupo de sincronía fueron evaluados contra el azar (AUC = 0.5) para encontrar ventanas de significancia. Adicionalmente, se llevó a cabo un análisis de matriz de generalización temporal, para investigar si respuestas neurales particulares recurrieron en otras ventanas temporales.

Representación de Topografías.

Los valores de voltaje dentro de una ventana temporal de 8 ms entorno al punto medio de las ventanas de significancia de desempeño del categorizador, obtenidas para cada grupo de sincronía, fueron recuperados y promediados para cada sujeto y canal. Para cada sujeto, los puntos de voltaje promediados fueron normalizados mediante la siguiente fórmula:

$$SN_i(n) = \frac{S_i(n)}{\max[S_1(n), \dots, S_j(n)]} \quad (2)$$

Donde el valor normalizado SN_i del voltaje promedio en el electrodo S_i para el sujeto n , es resultado de dividir el valor real de voltaje promedio en ese electrodo sobre el máximo de los valores absolutos obtenidos para los voltajes promediados de todos los electrodos $S_1, S_2, \dots, S_j$ del mismo sujeto.

Con estas topografías normalizadas para cada sujeto, se realizaron promedios grupales para ambos grupos de sincronización, con el fin de obtener topografías representativas para ambas condiciones.

Análisis de distancias euclidianas.

Para probar que las configuraciones de voltaje de las topografías promediadas a través de los participantes para cada grupo de sincronía fueran significativamente distintas entre sí en el tiempo de máxima significancia, se calculó la distancia euclidiana entre todos los electrodos de ambos grupos como:

$$dE = \sqrt{\sum_{i=1}^j \left(S_i^{highs} - S_i^{lows} \right)^2} \quad (3)$$

Donde $S_1, S_2, \dots, S_j$ representan las lecturas de voltaje promedio de cada uno de los electrodos, para cada grupo de sincronía.

Para demostrar que la distancia euclidiana de los electrodos calculada fuera significativamente distinta del azar, se construyeron distribuciones aleatorias de probabilidad mediante un método de remuestreo por *bootstrapping*, reasignado aleatoriamente las etiquetas de sincronía a los datos de voltaje originales para generar valores de distancias euclidianas sustitutos. Este procedimiento se repitió por 100,000 iteraciones, y los valores obtenidos fueron normalizados mediante transformación logarítmica. Después, se utilizó el valor en el percentil 95 como umbral de significancia para la comparación de las distancias euclidianas reales obtenidas.

Análisis estadísticos.

Todas las comparaciones entre grupos de señales (ERP estándar vs ERP atípico, curvas ROC y ondas de diferencia entre grupos de sincronía) se llevaron a cabo mediante pruebas no-paramétricas y no-pareadas de permutación de *clusters*. Para determinar que las señales obtenidas por el análisis de ERPs y por aprendizaje automático fueran significativamente distintas de cero o del azar, respectivamente, se dirigieron pruebas no-paramétricas y pareadas de permutación de *clusters*. Para evaluar diferencias en la media grupal para los valores de PLV obtenidos, el efecto de los grupos de sincronización alta o baja sobre la amplitud media y latencia de pico del componente MMN, y el efecto sobre el valor máximo de AUC para las curvas ROC, se llevaron a cabo pruebas no-paramétricas U de Mann-Whitney.

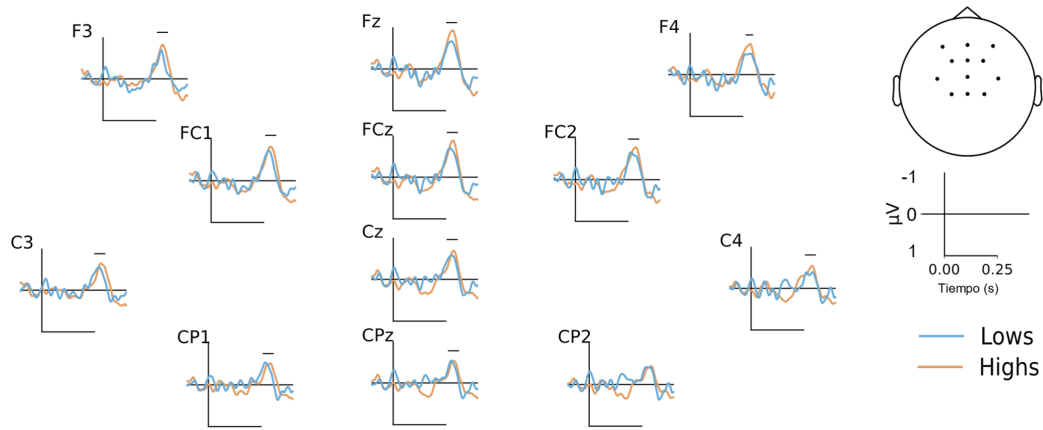
Todas las pruebas se evaluaron a un intervalo de confianza del 95%.

Resultados.

El potencial de disparidad (MMN) es generado en ambos grupos de sincronización para la modulación de ritmo.

Un total de 38 participantes completaron satisfactoriamente el paradigma oddball musical, de los cuales 19 fueron categorizados como de alta sincronía y 20 fueron de baja sincronía (**Anexos**, Fig. 1). Mediante el análisis de las señales cerebrales obtenidas y el cálculo de la onda de diferencia de los ERP oddball y estándar, nos fue posible reportar la presencia del componente MMN en ambos grupos de sincronización y en 11 de los 12 canales para la modulación de ritmo, pero no para la modulación de melodía (**Anexos**, Fig. 2). Análisis pareados de permutación de clusters contra cero, aplicados al conjunto de las ondas de diferencia obtenidas para cada grupo de sincronía y para cada canal, nos permitieron identificar, en cada caso excepto para CP2, una ventana temporal de voltaje significativamente más negativo (**Fig. 2a**), que coincidió con las latencias esperadas para la señal MMN frente a la modulación del ritmo de acuerdo a Putkinen et al. (2014) y Vaquero et al. (2021). Debido a que todos los canales significativos presentaron un cluster de valores significativos aproximadamente en la misma ventana temporal, se procedió a calcular el promedio a través de canales para cada grupo de sincronía, buscando nuevamente la ventana temporal de diferencia más significativa por medio de un análisis de permutación de clusters contra cero (**Fig. 2b**, izquierda; highs=0.245-0.309 s, $p=0.0006^{**}$; lows= 0.253-0.301 s, $p=0.003^{*}$). Interpretamos la caída de voltaje dentro de estas ventanas de significancia para cada grupo como el componente MMN, presente en una latencia similar a las reportadas anteriormente para la modulación del ritmo (Putkinen et al., 2014, Vaquero et al. 2021).

a



b

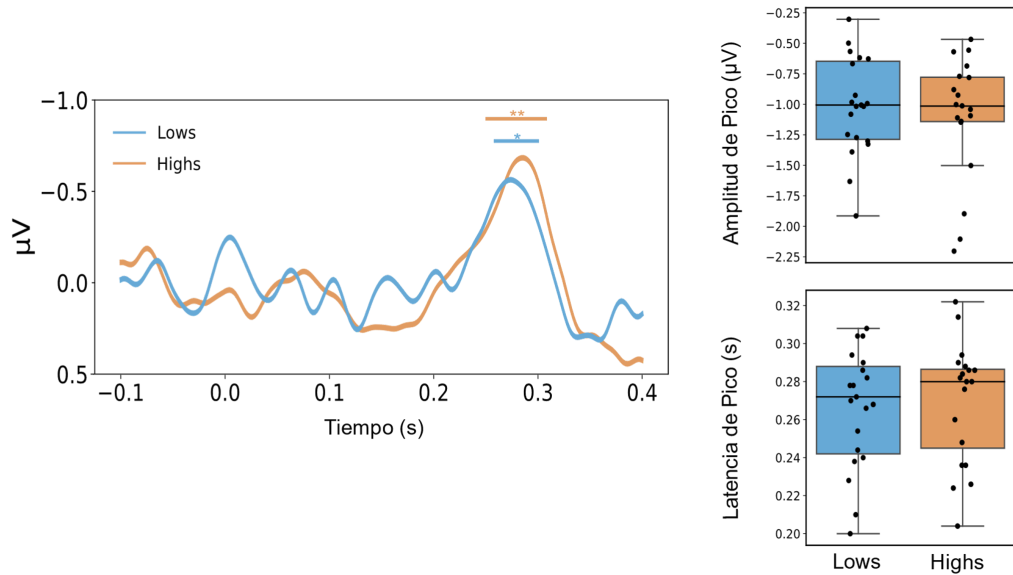


Figura 2; Ondas de diferencia para la modulación del ritmo. a. Se muestran los promedios grupales de las ondas de diferencia ($ERP_{odd} - ERP_{std}$) para la modulación del ritmo en el grupo de baja (azul) y alta (rojo) sincronía. La barra negra horizontal denota un pico significativo obtenido por un análisis de permutación no-pareado (ERP_{odd} vs ERP_{std}) presente en todos los canales y en ambos grupos. **b: Izquierda:** Promedio grupal a través de canales. Mediante un análisis de permutación de clusters, se encontraron ventanas donde la señal es significativamente distinta de cero para ambos grupos ($highs=0.245-0.309$ s, $p=0.0006^{**}$; $lows=0.253-0.301$ s, $p=0.003^{*}$). **Derecha:** Distribución de datos de latencia y amplitud del pico más negativo dentro de las ventanas de significancia para cada grupo (promedio a través de canales) en una ventana temporal de 200 a 350 ms. No hubo diferencias significativas entre grupos para los datos de amplitud ($AMP_{highs} = -1.09 \pm 0.39$ μV , $AMP_{lows} = -1.01 \pm 0.48$ μV , $p=0.85$) ni para la latencia media de pico ($LMP_{highs} = 0.269 \pm 0.03$ s, $LMP_{lows} = 0.265 \pm 0.03$ s, $p=0.77$) con una prueba U de Mann-Whitney. Así mismo, un análisis similar aplicado a cada canal por separado tampoco encontró diferencias significativas (Anexos, Tabla 1).

Análisis de ERPs: El estatus de sincronía no tiene efecto sobre la amplitud y latencia de pico de la señal MMN frente a cambios de modulación de ritmo.

Con el fin de buscar posibles diferencias en las señales obtenidas entre grupos de sincronía, se condujeron pruebas no-pareadas de permutación de clusters entre las ondas de diferencia de cada grupo. Estas pruebas fueron realizadas con los datos de voltaje para cada canal (datos no mostrados) y también con los datos de voltaje promediados a través de canales (**Fig. 2b**, izquierda). Sin embargo, ninguno de estos análisis arrojó clusters de valores significativos. Esta observación sugiere que las señales MMN no fueron distintas en función del grupo de sincronía. Para confirmar esta observación, buscamos cuantificar la señal MMN a través de la amplitud media de pico (AMP) y la latencia media de pico (LMP) para cada grupo de sincronía, independientemente por cada canal (**Tabla 1**, Anexos), y sobre el promedio a través de canales (**Fig. 2b**, derecha) con el objetivo de comparar la AMP y LMP entre grupos. Para este propósito, utilizamos las ondas de diferencia para buscar el pico de voltaje más negativo, junto con su latencia, para cada sujeto. Debido a que la inspección de las ondas individuales mostró que para algunos sujetos los picos de voltaje estaban fuera de la ventana de significancia obtenida para el promedio de los datos, elegimos una ventana temporal más amplia, de 0.2 s a 0.35 s para la búsqueda de estos picos. Sin embargo, mediante este análisis tampoco se encontraron diferencias significativas (**Fig. 2b**, derecha) entre grupos para la amplitud media de pico ($AMP_{\text{highs}} = -1.09 \pm 0.39 \mu\text{V}$, $AMP_{\text{lows}} = -1.01 \pm 0.48 \mu\text{V}$, $p=0.855$) ni para la latencia media de pico ($LMP_{\text{highs}} = 0.26 \pm 0.03 \text{ s}$, $LMP_{\text{lows}} = 0.26 \pm 0.03 \text{ s}$, $p=0.7786$) en la señal promedio, ni entre canales independientes (**Tabla 1**, Anexos).

El análisis de aprendizaje automático revela posibles diferencias entre grupos de sincronía en la duración de la respuesta neural frente a cambios de modulación de ritmo.

Como segunda parte de nuestro análisis, nos propusimos entrenar un algoritmo de aprendizaje automático con los datos de los ERPs estándares y atípicos de cada sujeto, con el objetivo de obtener curvas de desempeño (ROC AUC). El promedio grupal del desempeño del categorizador se muestra, para cada grupo de sincronía, en la Figura 2a. Un análisis pareado de permutación de clusters contra el azar ($AUC=0.5$) demostró que el categorizador tuvo un desempeño distinto del azar en ventanas temporales que se aproximan a las obtenidas para la señal MMN mediante el análisis de ERPs para ambos grupos de sincronía (**Fig. 3a**, Highs: 0.244-0.284 s, $p=0.0006^{**}$; Lows: 0.236-0.264 s, $p=0.01^{*}$). No hubo diferencias significativas entre el valor máximo de desempeño del categorizador obtenido dentro de las ventanas de significancia específicas para cada grupo de sincronía

(**Fig. 3b**; Max. $AUC_{\text{highs}} = 0.60 \pm 0.04$, Max. $AUC_{\text{lows}} = 0.57 \pm 0.04$, $p=0.07$). Sin embargo, un análisis no-pareado de permutación encontró que hay diferencias significativas entre los desempeños del categorizador en función del grupo de sincronía para una ventana de los 0.268 a 0.288 s (**Fig. 3a**, $p=0.0459^*$). Este hallazgo podría interpretarse como una posible diferencia en la duración de la señal MMN entre grupos de sincronía, con el grupo de baja sincronía presentando un aumento más temprano en el voltaje respecto de los sincronizadores de alta, o bien, pudiera sugerir que los sincronizadores de alta tienen una respuesta más prolongada, o un segundo tipo de respuesta adicional y distinta.

Para explorar estas posibilidades, se llevó a cabo un segundo análisis de aprendizaje automático, esta vez aplicando una matriz de generalización temporal al mismo categorizador y para los datos de cada sujeto. Esta aproximación permite conocer la dinámica temporal cuando se aplica un mismo modelo de regresión, generado a un tiempo t , en todos los demás puntos en el tiempo. La media grupal de las matrices de generalización para cada estatus de sincronía se muestran en la Figura 2c. Mediante este análisis se observa que, mientras que el grupo de baja sincronía muestra una estructura temporal con un máximo de desempeño en torno a los 0.236 a 0.264 s, el grupo de alta sincronía muestra una respuesta más robusta, consistente y sostenida en el tiempo, que se extiende hasta en torno a los 0.284 s (**Fig. 3c**). Además, las estructuras temporales de ambas matrices son continuas en el tiempo, aunque parecen mostrar fluctuaciones verticales y horizontales en el desempeño del categorizador. Esto sugiere que, para los sincronizadores de alta, el mismo proceso se extiende en una ventana temporal más amplia, que va de los 0.240 a los 0.288 s, mientras que para los sincronizadores de baja, la estructura temporal es más pequeña y culmina en torno a los 0.264 s.

Guiándonos con estas observaciones, quisimos también evaluar el grado de similitud entre el tipo de mecanismos actuando en las ventanas temporales de significancia descritos entre grupos de sincronía, a través de la topografía de los electrodos. Para ello, tomamos una ventana de 10 ms en torno a la media de la ventana de significancia para cada grupo de sincronía, de acuerdo a lo descrito por el categorizador ($t_H = 0.264 \pm 0.05$ s para highs y $t_L = 0.250 \pm 0.05$ s para lows). Los valores de voltaje registrados en estas vetanas fueron obtenidos de las ondas de diferencia de cada sujeto, y fueron promediados para obtener un valor de voltaje representativo por cada canal, y de forma separada para ambos grupo de sincronía. Los promedios grupales de estas respuestas individuales para cada grupo de sincronía y para cada canal se muestran en las topografías de la Figura 3d.

Por último, se evaluó si la configuración de voltaje de los sensores era diferente entre grupos de sincronía. Para ello, se calculó la distancia euclidiana entre los sensores de highs y lows de acuerdo a la ecuación 2. El valor de distancia euclidiana observado (dE_{OBS}) se comparó contra el percentil 95 (dE_{95}) de una distribución de probabilidad aleatoria generada al re-asignar al azar las etiquetas de sincronía de los datos de voltaje y calcular las distancias euclidianas de la primera condición 10000 veces, normalizando los resultados con transformación logarítmica. Mediante esta aproximación, encontramos que la distancia euclidiana de las topografías entre grupos de sincronía no es significativamente distinta del azar (**Fig. 3e**; $\log(dE_{OBS}) = -0.1172$, $\log(dE_{95}) = 0.2111$). Apoyándonos en estas observaciones, nuestros resultados sugieren que existe un mecanismo común entrando en juego durante la detección automática de cambios en patrones rítmicos para ambos grupos de sincronía, que continúa en una respuesta prolongada entre los 0.268 a 0.288 s sólo en los sujetos de alta sincronía. La interpretación de la matriz de generalización temporal no apoya la idea de que las topografías entre grupos sean sustancialmente distintas durante las respuestas temprana y tardía. Más bien, la misma respuesta parece sostenerse por más tiempo para los sincronizadores altos, sobre todo en los electrodos a la izquierda de la línea media (**Fig. 2a**).

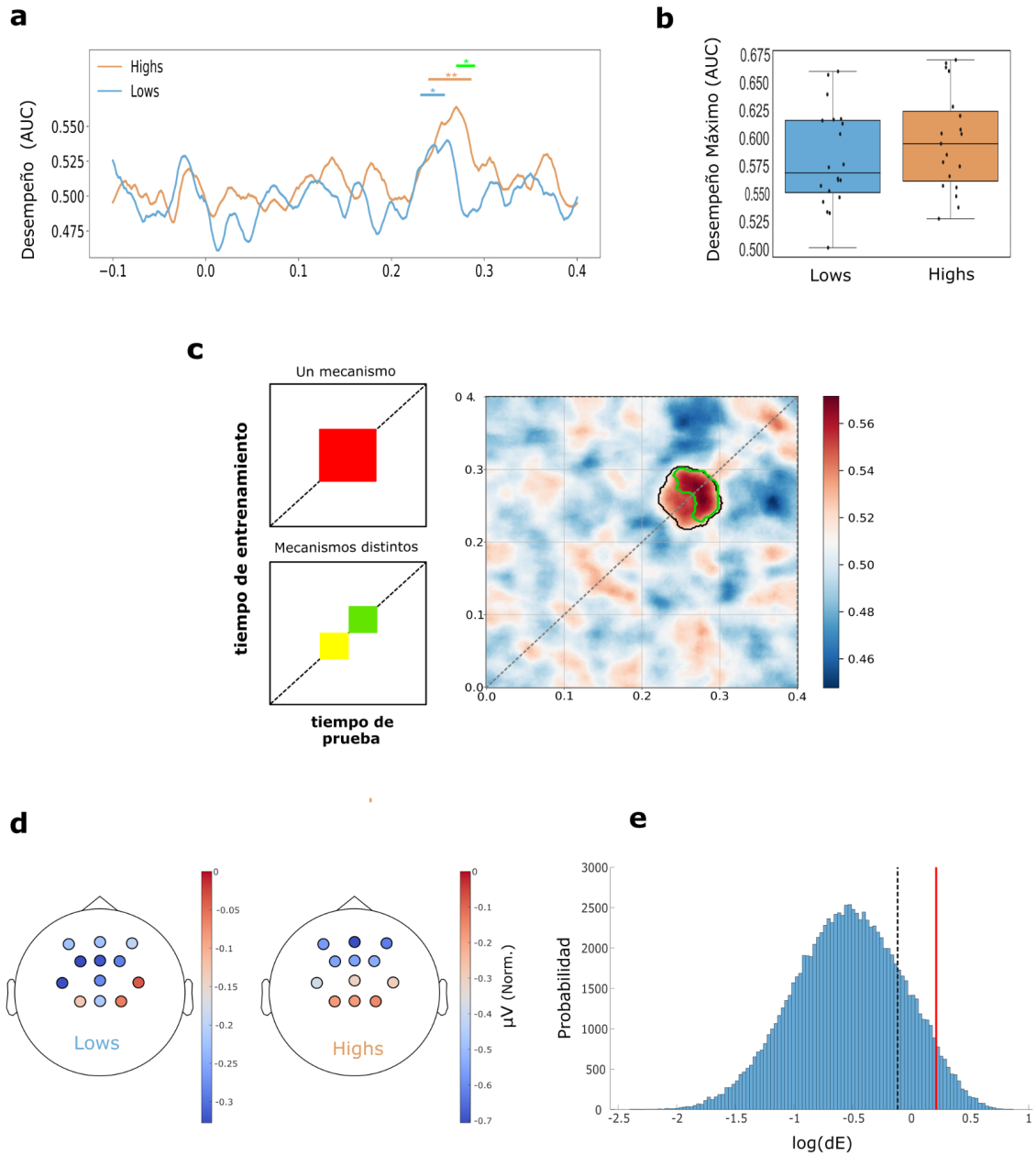


Figura 3; El aprendizaje automático revela diferencias en la duración de la respuesta frente a cambios de modulación de ritmo entre grupos de sincronía. **a.** Puntajes promedio (a través de sujetos) de área bajo la curva (AUC) obtenidos mediante un modelo de regresión logística a 5 iteraciones. Se muestran las ventanas de significancia estadística contra el azar (AUC=0.5) evaluadas mediante un análisis pareado de permutación de clusters para highs (rojo, $p=0.0006^{**}$) y lows (azul, $p=0.01^{*}$). También, se muestra la ventana de significancia obtenida mediante un análisis no-pareado de permutación (highs vs lows, verde, $p=0.0459^{*}$). **b.** Puntajes de desempeño máximo del categorizador dentro de las ventanas de significancia correspondientes a cada grupo de sincronía ($p=0.07$, U de Mann-Whitney). **c. Izquierda.** Esquematización de dos tipos de estructuras temporales reveladas por una matriz de generalización temporal, al entrenar un algoritmo de categorización en un tiempo (eje vertical) y entrenarlo en todos los demás puntos temporales (eje horizontal): una estructura continua que revela un mismo mecanismo neural en el tiempo (arriba, cuadro rojo), y dos estructuras discontinuas, que implican dos mecanismos neurales distintos (abajo, cuadros amarillo y verde). **Derecha.** Matriz de generalización temporal para sincronizadores de alta; Se muestran las áreas de diferencia significativa resultantes de un análisis de permutación contra el azar (línea negra) y entre grupos (línea verde), revelando un solo mecanismo para ambos

grupos, extendido en el tiempo para los sincronizadores altos. **d.** Topografías normalizadas de la respuesta de voltaje promediadas en una ventana $t_w = 0.264 \pm 0.005$ s para highs y $t_l = 0.250 \pm 0.005$ s para lows. **e.** Distribución de probabilidad de distancias euclidianas entre las topografías obtenidas por re-muestreo con intercambio de etiquetas de sincronía. Se muestra el valor observado de distancia euclidiana ($\log(dE_{OBS}) = -0.1172$, línea discontinua vertical negra) y el percentil 95 de la distribución ($\log(dE_{95}) = 0.2111$, línea continua vertical roja).

Discusión.

Mediante este trabajo fue posible detectar exitosamente respuestas de tipo MMN frente a la modulación del ritmo en estímulos musicales, pero no frente a la modulación de melodía (**Anexos**, Fig. 2). Desconocemos si esta ausencia del componente MMN frente a cambios de melodía es producto de un problema de diseño. Una posibilidad es que, dado el contexto musical de los estímulos, las variaciones melódicas inducidas en el presente experimento no hayan sido lo suficientemente infrecuentes como para permitir el establecimiento de un patrón marcado, resultando en una señal de tipo MMN de muy baja amplitud, que no fue posible identificar en nuestros análisis. Sin embargo, esto contrasta con los resultados de Putkinen et al. (2014) y Vaquero et al. (2021), que sí encontraron una respuesta de modulación de melodía utilizando un paradigma parecido. Por otro lado, una diferencia sustancial entre el presente paradigma y el paradigma utilizado en los trabajos previos es que nuestro experimento no indujo cambios de clave en la melodía, que se mantuvo siempre en clave de do mayor. Dado que los cambios de nota respetaron siempre la misma clave y nuestro set de posibles notas a variar se restringió consecuentemente a 4 notas, es una posibilidad que la modulación de melodía no haya sido lo suficientemente novedosa para evocar una respuesta de tipo MMN robusta, sobre todo a lo largo de los cuatro bloques del experimento. Además, en el presente trabajo tanto las modulaciones de melodía como de ritmo ocurrieron en frecuencias parecidas, por lo que la ausencia de la respuesta MMN frente a modulación de melodía sugiere que la sensibilidad en la detección de cambios en estímulos musicales puede ser sustancialmente distinta dependiendo de la característica variada.

Mediante un análisis clásico de potenciales relacionados a eventos, no encontramos diferencias en la amplitud y la latencia de la señal MMN para la modulación de ritmo entre grupos de sincronización entre canales (**Anexos**, Tabla 1) ni a través de canales (**Fig. 2b**). Por otro lado, mediante un enfoque de aprendizaje automático, los resultados del presente trabajo muestran tiempos de distinción del categorizador entre condiciones que corresponden a las esperadas para la señal MMN, sugiriendo que ambos métodos permiten investigar un mismo fenómeno. Además, el aprendizaje automático también permitió encontrar diferencias entre grupos, con los sincronizadores de alta presentando diferencias en el tiempo de duración de la distinción. Dada su capacidad para medir cuán

generalizable es una observación a través de un set de datos (en este caso, las épocas estándares y *oddballs*) mediante validación cruzada, el aprendizaje automático representa un análisis más robusto. Además, este tipo de análisis incluye información sobre la topología de la activación que no se tiene en cuenta en los análisis de ERPs clásicos que estudian el promediado a través de los canales o en canales independientes. El análisis de matrices de generalización temporal (**Fig. 3c**), junto con la interpretación de las topografías (**Fig. 2a**, **Fig. 3d**, **Fig. 3e**), apuntan a que un solo mecanismo ocurre para ambos grupos de sincronía. Para los sincronizadores altos, este mecanismo muestra una respuesta sostenida por aproximadamente 20 milisegundos más. En la literatura clásica de la respuesta MMN, han sido descritas caídas negativas de voltaje que ocurren después de la presentación de este componente, y que podrían representar el direccionamiento involuntario de la atención hacia el evento disruptivo (Näätänen et al., 1992) o bien representar a un componente de tipo N2b (Sasaki et al., 2000). Sin embargo, el análisis clásico de ERPs del presente trabajo no encontró la presencia de estos componentes tardíos y accesorios a la respuesta MMN en ninguno de los dos grupos experimentales. Por extensión, el análisis de aprendizaje automático con la matriz de generalización temporal tampoco describió la presencia de un proceso tardío distinto en la respuesta de los sincronizadores altos. Estas evidencias indican que la presencia de respuestas accesorias no explican la prolongación temporal del desempeño del categorizador observada en el grupo de alta sincronía, lo que nos lleva a suponer que el mismo sustrato neural se encuentra activo por más tiempo en este grupo.

Por otro lado, resulta interesante mencionar que la comparación de los desempeños máximos del categorizador en las ventanas de interés se puede interpretar como marginalmente significativa (**Fig. 3b**, $p=0.07$) entre grupos de sincronía. Esto puede sugerir que existen diferencias en el nivel de distinción entre los estímulos atípicos y estándares entre grupos de sincronía, que no llegan a ser identificados por la potencia estadística. Resulta importante mencionar que el descubrimiento de diferencias en la respuesta neural en individuos sanos mediante EEG es siempre un desafío, debido a la relación señal ruido del método. Sin embargo, los resultados muestran una tendencia de los sincronizadores de alta a tener amplitudes de señal más negativas respecto de los sincronizadores de baja en las ventanas temporales de interés (**Fig. 2a**, **Fig. 2b**).

En conjunto, estas observaciones revelan diferencias en la detección automática de cambios de ritmo entre grupos de sincronía, que podrían ser consecuencia de las diferencias estructurales y/o funcionales conocidas entre estos grupos (Assaneo et al., 2019; Kern et al., 2021; Orpella et al., 2022). Concretamente, los sincronizadores de alta muestran una ventana más prolongada de respuesta

durante la detección de cambios, que podría relacionarse a una mayor conectividad temporofrontal mediada por el LAF, como se reportó en Vaquero et al. (2021).

Se ha propuesto que mecanismos de predicción temporal estén detrás de las respuestas de tipo MMN (Wacogne et al., 2012). De igual manera, resultados experimentales sugieren que mecanismos de predicción temporal basados en procesos oscilatorios entre la corteza auditiva y motora, establecidos por el fascículo arqueado, ocurren durante la percepción de estímulos rítmicos de habla, y que el acoplamiento estrecho entre estos sistemas funciona como un facilitador de la percepción en sincronizadores altos (Assaneo et al., 2020). Las observaciones del presente trabajo demuestran que la capacidad de sincronización auditivo-motora del habla tiene un efecto sobre la detección automática de cambios en patrones rítmicos musicales, que muestra diferencias en la duración de la respuesta neural entre grupos de sincronía. Este resultado sugiere que la vía auditiva dorsal puede jugar un papel clave en distintos procesos que hacen uso de la codificación predictiva, como lo son la detección de cambios en patrones rítmicos y la percepción del habla, y que podrían compartir sustratos neurales generadores.

Conclusiones

El presente trabajo muestra diferencias en la duración de la respuesta neural durante la detección automática de cambios en patrones rítmicos entre grupos de sincronización auditivo-motora del habla. El grupo de los sincronizadores altos, un subgrupo de la población general caracterizado por mayor conectividad frontotemporal, muestra una respuesta más prolongada en el tiempo. Estos resultados sugieren que hay diferencias individuales en respuestas neurales automáticas antes no descritas para la población general, volviéndose necesario explorar sus posibles implicaciones cognitivas.

Referencias

- Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, 4(2).
- Assaneo, M.F., Ripollés, P., Orpella, J., Lin, W.M., Diego-Balaguer, R., & Poeppel, D. (2019) Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience*, 22:627-632.
- Ding N., Patel A.D., Chen L., Butler H., Luo C., & Poeppel D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 181-187.
- Doelling, K.B., Arnal, L.H., Ghitza, O. & Poppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, 85:761-768.
- Doelling, K.B. & Poppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, E6233–E6242.
- Clerc, M., Bougrain, L., & Lotte, F. (2016). Functional Neuroimaging. En Clerc, M., Bougrain, L., Lotte, F. (Eds.). *Brain-Computer Interfaces 1: Foundations and Methods* (pp. 31-33). *John Wiley & Sons, Inc.*
- Fitzgerald, K., & Todd, J. (2020). Making Sense of Mismatch Negativity. *Frontiers in Psychiatry*, 11(468).
- Ford, J.M., & Hillyard, S.A. (1981). Event-Related Potentials (ERPs) to interruptions of a Steady Rhythm. *Psychophysiology*, 18:322-330.
- Fujii, S., & Wan, C.Y. (2014) The role of rhythm in speech and language rehabilitation: the SEP hypothesis. *Frontiers in Human Neuroscience*, 777(8).
- Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, 120(3), 453–463.
- Goldstein, E.B., Cacciamani, L. (2021). Chapter 1: Introduction to Perception. En Glodstein, E.B., Cacciamani, L. (Eds.). *Sensation and Perception*, 11th Edition (pp. 6-9). *Cengage learning, Inc.*
- Gómez-Varela, I.M. (2023). *Modulación del Aprendizaje Estadístico Mediante la Variación de la Estructura Rítmica del Estímulo* [Tesis de maestría publicada]. *Universidad Nacional Autónoma de México*.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L. & Hämäläinen, M. (2013).

- MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(261).
- Jackson, A. F., & Bolger, D. J. (2014). The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, 51(11):1061–1071.
- King, J.R., Gwilliams, L., Holdgraf, C., Sassenhagen, J., Barachant, A., Engemann, D., Larson, E. & Gramfort, A. (2018). Encoding and Decoding Neuronal Dynamics: Methodological Framework to Uncover the Algorithms of Cognition. *HAL open science*.
- Kisley, M.A., Davalos, D.B., Layton, H.S., Pratt, D., Ellis, J.K., & Seger, C.A. (2004). Changes in temporal deviance modulate mismatch negativity amplitude in humans. *Neuroscience Letters*, 358:197-200.
- Kotz, S.A., & Schwartz, M. (2010). Cortical Speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, 14(9):392-399.
- Kropotov, J. (2009). Methods: Neuronal Networks and Event-Related Potentials. En Kropotov, J. (Ed.). Quantitative EEG, Event-Related Potentials and Neurotherapy (pp. 357-360). *Academic Press*.
- Lohvansuu, K., Torppa, M., Ahonen, T., Eklund, K., Hämäläinen, J.A., Leppänen P.H.T., & Lyytinen, H. (2021). Unveiling the Mysteries of Dyslexia—Lessons Learned from the Prospective Jyväskylä Longitudinal Study of Dyslexia. *Brain Sciences*, 427(11).
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190.
- Millidge, B., Seth, A.K. & Buckley, C.L. (2022). Predictive coding: A theoretical and experimental review. arXiv. [Preprint].
- Näätänen, R. (1992) Attention and Event Related Potentials. En Näätänen, R. (Ed.). Attention and Brain Function, *Routledge*.
- Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing. *Clinical Neurophysiology*, 2544-2590.
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, 115(1), 140–144.
- Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*, 24(4), 375–425.

- Näätänen, R., Jiang, D., Lavikainen, J., Reinikainen, K., Paavilainen, P. (1993). Event-related potentials reveal a memory trace for temporal features. *NeuroReport*, 5(3), 310–312.
- Nagai, T., Tada, M., Kirihaara, K., Araki, T., Jinde, S., & Kasai, K. (2013). *Mismatch Negativity as a “Translatable” Brain Marker Toward Early Intervention for Psychosis: A Review*. *Frontiers in Psychiatry*, 115(4).
- Olejniczak, P. (2006). Neurophysiologic Basis of EEG. *Journal of Clinical Neurophysiology*, 23(3):186–189.
- Pakarinen, S., Lovio, R., Huotilainen, M., Alku, P., Näätänen, R., & Kujala T. (2009). Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biological Psychology*, 82(3), 0–226.
- Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20(1), 159–172.
- Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N. & Huotilainen, M. (2014) Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiology of Learning and Memory*, 110(2014):8-15.
- Repp, B.H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6).
- Repp, B.H., & Su, Y.H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20:403-452.
- Schröger, E. (1998) Measurement and interpretation of the mismatch negativity. *Behaviour Research Methods, Instruments & Computers*, 30(1):131-145.
- Sasaki, T., Campbell, K. B., Gordon Bazana, P., & Stelmack, R. M. (2000). Individual differences in mismatch negativity measures of involuntary attention shift. *Clinical Neurophysiology*, 111(9), 1553–1560.
- Shelton, J., & Kumar, G.P. (2010). Comparison between auditory and visual simple reaction times. *Neuroscience & Medicine*, 1:30-32.
- Spratling, M.W. (2016). Predictive coding as a model of cognition (Research Report). *Cognitive Processing*, Springer.
- Sur, S., & Sinha, V.K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, 18(1):70-73.
- Sussman, E. S., Chen, S., Sussman-Fort, J., & Dinces, E. (2013). The Five Myths of MMN: Redefining How to Use MMN in Basic and Clinical Research. *Brain Topography*, 27(4), 553–564.
- Sussman, E., Ritter, W., & Vaughan, H. G. (1998). Predictability of stimulus deviance and the mismatch negativity. *NeuroReport*, 9(18), 4167–4170.

- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, 226(1), 1-4
- Tervaniemi, M., Maury, S., & Näätänen, R. (1994). Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *NeuroReport*, 5(7), 844–846.
- Vaquero, L., Ramos-Escobar, N., Cucurell, D., François, C., Putkinen, V., Segura, E., Huottilainen, M., Penhune, V., & Rodríguez-Fornells, A. (2021). Arcuate fasciculus architecture is associated with individual differences in pre-attentive detection of unpredicted music changes. *NeuroImage*, 229, 117759.
- Vuust, P., Liikala, L., Näätänen, R., Brattico, P., & Brattico, E. (2016). Comprehensive auditory discrimination profiles recorded with a fast parametric musical multi-feature mismatch negativity paradigm. *Clinical Neurophysiology*, 127(4), 2065–2077.
- Wacongne, C., Changeux, J.P. & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *The Journal of Neuroscience*, 32(11):3665-3678.
- Whitton, S.A. & Jiang, F. (2023). Sensorimotor synchronization with visual, auditory and tactile modalities. *Psychological Research*.
- Woodman, G. F. (2010). A brief introduction to the use of Event-Related Potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, 72(8):2031–2046.
- Yu, X., Liu, T., & Gao, D. (2015). The Mismatch Negativity: An Indicator of Perception of Regularities in Music. *Behavioural Neurology*, 2015, 1–12.

ANEXOS

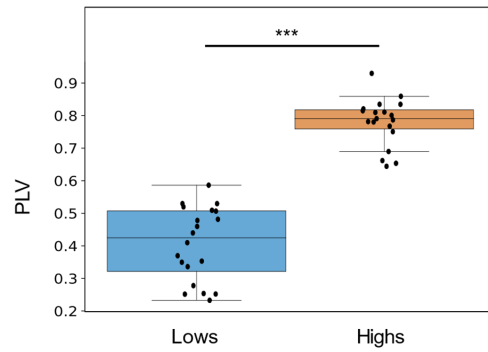


Figura 1; Distribución de valores de acoplamiento de fase (PLV) en la cohorte experimental. Se utilizó un valor de corte de $PLV=0.62$, de acuerdo con lo descrito por Gómez-Varela (2023), para categorizar a los participantes como sincronizadores de alta y de baja. Los valores de PLV entre ambos grupos presentaron diferencias muy significativas ($PLV_{\text{Lows}} = 0.4066 \pm 0.1097$, $PLV_{\text{Highs}} = 0.7803 \pm 0.0716$, $p < 0.0001$) en una prueba U de Mann-Whitney.

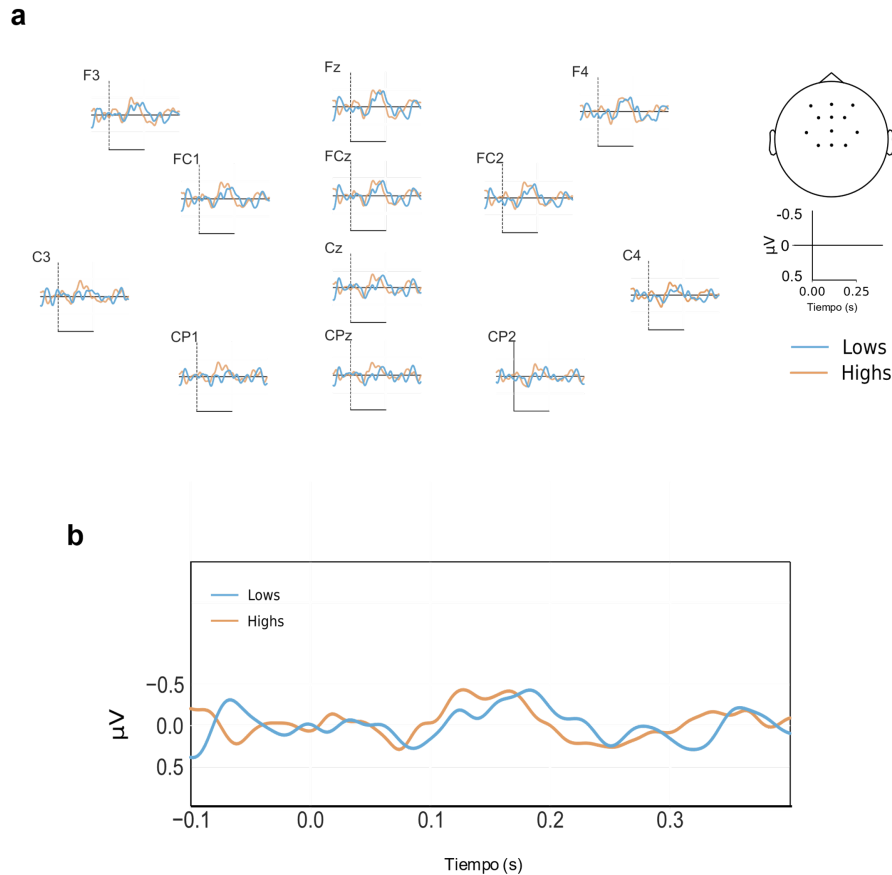


Figura 2; Ondas de diferencia para la modulación de melodía en grupos de sincronía. **a.** Un análisis pareado de permutación no encontró que las ondas de diferencia para la modulación de melodía fueran significativamente distintas de cero en ninguna ventana de tiempo para ningún canal, rechazándose la presencia de una señal MMN, que esperaba encontrarse en latencias de los 0.150 a 0.250 s de acuerdo con lo reportado por Vaquero et al. (2021). **b.** Onda de diferencia promedio a través de canales y sujetos para la modulación de melodía.

Tabla 1. Resultados de las pruebas U de Mann-Whitney para la comparación de las latencias medias de pico (LMP) y amplitudes medias de pico (AMP) entre grupos de sincronía, para cada canal. Ninguno de los canales mostró diferencias significativas entre grupos.

<i>Electrodo</i>	<i>LMP (valor p)</i>	<i>AMP (valor p)</i>
F3	0.184	0.787
Fz	0.989	0.626
F4	0.542	0.957
FCz	0.807	0.828
FC1	0.456	0.807
FC2	0.968	0.705
C3	0.093	0.892
Cz	0.914	0.21 3
C4	0.935	0.914
CPz	0.377	0.416
CP1	0.322	0.745