



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
Maestría en Ciencias (Neurobiología)

«Caracterización de la sincronía auditivo-motora para  
distintos estímulos y efectores»

TESIS QUE PARA OBTENER EL GRADO DE:

Maestra en Ciencias

PRESENTA:

Cecilia Paulina Mares Román

TUTORA PRINCIPAL

Dra. María Florencia Assaneo  
Instituto de Neurobiología

COMITÉ TUTOR:

Dra. María Magdalena Giordano Noyola  
Instituto de Neurobiología  
Dr. Luis Alonso Lemus Sandoval  
Instituto de Fisiología Celular

QUERÉTARO, QRO., AGOSTO 2022



Universidad Nacional  
Autónoma de México



**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.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Instituto de Neurobiología

Los miembros del Comité Tutorial certificamos que la tesis elaborada por: *Cecilia Paulina Mares Román*, cuyo título es: «*Caracterización de sincronía auditivo-motora para distintos estímulos y efectores*» se presenta como uno de los requisitos para obtener el grado de Maestra en Ciencias (Neurobiología) y cumple con los criterios de originalidad y calidad requeridos por la Coordinación General de Estudios de Posgrado de la Universidad Nacional Autónoma de México.

Presidente: \_\_\_\_\_

Secretario: \_\_\_\_\_

Vocal: \_\_\_\_\_

Suplente: \_\_\_\_\_

Suplente: \_\_\_\_\_

\_\_\_\_\_

Coordinadora del Programa

## Resumen

Se conoce como sincronización auditivo-motora a la habilidad de acoplar de forma temporal la realización de un gesto motor con un estímulo auditivo repetitivo. Esta habilidad es de naturaleza innata y ha sido relacionada con el habla en varios trabajos. El SSS-Test (Test de Sincronización Espontánea del Habla) es una prueba conductual con la cual se han distinguido dos grupos en la población (llamados sujetos de alta y baja sincronía), con diferencias en sus habilidades de sincronía y en características cerebrales a nivel estructural y funcional. Debido a que no se tiene conocimiento que una distinción de este tipo haya sido reportada en la literatura sobre sincronización auditivo-motora, con este trabajo se buscó determinar si esta distinción entre grupos se extiende a otros dominios fuera del habla. Para ello, se evaluó el nivel de sincronización auditivo-motora para distintas combinaciones de estímulos y gestos motores, relacionados con el habla (tren de sílabas/susurrar la sílaba «ta») o no (tren de tonos/aplaudir). Se encontró que el empleo de un estímulo compuesto por unidades acústicas que no se repiten de forma consecutiva, como el tren de sílabas empleado en el SSS-Test, es lo que permite distinguir a los dos grupos en la población. Este hallazgo es muy revelador puesto que los estímulos empleados en los experimentos clásicos de sincronización auditivo-motora utilizan la repetición de tonos o clicks, pudiendo ser esta la razón por la cual no se había reportado una distinción entre grupos, pudiendo tener consecuencias en las observaciones hasta ahora publicadas. Además, fue posible producir un efecto de mejoría en la sincronía de los participantes previamente clasificados como sujetos de baja al realizar con anterioridad una tarea de sincronización de forma efectiva. Queda pendiente de ser estudiado si esta mejoría tiene efectos a nivel cognitivo en tareas relacionadas con el aprendizaje de palabras y la percepción del habla.

## Abstract

The ability to temporarily align the performance of a motor gesture to a repetitive auditory stimulus is known as auditory-motor synchrony. This ability is innate in nature and has been related to speech in various works. The SSS-Test (Spontaneous Speech Synchronization Test) is a behavioral test which has distinguished two groups in the population (called high and low synchronizers), with differences on their synchrony abilities and at structural and functional brain level. Since it is not known that a distinction of this type has been reported in the literature on auditory-motor synchronization, this work aimed to determine whether this distinction between groups extends to domains other than speech. To do so, the level of auditory-motor synchronization for different combinations of stimuli and motor gestures was evaluated, related to speech (train of syllables/whispering the syllable «ta») or not (train of tones/clapping). It was found that the use of a stimulus made up of acoustic units that are not repeated consecutively, such as the train of syllables used in the SSS-Test, is what makes it possible to distinguish the two groups in the population. This finding is highly revealing since the stimuli used in the classic auditory-motor synchronization experiments use the repetition of tones or clicks, which may be the reason why a distinction between groups had not been previously reported, and may have consequences in observations published so far. In addition, it was possible to produce an improvement effect in the synchrony of the participants previously classified as low synchronizers by first performing a synchronization task effectively. It remains to be studied whether this improvement has effects at a cognitive level in tasks related to word learning and speech perception.

## Agradecimientos

Al Instituto de Neurobiología y al Instituto de Fisiología Celular.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) CVU: 1086447.

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) Protocolo 096.H.

Al Programa de Apoyo a los Estudios de Posgrado (PAEP).

A la Coordinación General de Estudios de Posgrado de la UNAM.

A la Coordinación del Programa de Maestría en Ciencias (Neurobiología): Dra. Maricela Luna Muñoz

A mi Tutora, Florencia Assaneo, por su enseñanza y pasión por la ciencia.

Al HabLab: Lizcano, Ileri, Lily, Richi, Montse y Mora.

A mi Comité Tutor y Jurado, por su apreciable orientación y acompañamiento.

A mis profesores, por su enseñanza y trabajo.

Al Dr. Leopoldo González Santos, por su asesoría y orientación.

A todas y todos los que participaron en los experimentos.

A la Orden del Puma:

- A Marina y Carlos, por enseñarme lo bonito que es tener amigos incondicionales.
- A Ileri, por escucharme y compartir nuestras ansiedades.
- A Barbie, por ser auténticamente hermosa.
- A Poblano, por estar conmigo incluso cuando todo vlv.
- A Denisse, por tu risa y tu sonrisa.
- A Lupita, por ser tan bella persona y siempre echarme ánimos.
- A Angel y Zac, por hacerme Crayoli honoraria.
- A Laura, por ser mi madrina.

A Paty.

A Juan José Venegas Márquez, por estar conmigo en esto.

A mi Familia elegida: Fer Bb (juntas en las mejores momentos), Mari, Caro, Abril, Mariana, Coni, Diana y los demás.

A Pucho y Scrappy Coco, por ser los mejores terapeutas.

A mí, por haber llegado hasta aquí.

# Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Antecedentes y Justificación</b>	<b>2</b>
<b>3. Hipótesis y Objetivos</b>	<b>5</b>
<b>4. Metodología</b>	<b>6</b>
4.1. Participantes . . . . .	6
4.2. Equipo . . . . .	7
4.2.1. Hardware . . . . .	7
4.2.2. Software . . . . .	7
4.3. Tarea de sincronización . . . . .	7
4.4. Estímulo auditivo . . . . .	8
4.5. Phase-Locking value (PLV) . . . . .	10
4.6. Diseños experimentales . . . . .	10
4.6.1. Experimento 1 . . . . .	11
4.6.2. Experimento 2 . . . . .	11
4.6.3. Experimento 3 . . . . .	12
4.7. Análisis estadístico . . . . .	13
4.7.1. Random Forest Clustering . . . . .	13
4.7.2. ANOVA de Medidas Repetidas . . . . .	13
4.7.3. Test no paramétrico circular . . . . .	13
4.7.4. T-Test para Muestras Pareadas . . . . .	13
<b>5. Resultados</b>	<b>15</b>
5.1. Resultados del Experimento 1 . . . . .	15
5.2. Resultados del Experimento 2 . . . . .	18
5.3. Resultados del Experimento 3 . . . . .	20
<b>6. Discusión</b>	<b>23</b>
6.1. Dos grupos con cualidades distintas . . . . .	23
6.2. Efectos del estímulo . . . . .	23
6.3. Efectos del gesto motor . . . . .	24
6.4. Mejora temporal de la sincronía . . . . .	25
<b>7. Conclusiones</b>	<b>26</b>
<b>Referencias</b>	<b>27</b>

# 1. Introducción

La *sincronización auditivo-motora* es la capacidad de acoplar temporalmente un ritmo motor (interno) con un ritmo auditivo (externo) [1, 2]. Nosotros, los seres humanos, estamos acostumbrados a realizar actividades que involucran esta sincronización, como lo son el baile, tocar un instrumento musical o seguir el ritmo de una canción con el pie [3, 4]. Además, esta habilidad cognitiva tiene una naturaleza innata y hasta trivial: los bebés la desarrollan sin un entrenamiento explícito [5], y se modifica conforme envejecemos [6, 7].

De forma reveladora, diversos trabajos previos relacionan el nivel de sincronización auditivo-motora con el desempeño en tareas relacionadas con el habla. Por ejemplo, en un grupo de niños de edad preescolar se observó que su nivel de ritmicidad al tocar las congas a la par con su docente predice su desempeño en tareas relacionadas con la lectura y la percepción fonológica [8]. Es destacable mencionar que únicamente aquellas especies capaces de aprender sus propias vocalizaciones —especies que necesitan de un tutor conespecífico para aprender a vocalizar— pueden efectuar la sincronización auditivo-motora; siendo éste el caso de los seres humanos, algunas aves, elefantes, focas, entre otros [9].

Para que el habla sea comprendida, es crucial que tenga ritmicidad, es decir, que sea presentada con cierta regularidad temporal. La frecuencia de producción silábica promedio entre distintos lenguajes está en torno a los 4.5 Hz (i.e. 4.5 sílabas/segundo) [10]. En [11] se demostró que la producción rítmica del habla mejora la percepción de la misma; en otros trabajos se ha mostrado que al modificar la ritmicidad del habla, se disminuye su inteligibilidad [12, 13]. Curiosamente, en [14] se observó que cuando los participantes escuchan de forma pasiva varios trenes de sílabas presentados con distintas frecuencias, existe un aumento en la sincronía entre regiones de la corteza auditiva y la corteza motora (relacionadas con el habla) alrededor de los 4.5 Hz; es decir, incluso sin realizar movimiento, al escuchar habla se manifiesta una preferencia por la frecuencia natural de la misma.

Acorde con los datos anteriormente descritos, resulta natural hipotetizar la existencia de una relación entre nuestra capacidad de hablar con la sincronización auditivo-motora. Sin embargo, la existencia de dicha relación se encuentra poco estudiada en la literatura actual, la cual está enfocada principalmente en la sincronización del movimiento relacionada con la música y el *finger tapping* (golpeteo con el dedo) [15, 16].

En resumen, la sincronización auditivo-motora es un fenómeno complejo, restringido a algunas especies animales y relacionado con capacidades cognitivas exclusivas del ser humano, como lo es el habla. No obstante, al momento no se cuenta con una descripción completa a nivel comportamental; así mismo, las bases neuronales de este fenómeno no están enteramente descritas. Además, de manera histórica la sincronización auditivo-motora ha sido estudiada sin tomar en consideración las diferencias individuales entre los participantes.

## 2. Antecedentes y Justificación

En un trabajo reciente publicado por el equipo de trabajo que conforma este proyecto [1, 17], se evaluó la sincronización espontánea del habla en un conjunto de sujetos, encontrándose una distribución bimodal en la población. En este tipo de distribución las frecuencias de los valores de la muestra tienen dos picos definidos, aunque algunas partes de la distribución pudieran estar sobrelapadas [18], lo cual implica que estos elementos no están claramente clasificados en alguno de los dos grupos.

Esta significativa distinción se consiguió a través un test conductual, llamado *Test de Sincronización Espontánea del Habla*, o *SSS-Test* (por sus siglas en inglés, *Spontaneous Speech Sincronization Test*), en el cual los sujetos escucharon un tren de sílabas isócronas mientras se les solicitó susurrar de manera continua y repetida la sílaba «ta». De manera interesante, se observó que una parte de la población alineó espontáneamente su propio ritmo de producción silábica con el ritmo del tren de sílabas escuchado (*sujetos de alta sincronía*), mientras que la otra parte no lo hizo (*sujetos de baja sincronía*). En la figura 1 se representa en a: un esquema de este Test y en b: un histograma con los resultados, donde es visible la distribución bimodal de los mismos. Se observó que al realizar el SSS-Test nuevamente después de transcurrir varias semanas, los participantes volvían a ser clasificados en el mismo grupo; esto prueba que ser sujeto de baja o alta sincronía es una característica estable en los participantes.

Lejos de ser ésta tan solo una descripción peculiar de un fenómeno psicofísico, la pertenencia a uno de los dos grupos (alta o baja sincronía) predice características tanto estructurales como funcionales a nivel cerebral en los sujetos. Por un lado en el mismo trabajo [1] se observaron diferencias en la densidad de los tractos de materia blanca que conectan áreas frontales (control motor) con temporales (audición), las cuales pertenecen al fascículo arqueado [19, 20]; se encontró que los sujetos de alta sincronía presentan un mayor volumen en el fascículo arqueado que los de baja sincronía. Por otro lado, mediante un protocolo de magnetoencefalografía, se registró la actividad cerebral durante la escucha pasiva de trenes isócronos de sílabas; se halló que para los sujetos de alta existe una mayor sincronización entre el estímulo acústico y la actividad cerebral en áreas frontales relacionadas con la producción de habla.

Además, se advirtió que los sujetos de alta sincronía se desempeñaron mejor en una tarea de aprendizaje fonológico de palabras con respecto a los de baja. Esto coincide con publicaciones previas que muestran que el volumen del fascículo arqueado correlaciona con un mejor desempeño en el aprendizaje de palabras [21, 22].

Para entender cuál es el mecanismo funcional que subyace a este beneficio cognitivo, es fundamental entender cómo son las diferencias individuales en la sincronía de los participantes a un nivel más general. Una pregunta que permanece abierta después de la publicación de [1] es la siguiente:

*¿Es esta distribución bimodal de la población exclusiva de la percepción y producción del habla o se observará lo mismo al tener estímulos y efectores distintos al habla, pero relacionados con la percepción auditiva y la ejecución de movimiento?*

Con el objetivo de dar respuesta a lo anterior, se buscó literatura relacionada con el análisis de la sincronía auditivo-motora, excluyendo al habla. El estándar de oro para estudiar la sincronización auditivo-motora es el *finger tapping*, o golpeteo con el dedo. Con este test se evalúa la coordinación rítmica entre la percepción de estímulos auditivos, como tonos o clicks, y la ejecución de un movimiento de golpeteo con un dedo de la mano, sobre una superficie.

Pese a que la literatura existente en torno al finger tapping es muy extensa [2, 23], hasta la fecha no ha sido reportada alguna distribución bimodal de la población que distinga a las personas de acuerdo a sus habilidades de sincronización. De acuerdo a lo observado en [1], es posible que esto se deba a las diferencias en las características acústicas de los estímulos empleados al realizar finger tapping, o propiamente a efectos relacionados con la naturaleza del efector empleado. Por ejemplo, la frecuencia empleada en estos experimentos es diferente a la empleada en el SSS-Test; en finger tapping usualmente se utilizan frecuencias de 2 Hz (i.e. 2 golpeteos/segundo) [24, 25, 26], mientras que para el habla la frecuencia promedio es de 4.5 Hz (i.e. 4.5 sílabas/segundo) [10].

Debido a lo anterior, nuestro interés se ha centrado en determinar cómo varían las características de sincronización auditivo-motora individuales en los participantes para distintas combinaciones de estímulo - efector relacionados o no con el habla.

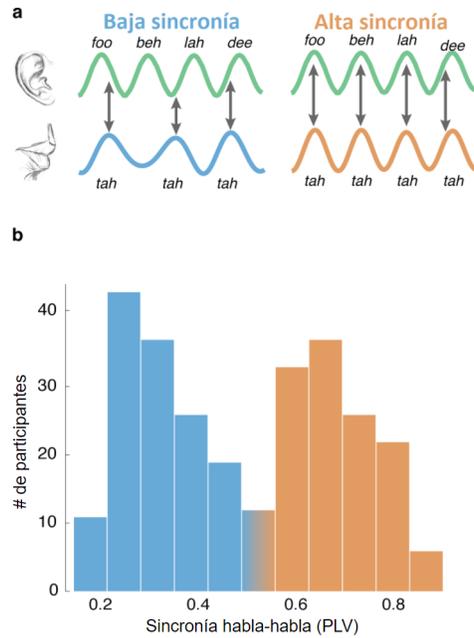


Figura 1: a: Representación esquemática del SSS-Test; mientras algunos participantes alinean espontáneamente su producción silábica con la percibida (sujetos de alta sincronía, en anaranjado), otros no lo hacen (sujetos de baja sincronía, en azul). Los trazos de distintos colores ilustran los cambios de amplitud de las ondas sonoras (en cada máximo se ubica una sílaba); el color verde representa a las sílabas percibidas, mientras que el azul y el anaranjado, las sílabas producidas. Nótese que para los sujetos de alta sincronía las flechas verticales muestran una coincidencia entre las sílabas percibidas y las producidas, mientras que para los sujetos de baja sincronía no existe dicha coincidencia. b: Histograma con los resultados de una cohorte de 273 participantes que realizaron el SSS-Test. Se observa una distribución bimodal de la población. Figura adaptada de [1].

### **3. Hipótesis y Objetivos**

#### **Hipótesis principal**

La distribución bimodal que divide a la población en sujetos de alta sincronía y de baja sincronía, es propia del habla y no se repite cuando se utilizan estímulos y efectores distintos.

#### **Hipótesis alternativa**

La distribución bimodal de la población se conserva con estímulos y efectores distintos al habla, pero no había sido observada por sesgos en las mediciones existentes ya que los estudios han sido enfocados en músicos o utilizando una frecuencia diferente para la presentación de estímulos auditivos.

#### **Objetivos**

- Determinar para un conjunto de individuos el nivel de sincronización auditivo-motora para distintas combinaciones de estímulos (i.e. sílabas/tonos) y efectores (i.e. articuladores del tracto vocal repitiendo una sílaba/aplausos).
- Estudiar la distribución de los valores de sincronización obtenidos y ver si la forma funcional de esta distribución se modifica por la naturaleza de los estímulos acústicos y/o el efector empleado.

## 4. Metodología

### 4.1. Participantes

#### Características generales:

- 64 hablantes nativos de español mexicano; 36 mujeres, 28 hombres; edad promedio: 27 años; rango de edades: 21 – 48 años.
- Todos los participantes respondieron una encuesta para conocer datos demográficos (edad, sexo, lateralidad) y su nivel de entrenamiento en música, en danza, idiomas, videojuegos y deportes.
- Todos los participantes firmaron un consentimiento informado aprobado por el Comité de Ética del Instituto de Neurobiología de la UNAM (protocolo 096.H).
- Se retribuyó económicamente a todos los participantes por su participación. Para el experimento 1 y 2, se ofrecieron \$50.00 (cincuenta pesos 00/100 M.N.) a cada uno. Para el experimento 3, se ofrecieron \$100.00 (cien pesos 00/100 M.N.) a cada uno.
- Los participantes realizaron el experimento en una habitación con aislamiento acústico; se sentaron enfrente de una computadora, se colocaron unos audifonos intrauriculares y se registró su respuesta con el micrófono propio de la computadora.
- Previamente en el laboratorio, los participantes realizaron el SSS-Test clásico con lo cual se conocía a qué grupo pertenecían (alta o baja), respecto de sus valores de sincronía del habla.

**Experimento 1:** Completado por una cohorte de 54 participantes, de los cuales fueron excluidos 3 sujetos por no completar correctamente el experimento. Se conservaron los datos de 51 participantes (26 mujeres, 25 hombres; edad promedio: 27 años; rango de edades: 21 – 37 años).

**Experimento 2:** Completado por una cohorte de 22 participantes, de los cuales 13 participaron también en el Experimento 1. Fueron excluidos 6 sujetos por no completar correctamente el experimento. Se conservaron los datos de 16 participantes (10 mujeres, 6 hombres; edad promedio: 28 años; rango de edades: 23 – 48 años).

**Experimento 3:** Completado por una cohorte de 18 participantes, de los cuales 13 participaron también en el Experimento 1. Fueron excluidos 3 sujetos por no completar correctamente el experimento. Se conservaron los datos de 15 participantes (8 mujeres, 7 hombres; edad promedio: 27 años; rango de edades: 23 – 38 años).

## 4.2. Equipo

### 4.2.1. Hardware

Para realizar los experimentos se utilizó una computadora MAC de escritorio, de 21.5 pulgadas, versión 2017; cuenta con un procesador es un Intel Core i5 de dos núcleos y 2.3 GHz y una memoria de 16 Gb, 2133 MHz DDR4.

El micrófono empleado para registrar las respuestas de los participantes (voz y aplausos) se encuentra integrado en la computadora. Para los estímulos, se emplearon los audífonos especializados ER1, de la marca ETYMOTIC [27], los cuales están diseñados para reproducir en el tímpano la misma respuesta en frecuencias que existe el habla, tal como sería producido en una situación en vivo; al extremo de éstos se colocaron auriculares de espuma de un solo uso (modelo ER1-14A).

### 4.2.2. Software

Para la síntesis de los audios compuestos por trenes de sílabas se utilizó el programa mBrola [28].

Se empleó el software de MATLAB R2020b [29] para programar y presentar los experimentos; para esto último además se hizo uso de Psychtoolbox v3.0.17 [30], que permitió controlar el audio y micrófono de la computadora de una forma precisa.

Para el análisis de las características acústicas de las respuestas producidas por los participantes, se hizo uso de Praat v6.1.34 [31].

Para el análisis estadístico se empleó MATLAB R2020b así como JASP v0.14.1 [32].

El procesamiento de las figuras aquí reportadas se realizó con el software de Inkscape v1.1 [33].

## 4.3. Tarea de sincronización

Partiendo del diseño experimental ya reportado del SSS-Test [1], se construyó la Tarea de sincronización; los pasos son los mismos que en el SSS-Test, variando únicamente los estímulos y efectores empleados. Dependiendo del protocolo, se utilizaron como efectores (o gestos motores): *i*) los articuladores del tracto vocal repitiendo continua y repetidamente la sílaba «ta» (Habla), o *ii*) ambas manos aplaudiendo (Aplausos). A lo largo de este escrito se referirá al efector o gesto motor indistintamente. La descripción de los estímulos empleados se expone a detalle en la sección 4.4.

Cada bloque de la Tarea de sincronización, indistinto a la combinación estímulo-efector, constó de los pasos que se exponen en los siguientes párrafos; los pasos 2 y 3 son repetidos en dos ocasiones. En la figura 2a se esquematiza con un diagrama esta serie de pasos.

1. *Ajuste de volumen*: Los sujetos escucharon un audio que contiene el estímulo acústico invertido temporalmente. Se les pidió que incrementaran el volumen lo más alto posible, haciendo hincapié en que no debían sentir molestia alguna.

2. *Ejemplo de repetición:* Se presentó un audio con trenes del sonido correspondiente al efector (i.e. aplausos o sílaba «ta»), con una frecuencia de 4.3 Hz. Más adelante se describe que es la misma frecuencia que se emplea al comienzo del siguiente paso. Se les pidió que escucharan con atención el audio de ejemplo y que luego realizaran continuamente el gesto motor seleccionado para ese bloque, de forma tal que ensayaran la ejecución del mismo. La duración del ejemplo fue de 10 s.
3. *Tarea de repetición:* Se reprodujo el estímulo y se solicitó de manera explícita a los participantes que sincronizaran el efector correspondiente con el estímulo recibido. La producción del efector fue registrada en un archivo de audio. La duración de este bloque fue de 50 s.

Los estímulos y efectores empleados se especifican en cada experimento.

#### 4.4. Estímulo auditivo

A lo largo de este trabajo, el estímulo empleado fue variando de acuerdo al protocolo del que se trataba. Es importante destacar que la frecuencia a la que se presentó el estímulo se mantuvo estable, lo que cambió fue la unidad acústica que lo compone (sílabas o tonos). En todos los casos, las unidades acústicas se presentaron con una frecuencia de 4.3 Hz (i.e. unidades/segundo), que incrementó en pasos de 0.1 Hz, hasta alcanzar los 4.7 Hz. Cada frecuencia de presentación de estímulos se mantuvo fija durante 10 s (ver figura 2b).

A continuación se describen a detalle los distintos estímulos empleados en este trabajo. A manera de ejemplo, en la figura 2c se muestran fragmentos de las señales acústicas correspondientes a cada estímulo.

**Sílabas:** Tren de 16 sílabas distintas, sintetizadas a través de mBrola, aleatoriamente presentadas de forma que no se repitiera la misma sílaba de forma consecutiva.

**Metronomo:** Tren de tonos de 1 kHz, sintetizado con MATLAB. El ataque de cada tono es pronunciado, es decir, tiene un inicio abrupto en el tiempo.

**Varios tonos:** Tren de tonos entre los 0.5 y los 2 kHz, logarítmicamente separados; los tonos fueron sintetizados con MATLAB y concatenados de forma aleatoria, sin que se repitieran tonos con la misma frecuencia de forma consecutiva. El ataque de cada tono está suavizado, modulado por una función senoidal, lo que le confiere un inicio suave en el tiempo.

Las frecuencias de los tonos fueron las siguientes: 0.50, 0.548, 0.601, 0.659, 0.723, 0.793, 0.870, 0.954, 1.047, 1.148, 1.259, 1.381, 1.515, 1.662, 1.823, 2.00 kHz.

**Mismo tono:** Tren de tonos de 1 kHz, sintetizado a través de MATLAB. El ataque de cada tono está suavizado, modulado también por una función senoidal.

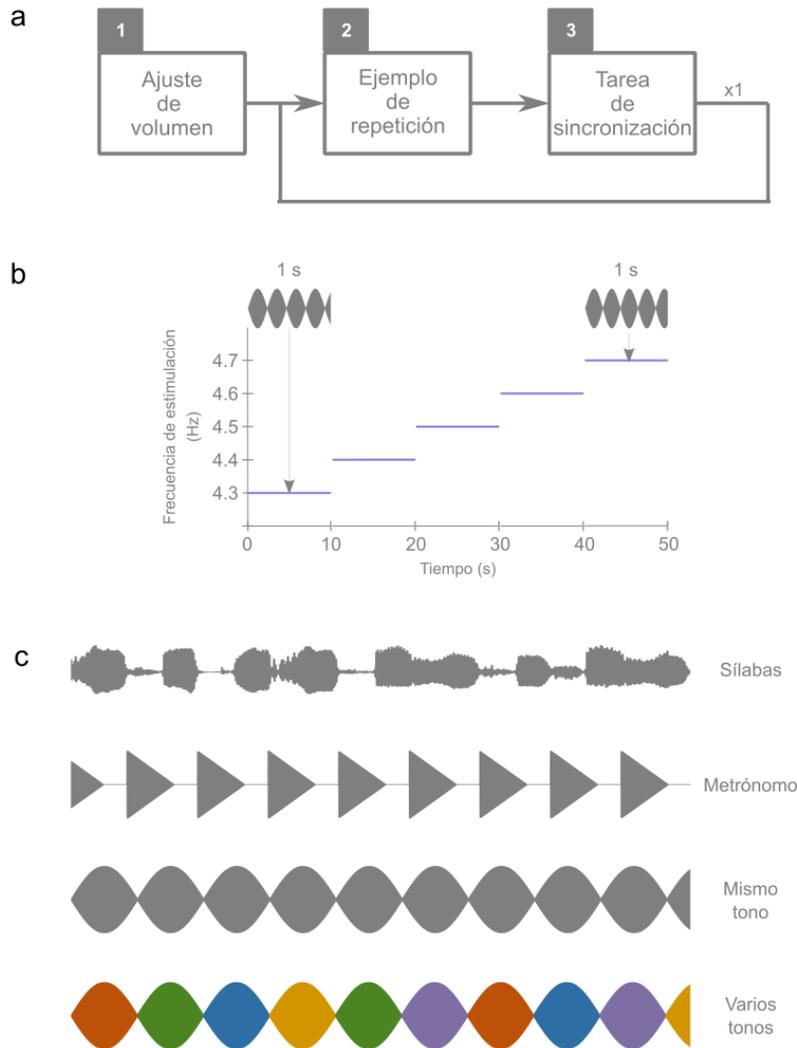


Figura 2: a: Diagrama que esquematiza un bloque con los pasos de una Tarea de sincronización tipo; el segundo y el tercer paso se repiten en dos ocasiones. b: Esquema del incremento en la frecuencia de estimulación; se comienza con 4.3 Hz y se incrementa 0.1 Hz cada 10 s, hasta alcanzar los 4.7 Hz. c: Se muestran como ejemplo extractos de las señales acústicas correspondientes a todos los estímulos empleados en este trabajo; para las *Sílabas*, se pueden observar variaciones en cada unidad acústica; en el caso del *Metrónomo*, el tono de 1 kHz presenta un ataque pronunciado; en *Mismo tono* (1 kHz) y *Varios tonos* (0.5 – 2 kHz), el ataque de cada tono está modulado por una función senoidal, por lo cual no tiene un inicio abrupto.

#### 4.5. Phase-Locking value (PLV)

Para medir la sincronización entre el estímulo y la respuesta producida por el participante (registrada en un archivo de audio), se realizó un análisis para cada uno de los bloques de la Tarea de Sincronización; con este análisis se obtuvo el *Valor de Loqueo de Fase*, o *Phase-Locking Value* (PLV, por sus siglas en inglés), descrito en [1]. Para obtener este valor se realizan los siguientes pasos:

1. Se calcula la envolvente de la señal utilizada como estímulo y de la respuesta registrada por el micrófono. Esto se consigue encontrando el valor absoluto de la transformada de Hilbert de cada una de ellas.
2. Posteriormente, se aplica a cada envolvente un filtro pasa-bandas, con frecuencias de corte entre los 3.3 y 5.7 Hz. Esto se hace para conservar únicamente la banda que está en torno a las frecuencias con la cuales se presentaron los estímulos (entre 4.3 y 4.7 Hz).
3. Luego se vuelve a aplicar la transformada de Hilbert a las envolventes filtradas, con lo que es posible encontrar la fase de ambas señales a lo largo del tiempo.
4. Estas fases se comparan en ventanas de 5 s, con corrimientos de 2 s. A raíz de esto, se obtiene un PLV por ventana.
5. Por último, se promedian todos los PLV obtenidos a lo largo de las señales, con lo cual se tiene un PLV promedio para cada respuesta registrada por sujeto.

El PLV se calcula de acuerdo a la siguiente fórmula:

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

donde

$t$  es el tiempo discreto.

$T$  es el número total de puntos del vector de tiempo.

$\theta_1$  y  $\theta_2$  son las fases del estímulo y de la respuesta, respectivamente, y son funciones que varían a lo largo del tiempo.

Por medio de este análisis se logró determinar el grado de sincronía de cada respuesta con el estímulo correspondiente.

#### 4.6. Diseños experimentales

Se realizó un experimento principal (Experimento 1) y dos experimentos secundarios (Experimento 2 y 3); estos últimos surgieron a partir de los resultados del Experimento 1.

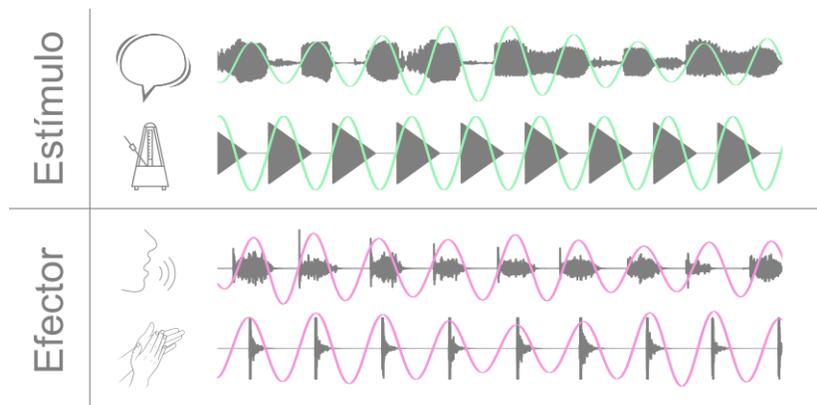


Figura 3: Ejemplo de las señales acústicas correspondientes a los estímulos (Sílabas y Metrónomo) y efectores (Habla y Aplausos). En color gris se muestran las señales crudas a través del tiempo; se muestran las envolventes filtradas de los estímulos (verde) y de los efectores (rosa).

#### 4.6.1. Experimento 1

En este experimento se realizó una prueba conductual en la cual se estudiaron las características generales de sincronización de los participantes. Para ello, se usaron como estímulos las Sílabas y el Metrónomo y como efectores, el Habla (como en el SSS-Test) y los Aplausos. Cabe aclarar que el efector más natural de acuerdo con la literatura de sincronización sensoriomotora previamente descrita es el dedo, realizando golpeteos con el mismo [2, 23]. Sin embargo, pruebas preliminares utilizando este efector nos mostraron que no es posible para los sujetos ejecutar este movimiento con una frecuencia tan alta como la planteada en este protocolo (la tarea se vuelve sumamente incómoda). Debido a lo anterior, decidimos utilizar como efector alternativo los aplausos con las manos [34]; ha sido reportado que la frecuencia promedio de producción de aplausos donde los participantes se sienten cómodos está en torno a los 4 Hz (i.e. 4 aplausos/segundo) [35].

Mediante un procesamiento de las señales acústicas correspondientes al estímulo y el gesto motor, se obtuvo la envolvente de las mismas y luego se calculó el PLV entre ambas; en la figura 3 se muestra a manera de ejemplo un fragmento de las señales correspondientes a los estímulos y efectores; en color verde se observan las envolventes de los estímulos y en rosa las de los efectores.

El experimento constó de cuatro sesiones de tareas de sincronización, una por cada combinación estímulo-efector, que se repitieron en dos ocasiones. El orden en que se presentaron los estímulos permaneció fijo y se muestra en el panel a de la figura 4.

#### 4.6.2. Experimento 2

En la figura 4b se muestra un esquema de las dos sesiones que conformaron este experimento. Cada sesión constó de dos repeticiones de la Tarea de sincronización, con el estímulo y efector correspondientes. Entre sesiones hubo un descanso de entre 15 minutos a 1 día, dentro del cual

los participantes realizaron alguna actividad no relacionada con el experimento.

**Sesión 1: Varios tonos** En esta sesión se utilizó el estímulo de Varios tonos, descrito previamente en la página 8. Los participantes fueron instruidos para susurrar continua y repetidamente la sílaba «ta» en sincronía con el estímulo; por tanto, el efector de esta sesión fue el Habla.

**Sesión 2: Mismo tono** Para la última sesión se utilizó el estímulo de Mismo tono, también descrito anteriormente. El efector utilizado fue el mismo que el de la Sesión 2 de este experimento (Habla).

### 4.6.3. Experimento 3

Los participantes completaron tres sesiones en este experimento, espaciadas por un tiempo que abarcó entre 15 minutos a 1 día. En cada sesión, los participantes realizaron primero una tarea de percepción (Sesión A) o de sincronización (Sesión B), para después completar el SSS-Test Clásico (Sílabas-Habla); en la figura 14c se esquematizan las sesiones, así como los estímulos y efectores empleados. Es importante recalcar que la Sesión A fue presentada en primer lugar a todos los participantes. Las Sesiones B y C fueron presentadas en orden aleatorio a los participantes. A continuación se describen a detalle las características de las sesiones.

**Sesión A: Sincronización de Habla a Metrónomo** En esta sesión, los participantes repitieron el bloque de sincronización de Metrónomo-Habla (susurrar la sílaba «ta» en sincronía con el tren de tonos de 1 kHz), realizado en el Experimento 1; al igual que en éste, se realizaron dos repeticiones del bloque. Al terminar, los sujetos realizaron el SSS-Test Clásico de nueva cuenta.

**Sesión B: Sincronización de Aplausos a Metrónomo** Para esta sesión se realizaron dos repeticiones del bloque de sincronización de Metrónomo-Aplausos (aplaudir en sincronía con el tren de tonos de 1 kHz), mismo que había sido ejecutado en el Experimento 1. De igual forma, al concluir las dos repeticiones, los participantes hicieron el SSS-Test Clásico.

**Sesión C: Escucha Pasiva** Se pidió a los participantes que escucharan con atención el estímulo de Metrónomo mientras no realizaban ningún movimiento. Se les solicitó colocar las palmas de sus manos sobre el escritorio para evitar que las movieran. La presentación del estímulo se realizó en tres ocasiones consecutivas, entre las cuales los participantes debían responder la siguiente pregunta, presionando en el teclado la letra correspondiente: *¿Te parece que el ritmo se aceleró (presiona A), que se desaceleró (presiona D) o que quedó igual (presiona I)?* Al finalizar, los participantes realizaron una vez más el SSS-Test Clásico.

## **4.7. Análisis estadístico**

### **4.7.1. Random Forest Clustering**

Éste es un algoritmo de particionamiento duro, no supervisado, que busca agrupar los datos en varios clústers, donde cada observación pertenece a un solo grupo; el algoritmo genera una matriz de proximidad que calcula un estimado de la distancia entre las observaciones, basado en la frecuencia de las observaciones que quedaron en el mismo nodo [36, 32]. El resultado que arroja este test es el número de clústers que mejor ajusta a los valores de sincronización obtenidos en el experimento 1. Se hizo uso de JASP para realizar este procedimiento.

### **4.7.2. ANOVA de Medidas Repetidas**

El análisis de varianza (ANOVA) está definido como un método para determinar la cantidad de varianza en un conjunto de datos, dividido en uno o más componentes; se asocia a cada uno de estos componentes una fuente específica de varianza, de tal forma que es posible cerciorarse de la magnitud de las contribuciones de cada una de las fuentes a la varianza total [37].

Para el análisis del experimento 1 se realizó un ANOVA de Medidas Repetidas, con dos niveles (Efactor y Estímulo). Para el nivel 1 se utilizaron como factores intra-sujeto el Habla y los Aplausos, mientras que para el nivel 2, las Sílabas y el Metrónomo. El factor inter-sujetos se mantuvo en Grupo.

Dados los resultados arrojados por el ANOVA, se realizaron dos análisis post hoc, corregidos por múltiples comparaciones, empleando el método de Holm.

### **4.7.3. Test no paramétrico circular**

Para analizar las fases se empleó un test no paramétrico para múltiples muestras con medianas iguales, específico para realizar estadística circular, que pertenece al Circular Statistics Toolbox de MATLAB [38].

### **4.7.4. T-Test para Muestras Pareadas**

Este tipo de test permite estimar el tamaño del efecto y probar la hipótesis nula de que la media de la diferencia entre dos observaciones pareadas es igual a cero [32]. Este tipo de análisis se realiza para los experimentos 2 y 3, debido a que las distintas condiciones son aplicadas a los mismos sujetos.

El objetivo de los tests para muestras pareadas es eliminar el máximo número de fuentes de variación no relacionadas con el experimento [37].

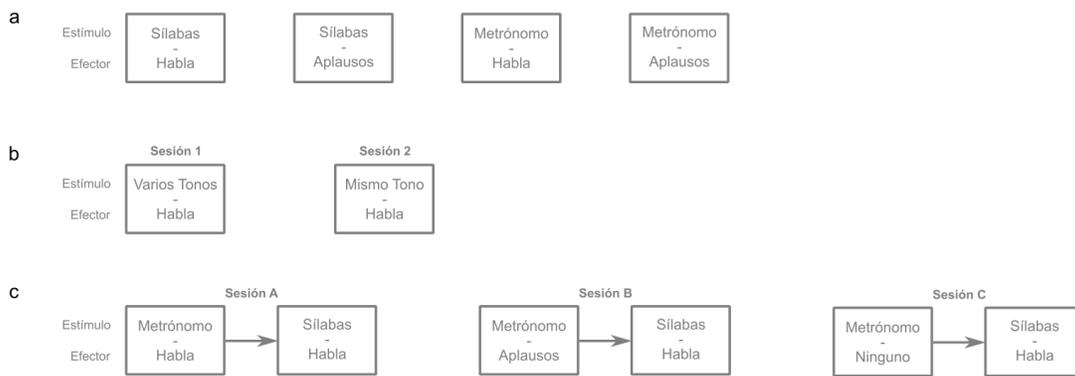


Figura 4: Esquemas que muestran las combinaciones de estímulo-efector que se realizaron en cada una de las sesiones de los tres experimentos que conforman este trabajo. a: Muestra las sesiones que conformaron al Experimento 1; b: Se exponen las dos sesiones que se realizaron en el Experimento 2; c: Se presentan las tres sesiones que compusieron al Experimento 3.

## 5. Resultados

### 5.1. Resultados del Experimento 1

Para constatar que la población general presenta dos comportamientos cuantitativamente distintos en cuanto a sus características de sincronía auditivo-motora, se aplicó a los valores de sincronización obtenidos en el experimento 1 un algoritmo de clasificación llamado *Random Forest Clustering*. Como resultado, se corroboró la naturaleza dicotómica de las características de sincronización para distintos estímulos y efectores; más precisamente, el algoritmo arrojó que el modelo que mejor aproxima estos datos es uno que asume la presencia de dos grupos ( $R^2 = 0.503$ , N.Clusters = 2). Es llamativo que la clasificación obtenida por este método coincide perfectamente con la obtenida a través del SSS-Test; es decir, todos los sujetos previamente clasificados como sujetos de baja sincronía pertenecen al clúster 1, mientras que los de alta lo hacen al clúster 2 (ver figura 5a).

En el panel b de la figura 5 se muestra la distribución del PLV de los participantes del Experimento 1, para cada combinación posible de estímulo-efector. Los puntos de color anaranjado representan a los sujetos de alta sincronía, mientras que en azul se representan los de baja sincronía, de acuerdo a los resultados obtenidos con el algoritmo de clasificación. Se observa en la figura que únicamente para las dos primeras combinaciones existe una separación en dos grupos de los valores de sincronización (i.e. cuando permanece fijo el estímulo de Sílabas); en las dos combinaciones donde el estímulo es el Metrónomo se pierde esta separación.

Una vez confirmada la existencia de dos grupos, se procedió a evaluar el efecto del tipo de estímulo y efector sobre las capacidades de sincronización de cada uno de los grupos (alta y baja sincronía); se calculó un análisis ANOVA de Medidas Repetidas. En éste se utilizaron como factores intra-sujeto para el nivel 1: Efector, con Habla ( $H$ ) y Aplausos ( $A$ ); para el nivel 2: Estímulo, con Sílabas ( $S$ ) y Metrónomo ( $M$ ). El factor inter-sujetos fue el Grupo (alta o baja sincronía). Este análisis arrojó los siguientes resultados:

1. Existe un efecto significativo de Grupo ( $p < 0.001$ ,  $F = 269.096$ ).
2. Existe un efecto significativo de Estímulo ( $p < 0.001$ ,  $F = 68.077$ ).
3. Existen efectos significativos en la interacción Estímulo  $\times$  Grupo ( $p < 0.001$ ,  $F = 29.283$ ).
4. Existen efectos significativos en la interacción Efector  $\times$  Grupo ( $p = 0.006$ ,  $F = 8.288$ ).

Debido a los resultados anteriores, se realizó un análisis post hoc para observar cómo es la interacción Estímulo  $\times$  Grupo y Efector  $\times$  Grupo, la cual arrojó los resultados que se muestran en los paneles c y d de la figura 5 y se describen a continuación:

1. *Interacción Estímulo  $\times$  Grupo*: Ésta se debe a que existen efectos significativos entre los sujetos de alta sincronía (dif.media $_{S-M} = -0.058$ ,  $S.E._{S-M} = 0.026$ ,  $t_{S-M} = -2.213$ ,

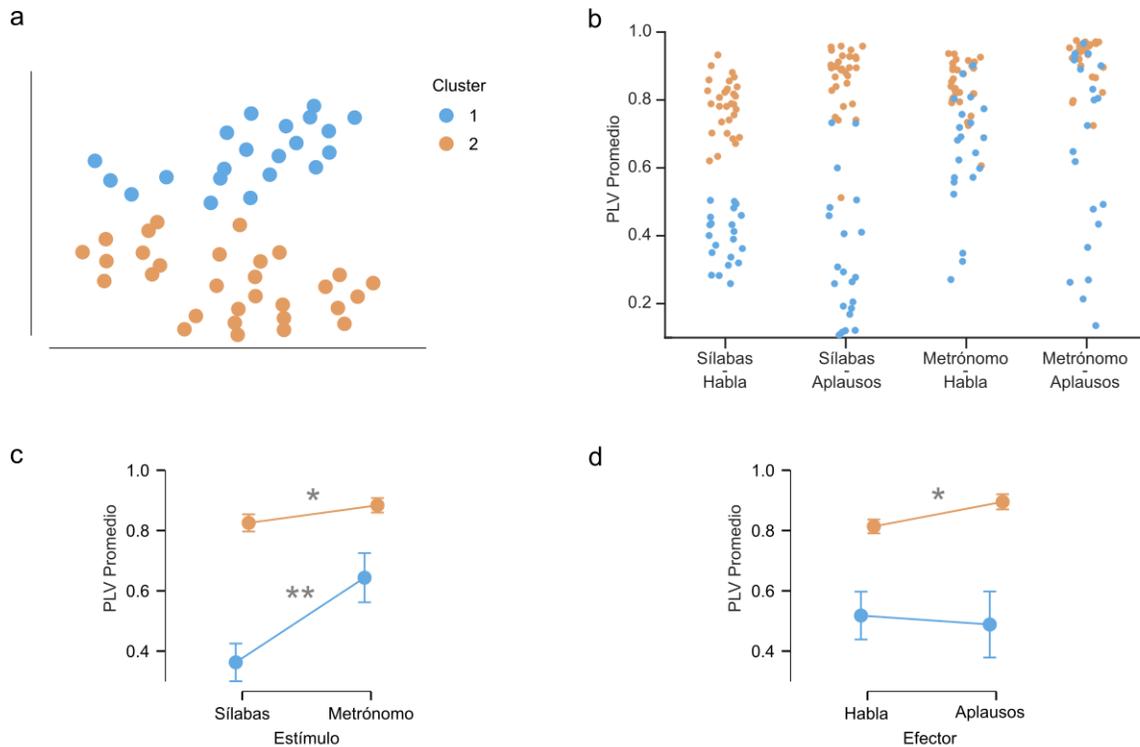


Figura 5: a: Gráfico de tipo t-SNE (*t-distributed stochastic neighbor embedding*, por sus siglas en inglés) en donde se visualiza la clasificación de cada participante dentro de los dos clústers que arroja el algoritmo Random Forest Clustering; en azul se muestran los sujetos que fueron asignados al clúster 1 (previamente clasificados como de baja sincronía) y en anaranjado los que fueron asignados al clúster 2 (previamente clasificados como de alta sincronía). b: Distribución del PLV (eje  $y$ ) de los participantes obtenido en el experimento 1, para las cuatro combinaciones posibles de estímulo-efector (eje  $x$ ). Resultados del análisis post hoc que muestran la interacción c: Estímulo  $\times$  Grupo (c) y d: Efecto  $\times$  Grupo.  $**p < 0.001$ ,  $*p = 0.006$ .

$p_{S-M} = 0.032^1$ ) y en los de baja sincronía (dif.media $_{S-M} = -0.281$ ,  $S.E._{S-M} = 0.032$ ,  $t_{S-M} = -8.907$ ,  $p_{S-M} < 0.001$ ). En ambos grupos, los participantes tienen valores de sincronización mayores cuando son estimulados con el Metrónomo. Sin embargo, en el caso de los sujetos de baja sincronía, esta diferencia está más pronunciada.

2. *Interacción Efecto × Grupo*: Ésta se da por los efectos significativos entre los sujetos de alta sincronía (dif.media $_{H-A} = -0.081$ ,  $S.E._{H-A} = 0.025$ ,  $t_{H-A} = -3.290$ ,  $p_{H-A} = 0.004$ ), no siendo así para los de baja (dif.media $_{H-A} = 0.030$ ,  $S.E._{H-A} = 0.030$ ,  $t_{H-A} = 1.001$ ,  $p_{H-A} = 0.322$ ). Únicamente los sujetos de alta sincronía tienen una ventaja cuando se les pide realizar Aplausos, respecto al Habla.

En vista de que tanto sujetos de alta como de baja se sincronizan mejor cuando el estímulo es el Metrónomo, se realizó un análisis de sus fases en las condiciones de Metrónomo-Habla y Metrónomo-Aplausos, para determinar si existen diferencias significativas entre los grupos que indiquen un comportamiento distinto. Es importante mencionar que para este análisis solo se incluyeron los datos de los participantes que se lograron sincronizar con el Metrónomo (i.e. cuyo  $PLV \geq 0.5$ ), ya que el valor de la fase media carece de significado si no existe sincronía. En el panel a de la figura 6 se muestran, en grados, las fases para los sujetos de alta sincronía; la fase media para la combinación Metrónomo-Habla es de  $3.769^\circ$  (izquierda), mientras que para Metrónomo-Aplausos es de  $28.960^\circ$  (derecha). El panel b muestra, también en grados, las fases de los sujetos de baja sincronía; la fase media en la condición Metrónomo-Habla es de  $15.216^\circ$  (izquierda) y en la condición de Metrónomo-Aplausos es de  $53.950^\circ$  (derecha). Al efectuar un test no paramétrico para múltiples muestras se arrojó que existen diferencias significativas entre los grupos para la combinación Metrónomo-Aplausos ( $p < 0.05$ ), no siendo así para Metrónomo-Habla ( $p = 0.11$ ). Esto nos permite observar que los sujetos de baja sincronía suelen atrasarse al aplaudir más que los de alta.

En suma, las diferencias entre los sujetos de alta y baja sincronía, reportadas en [1] se ven complementadas con los hallazgos realizados en este experimento:

1. Ambos grupos tienen una preferencia por el estímulo Metrónomo, respecto de Sílabas, pero los sujetos de baja muestran diferencias más pronunciadas.
2. Únicamente los sujetos de alta sincronía muestran una preferencia por los Aplausos respecto del Habla.
3. Ambos grupos atrasan la producción de su gesto motor respecto del Metrónomo (contrario a lo reportado en la literatura sobre finger tapping [2, 23]), pero únicamente los sujetos de baja se atrasan significativamente más al aplaudir, respecto a los de alta.

Lo anterior nos invita a investigar cuáles son las características en el estímulo de Metrónomo que en general permite a todos los participantes sincronizarse, perdiéndose la clara distinción

---

<sup>1</sup>Todos los valores de  $p$  reportados en el análisis post hoc han sido corregidos por múltiples comparaciones por utilizando el método de Holm.

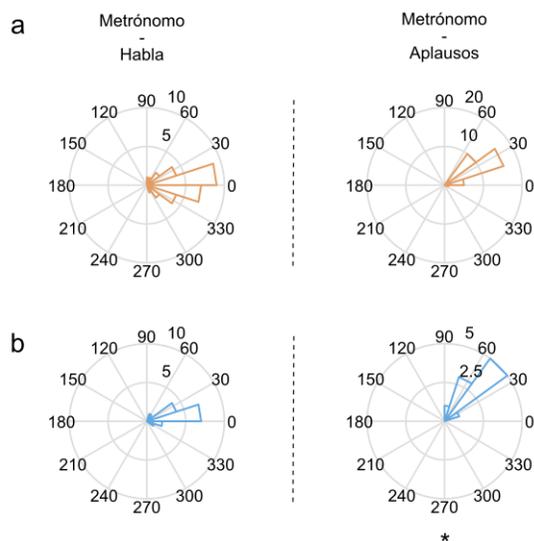


Figura 6: Diagramas de roseta que muestran las fases en grados de los participantes de alta sincronía en las condiciones de a: Metrónomo-Habla (izquierda) y Metrónomo-Aplausos (derecha). Las fases de los participantes de baja sincronía se muestran en b para Metrónomo-Habla (izquierda) y Metrónomo-Aplausos (derecha). \*  $p < 0.05$ .

entre sujetos de alta y baja. ¿Es el ataque pronunciado de cada una de las unidades que componen este estímulo o es la repetición de la misma unidad acústica a lo largo del estímulo? Además, nos preguntamos si es posible mejorar temporalmente las capacidades de sincronía en los sujetos de baja al realizar con antelación una tarea de sincronización donde en efecto sean capaces de sincronizarse.

Para dar respuesta a estos cuestionamientos, se diseñaron los experimentos 2 y 3, cuyos resultados serán expuestos en las subsecciones posteriores.

## 5.2. Resultados del Experimento 2

Por medio de este experimento se determinó qué característica acústica en un estímulo como el Metrónomo impide distinguir entre sujetos de alta y baja sincronía. Para ello se crearon dos nuevos estímulos, variando esencialmente dos parámetros: el ataque se suavizó y se repitió la misma unidad acústica durante todo el estímulo o ésta fue variando a lo largo del estímulo; esto permitió explorar si es lo pronunciado del ataque o es la repetición de la misma unidad acústica, lo que favorece la sincronía. Se pidió a un grupo de participantes de baja sincronía que acoplaran el ritmo de producción silábica (ver figura 7a) con el del estímulo correspondiente (Mismo tono o Varios tonos); posteriormente se compararon sus valores de sincronía con los obtenidos en la combinación de Sílabas-Habla del experimento 1, misma que recupera la distinción entre los dos grupos (SSS-Test).

Se realizaron dos T-tests para Muestras Pareadas que arrojaron los siguientes resultados:

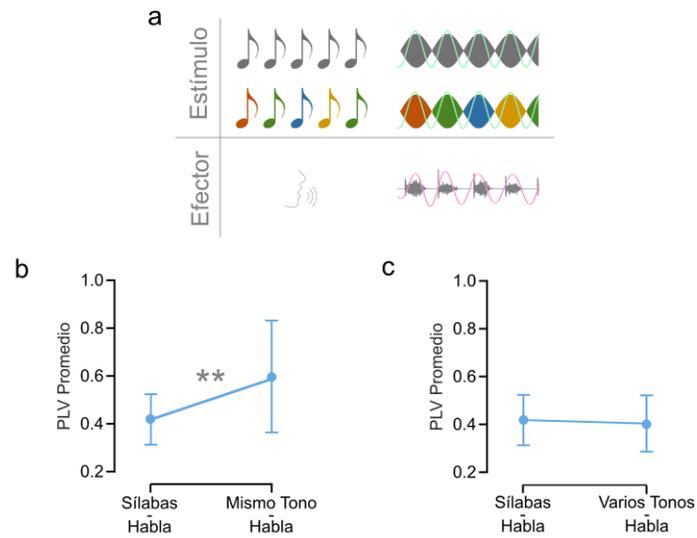


Figura 7: a: Esquema que muestra los estímulos empleados en el experimento 2, Mismo tono y Varios tonos, así como el efector empleado, Habla; en color verde se muestra la envolvente de los estímulos y en rosa la del efector. b: Resultados de la prueba t para muestras pareadas donde se comparan los valores de sincronía de la combinación Sílabas-Habla (SSS-Test) con Mismo Tono-Habla. c: Resultados de la prueba t para muestras pareadas de la comparación entre la condición Sílabas-Habla y Varios Tonos-Habla; b: Se compara el SSS-Test con la combinación de Mismo tono en sincronía con el Habla.  $** p < 0.001$ .

1. Al comparar la condición de Sílabas-Habla (*S.H.*) con Mismo tono-Habla (*Mt.H.*), se encontró que existen diferencias significativas en el desempeño de los participantes ( $t_{S.H.-Mt.H.} = -4.077$ ,  $p_{S.H.-Mt.H.} < 0.001$ ). Los sujetos de baja pueden aumentar su sincronía cuando se repite la misma unidad acústica durante todo el estímulo (ver figura 7b).
2. En la comparación entre Sílabas-Habla (*S.H.*) y Varios tonos-Habla (*Vt.H.*) no se hallaron diferencias significativas ( $t_{S.H.-Vt.H.} = 0.418$ ,  $p_{S.H.-Vt.H.} = 0.682$ ), lo cual apoya lo observado con anterioridad. Esto es, para que los participantes de baja tengan un mejor PLV en una Tarea de sincronía es suficiente con que se repita la misma unidad a lo largo de todo el estímulo (ver figura 7c).

Se infiere que, debido a lo anterior, la distinción entre sujetos de alta o de baja sincronía no había sido reportada en la literatura de sincronización auditivo-motora considerando que clásicamente se utilizan estímulos de tipo Metrónomo, los cuales repiten la misma unidad acústica durante todo el experimento.

### 5.3. Resultados del Experimento 3

A través de este experimento, se propuso definir si es posible mejorar las capacidades de sincronización de los sujetos de baja sincronía de forma temporal. Para este propósito, se solicitó a un grupo de participantes de baja sincronía que realizaran el SSS-Test después de haber completado una Tarea de sincronización —usando el mismo efector (post Metrónomo-Habla, *pM.H.*) o un efector diferente al de Habla (post Metrónomo-Aplausos, *pM.A.*)— y se compararon sus PLVs con los obtenidos en la condición Sílabas-Habla (*S.H.*) del experimento 1 (SSS-Test). Para descartar si la Escucha Pasiva (*E.P.*) del Metrónomo era suficiente para facilitar la sincronía, se pidió a los mismos participantes que efectuaran el SSS-Test al finalizar una Tarea de percepción en la cual escucharon el estímulo sin realizar ningún movimiento.

Para visualizar si alguna de las condiciones proporciona diferencias respecto del valor de sincronía obtenido en el SSS-Test, se realizaron tres T-tests para Muestras Pareadas. Estas pruebas arrojaron los siguientes resultados:

1. Existen diferencias estadísticamente significativas en el PLV obtenido en el SSS-Test posterior a la tarea de sincronización con Metrónomo-Aplausos ( $t_{S.H.-pM.A.} = -3.025$ ,  $p_{S.H.-pM.A.} = 0.009$ ) (Ver figura 8a).
2. De misma cuenta, hay diferencias significativas en los valores de sincronía del SSS-Test posterior a la tarea de sincronización con Metrónomo-Habla ( $t_{S.H.-pM.H.} = -2.469$ ,  $p_{S.H.-pM.H.} = 0.027$ ) (Ver figura 8b).
3. La diferencia del PLV obtenido en el SSS-Test después de haber realizado la tarea de Escucha Pasiva no fue significativa ( $t_{S.H.-E.P.} = -1.363$ ,  $p_{S.H.-E.P.} = 0.194$ ) (Ver figura 8c).

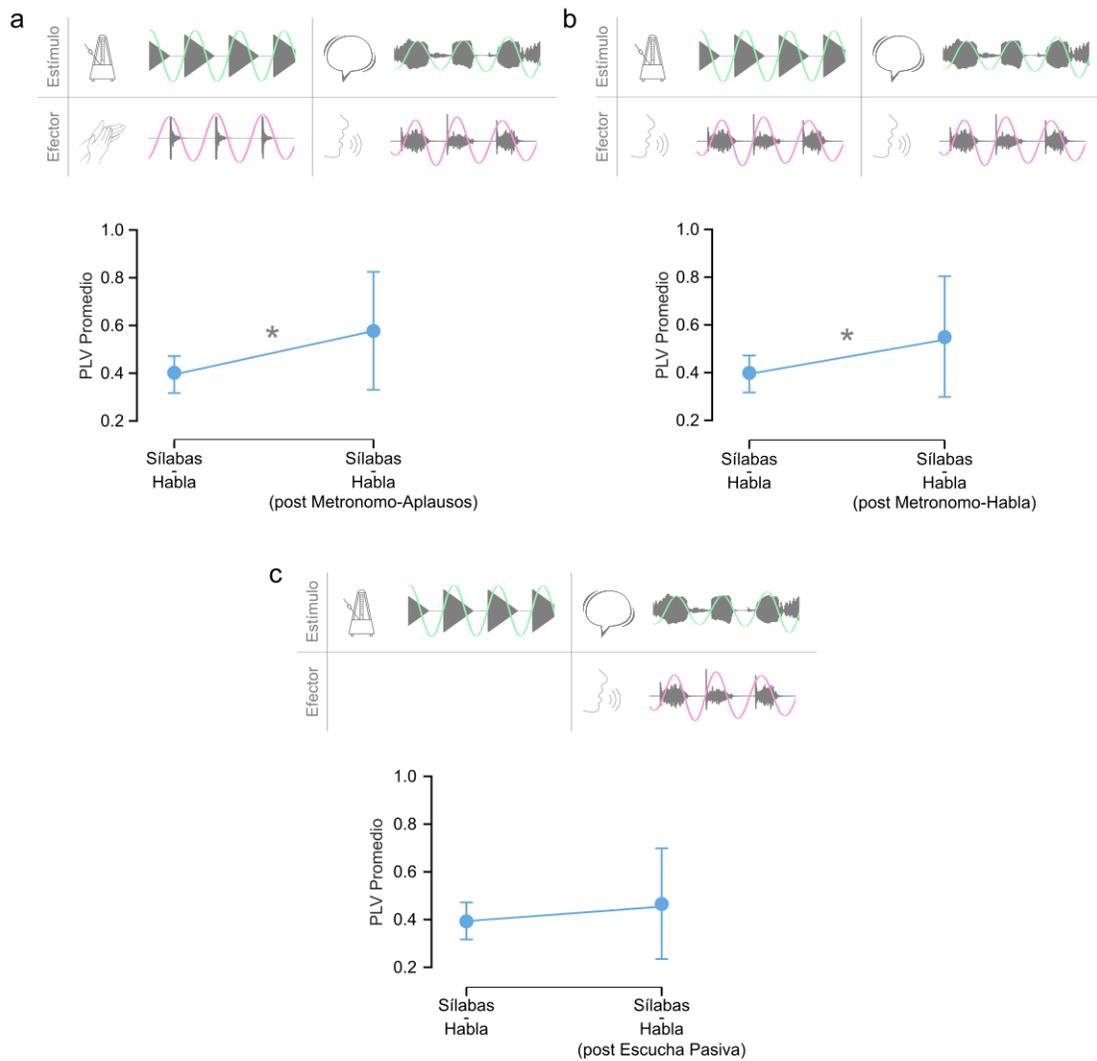


Figura 8: a: En la parte superior se muestra un esquema con la tarea de sincronización Metrónomo-Aplausos seguida por el SSS-Test del experimento 3; en la parte inferior se grafican los resultados de la prueba t que compara el PLV obtenido en el SSS-Test con el obtenido en el mismo test posterior a la sincronización con Metrónomo-Aplausos. b: En la parte superior se presenta el esquema de la tarea de sincronización Metrónomo-Habla seguida por el SSS-Test, realizado en el experimento 3; en la parte inferior se muestran los resultados de la prueba t donde se compara el PLV obtenido en el SSS-Test y el alcanzado después de realizar la tarea de sincronización con Metrónomo-Habla. c: En la parte superior se muestra un esquema de la Escucha Pasiva seguida por el SSS-Test, efectuados en el experimento 3; en la parte inferior se presentan los resultados de la prueba t que compara los valores de sincronía del SSS-Test con aquellos obtenidos después de realizar la tarea de percepción (escucha pasiva). Para las condiciones de b: Metrónomo-Aplausos y c: Metrónomo-Habla, las diferencias son significativas; esto no sucede en la condición de a: Escucha Pasiva.  $*p < 0.05$ .

Gracias a estos resultados podemos afirmar que los sujetos de baja pueden aumentar su sincronía en el SSS-Test de manera transitoria si antes realizan una Tarea de sincronización con un estímulo como el Metrónomo, no importando si el efector que utilizan está o no relacionado con el Habla.

## 6. Discusión

### 6.1. Dos grupos con cualidades distintas

Como ya hemos mencionado, la clasificación de la población en dos grupos distinguidos en sus habilidades de sincronización no se había reportado con anterioridad en la vasta literatura sobre sincronización auditivo-motora [2, 23]; las diferencias entre ambos se extienden a través de distintos efectores y estímulos, sugiriendo que existen diferencias generales en la capacidad de sincronización en la población. Es sobresaliente la clasificación que el algoritmo Random Forest Clustering proporcionó debido a que, sin tomar en cuenta la previa clasificación del SSS-Test, corrobora que en efecto existen dos grupos diferenciados en la población [36]. De hecho, es todavía más impactante que la clasificación que hizo de los 51 participantes coincide perfectamente con la que proporcionó el SSS-Test.

### 6.2. Efectos del estímulo

Anteriormente se señaló que el estímulo influye en la aparición de diferencias significativas en la sincronía de ambos grupos (alta y baja): el Metrónomo favorece la sincronía respecto de las Sílabas. Las diferencias entre estos estímulos son dos: el Metrónomo presenta unidades acústicas que se repiten a lo largo de toda la señal y tienen un ataque pronunciado, mientras que las Sílabas presentan 16 unidades acústicas diferentes entre sí, con la peculiaridad de que no se repite la misma sílaba de forma consecutiva y el ataque de cada una es suave. En el experimento 2 se comprobó que *es la presentación repetitiva de la misma unidad acústica a lo largo del estímulo (y no el ataque pronunciado de las unidades acústicas repetidas) lo que favorece la sincronía*, dado que solo con el estímulo Mismo tono, con ataque suavizado, un subgrupo de participantes de baja se logró sincronizar (ver figura 7b y c).

El hecho de que el uso del Metrónomo tenga un efecto con el que se pierde la distinción de los dos grupos es clave: hasta donde es de nuestro conocimiento, todos los experimentos clásicos donde se estudia la sincronización auditivo-motora utilizan precisamente estímulos en los que la misma unidad acústica es repetida [2, 23]. Probablemente esta sea la razón por la cual previamente no se notificó sobre la existencia de dos grupos en la población, con diferentes cualidades en su capacidad de sincronizarse.

Hay que tener presente que el incremento en la mejoría de la sincronización con el Metrónomo es más acentuado en los sujetos de baja sincronía respecto de los de alta. Sería interesante estudiar en el futuro cuál es el origen de este efecto, cuál es la relación que guarda con las diferencias en la estructura y funcionalidad del cerebro de sujetos de alta y baja sincronía, reportadas en [1] y a partir de qué punto un estímulo puede o no propiciar la sincronía en los sujetos de baja.

### 6.3. Efectos del gesto motor

De acuerdo a lo observado, existe una mejor disposición de los sujetos de alta para sincronizarse con Aplausos respecto del Habla, mientras los sujetos de baja no presentan diferencias entre los efectores (ver figura 5d). Esto tal vez esté relacionado con las diferencias en la retroalimentación táctil y cinética que se obtiene al realizar ambos gestos. La superficie que entra en contacto entre las palmas de las manos al aplaudir es mayor que la de los elementos articuladores del habla; por otro lado, la cantidad de movimiento que se realiza al aplaudir es mayor que la efectuada al hablar. Se sabe que el sistema motor está íntimamente relacionado con la predicción temporal pues refina la percepción sensorial, mejorando la atención a eventos relevantes en el tiempo [39]. Posiblemente el hecho de que la información motora relacionada a los Aplausos sea más intensa que la del Habla, guarde relación con la mejor sincronía al aplaudir en aquellos participantes que pueden sincronizarse con una variedad de estímulos (los sujetos de alta). Las causas que suscitan este comportamiento quedan pendientes de ser estudiadas de forma posterior a este trabajo.

En el análisis de las diferencias en las fases de los participantes en las condiciones con Metrónomo —el único estímulo que permite tanto a sujetos de alta como de baja sincronizarse realmente— se puede observar que el signo de las fases medias indica en todos los casos una asincronía positiva entre la producción del gesto motor (Habla o Aplausos) respecto del estímulo (Metrónomo). Se observa que solo hay diferencias significativas entre grupos en las fases para la combinación Metrónomo-Aplausos: los sujetos de baja sincronía se atrasan más al aplaudir que los de alta.

Está reportado en experimentos de finger tapping que los participantes tienden a realizar el golpeteo con el dedo en anticipación al estímulo, como una predicción motora de un evento auditivo, lo cual se suele llamar asincronía negativa [2, 23], aunque en este trabajo nos referimos a ella como la fase entre el gesto motor y la señal auditiva. Algunos trabajos sugieren que esta respuesta anticipadora está relacionada con lograr la sincronía entre la percepción de la señal acústica y de la retroalimentación táctil y cinética que proviene del golpe con el dedo [40]; esto es, la fase entre la respuesta motora y el estímulo parece estar provocada por las diferencias en la conducción nerviosa de la señal acústica y la retroalimentación táctil y cinética. No obstante hablamos de asincronías positivas en la producción de Sílabas o Aplausos en estos experimentos, es esencial recalcar que aún se trata de un gesto de predicción si las asincronías tienen una duración menor al tiempo de reacción (de aproximadamente 150 ms) [23].

Es interesante que en estos resultados todos los participantes presenten asincronías positivas, lo cual es consistente con la literatura; se ha reportado que las asincronías positivas suceden cuando el gesto motor se acerca a los límites biomecánicos del efector [41, 23]. Por ejemplo, en finger tapping la asincronía tiende a desaparecer (se acerca a cero) cuando la velocidad del movimiento se incrementa [2]. La explicación a este fenómeno queda pendiente para trabajos posteriores.

Por otro lado, la tendencia de los participantes de baja a tener una mayor fase (i.e. una

asincronía positiva mayor) cuando aplauden con el Metrónomo podría deberse a que las limitaciones de la periferia tienen un impacto más grande en ellos por las diferencias estructurales en su cerebro. De igual manera, la descripción particular de esto escapa de los objetivos de este trabajo.

#### 6.4. Mejora temporal de la sincronía

En los resultados presentados en los paneles a y b de la figura 8 observamos que sincronía de los participantes de baja en el SSS-Test aumenta si previamente se realiza una tarea de sincronización de forma efectiva. Teniendo en cuenta que en [1] se reportó que la clasificación del SSS-Test (alta o baja sincronía) se replica en los participantes al realizar esta prueba en sesiones separadas por varias semanas, este hallazgo resulta muy atractivo. Está reportado que la información motora produce beneficios en la percepción temporal [39]. El incremento en los valores de sincronía que presentaron los participantes de baja en el experimento 3 podría deberse a este tipo de mecanismos de mejoramiento en la predicción temporal relacionados con el gesto motor.

No obstante se ha mostrado que durante la escucha pasiva de estímulos rítmicos se reclutan áreas motoras [42, 43], parece ser que esta actividad provocada por la misma no tiene la fuerza suficiente para contribuir en la mejora temporal de la sincronía de los sujetos de baja (figura 8c).

Otro aspecto interesante de estos resultados se encuentra al observar que no hay diferencias significativas en la mejoría cuando los participantes realizan la tarea de sincronía con el mismo efector del SSS-Test (Habla) o con uno diferente (Aplausos). Estos resultados apoyan la teoría sobre la existencia de un cronómetro central, que permite realizar predicciones temporales sensoriomotoras de forma independiente al efector motor, incluyendo los efectos de arrastre (i.e. la mejoría proporcionada a los sujetos de baja) [44].

Queda por evaluar si esta mejora en la sincronización de los sujetos de baja se extiende a capacidades cognitivas en las cuales ya han sido reportadas diferencias entre los dos grupos [45], tales como el Aprendizaje Estadístico de Palabras (*Statistical Learning*, o *SL*).

## 7. Conclusiones

Este trabajo permitió observar que la población general se distingue en efecto en dos grupos, con características cualitativamente diferentes en cuanto a sus habilidades de sincronía. Es muy importante recalcar que fue el estudio de las capacidades individuales de sincronización lo que permitió determinar que los sujetos de alta y baja tienen un comportamiento distinto que se extiende a otros dominios, incluyendo el habla, mismo que no había sido reportado antes tras años de estudio de la sincronización auditivo-motora.

Incluir estas diferencias individuales en trabajos futuros podría ser crucial para observar fenómenos que se enmascaran al asumir la igualdad de cualidades en la población general.

## Referencias

- [1] M. F. Assaneo, P. Ripollés, J. Orpella, W. M. Lin, R. de Diego-Balaguer, and D. Poeppel, “Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning,” *Nature Neuroscience*, vol. 22, no. 4, pp. 627–632, 2019.
- [2] B. H. Repp, “Sensorimotor synchronization: A review of the tapping literature,” *Psychonomic Bulletin and Review*, vol. 12, no. 6, pp. 969–992, 2005.
- [3] C. Lezama-Espinosa and H. L. Hernandez-Montiel, “Neuroscience of the auditory-motor system: How does sound interact with movement?” *Behavioural Brain Research*, vol. 384, no. January, p. 112535, 2020. [Online]. Available: <https://doi.org/10.1016/j.bbr.2020.112535>
- [4] S. D. Bella, M. Berkowska, and J. Sowiński, “Moving to the beat and singing are linked in humans,” *Frontiers in Human Neuroscience*, vol. 9, no. DEC, p. 663, dec 2015. [Online]. Available: [www.frontiersin.org](http://www.frontiersin.org)
- [5] W. S. Condon and L. W. Sander, “Neonate movement is synchronized with adult speech: Interactional participation and language acquisition,” *Science*, 1974.
- [6] M. S. Cousins, C. Corrow, M. Finn, and J. D. Salamone, “Temporal measures of human finger tapping: Effects of age,” *Pharmacology Biochemistry and Behavior*, vol. 59, no. 2, pp. 445–449, feb 1998.
- [7] T. Goossens, C. Vercammen, J. Wouters, and A. van Wieringen, “Aging affects neural synchronization to speech-related acoustic modulations,” *Frontiers in Aging Neuroscience*, 2016.
- [8] K. Woodruff Carr, T. White-Schwoch, A. T. Tierney, D. L. Strait, and N. Kraus, “Beat synchronization predicts neural speech encoding and reading readiness in preschoolers,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 40, pp. 14559–14564, oct 2014. [Online]. Available: <https://www.pnas.org/content/111/40/14559><https://www.pnas.org/content/111/40/14559.abstract><http://www.pnas.org/lookup/doi/10.1073/pnas.1406219111>
- [9] A. D. Patel, “The Evolutionary Biology of Musical Rhythm: Was Darwin Wrong?” *PLoS Biology*, 2014.
- [10] N. Ding, A. D. Patel, L. Chen, H. Butler, C. Luo, and D. Poeppel, “Temporal modulations in speech and music,” *Neuroscience and Biobehavioral Reviews*, vol. 81, pp. 181–187, 2017. [Online]. Available: <https://doi.org/10.1016/j.neubiorev.2017.02.011>

- [11] M. F. Assaneo, J. M. Rimmele, Y. Sanz Perl, and D. Poeppel, “Speaking rhythmically can shape hearing,” *Nature Human Behaviour*, 2020. [Online]. Available: <http://dx.doi.org/10.1038/s41562-020-00962-0>
- [12] D. Henao, M. Navarrete, M. Valderrama, and M. Le Van Quyen, “Entrainment and synchronization of brain oscillations to auditory stimulations,” *Neuroscience Research*, vol. 156, pp. 271–278, 2020. [Online]. Available: <https://doi.org/10.1016/j.neures.2020.03.004>
- [13] D. Poeppel and M. F. Assaneo, “Speech rhythms and their neural foundations,” *Nature Reviews Neuroscience*, vol. 21, no. 6, pp. 322–334, 2020. [Online]. Available: <http://dx.doi.org/10.1038/s41583-020-0304-4>
- [14] M. F. Assaneo and D. Poeppel, “The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm,” *Science Advances*, vol. 4, no. 2, pp. 1–10, 2018.
- [15] R. J. Zatorre, J. L. Chen, and V. B. Penhune, “When the brain plays music: Auditory-motor interactions in music perception and production,” *Nature Reviews Neuroscience*, vol. 8, no. 7, pp. 547–558, 2007.
- [16] H. Merchant, J. Grahm, L. Trainor, M. Rohrmeier, and W. T. Fitch, “Finding the beat: A neural perspective across humans and non-human primates,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, no. 1664, 2015.
- [17] F. Lizcano-Cortés, I. Gómez-Varela, C. Mares, P. Wallisch, J. Orpella, D. Poeppel, P. Ripollés, and M. F. Assaneo, “Speech-to-Speech Synchronization protocol to classify human participants as high or low auditory-motor synchronizers,” *STAR Protocols*, vol. 3, no. 2, p. 101248, jun 2022. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2666166722001289>
- [18] M. J. P. Hale W. G., Saunders V. A., “Collins Dictionary of Biology,” 2005. [Online]. Available: <https://medical-dictionary.thefreedictionary.com/bimodal+distribution>
- [19] J. K. Rilling, M. F. Glasser, T. M. Preuss, X. Ma, T. Zhao, X. Hu, and T. E. Behrens, “The evolution of the arcuate fasciculus revealed with comparative DTI,” *Nature Neuroscience*, vol. 11, no. 4, pp. 426–428, apr 2008. [Online]. Available: <http://www.nature.com/natureneuroscience>
- [20] N. Geschwind, “The Organization of Language and the Brain,” Tech. Rep. 3961, 1970.
- [21] P. Ripollés, D. Biel, C. Peñaloza, J. Kaufmann, J. Marco-Pallarés, T. Noesselt, and A. Rodríguez-Fornells, “Strength of temporal white matter pathways predicts semantic learning,” *Journal of Neuroscience*, vol. 37, no. 46, pp. 11 101–11 113, nov 2017. [Online]. Available: <https://www.jneurosci.org/content/37/46/11101><https://www.jneurosci.org/content/37/46/11101.abstract>

- [22] D. López-Barroso, M. Catani, P. Ripollés, F. Dell’Acqua, A. Rodríguez-Fornells, and R. De Diego-Balaguer, “Word learning is mediated by the left arcuate fasciculus,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 32, pp. 13 168–13 173, aug 2013. [Online]. Available: [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301696110/-/DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301696110/-/DCSupplemental).[www.pnas.org/cgi/doi/10.1073/pnas.1301696110](http://www.pnas.org/cgi/doi/10.1073/pnas.1301696110)
- [23] B. H. Repp and Y. H. Su, “Sensorimotor synchronization: A review of recent research (2006-2012),” *Psychonomic Bulletin and Review*, vol. 20, no. 3, pp. 403–452, 2013.
- [24] D. Rose, D. J. Cameron, P. J. Lovatt, J. A. Grahn, and L. E. Annett, “Comparison of spontaneous motor tempo during finger tapping, toe tapping and stepping on the spot in people with and without parkinson’s disease,” *Journal of Movement Disorders*, vol. 13, no. 1, pp. 47–56, jan 2020. [Online]. Available: [/pmc/articles/PMC6987525/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6987525/](https://pubmed.ncbi.nlm.nih.gov/32011111/)
- [25] D. Moelants, “Dance music, movement and tempo preferences,” in *Proceedings of the 5th Triennial ESCOM Conference*, K. R., Ed., 2003. [Online]. Available: <http://www.clubdjforum.com/genreintro.html>
- [26] B. H. Repp, “Rate limits of sensorimotor synchronization,” *Advances in Cognitive Psychology*, 2006. [Online]. Available: <http://ac-psych.org/en/download-pdf/page/14/volume/2/issue/2/id/16>
- [27] Etymotic, “ER1 Insert Earphones.” [Online]. Available: <https://www.etymotic.com/product/er1-insert-earphones/>
- [28] T. Dutoit, V. Pagel, N. Pierret, E. Bataille, and O. van der Vrecken, “MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes,” *International Conference on Spoken Language Processing, ICSLP, Proceedings*, vol. 3, pp. 1393–1396, 1996.
- [29] The MathWorks Inc., “MATLAB R2020b[Computer software].” [Online]. Available: <https://www.mathworks.com/products/matlab.html>
- [30] D. Brainard, “The psychophysics toolbox,” *Spatial Vision*, vol. 10, pp. 433–436, 1997. [Online]. Available: <http://faculty.washington.edu/jpalmer/files/forMJ/NewPapersForMJ/brainard1997PsychToolbox.pdf>
- [31] Paul Boersma and David Weenink, “PRAAT: doing phonetics by computer,” 2022. [Online]. Available: <http://www.praat.org>
- [32] JASP Team, “JASP (Version 0.14.1)[Computer software],” 2020. [Online]. Available: <https://jasp-stats.org/>
- [33] Inkscape Project, “Inkscape.” [Online]. Available: <https://inkscape.org>

- [34] P. Tranchant, D. T. Vuvan, and I. Peretz, “Keeping the beat: A large sample study of bouncing and clapping to music,” *PLoS ONE*, vol. 11, no. 7, jul 2016.
- [35] B. H. Repp, “The sound of two hands clapping: An exploratory study,” *The Journal of the Acoustical Society of America*, vol. 81, no. 4, p. 1100, jun 1987. [Online]. Available: <https://asa.scitation.org/doi/abs/10.1121/1.394630>
- [36] L. Breiman, “Random Forests,” *Machine Learning 2001 45:1*, vol. 45, no. 1, pp. 5–32, oct 2001. [Online]. Available: <https://link.springer.com/article/10.1023/A:1010933404324>
- [37] W. W. Daniel and C. L. Cross, *Biostatistics: A Foundation for Analysis in the Health Sciences*, 10th ed., Wiley, Ed., 2013.
- [38] P. Berens, “CircStat: A MATLAB Toolbox for Circular Statistics,” *Journal of Statistical Software*, vol. 31, no. 10, pp. 1–21, sep 2009. [Online]. Available: <https://www.jstatsoft.org/index.php/jss/article/view/v031i10>
- [39] B. Morillon, C. E. Schroeder, and V. Wyart, “Motor contributions to the temporal precision of auditory attention,” *Nature Communications*, vol. 5, 2014.
- [40] G. Aschersleben, “Temporal Control of Movements in Sensorimotor Synchronization,” *Brain and Cognition*, vol. 48, no. 1, pp. 66–79, feb 2002.
- [41] S. Fujii, M. Hirashima, K. Kudo, T. Ohtsuki, Y. Nakamura, and S. Oda, “Synchronization Error of Drum Kit Playing with a Metronome at Different Tempi by Professional Drummers,” *Music Perception*, vol. 28, no. 5, pp. 491–503, jun 2011. [Online]. Available: <https://online.ucpress.edu/mp/article/28/5/491/62515/Synchronization-Error-of-Drum-Kit-Playing-with-a>
- [42] J. A. Grahn and M. Brett, “Rhythm and Beat Perception in Motor Areas of the Brain,” *Journal of Cognitive Neuroscience*, vol. 19, no. 5, pp. 893–906, may 2007. [Online]. Available: <https://direct.mit.edu/jocn/article/19/5/893/4337/Rhythm-and-Beat-Perception-in-Motor-Areas-of-the>
- [43] J. L. Chen, V. B. Penhune, and R. J. Zatorre, “Listening to Musical Rhythms Recruits Motor Regions of the Brain,” *Cerebral Cortex*, vol. 18, no. 12, pp. 2844–2854, dec 2008. [Online]. Available: <https://academic.oup.com/cercor/article/18/12/2844/362955>
- [44] J. M. Ross and R. Balasubramaniam, “Physical and neural entrainment to rhythm: Human sensorimotor coordination across tasks and effector systems,” *Frontiers in Human Neuroscience*, vol. 8, no. AUG, pp. 1–6, 2014.
- [45] M. F. Assaneo, J. Orpella, P. Ripollés, L. Noejovich, D. López-Barroso, R. de Diego-Balaguer, and D. Poeppel, “Population-level differences in the neural substrates

supporting Statistical Learning” *bioRxiv*, p. 2020.07.03.187260, jul 2020. [Online].  
Available: <https://doi.org/10.1101/2020.07.03.187260>