



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

MAESTRÍA EN CIENCIAS (NEUROBIOLOGÍA)
INSTITUTO DE NEUROBIOLOGÍA

MODULACIÓN DEL APRENDIZAJE ESTADÍSTICO
MEDIANTE LA VARIACIÓN DE LA ESTRUCTURA RÍTMICA
DEL ESTÍMULO

TESIS

Que para optar por el grado de:
MAESTRA EN CIENCIAS

PRESENTA:

IRERI MELISSA GÓMEZ VARELA

TUTORA PRINCIPAL:

DRA. MARÍA FLORENCIA ASSANEO (INB, UNAM)

COMITÉ TUTOR:

DR. HUGO MERCHANT NANCY (INB, UNAM)

DRA. ELIA HAYDEE CARRASCO ORTIZ (UAQ)

Querétaro, México, Enero 2023.



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.

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. Contó principalmente con la dirección de la Dra. M. Florencia Assaneo, titular del laboratorio.

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

- Dr. Hugo Merchant Nancy (INB, UNAM).
- Dra. Elia Haydee Carrasco Ortiz (UAQ).

Dicha investigación fue financiada por el programa PAPIIT-2021 IA202921 y por la beca otorgada por el CONACYT (CVU: 1045881).

Índice General

Antecedentes	10
General:.....	15
Específicos:.....	15
Diseño experimental.....	16
Diseño general de los estímulos	18
Características fonotácticas de los pseudolenguajes	20
Síntesis de los estímulos auditivos	20
Síntesis del estímulo auditivo: concatenación de sílabas no coarticuladas.....	20
Síntesis del estímulo auditivo: sílabas coarticuladas utilizando Mbrola.	21
Condición no isócrona.....	22
Análisis estadísticos	23
Participantes	25
Resultados del primer piloto	26
Resultados del segundo piloto	30
Delimitación de grupos de sincronía.....	32
Modelos estadísticos resultantes	34
Primer experimento (frecuencias planas)	35
Análisis post-hoc	35
Segundo experimento (duraciones planas)	37
Análisis post-hoc	38
Figuras.....	46
Tablas	47

Apéndice A. Valor de fase de loqueo o Phase Locking Value48
Apéndice B. Lista de pseudopalabras y palabras falsas de cada pseudolenguaje.....49

Agradecimientos

A mi comité tutor, cuya guía y reconocimiento me permitió creer en mí misma.

A Flor, por darme un lugar seguro, no sólo para aprender, sino también para seguir adelante cuando mi paso era vacilante. A todo Hablab, un equipo entrañable (¡arriba los highs!). Pero en particular a Cecy, por ser espejo y refugio; por acompañarnos en nuestro andar.

A la Orden del Puma, por las aventuras —secretas y no tan secretas—, las alegrías compartidas, y el apoyo en las tristezas. Su bella amistad llegó más que oportunamente, más que necesariamente; sin ustedes no lo hubiera logrado (¡sobrevivimos un posgrado en pandemia!). A Ana, por ser mi primera cómplice en este mundillo fascinante llamado INB. A Daniela y a Jared por cargarme en sus hombros y dejarme llorar de felicidad. Pero en especial a Dani, por ser tan tú y dejarme ser tan yo.

A Omar, por alentarme a emprender este camino, por los años recorridos y permanecer aún a la distancia. A Leonora, la mejor y más tierna compañera.

A Claudia y Pablo, Alfredo y Maricruz: quererme y recibirme como una de los suyos fue el salvavidas que me permitió volver a empezar. A mis abuelos, los modelos a seguir más llenos de cariño que podría pedir. A Edgar, por ser el primero en hacerme pensar en mi futuro. A Iván, por hacerme disfrutar más el presente.

Y, sobre todo, a mis papás. Gracias, por tanto, tanto apoyo y, por tanto, tanto amor.

Resumen

El aprendizaje estadístico se describe como la habilidad implícita que posee el cerebro para extraer regularidades y patrones del ambiente. Se ha encontrado en distintas modalidades sensoriales (visual, auditivo, etc.), tanto en humanos como en otros animales (por ejemplo, primates no humanos y ratas), y es una capacidad que se manifiesta desde la infancia temprana y se mantiene a lo largo de la vida. Diversas investigaciones apuntan que se trata de una de las herramientas que permiten la adquisición de lenguaje, y uno de los protocolos más utilizados que exploran esta capacidad cognitiva es el Test de Aprendizaje Estadístico de Palabras. Este consiste en la exposición auditiva de un tren continuo de pseudopalabras trisilábicas concatenadas aleatoriamente, sin señales acústicas que les delimiten, para posteriormente evaluar su aprendizaje por medio de la elección forzada entre dos opciones.

De manera clásica, este test presenta sílabas isócronas, es decir que todas poseen la misma duración. Sin embargo, el habla natural se caracteriza por la variabilidad temporal de la producción silábica; a través de distintos idiomas se ha observado una frecuencia silábica entre los 2 y 8 Hz. Por lo anterior, la presente investigación se enfoca en evaluar la validez ecológica de una prueba de aprendizaje estadístico de palabras. Para ello, construimos nuestras propias pseudopalabras en español, aplicamos variaciones a las duraciones de las sílabas que componen el estímulo, y estudiamos cómo esto modula el desempeño de los/las participantes.

Además, para incluir diferencias individuales, introducimos a nuestras observaciones la capacidad de sincronización auditiva-motora del habla de cada participante al aplicarles el Test de Sincronización Espontánea del Habla. Durante esta prueba se presenta un tren rítmico de sílabas aleatorias, mientras que los/las participantes a la par deben susurrar continuamente la sílaba “tah”, lo que permite categorizar a la población como sincronizadores altos (que espontáneamente alinean su producción silábica al ritmo percibido), o sincronizadores bajos (cuya producción silábica no se alinea).

Los resultados obtenidos nos permiten aseverar que existe un aprendizaje robusto frente a las diferentes condiciones de periodicidad de los estímulos. Más aún, el grupo de sincronización alta

tiene una ventaja significativa ante duraciones silábicas isócronas, y esta ventaja se pierde ante variaciones rítmicas del estímulo. Además, encontramos que esta ventaja otorgada por la sincronía de los estímulos a los sincronizadores de alta se ve facilitada cuando los estímulos se construyen según los patrones fonotácticos de la lengua materna de los participantes.

Summary

Statistical learning is described as the brain's implicit ability to extract regularities and patterns from the environment. It has been found in different sensory modalities (visual, auditory, etc.), both in humans and other animals (such as non-human primates and rats), and it is a capacity that manifests in early childhood and is maintained throughout life. Various investigations indicate that is one of the tools that allow language acquisition, and one of the most used protocols that explore this cognitive ability is Words Statistical Learning Test, which consists of the auditory exposition to a continuous stream of trisyllabic pseudowords, randomly concatenated without any acoustic signals between them, to later evaluate their learning by means of a two-forced choice task.

Classically, this test presents isochronous syllables, that is, they all have the same duration. However, natural speech is characterized by temporal variability in syllabic production; through different languages, it has been observed a syllabic frequency between 2 and 8 Hz. Therefore, this research focuses on evaluating the ecological validity of a statistical word learning test. To achieve this, we constructed our own pseudowords in Spanish, we applied rhythmic variations to the stimuli syllables, and studied how this modulates the participants' performance.

In addition, to include individual variability, we introduced the auditory-motor speech synchronization capacity of each participant to our observations by applying the Spontaneous Speech Synchronization Test. During this test, a rhythmic train of random syllables is presented, while the participants must continuously whisper the syllable "tah" at the same time, which allows us to categorize the population as high synchronizers (those who spontaneously align their syllabic production to the rhythm perceived), or low synchronizers (whose syllabic production does not line up).

The results obtained allow us to assert that there is a robust learning against the different conditions of stimuli periodicity. Furthermore, the high-synchronized group has a significant advantage over isochronous syllable durations, and this advantage is lost over stimulus rhythmic variations. Additionally, we found that this advantage that stimulus synchrony confers on high

synchronizers is facilitated when the stimuli are constructed according to the phonotactic patterns of the participants' native language.

Introducción

La principal herramienta de comunicación que posee la humanidad es el habla, compuesta por una señal acústica generada por una compleja interacción entre el tracto vocal, la lengua, los labios y las cuerdas vocales. La evolución del habla requirió de la capacidad del aprendizaje vocal para adquirir y repetir vocalizaciones a través de la imitación, además de un tracto vocal con un amplio rango fonético (Fitch *et al.*, 2000), de forma tal que nuestras habilidades lingüísticas y las estructuras que le sostienen se fueran complejizando.

Entre las características de la producción del habla fluida, se ha observado que la duración de sus unidades no es perfectamente *isócrona* (del griego *isos* = igual, y *khronos* = tiempo), y que carece de límites acústicos claros entre sílabas y entre palabras. No obstante, mantiene cualidades rítmicas en sus patrones temporales, con una variabilidad estable —entre 2 y 8 Hz— y que se conserva a lo largo del mundo en diferentes idiomas (Ding *et al.*, 2017). Adicionalmente, esta característica pseudo-rítmica del habla facilita la inteligibilidad del lenguaje para quien recibe el mensaje: la sincronización entre la actividad de la corteza auditiva y la frecuencia de la señal recibida propicia la comprensión del habla (Ahissar *et al.*, 2001). Por ende, tanto la producción como la percepción del habla presentan ciertas características rítmicas, las cuales a pesar de que se han descrito ampliamente (Poeppel & Assaneo, 2020), hasta el momento se desconocen los mecanismos precisos que les dan origen.

Un protocolo cognitivo clásico que hace uso de estímulos rítmicos para estudiar la adquisición del habla es el que se conoce como *aprendizaje estadístico de palabras* (Saffran, J. *et al.*, 1996), durante el cual se presentan secuencias isócronas de sílabas, cuyas probabilidades de transición entre ellas pueden ser extraídas de forma espontánea. Es decir, los participantes son capaces de detectar y recordar patrones de las sílabas cuya probabilidad de aparición en el estímulo fue mayor.

En el presente trabajo se estudió qué efectos tienen los cambios de temporalidad entre los estímulos silábicos en el aprendizaje estadístico de palabras. Para ello, se evaluó el desempeño de individuos frente a una tarea de aprendizaje estadístico en el que la presentación de sílabas sea lo

más similar posible a la señal producida durante el habla fluida, eliminando la regularidad temporal en el estímulo acústico.

Además, se incluyó el nivel de sincronización auditivo-motora del habla para observar la variabilidad individual en el desempeño de los participantes. Esta variable fue estimada a partir del Test de Sincronización Espontánea del Habla, propuesto recientemente por Assaneo *et al* (2019), que permite categorizar a los participantes en altos y bajos sincronizadores según su capacidad de alinear su producción de habla a un estímulo externo —compuesto por una sucesión rítmica de sílabas aleatorias—. De este modo, se buscó conocer las diferencias en el desempeño que produce la temporalidad silábica en el aprendizaje de palabras en relación a las habilidades de sincronización individual.

Antecedentes

El aprendizaje estadístico es un mecanismo que ocurre a través de diferentes modalidades sensoriales y dominios cognitivos, y ha sido estudiado ampliamente tanto en humanos como en otras especies animales, por ejemplo: en algunos pájaros cantores (Fehér *et al.*, 2017; Takahasi *et al.*, 2010), en primates no humanos (Goujon *et al.*, 2013) y en ratas (Toro *et al.*, 2005). Dicho mecanismo favorece el aprendizaje mediante el reconocimiento de patrones recurrentes de los estímulos sensoriales, por su probabilidad transicional y la frecuencia de coocurrencia de estos. La probabilidad transicional hace referencia a la probabilidad condicional de que un elemento prediga la aparición de otro, mientras que la frecuencia de coocurrencia indica qué tan seguido dos elementos aparecen juntos (Santolin & Saffran, 2018).

En el caso de la especie humana, el aprendizaje estadístico ha sido un tema relevante para el estudio de los procesos involucrados en la adquisición del lenguaje. Saffran y colaboradores (1996), demostraron que infantes de 8 meses de edad son capaces de aprender y recordar grupos particulares de cadenas de palabras por la probabilidad de transición de sus sílabas. Esta segmentación de las palabras en sílabas y su probabilidad transicional dieron origen al término de aprendizaje “estadístico” de palabras (Plante & Gómez, 2018) y un protocolo utilizado para su estudio se muestra en la Fig. 1.



Figura 1. Representación esquemática de un bloque de aprendizaje estadístico de palabras. Izquierda: Fase de Aprendizaje; los participantes escuchan un audio de dos minutos donde se repiten continuamente cuatro pseudopalabras trisilábicas (indicadas en distintos colores en la figura), concatenadas aleatoriamente y sin señales acústicas que les delimiten. Derecha: Fase de Prueba; el aprendizaje es evaluado a partir de un protocolo del tipo elección forzada entre dos alternativas con 8 ensayos. Aquí, se le presenta al participante de forma acústica una pseudopalabra *correcta* (arriba), constituida con sílabas cuya probabilidad de transición es de 1/1, y una *incorrecta* (abajo), cuya probabilidad de transición de sus sílabas es de 1/3. Modificado de: Assaneo *et al.* (2020) Population-level differences in the neural substrates supporting Statistical Learning.

En las últimas décadas, distintos grupos de investigación han estudiado este protocolo y los procesos neuronales que lo subyacen, debido a que se lo considera uno de los mecanismos básicos que permiten el aprendizaje del lenguaje durante el desarrollo (López-Barroso *et al.*, 2013; López-Barroso *et al.*, 2015; Rodríguez-Fornells *et al.*, 2009). Es importante notar que en todos los casos el diseño experimental consta de una presentación perfectamente rítmica del estímulo auditivo (esto es, una exposición isócrona de sílabas), pese a que el habla natural posee ciertas regularidades temporales y no está compuesta por una sucesión isócrona de sílabas, sino que su producción muestra variaciones temporales (ver siguiente párrafo para más información). Por lo tanto, para asegurar que el aprendizaje estadístico es uno de los mecanismos fundamentales para la adquisición del lenguaje oral, es necesario probar que tal aprendizaje es posible al ingresar cierta variabilidad en la duración de las unidades del estímulo auditivo.

Para contextualizar la importancia de la variabilidad temporal en la producción del habla, es importante visualizar la onda sonora que le compone (Fig. 2); se ha observado que su amplitud presenta una sucesión de incrementos y decrementos, lo que llevó a desarrollar múltiples investigaciones sobre la presencia de regularidades temporales en la evolución temporal de la amplitud de la señal, o como se le define: la *envolvente del habla*. Diferentes estudios revelaron que la envolvente conserva una estructura temporal estable entre distintos idiomas: ~2–8 Hz, lo

que indica que durante la producción de habla la frecuencia silábica media es de entre 2 y 8 sílabas por segundo, y que probablemente esto refleja propiedades intrínsecas de sistemas de percepción y de producción en el cerebro humano (Poeppel & Assaneo, 2020). Aunado a esto, las regularidades temporales en el habla poseen un correlato lingüístico, donde cada uno de los ciclos de la envolvente se corresponde aproximadamente a una sílaba, marcando la importancia de la sílaba como unidad de estudio en los procesos del habla.

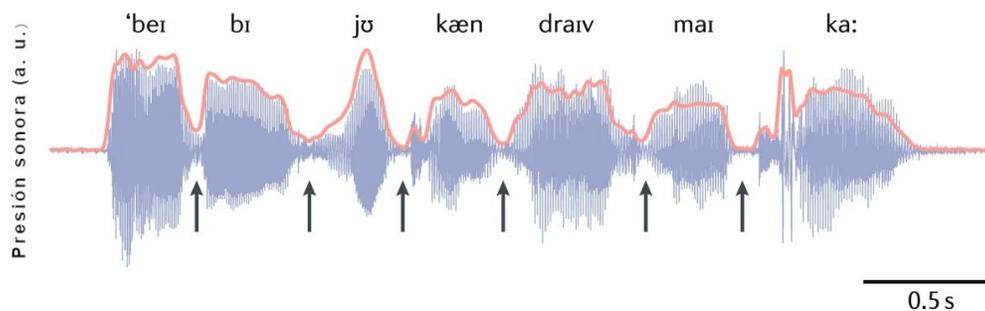


Figura 2. Evolución de la onda de amplitud a través del tiempo, o la envolvente del habla (en color rojo pálido). Los ciclos en la envolvente (señalados con flechas), se corresponden aproximadamente a los límites silábicos, que aún con variaciones temporales, poseen cierta regularidad. Los símbolos arriba representan la transcripción del Alfabeto Fonético Internacional (IPA; International Phonetic Alphabet) para el segmento del habla correspondiente. Modificado de: Poeppel & Assaneo, (2019). *Speech Rhythms and their Neural Foundations*.

Cabe mencionar que al realizar estudios cognitivos es importante considerar diferencias individuales que puedan modular el fenómeno a investigar, y en el caso particular del aprendizaje estadístico de palabras los trabajos más actuales muestran que el nivel de interacción rítmica auditivo-motora predice el desempeño individual de los/las participantes en esta tarea. En una investigación reciente, Assaneo y colaboradores (2019) desarrollaron la prueba de Sincronización Espontánea del Habla, conocido como SSS-test (por sus siglas en inglés, *Spontaneous Speech Synchronization test*). En esta prueba, los/las participantes atienden durante un minuto a un tren rítmico de sílabas, al mismo tiempo que susurran de manera continua y repetida la sílaba “ta”. Los resultados obtenidos en este estudio revelaron que la población general se puede separar en dos grupos: (i) personas con **alta** sincronía, que alinean automáticamente su producción silábica al ritmo de los estímulos entrantes, y (ii) personas de **baja** sincronía, cuyo ritmo de producción silábica no se ve modificado por la presencia del ritmo externo.

De especial interés para el presente trabajo, Assaneo y colegas demostraron que personas de alta sincronía tienen un mejor desempeño en una tarea de aprendizaje estadístico de palabras que las de baja sincronía. Sin embargo, queda por dilucidar de dónde proviene tal ventaja en el aprendizaje frente al nivel de sincronía individual y, más aún, si esta ventaja está dada por la estructura rítmica del estímulo.

Adicionalmente, en otro trabajo reciente, Elazar y colaboradores (2022) encontraron que, ante una tarea de aprendizaje estadístico de palabras, las frecuencias de co-ocurrencia de las sílabas que componen el estímulo dentro del lenguaje nativo de los participantes tienen peso en el desempeño. Y si bien la literatura previa sobre aprendizaje estadístico no hace distinción respecto a la *fonotáctica* —la estructura permitida de combinaciones de fonemas dentro de una lengua—, estos resultados indican que las características fonotácticas del idioma nativo de los y las participantes pudieran intervenir en el nivel de aprendizaje. Es decir, mientras más similares sean las combinaciones silábicas de los estímulos a las encontradas de manera natural en el habla, más alto será su aprendizaje. Por consiguiente, es importante considerar las características fonotácticas a la hora de diseñar los estímulos para una tarea de aprendizaje estadístico.

Justificación

Reuniendo la evidencia experimental descrita previamente, surgen las preguntas de investigación directrices de este trabajo:

Al aplicar una prueba donde se replican las condiciones naturales del lenguaje, esto es, en el que las sílabas sean presentadas con variaciones temporales cuya señal sea más parecida a la que se produce durante el habla fluida:

(i) ¿En qué grado se modifican los niveles de aprendizaje estadístico de palabras? ¿Es este efecto modulado por las características fonotácticas de los estímulos?

(ii) ¿Se mantiene la diferenciación entre sujetos de alta y baja sincronía cuando se remueve la sincronía del estímulo acústico?

Responder estas preguntas permitirá entender mejor el rol que juega la ritmicidad en los procesos de aprendizaje, con lo que se podrían plantear y diseñar estrategias a futuro para su aplicación en casos de dificultades en el aprendizaje, problemas de lenguaje, entre otros.

Hipótesis

El aprendizaje estadístico de palabras es robusto a una disminución en el nivel de sincronía en los estímulos acústicos, independientemente de las características fonotácticas de los estímulos.

El desempeño de los sujetos de alta sincronización auditivo-motora es más susceptible a cambios en la estructura rítmica de los estímulos que el de los de baja.

Objetivos

General:

Contrastar el desempeño de participantes con diferente nivel de sincronización auditivo-motora del habla ante una prueba de aprendizaje estadístico con pseudopalabras diseñadas a partir de sílabas que por su posición silábica sean de frecuencia alta en el idioma español. Determinar cómo se modula su aprendizaje al modificar la estructura rítmica y las características fonotácticas del estímulo.

Específicos:

- Diseñar y estandarizar estímulos acústicos óptimos para una tarea de aprendizaje estadístico de palabras en hablantes nativos de español, controlando sus características fonotácticas. Modificar los estímulos seleccionados para que posean características temporales similares a las del habla natural y ser aplicados en una condición experimental no isócrona.

- Evaluar el desempeño en una tarea de aprendizaje estadístico de palabras ante diferencias fonotácticas en dos condiciones distintas: estímulos isócronos y estímulos no isócronos. Categorizar el nivel de sincronización auditivo-motora del habla de los participantes con la aplicación de la prueba de sincronización espontánea del habla SSS-test.

- Estimar la susceptibilidad de los sujetos de alta y baja sincronía auditivo-motora a los cambios rítmicos y fonotácticos en los estímulos para contrastar y determinar las diferencias en su desempeño.

Material y métodos

Diseño experimental

La prueba de aprendizaje estadístico de palabras aplicada en este estudio está estructurada por cuatro diferentes bloques, cada uno conformado por un audio de dos minutos de una concatenación aleatoria de cuatro distintas pseudopalabras trisilábicas, sin señales acústicas entre ellas. A cada bloque lo llamaremos pseudolenguaje —del 1 al 4—. Su presentación (ver figura 3) inicia con una fase de aprendizaje, donde uno de los bloques es asignado pseudo-aleatoriamente (contra-balanceando la presentación de los bloques), para después en una fase de prueba de 8 ensayos evaluar el aprendizaje a partir de la elección forzada entre dos alternativas: una *pseudopalabra* “correcta”, constituida con sílabas cuya probabilidad de transición es de 1/1, y una *palabra falsa* o “incorrecta”, cuya probabilidad de transición de sus sílabas es de $\frac{1}{3}$ (ver figura 1). Esto se repite hasta completar los 4 pseudolenguajes, alternando entre condición isócrona y condición no isócrona (cuya construcción se explica en la sección del diseño general de los estímulos). Al final del test de aprendizaje estadístico, se aplicó el SSS-test para determinar el nivel de sincronización auditiva-motora del habla (protocolo descrito más adelante).

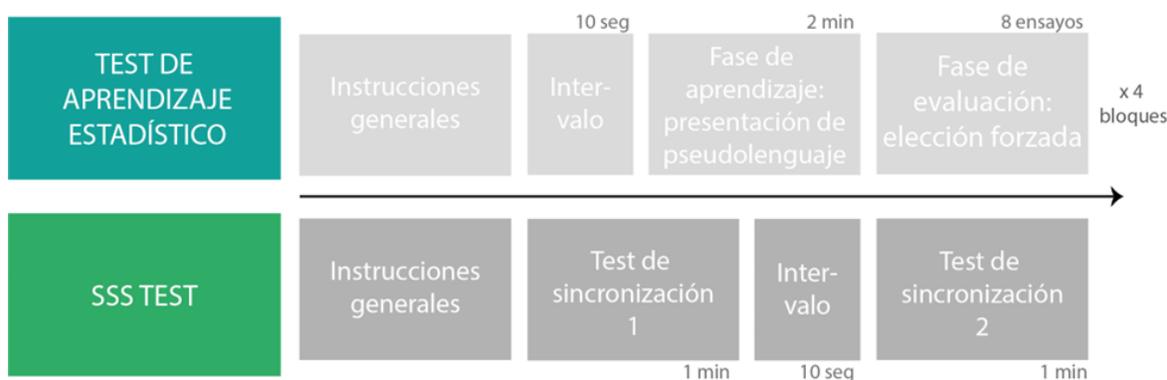


Figura 3. Diseño experimental. Los recuadros representan las distintas fases de los experimentos, y la flecha indica su orden de aparición (de izquierda a derecha). Arriba, se muestra el protocolo para el test de aprendizaje estadístico: dadas las instrucciones, los audios de los pseudolenguajes se presentan pseudo-aleatoriamente en bloques de familiarización/aprendizaje de 2 minutos. A continuación, una fase de evaluación donde el participante realiza 8 ensayos de elección forzada entre dos alternativas (escoger entre dos opciones aquella que le resulte más familiar en relación al audio). En cada ensayo se presenta de forma escrita y auditiva una combinación silábica incorrecta y una

de las pseudopalabras del lenguaje. Antes de cada inicio de otra etapa de aprendizaje, hay intervalos de 10 segundos. Al finalizar el test de aprendizaje estadístico, se aplica uno de sincronización espontánea del habla (abajo), donde se da la instrucción de susurrar durante un minuto la sílaba “ta” continua y simultáneamente a un audio con un tren silábico aleatorio. Esta tarea se realiza en dos bloques, con un intervalo de 10 segundos entre ellos.

El diseño del SSS test consiste en: (i) ajuste del volumen, donde el o la participante deben escuchar un tren de sílabas aleatorias para ajustar el volumen del audio de tal forma que no pueda ser capaz de percibir su propia voz al momento de susurrar. Posteriormente se prosigue a un (ii) ensayo de prueba, donde los participantes son instruidos para susurrar la sílaba ‘tah’ al mismo ritmo de un audio con un tono muestra por 10 segundos. Finalmente, (iii) la prueba de percepción silábica, donde los participantes deben susurrar ‘tah’ mientras escuchan un tren de sílabas rítmicas por un minuto (Fig. 4). Los pasos 2 y 3 se repiten en dos ocasiones, obteniendo así dos grabaciones de un minuto de duración en la prueba de percepción silábica para cada participante.

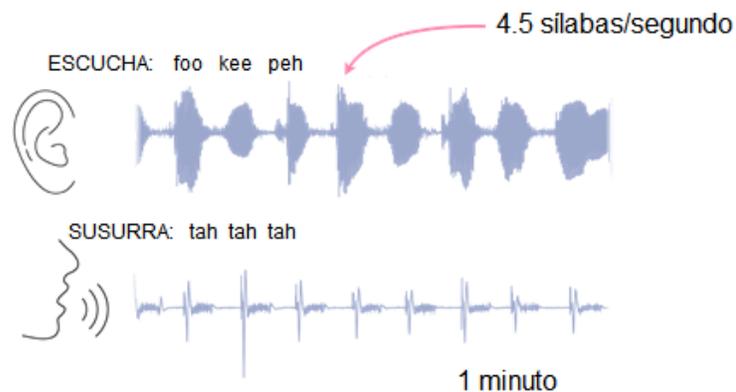


Figura 4. Test de Sincronización Espontánea del Habla (SSS test). Representación de señales acústicas percibidas de los estímulos presentados (arriba) y la señal producida al susurrar la sílaba ‘tah’ (abajo) durante la prueba de percepción silábica del SSS test. Modificado de: Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019) Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature neuroscience*, 22(4), 627-632.

Debido a las medidas de distanciamiento social a causa de la pandemia de COVID-19, todo el experimento fue diseñado para realizarse a distancia por medio de la aplicación web *Gorilla*

(<https://gorilla.sc/>), y evitar llevar a cabo los registros presencialmente. Únicamente fue necesario que los y las participantes tuvieran a su alcance una computadora o laptop con acceso a internet, audífonos y un micrófono que permitiera grabar durante el test de sincronización.

Diseño general de los estímulos

Para la construcción de las *pseudopalabras* —palabras sin significado léxico—, se escogieron sílabas según sus frecuencias de aparición respecto a la posición dentro de las palabras para el idioma español, tanto escrito como hablado. La frecuencia posicional de una sílaba en un lenguaje dado se establece al tomar en cuenta las palabras que comparten una sílaba en determinada posición, y existen dos modos estándar para medirla: las frecuencias *type* y las frecuencias *token*. La frecuencia silábica *type* contabiliza el número de palabras en las que aparece una sílaba en una posición particular dentro del diccionario del lenguaje estudiado. En cambio, la frecuencia *token* le otorga peso a cada una de estas palabras a partir de cuán comunes sean a través de distintos corpus, es decir: cuántas veces se repiten estas dentro de una compilación de muestras orales y textos escritos (Baquero *et al.*, 2019; Conrad *et al.*, 2008).

En el caso del diseño de los estímulos de este experimento se escogió la frecuencia silábica *token* para el castellano, utilizando la aplicación online *Syllabarium, Complete Statistics for Basque and Spanish Syllables* (<https://www.bcbl.eu/syllabarium/index.php>), tanto para su construcción como para el control de su información fonotáctica. A partir de esta base de datos, se extrajeron sílabas de dos fonemas cuya conformación fuese consonante-vocal. Adicionalmente, se agruparon aquellas con sonidos idénticos según la fonología del Español Mexicano (por ejemplo “yo” y “llo”), y se sumaron las frecuencias respectivas a cada posición silábica, excluyendo sílabas con la letra “h”.

Se examinó la distribución de las posiciones silábicas de todas las sílabas que cumplieron los criterios anteriores, detectando que existe un alto número de sílabas con baja frecuencia de aparición, y que la cantidad de sílabas se va reduciendo a mayor frecuencia (Fig. 5). Debido a que las frecuencias altas se alejan considerablemente del resto de los datos y que las frecuencias más bajas no representan interés para este diseño experimental que busca la replicación de características habituales en el idioma español, se decidió dejar fuera los valores extremos, de modo tal que el rango de frecuencias *token* elegido fue de 700 a 6000. Así, se seleccionaron 48 sílabas

únicas para completar el total de 16 pseudopalabras trisilábicas que integrarían los pseudolenguajes, evitando su repetición entre las distintas pseudopalabras y los distintos bloques.

Una vez que las sílabas fueron seleccionadas, estas fueron concatenadas aleatoriamente para crear las pseudopalabras (ver pseudolenguajes en apéndice B), y a partir de ellas construir las palabras falsas con la última sílaba de una pseudopalabra y las primeras dos sílabas de otra.

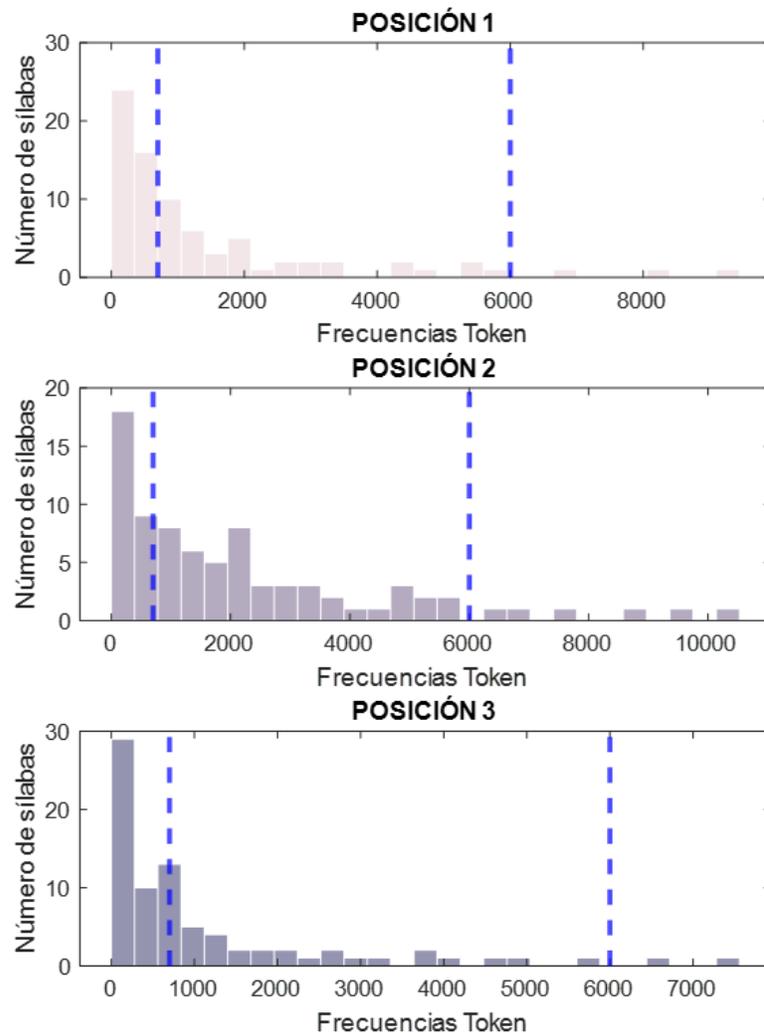


Figura 5. Frecuencias silábicas token para distintas posiciones. Histogramas de frecuencias silábicas token. Es decir, cantidad de sílabas que aparecen en una posición dada a través de distintos corpus. **a)** Posición 1, **b)** Posición 2, **c)** Posición 3. Las líneas punteadas representan los límites del rango usado para la construcción de los estímulos: entre 700 y 6000 frecuencias token.

Características fonotácticas de los pseudolenguajes

Para conocer las *características fonotácticas* de los pseudolenguajes, se calculó la probabilidad de aparición de las sílabas por su posición en cada pseudopalabra respecto del corpus consultado para su construcción. Es decir, en este trabajo definimos a las características fonotácticas de nuestros estímulos como la probabilidad que una sílaba aparezca en una posición y no en otra, para estimar qué tan similares pueden ser a las combinatorias silábicas permitidas en castellano. Esto se hizo tomando las frecuencias token de cada sílaba de las pseudopalabras en su posición dada (o *correcta*), dividiéndolas entre el *total* de frecuencias token de esa misma sílaba en todas las posibles posiciones (i.e., posición uno, dos y tres). Posteriormente, se obtuvo la media de los valores obtenidos a través de todas las sílabas que componen cada pseudolenguaje.

Del procedimiento anterior, los resultados para los pseudolenguajes fueron: pseudolenguaje 1=0.385, 2=0.472, 3=0.399, 4=0.384. Para el protocolo principal, se obtuvieron los valores z de los pseudolenguajes para ser utilizados en los análisis estadísticos correspondientes, resultando en una escala de -1 a 1.

Síntesis de los estímulos auditivos

Terminado el proceso de construcción de las pseudopalabras, el siguiente paso consistió en la síntesis de los estímulos auditivos. Dentro de la literatura existen dos métodos que generalmente se utilizan para la síntesis de estímulos auditivos relacionados con el habla: (i) la concatenación de sílabas sintetizadas de forma aislada a través de algún algoritmo tipo *text to speech* y posteriormente concatenadas, y (ii) la generación de un flujo silábico continuo y coarticulado utilizando el *software Mbrola*. En este trabajo se sintetizaron los estímulos mediante ambos métodos y se compararon con un piloto previo al experimento para conocer sus valores de aprendizaje y seleccionar aquel que maximice el desempeño de la tarea.

Síntesis del estímulo auditivo: concatenación de sílabas no coarticuladas.

Las sílabas elegidas fueron sintetizadas a audios wav por medio de la página web *Wideo* (<https://wideo.co/text-to-speech/>). Luego, utilizando el software *Praat* se normalizó el *pitch* (la frecuencia fundamental de la voz, responsable de otorgar el tono) y se controló la duración de las sílabas de acuerdo con la condición experimental isócrona, comprimiendo o expandiendo la

duración de todas a 250 ms. Subsiguientemente, la construcción de las pseudopalabras se hizo concatenando los audios de las sílabas correspondientes por medio de la plataforma *Matlab*. Con el mismo procedimiento, se fabricaron las *palabras falsas* para la fase de evaluación del aprendizaje de elección forzada, construidas con la última sílaba de una pseudopalabra y las primeras dos sílabas de otra.

Una vez sintetizadas las pseudopalabras, estas a su vez fueron concatenadas aleatoriamente para armar el audio de dos minutos de duración para el experimento, sin la repetición contigua de palabras. Este procedimiento genera una concatenación continua de las pseudopalabras sin coarticulación de las sílabas que componen el audio (Fig. 6a).

Síntesis del estímulo auditivo: sílabas coarticuladas utilizando Mbrola.

El algoritmo llamado *Mbrola* (Dutoit, 1997), es un sintetizador de alta calidad para audios de habla basado en la concatenación de dífonos, tomando como input una lista de fonemas y su información prosódica (<https://github.com/numediart/MBROLA>). Por lo tanto, este algoritmo permite controlar de forma precisa distintas variables de los estímulos generados (tales como volumen y pitch), y genera flujos de habla coarticulados.

Según lo mencionado, este método de síntesis requiere establecer primeramente el número de fonemas necesarios y su duración individual para completar la duración total del audio. Para los estímulos isócronos se requirieron listas de las pseudopalabras correspondientes a cada bloque, generadas aleatoriamente en *Matlab*. Tales listas integran la concatenación de 960 fonemas con duración de 125 ms cada uno, para que las sílabas fueran de 250 ms y completar 2 minutos en total. Además, *Mbrola* requiere que se precise en estas listas tanto el pitch como el punto de máxima amplitud para cada fonema. En este caso, se fijó la máxima amplitud a la mitad del fonema y el valor del pitch se mantuvo constante a 200 Hz. Los audios fueron sintetizados ingresando las listas generadas, y utilizando el alfabeto fonético de la voz masculina en español ‘*es2*’ de *Mbrola*. El resultado es una serie fluida de sílabas coarticuladas (ver Fig. 6b).

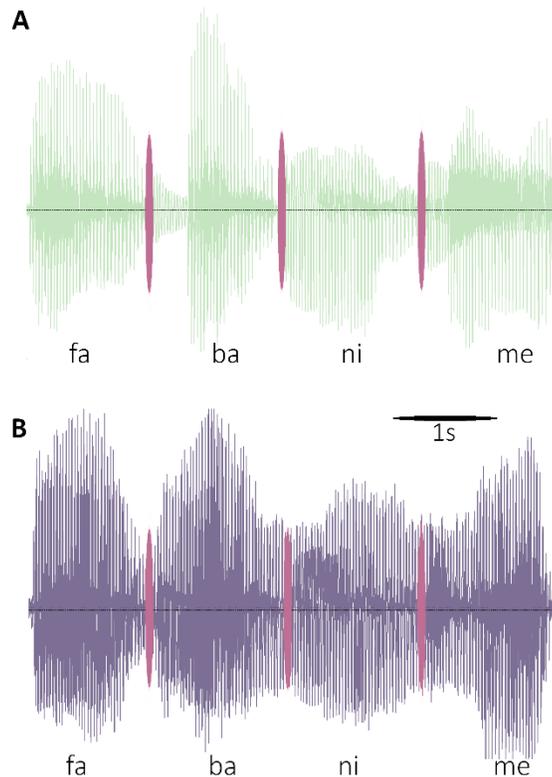


Figura 6. Comparación de los estímulos acústicos. Señal acústica correspondiente un segmento del pseudolenguaje 1, sintetizado por medio de: la concatenación de sílabas independientes (panel a, arriba) o por coarticulación de sílabas utilizando el programa Mbrola (panel b, abajo). Las letras representan las sílabas del pseudolenguaje, y las líneas magenta marcan aproximadamente los límites entre sílabas.

Condición no isócrona

La elección de las nuevas duraciones silábicas para la condición no isócrona para dos versiones de experimentos se llevó a cabo de la siguiente manera: **(i)** para el primer experimento, a partir de una distribución que se corresponde con la de un histograma plano en el espacio de las frecuencias (Fig. 7a), y **(ii)** para el segundo experimento, a partir de un histograma plano de duraciones (Fig. 7b). Ambas distribuciones están delimitadas por rangos encontrados en el habla natural en distintas lenguas: frecuencias entre 2 y 8 Hz, y duraciones entre 0.144 y 0.45 segundos (Ding, 2017). Las sílabas fueron asignadas aleatoriamente dentro de estos rangos.

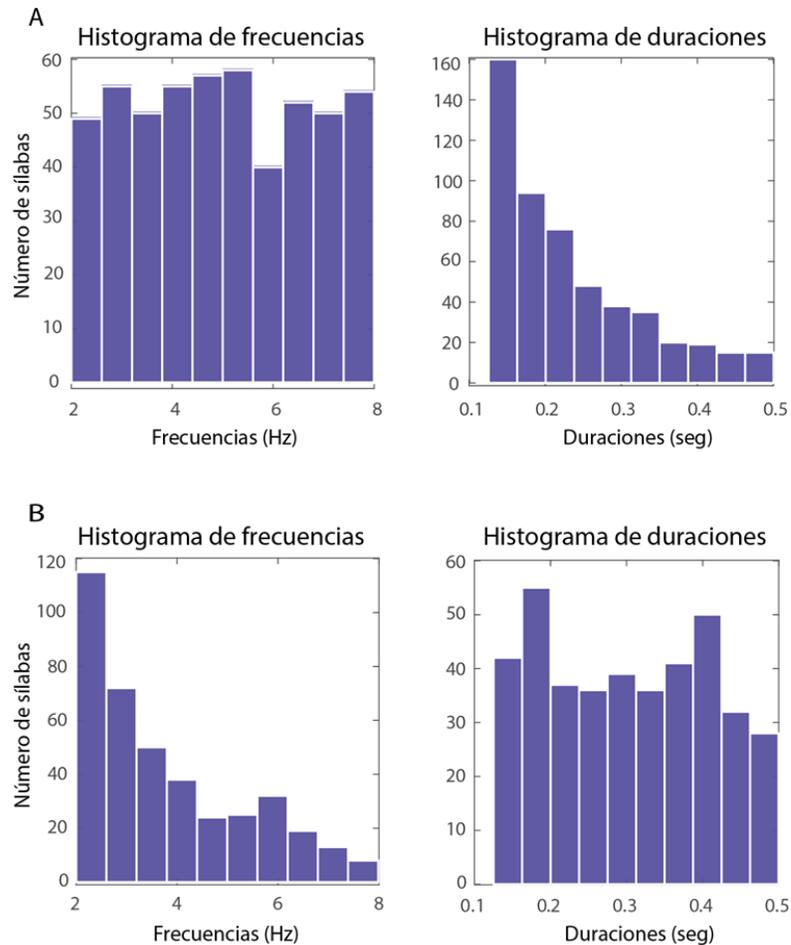


Figura 7. Distribuciones de frecuencias y duraciones para la condición no isócrona. En el eje X se muestran rangos de frecuencias de (2 a 8 Hz, a la izquierda) y duraciones (de 0.14 y 0.45 segundos, a la derecha) que fueron consideradas para la síntesis de los estímulos no isócronos para la prueba de Aprendizaje Estadístico de Palabras. En el eje Y la cantidad de sílabas que de manera aleatoria se asignaron dentro de estos rangos. Arriba en A, los histogramas para el primer experimento, y abajo en B para el segundo.

Análisis estadísticos

Dentro de las variables explicativas registradas durante los experimentos se nombraron: **fon**, como un regresor continuo de la fonotáctica de los pseudolenguajes o la probabilidad de aparición de las sílabas por su posición (0.385, 0.472, 0.399, 0.484), **HL** como el nivel de sincronización auditivo-motora de los participantes a estímulos auditivos externos obtenido con el SSS test (nivel de sincronización alta o baja), y **periodicidad** como el tipo de condición experimental (isócrona o

no isócrona). La variable dependiente fue **learn**, como el aprendizaje, calculando el acierto o error (1 o 0) en cada una de las respuestas de los participantes.

Tales variables se integraron a nuestras observaciones considerando lo reportado en otras investigaciones: **(i)** fonotáctica o *fon*, ya que Elazar y colaboradores (2022) reportaron diferencias de aprendizaje para distintos pseudolenguajes, donde aquellos estímulos con características fonotácticas más cercanas a reglas del habla natural logran aprenderse mejor **(ii)** el nivel de sincronización o *HL*, puesto que el trabajo reportado por Assaneo (2019) muestra que tener alta o baja sincronía predice el nivel de aprendizaje, y **(iii)** la condición de *periodicidad*, por ser el eje que rige la presente investigación.

Durante los pilotos, el nivel de aprendizaje de los/las participantes en los pseudolenguajes fue calculado utilizando la media de respuestas correctas durante la fase de evaluación en cada bloque. Para comprobar si la distribución media resultaba diferente al azar (0.5), se utilizó la prueba no paramétrica de *signos de dos colas*. Para los resultados del primer piloto, las medias del nivel de aprendizaje para ambos métodos de síntesis de los estímulos (text to speech contra Mbrola) fue comparada con la prueba *Mann-Whitney U* $p\text{Corr} < 0.05$, *corrección de Bonferroni* para tres comparaciones.

En cambio, ya que contamos con múltiples variables explicativas en el protocolo principal —fonotáctica, nivel de sincronía y periodicidad de los estímulos, además de sus posibles interacciones—, para el experimento fue importante seleccionar un análisis estadístico que permitiera incluir las variables mencionadas sin perder potencia estadística. Por ende, se realizaron análisis por medio de modelos lineales mixtos generalizados, que permiten observaciones a varios niveles sin perder datos ni potencia estadística al no utilizar medias, además de permitir respuestas binomiales (0 y 1 como respuestas del aprendizaje).

Para encontrar el conjunto de variables que mejor predicen las respuestas de los participantes, los modelos para cada experimento fueron seleccionados por medio de un paquete en *RStudio* llamado *buildmer* (Voeten, 2021). Este paquete permite hacer múltiples comparaciones entre distintos modelos por un doble proceso de eliminación paso a paso hacia atrás: al ingresar todas las variables e interacciones posibles, primero compara posibles combinaciones de estas hasta encontrar el mayor modelo que converja los datos. A partir del mayor modelo, nuevamente va

eliminando elementos uno a uno, comparando la información hasta encontrar el modelo que mejor prediga los datos. Para tal fin, en el presente trabajo se seleccionó el criterio de información *Akaike* corregido (AICc por sus siglas en inglés) como criterio de comparación entre modelos, y el ID de los participantes como factor random para agrupar los valores registrados. El criterio AIC minimiza la pérdida de información al calificar los modelos en función del porcentaje de variación que explique los datos, mientras “penaliza” modelos sobreparametrizados.

A partir de los datos ajustados por los modelos seleccionados, se llevó a cabo un análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) para identificar las variables con significancia más alta. Además, los análisis post-hoc incluyeron las medias marginales estimadas y las medias marginales estimadas para tendencias lineales, que modelan estadísticamente la media de la variable dependiente para cada nivel de las variables explicativas, y así contrastar tales efectos y niveles de significancia del modelo estadístico.

Participantes

La selección de participantes cubrió un rango de edad de 18 a 55 años, sin preferencia de sexo, cuya lengua materna fuera el castellano, y sin dificultades auditivas o del habla auto reportadas. Se invitó a estudiantes de campus de la Universidad Nacional Autónoma de México y la Universidad Autónoma de Querétaro, a través de medios electrónicos y colocación de volantes. Además, a través de la página del laboratorio y en redes sociales, se realizó una invitación abierta para el experimento en línea. Todos los y las participantes, tanto para el piloto como para el protocolo principal, respondieron el consentimiento informado y un cuestionario de antecedentes en música, danza, deportes, videojuegos e idiomas. Los procedimientos aplicados fueron aprobados por el comité de ética del Instituto de Neurobiología, UNAM, en el protocolo 096.

Primer Piloto: Aprendizaje de los estímulos.

Para corroborar que los estímulos cumplieran las condiciones óptimas para la prueba, se llevó a cabo un piloto donde se comparó el nivel de aprendizaje frente a ambos métodos de sintetización de estímulos: (i) *text2speech* y (ii) *Mbrola*. El objetivo del piloto consistió en evaluar la dificultad de la tarea, para asegurar que exista un aprendizaje significativo contra el azar, pero que no resulte saturado, y así se pueda obtener variabilidad en los valores de desempeño de los participantes.

Se obtuvieron datos de 20 participantes para los lenguajes sintetizados por el método de concatenación de sílabas independientes, y 24 para el método de coarticulación de *Mbrola*. Cada participante completó 4 bloques de aprendizaje estadístico. Como se explicó previamente, cada bloque de la prueba consiste en la presentación de un audio de dos minutos de un pseudolenguaje asignado aleatoriamente, y al concluir cada audio, se evalúa el aprendizaje por medio de la elección forzada entre una pseudopalabra y una palabra falsa. Todos los pseudolenguajes fueron presentados sólo en su condición isócrona.

Resultados del primer piloto

Las observaciones realizadas durante la aplicación de las pruebas piloto permitieron discriminar los pseudolenguajes con un alto índice de aprendizaje, pero que aún mantienen variabilidad en el desempeño. Para medir el nivel de aprendizaje general de los distintos pseudolenguajes, se calculó la media de aciertos obtenidos. Los resultados al promediar a través de los dos tipos de síntesis —*mbrola*, método por sílabas coarticuladas, o *text2speech*, concatenación de sílabas no coarticuladas—, mostraron que los pseudolenguajes 1, 2 y 3, presentan un aprendizaje significativamente por arriba del azar ($N = 44$, $p < 0.01$ test de signos de dos colas comparado contra 0.5).

No obstante, esto no se cumple para el cuarto pseudolenguaje (Fig. 8a; $N = 44$, $p = 0.23$ test de signos de dos colas comparado contra 0.5). Por esta razón, se construyó un nuevo pseudolenguaje siguiendo el mismo procedimiento descrito anteriormente. En este caso, la síntesis se realizó solo con *Mbrola* (ver los resultados descritos más adelante, que justifican esta elección)

y se evaluó el aprendizaje en una nueva cohorte de 30 sujetos. Este pseudolenguaje presenta las condiciones requeridas ya que su aprendizaje difiere significativamente del azar (Fig. 8b; $N = 30$, $p < 0.001$ test de signos de dos colas comparado contra 0.5). Además, su valor fonotáctico fue estimado con el mismo método que los pseudolenguajes anteriores al calcular la probabilidad de aparición de sus sílabas, cuyo resultado fue de 0.4843.

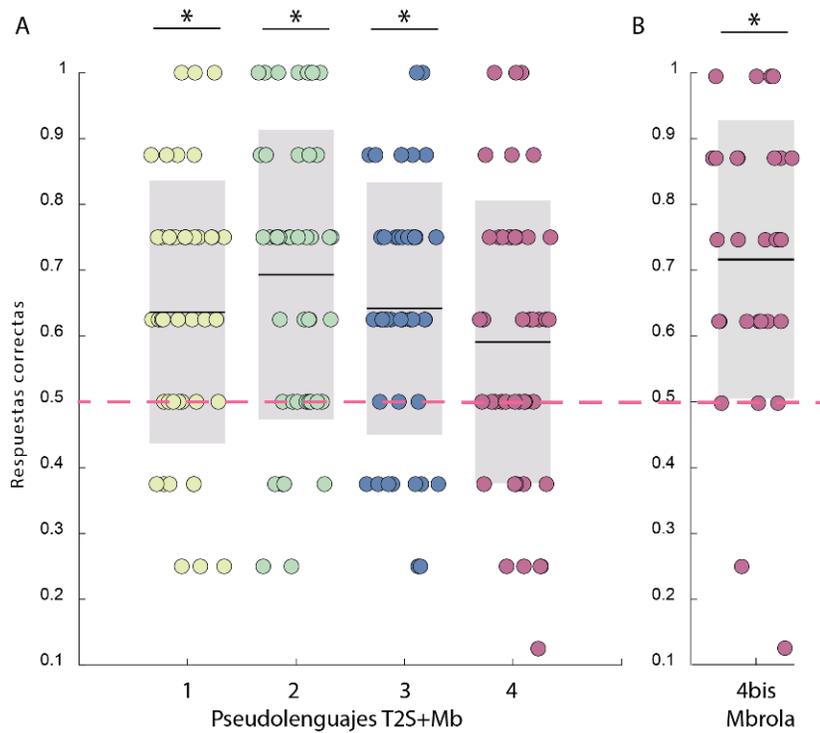


Figura 8. Aprendizaje para los distintos pseudolenguajes. A) Índice de respuestas correctas (respuestas correctas sobre el número total de ensayos) promediados a través de los distintos métodos de síntesis. **B)** Índice de respuestas correctas para el nuevo pseudolenguaje 4 sintetizado con Mbrola. En todos los paneles: los círculos representan participantes, las líneas negras el valor medio, las regiones sombreadas la desviación estándar, las líneas punteadas valor correspondiente al azar (0.5) y los asteriscos la significancia de los valores obtenidos (test de signos de dos colas * $p < 0.01$ y ** $p < 0.001$).

Por otro lado, para evaluar si el tipo de síntesis modula el aprendizaje, se realizó una comparación entre ambos métodos, calculando la media de respuestas correctas de todos los pseudolenguajes para cada método. De este modo, encontramos que existe una tendencia que favorece el aprendizaje en el método por Mbrola (Mann-Whitney U test $p = 0.055$ comparado entre métodos). Luego, se comparó el índice de respuestas correctas entre ambos métodos de síntesis para aquellos pseudolenguajes que presentan un aprendizaje significativo (lenguajes 1, 2 y 3). Como muestra la Fig. 9, existe un mayor número de respuestas correctas para los estímulos sintetizados por medio de Mbrola para el lenguaje 2 (Mann-Whitney U test $p_{Corr} = 0.014$, corrección de Bonferroni para tres comparaciones).

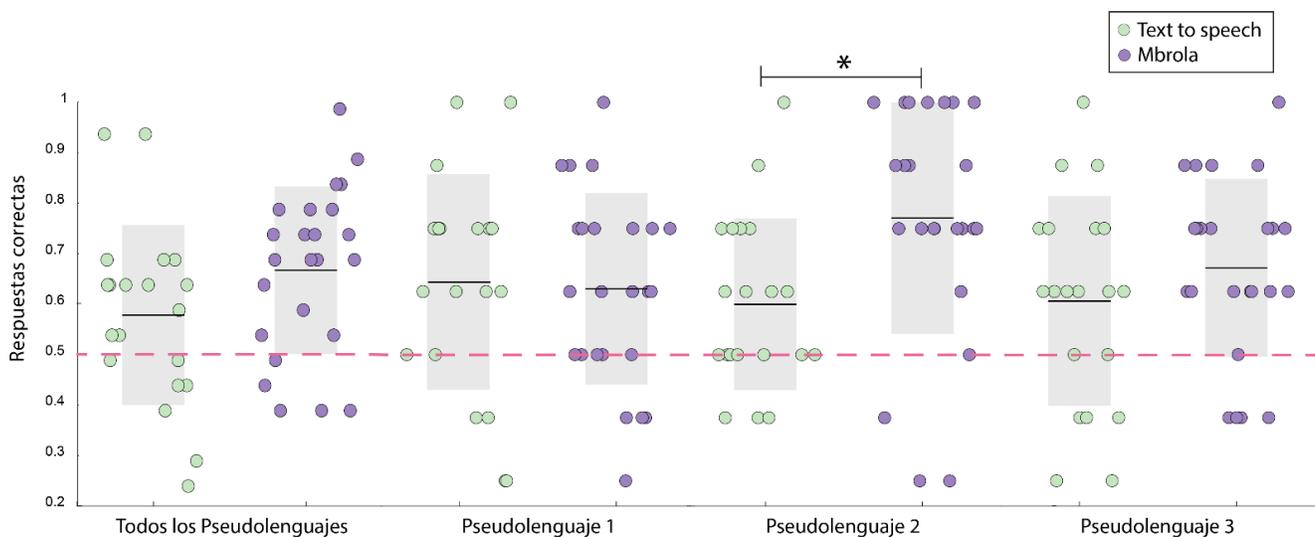


Figura 9. Nivel de aprendizaje entre métodos de síntesis de estímulos. Índice de respuestas correctas (respuestas correctas sobre el número total de ensayos). A la izquierda, en la sección *Todos los pseudolenguajes*, se encuentran promediadas las respuestas de todos los pseudolenguajes para cada método de síntesis. Las demás secciones (de izquierda a derecha a partir de la segunda) muestran la comparación para los pseudolenguajes que se aprendieron significativamente por encima de chance (1, 2 y 3). En todos los paneles: los círculos representan participantes, en verde el método de text2speech, en morado síntesis con Mbrola, las líneas negras dentro de los boxplot el valor medio, las regiones sombreadas la desviación estándar, las líneas punteadas valor correspondiente al azar (0.5) y los asteriscos la significancia de los valores obtenidos (* $p_{Corr} < 0.05$, corrección de Bonferroni).

Con base a estos resultados, donde uno de los pseudolenguajes tuvo diferencias significativas en el aprendizaje comparando ambos métodos, se determinó que todos los audios utilizados en este experimento fueran los sintetizados con el método de coarticulación del programa Mbrola. De esta manera, los pseudolenguajes seleccionados para el protocolo principal poseen los siguientes valores respecto a su información fonotáctica: 1 = 0.385, 2 = 0.472, 3 = 0.399 y 4bis = 0.484. Además, como se mencionó previamente, estos fueron transformados a sus valores z para realizar los contrastes necesarios en los análisis estadísticos del protocolo principal, resultando en una escala de -1 a 1.

Segundo Piloto: Comparación de fonotáctica.

Para comprobar si la fonotáctica juega un rol en el aprendizaje estadístico, favoreciendo el aprendizaje de ciertos pseudolenguajes por sus frecuencias token, se realizó un segundo piloto. El objetivo de este consistió en separar la fonotáctica del estímulo, y para ello se sintetizó otro audio con las mismas sílabas del pseudolenguaje 4bis = 0.484, esto por ser el que posee el valor más alto en fonotáctica. Para este piloto, todas las sílabas que integran el pseudolenguaje sin importar la pseudopalabra, fueron concatenadas de manera aleatoria y sintetizadas en un nuevo audio que se presentó a los sujetos en la fase de aprendizaje. Por otro lado, la fase de evaluación se mantuvo igual; presentando para la elección forzada nuevamente las pseudopalabras y las palabras falsas tal como en el protocolo principal.

Es importante señalar que este diseño pretende imposibilitar el aprendizaje estadístico, ya que la probabilidad de transición entre las sílabas presentadas durante la fase de aprendizaje es completamente aleatoria, eliminando la probabilidad de transición de 1/1 de las sílabas que componen las pseudopalabras. Por ende, las respuestas de los participantes deben reflejar su capacidad de reconocer estructuras fonotácticamente más probables dentro de su lengua materna durante la fase de evaluación.

Resultados del segundo piloto

Debido a los cambios en la transición de las sílabas en esta tarea, donde no hay palabras *correctas e incorrectas*, y los participantes deben de responder de manera implícita utilizando su conocimiento de la estructura fonotáctica de su lengua materna, la dificultad de la tarea es mayor. Esto se refleja en los resultados obtenidos, ya que hay muchos sujetos por debajo del azar. Sin embargo, la media de respuestas correctas para esta condición experimental se mantuvo significativamente por encima del azar ($N = 99$, $p = 0.001$ test de signos de dos colas comparado contra 0.5), como se observa en la figura 10. Esto demuestra que los participantes hacen uso de la fonotáctica del lenguaje al momento de responder durante la fase de evaluación.

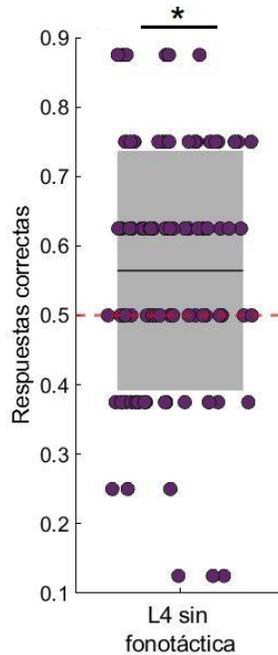


Figura 10. Nivel de aprendizaje para la condición sin fonotáctica. Índice de respuestas correctas (respuestas correctas sobre el número total de ensayos) promediados para el pseudolenguaje 4 sin fonotáctica. Media= 0.564, $p= 0.001$ test de signos de dos colas comparado contra 0.5. Los círculos representan participantes, la línea negra dentro del boxplot el valor medio, las regiones sombreadas la desviación estándar, las líneas punteadas valor correspondiente al azar (0.5) y el asterisco la significancia de los valores obtenidos (* $p<0.05$).

Este resultado indica que la información fonotáctica de las pseudopalabras presentadas durante la fase de evaluación está jugando un rol beneficioso para el aprendizaje estadístico. Esto se comprobó al observar que el desempeño de los y las participantes se mantuvo significativamente por encima del azar a pesar de haber eliminado la probabilidad 1/1 de transición de las sílabas que componen las pseudopalabras en la fase de aprendizaje.

De esta manera, terminó la evaluación de los estímulos diseñados para el protocolo principal, completando los primeros objetivos del presente trabajo. Los resultados obtenidos durante los experimentos de esta investigación se muestran a continuación.

Protocolo Principal: Resultados

Bajo el esquema del protocolo principal se realizaron dos experimentos. En cada uno, para el test de aprendizaje estadístico se probó el efecto de la condición isócrona y de una de las condiciones no isócronas descritas anteriormente: un primer experimento cuyas variaciones rítmicas de sílabas se eligieron con base a un histograma plano de frecuencias, y un segundo experimento a un histograma plano de duraciones.

Para el primer experimento —frecuencias planas—, se recolectaron los datos de 172 participantes, y para el segundo experimento —duraciones planas—, 152 participantes. Sin embargo, por dificultades propias del experimento en línea (p.e. falta de control de ruidos ambientales o fallas en el equipo de grabación), para el primer experimento sólo 114 participantes completaron correctamente tanto el test de aprendizaje estadístico como el test de sincronización espontánea (63 mujeres, edad media: 28.74, std \pm 6.97), y para el segundo experimento 98 participantes (60 mujeres, edad media: 28.79, std \pm 8.95).

Con intención de responder nuestras preguntas de investigación, y así conocer el desempeño de cada grupo de sincronización frente a las variaciones rítmicas en el test de aprendizaje estadístico, primero se categorizaron a los y las participantes como de alta o de baja sincronía, procedimiento que se describe a continuación.

Delimitación de grupos de sincronía

A la cohorte compuesta por todos los y las participantes que completaron correctamente ambas pruebas, se le calculó el valor de loqueo de fase, o PLV por sus iniciales en inglés (*Phase Locking Value*, ver fórmula en Apéndice A), que permite estimar la sincronización entre las fases de dos estímulos: en este caso, la envolvente del habla producida y la del audio presentado. Cuando el PLV tiende a 0, el nivel de sincronización es bajo, y a medida que se acerca a 1 el nivel de sincronización es mayor. Extraídos estos valores, se calculó un histograma de los datos, y se ajustó a este la distribución a dos gaussianas (Fig. 11).

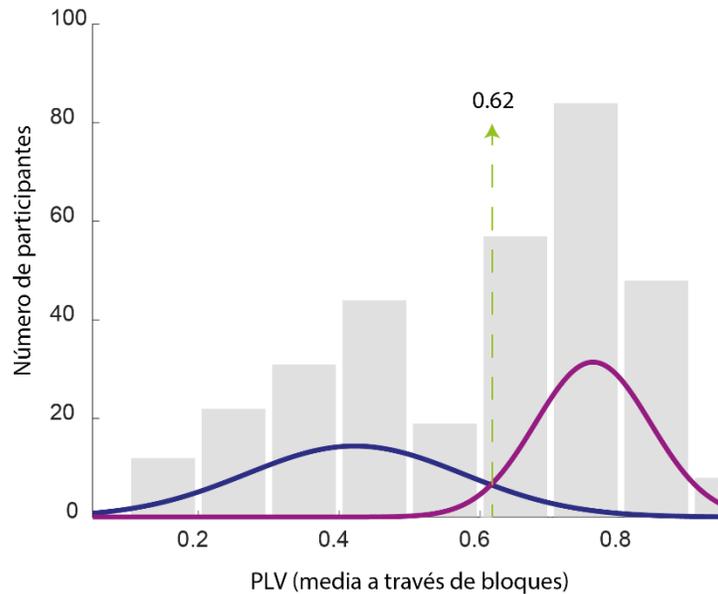


Figura 11. Histograma de PLV en el SSS test (media a través de bloques). En el eje x se esquematiza el promedio del PLV obtenido a través de los bloques de la prueba de SSS test, y en el eje y el número de participantes. Las líneas coloreadas representan la distribución ajustada a dos gaussianas para los grupos: la línea azul representa al grupo de sincronización bajo, y la morada al de alta. La línea verde punteada marca el punto de cruce entre ambos grupos en el 0.62 de PLV.

El procedimiento anterior permitió encontrar los criterios necesarios para categorizar a los y las participantes por su nivel de sincronización: para el grupo de baja sincronía, una media de 0.424, y una proporción de 0.458 del total de los datos, y para el nivel de sincronización alto una media de 0.764, y una proporción de 0.542. Para calcular la probabilidad de cada participante de pertenecer a un grupo u a otro se dividió su probabilidad relativa al grupo de sincronización en cuestión, entre la suma de los valores de probabilidad de ambos grupos.

Además, se encontró que el punto de cruce entre ambas gaussianas ocurre en el valor 0.62 de PLV. Sin embargo, los sujetos más cercanos a este punto tienen una probabilidad alta de pertenecer tanto a un grupo de sincronización como al otro, pudiendo ocasionar ruido a los datos. Por lo que la estrategia adoptada fue tomar únicamente los datos que tuvieran un índice de probabilidad arriba de 75% de pertenecer a cada grupo de sincronización, y cerciorarse que los participantes pertenecieran efectivamente a sus grupos correspondientes.

Con este método, para el primer experimento se clasificaron a 60 sincronizadores altos (con PLV de 0.655 para arriba) y 39 sincronizadores bajos (con PLV de 0.581 para abajo), y para segundo experimento 44 y 43, dejando fuera del análisis a participantes cuya probabilidad de pertenecer tanto a un grupo como a otro era menor al establecido.

Modelos estadísticos resultantes

Todas las variables e interacciones se incluyeron en el modelo (fon*HL*periodicidad) a partir del cual se realizó el doble proceso de eliminación paso a paso con la función buildmer de RStudio mencionado anteriormente; primero para obtener el mayor modelo que converja, y después para encontrar el mejor modelo explicativo. Los mejores modelos resultantes para cada experimento, seleccionados por el criterio de exclusión AIC, se muestran en la tabla 1.

Tabla 1

Modelos estadísticos seleccionados.

Experimento	Modelo	AIC
1. Frecuencias planas	learn ~ (1 ID) + Fon + HL + Fon*HL	3809.7
2. Duraciones planas	learn ~ (1 ID) + Fon + HL + Fon*HL + Periodicidad + HL*Periodicidad + Fon*Periodicidad + Fon*HL*Periodicidad	3520.7

Nota: Las variables integradas en cada modelo están representadas como: variable dependiente learn (aprendizaje), factor random ID (Id por participante). Variables explicativas: fon (fonotáctica: 1=0.385, 2=0.472, 3=0.399, 4=0.484), HL (nivel de sincronización espontánea: High vs. Low), periodicidad (variaciones temporales en el estímulo: Isócrono vs. No Isócrono). AIC = Criterio de Información de Akaike.

Primer experimento (frecuencias planas)

El mejor modelo para el experimento de frecuencias planas integra las variables de fonotáctica, nivel de sincronía y la interacción entre fonotáctica y nivel de sincronía. En la tabla 2 se muestran los resultados del análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) aplicado posteriormente para conocer los niveles de significancia de los datos ajustados de este modelo (utilizando el paquete *lme4* de RStudio).

Tabla 2

Análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) para el modelo estadístico seleccionado en Frecuencias Planas.

Variables	Chisq	Df	Pr (>Chisq)
(Intercepto)	123.019	1	<2e-16 ***
Fon	52.886	1	3.53e-13 ***
HL	1.498	1	0.221
Fon*HL	3.904	1	0.048 *

Nota: Las variables integradas están representadas como: fon (fonotáctica: 0.3854, 0.3993, 0.4717, 0.4843), orden (1, 2, 3 o 4), HL (nivel de sincronización espontánea -High vs. Low-), periodicidad (variaciones temporales en el estímulo -Isócrono vs. No Isócrono-). (Nivel de significancia: • $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Lo primero que llama la atención es que la periodicidad no se encuentra dentro del modelo seleccionado por buildmer, lo cuál sugiere que el grado de variabilidad temporal introducido en este experimento no es suficiente para producir un cambio en el desempeño de los y las participantes en comparación con la presentación isócrona clásica de los estímulos. Por otro lado, un efecto de fonotáctica y su interacción con el nivel de sincronía sí resultó significativo, y para entender mejor esta relación se realizaron los análisis post-hoc mostrados en la siguiente sección.

Análisis post-hoc

Como parte de los análisis más detallados a los datos ajustados por este modelo, a través de un análisis de comparación con medias marginales estimadas (con el paquete *emmeans* de Rstudio),

se comprobó que tanto el grupo de sincronización alta como el de baja obtuvieron valores significativos de aprendizaje sobre el azar (medias marginales estimadas grupo de sincronización alta=0.717 , $p < 0.001$; grupo de sincronización baja= 0.683, $p < 0.001$, corrección de Bonferroni para dos pruebas, valor de azar = 0.5).

A continuación, se exploró la interacción entre fonotáctica y nivel de sincronía con el análisis de medias marginales estimadas de tendencias lineales (paquete *emtrends* de Rstudio, ver figura 14). En la figura 12 se puede observar que existe, de manera general, una mejoría en el aprendizaje para ambos grupos de sincronización respecto a un aumento en los valores de la fonotáctica, es decir: a mayor información fonotáctica, mayor aprendizaje. Sin embargo, al contrastar ambas pendientes de aprendizaje de los grupos de sincronización en relación a la fonotáctica, se encontró una diferencia significativa entre ellos. Los sincronizadores de alta muestran tendencias lineales más pronunciadas que los de baja sincronía ($p = 0.048$) Esto indica que el grupo de alta sincronización extrae un beneficio mayor de las características fonotácticas de los pseudolenguajes.

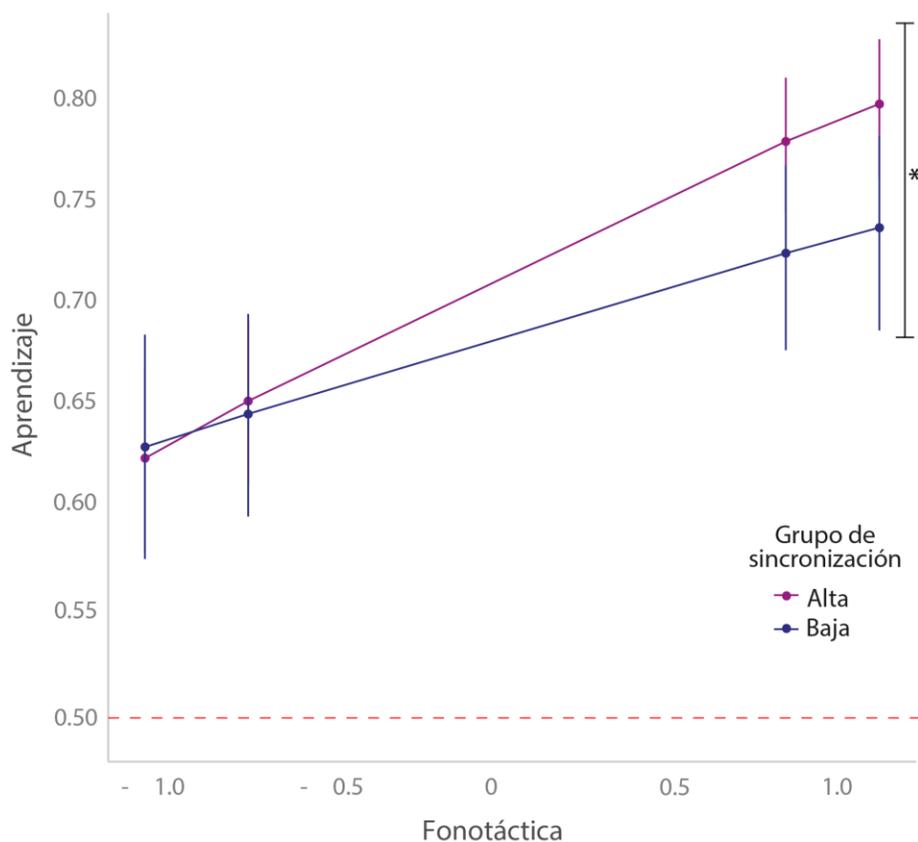


Figura 12. Aprendizaje para los grupos de sincronización en función de la fonotáctica para el experimento de frecuencias planas. En el eje x se esquematiza el nivel de fonotáctica de los pseudolenguajes como regresor continuo, y en el eje Y el nivel de aprendizaje obtenido. En color azul se muestra el desempeño del grupo de baja sincronización, en morado el grupo de alta, la línea punteada el valor de azar (0.5), y en asterisco el nivel de significancia entre las pendientes. (Nivel de significancia: * $p < 0.05$). Ambos grupos de sincronización obtuvieron un nivel significativo de aprendizaje contra el azar (medias marginales estimadas grupo de sincronización alta=0.717, $p < 0.001$; grupo de sincronización baja= 0.683, $p < 0.001$, corrección de Bonferroni para dos pruebas, valor de azar = 0.5).

Segundo experimento (duraciones planas)

En contraste, el modelo seleccionado para el segundo experimento está conformado por las variables de: fonotáctica, nivel de sincronía, la interacción entre fonotáctica y nivel de sincronía, periodicidad y la interacción entre nivel de sincronía y periodicidad. Y para conocer la significancia de los datos ajustados por este modelo, se realizó un Análisis de desviación (pruebas de chi-cuadrado de Wald tipo III, ver tabla 3).

Tabla 3.

Análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) para el modelo estadístico seleccionado en Duraciones Planas

Variablen	Chisq	Df	Pr (>Chisq)
(Intercepto)	38.9286	1	4.396e-10 ***
Fon	7.5569	1	0.005978 **
HL	0.0179	1	0.893623
Fon*HL	0.0006	1	0.979672
Periodicidad	6.5896	1	0.010257 *
HL*Periodicidad	6.3921	1	0.011463 *
Fon*Periodicidad	3.9366	1	0.047247 *
Fon*HL*Periodicidad	3.3429	1	0.067497 •

Nota: Las variables integradas están representadas como: fon (o fonotáctica: 0.385, 0.399, 0.472, 0.484), HL (nivel de sincronización espontánea -High vs. Low-), periodicidad (variaciones temporales en el estímulo – Isócrono vs. No Isócrono-). (Nivel de significancia: • $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Resalta que, a diferencia del experimento anterior, en esta ocasión el efecto de periodicidad sí aparece, además de tener una triple interacción entre fonotáctica, nivel de sincronía y periodicidad con valores cercanos a la significancia. En primera instancia, esto sugiere que las variaciones temporales en los estímulos de este experimento sí jugaron un rol en el aprendizaje, pero que además estas actúan de manera distinta frente a los grupos de sincronización. Para conocer cómo actuaron tales efectos, estos fueron explorados más detalladamente con los análisis post hoc correspondientes, descritos a continuación.

Análisis post-hoc

Tal como se hizo en el experimento anterior, el nivel aprendizaje fue probado contra el azar. Para ello, se utilizó el método de comparación por medio de medias marginales estimadas. Una vez más, observamos que en ambas condiciones experimentales existe un aprendizaje significativo contra el azar, sin importar nivel de sincronización (medias marginales estimadas grupo de sincronización alta: cond. isócrona = 0.720, $p < 0.001$, cond. no isócrona = 0.653, $p < 0.001$; grupo de sincronización baja: cond. isócrona = 0.658, $p < 0.001$, cond. no isócrona = 0.632, $p < 0.001$ corrección de Bonferroni para dos pruebas, valor de azar = 0.5).

Posteriormente, se investigó el origen de la tendencia a la significancia de la triple interacción entre fonotáctica, nivel de sincronización y periodicidad. En la figura 13 puede observarse el nivel de aprendizaje para cada grupo de sincronización, en cada condición de periodicidad, y en función del valor fonotáctico de los pseudolenguajes. Al comparar las pendientes estimadas de aprendizaje en función de la fonotáctica para cada condición de periodicidad, observamos que sólo en la condición isócrona existe una diferencia significativa entre los sujetos de alta y baja sincronía (condición isócrona: $p < 0.05$; condición no isócrona $p = 0.98$). Por otro lado, si comparamos las pendientes para cada grupo de sincronización en las distintas condiciones (i.e. isócrona y no isócrona), es el grupo de alta sincronización el que ve afectado su desempeño de manera significativa (grupo de sincronización alta: $p < 0.05$; grupo de sincronización baja $p = 0.563$). Esto resulta interesante, ya que nos indica que la periodicidad afecta únicamente el aprendizaje del grupo de alta sincronización: cuando los estímulos son isócronos los sincronizadores de alta logran extraer la información fonotáctica de manera más efectiva que los sincronizadores de baja, pero esta diferencia desaparece si se introduce variabilidad en la periodicidad de los estímulos.

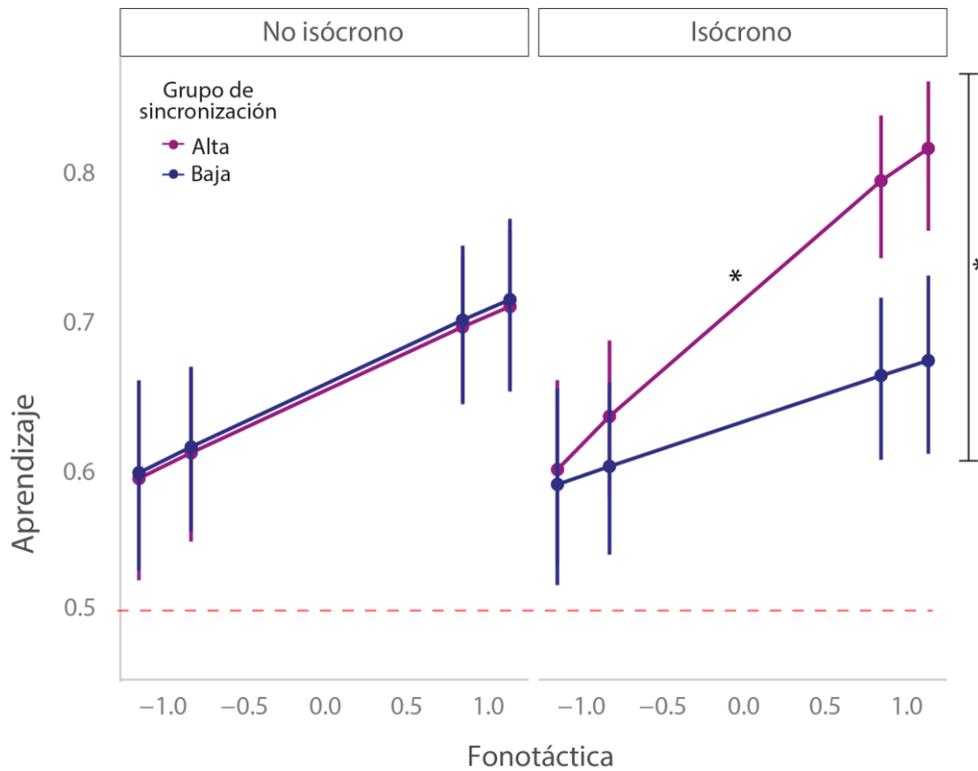


Figura 13. Aprendizaje para los grupos de sincronización en función de la fonotáctica para el experimento de duraciones planas. En el eje x se esquematiza el nivel de fonotáctica de los pseudolenguajes como regresor continuo, y en el eje Y el nivel de aprendizaje obtenido. En color azul se muestra el desempeño del grupo de baja sincronización, en morado el grupo de alta. Ambos grupos de sincronización obtuvieron un nivel significativo de aprendizaje contra el azar (medias marginales estimadas de tendencias lineales; grupo de sincronización alta: pendiente estimada = 7.71, $p < 0.001$; grupo de sincronización baja: pendiente estimada 4.40, $p < 0.05$; corrección de Bonferroni para cuatro pruebas, valor de azar = 0.5). Para la comparación entre pendientes * $p < 0.05$

Discusión

Una de las primeras observaciones realizadas a partir de los resultados obtenidos fue que para ambos experimentos se encontró un efecto dado por las características fonotácticas de los estímulos (ver figuras 12 y 13 en la sección anterior); al aumentar la información fonotáctica hubo a su vez un incremento en el nivel de aprendizaje, indistintamente de la condición de periodicidad. Es decir, las pseudopalabras se aprenden con mayor facilidad al ser más similares a las reglas del español, y esta capacidad de aprendizaje subsiste al cambiar la ritmicidad de los estímulos. Más aún, en los modelos estadísticos resultantes para cada experimento (ver Tabla 1), la interacción entre fonotáctica y nivel de sincronización es significativa, y si bien el grupo de sincronización baja también mejora su desempeño a mayor información fonotáctica, la diferencia entre grupos favorece significativamente al grupo de alta cuando existe cierta regularidad en los estímulos presentados.

Pero a pesar de las diferencias entre grupos de sincronización respecto a la fonotáctica de los estímulos, el aprendizaje estadístico de nuestros pseudolenguajes se mantuvo de manera general por encima del azar sin importar el nivel de sincronización auditivo-motora del habla y a través de condiciones experimentales de periodicidad. Lo anterior resulta relevante ya que, al ser considerado uno de los mecanismos que utilizamos para aprender a hablar, es necesario que el aprendizaje estadístico posea flexibilidad ante los cambios de ritmicidad que ocurren naturalmente en el habla, y que este aprendizaje sea posible más allá de las características individuales de sincronización.

Y si bien se comprobó la robustez del aprendizaje estadístico ante variaciones en la periodicidad, sólo una de las dos condiciones no isócronas produjo cambios importantes en el desempeño de los y las participantes, observándose una interacción entre periodicidad y nivel de sincronización: la condición basada en un histograma plano de duraciones. Por qué únicamente tal condición produce un efecto sobre el aprendizaje y no aquella basada en el histograma plano de frecuencias, podría deberse precisamente a la distribución de las duraciones de los estímulos. Mediante el procedimiento de extracción de duraciones basado en el histograma de frecuencias planas, la mayoría de las sílabas tienen duraciones cercanas a los 150 ms (ver figura 7a, en la sección de Material y Métodos), siendo así insuficiente para modificar el desempeño en la prueba, pues aún existe una duración que ocurre con mayor frecuencia que el resto. En cambio, en el segundo método basado en un histograma plano de duraciones, es imposible realizar ningún tipo

de predicción temporal debido a que no hay ninguna duración silábica que ocurra más que otra (Fig. 7b).

Es de esta segunda interrupción donde surge el cambio en el desempeño del grupo de sincronización alta, lo que lleva a otro de nuestros descubrimientos: el nivel de sincronización individual predice el desempeño ante variaciones rítmicas del estímulo. Lo anterior se confirma al observar que los cambios de periodicidad afectaron de manera distinta a los grupos de sincronización en el experimento (ver figura 13); para condiciones isócronas los y las participantes con un nivel alto de sincronización registraron una pendiente más pronunciada de aprendizaje en función de la fonotáctica en comparación de los de sincronización baja, mientras que en una condición no isócrona donde existe mayor variabilidad de duraciones silábicas —segundo experimento de duraciones planas—, este beneficio de la isocronía desaparece y el desempeño de ambos grupos se mantiene al mismo nivel.

Esto podría indicar que la ventaja reportada para las personas con alta sincronización ante una prueba de aprendizaje estadístico viene dada por la isocronía de los estímulos, y es posible que una alta capacidad de sincronización auditiva-motora a estímulos rítmicos facilite tanto la extracción de regularidades externas, así como de las características fonotácticas del lenguaje. Esto no sólo reproduce lo reportado previamente por Assaneo y colaboradores (2019) sobre el mejor desempeño del grupo de sincronización alta, sino que expande el entendimiento de que al poseer un mayor nivel de sincronización se logra extraer significativamente mejor la información fonotáctica de estímulos auditivos silábicos.

Conclusiones

Los resultados obtenidos nos permiten confirmar las hipótesis que orientaron este trabajo: (i) el aprendizaje estadístico de palabras es robusto a una disminución en el nivel de sincronía en los estímulos acústicos y (ii) el desempeño de los sujetos de alta sincronización auditivo-motora es más susceptible a cambios en la estructura rítmica de los estímulos, que el de los de baja. El aprendizaje estadístico de palabras se favorece de la información fonotáctica, y si bien la isocronía no es necesaria para este tipo de aprendizaje, la ritmicidad de los estímulos facilita su aprendizaje cuando se posee un nivel alto de sincronización. Esto no sólo confirma las hipótesis de este trabajo, sino que ayuda a dilucidar el rol que juega la ritmicidad de los estímulos en la adquisición de palabras, la importancia de las reglas del lenguaje para la construcción de estímulos, y a abrir nuevas perspectivas para el estudio de las características de la sílaba como unidad básica del habla.

Referencias bibliográficas

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, 98(23), 13367-13372.

Anwyl-Irvine, A.L., Massoné J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019) Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods*. Doi: <https://doi.org/10.3758/s13428-019-01237-x> Página web: <https://gorilla.sc/>

Assaneo, M. F., Orpella, J., Ripollés, P., Noejovich, L., López-Barroso, D., de Diego-Balaguer, R., & Poeppel, D. (2020). Population-level differences in the neural substrates supporting Statistical Learning. *bioRxiv*.

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature neuroscience*, 22(4), 627-632.

Baquero Castellanos, S. I. L. V. I. A., Gallo, A. L., & Müller, O. (2019). Evidencia para un efecto facilitador de la frecuencia silábica en la lectura de palabras en escolares hispanohablantes de 2° y 5° grado. *RLA. Revista de lingüística teórica y aplicada*, 57(2), 165-182.

Conrad, M., Carreiras, M., & Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes*, 23(2), 296-326

Christiansen, M. H. (2019). Implicit statistical learning: A tale of two literatures. *Topics in Cognitive Science*, 11(3), 468-481.

De Marco, A., Esperón, A., Sabre, F. (2012) Wideo. <https://wideo.co/text-to-speech/>

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 81, 181-187.

Duñabeitia, J. A., Cholin, J., Corral, J., Perea, M., & Carreiras, M. (2010). SYLLABARIUM: An online application for deriving complete statistics for Basque and Spanish orthographic syllables. *Behavior research methods*, 42(1), 118-125. Página web: <https://www.bcbl.eu/syllabarium/index.php>

Dutoit, T. (1997). *An introduction to text-to-speech synthesis* (Vol. 3). Springer Science & Business Media.

Elazar, A., Alhama, R. G., Bogaerts, L., Siegelman, N., Baus, C., & Frost, R. (2022). When the “Tabula” is Anything but “Rasa:” What Determines Performance in the Auditory Statistical Learning Task?. *Cognitive Science*, 46(2), e13102.

Fehér, O., Ljubičić, I., Suzuki, K., Okanoya, K., & Tchernichovski, O. (2017). Statistical learning in songbirds: from self-tutoring to song culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160053.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in cognitive sciences*, 4(7), 258-267.

Goujon, A., & Fagot, J. (2013). Learning of spatial statistics in nonhuman primates: contextual cueing in baboons (*Papio papio*). *Behavioural Brain Research*, 247, 101-109.

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing statistical learning implicitly: A novel chunk-based measure of statistical learning. In the 39th Annual Conference of the Cognitive Science Society (CogSci 2017) (pp. 564-569). Cognitive Science Society.

López-Barroso, D., Ripollés, P., Marco-Pallarés, J., Mohammadi, B., Muent, T. F., Bachoud-Levi, A. C., ... & de Diego-Balaguer, R. (2015). Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *Neuroimage*, 110, 182-193.

López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences*, 110(32), 13168-13173.

Plante, E., & Gómez, R. L. (2018). Learning without trying: The clinical relevance of statistical learning. *Language, speech, and hearing services in schools*, 49(3S), 710-722.

Poepfel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 1-13.

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3711-3735.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.

Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences*, 22(1), 52-63.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current opinion in neurobiology*, 20(2), 172-176.

Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and prosodic cues for song segmentation learning by Bengalese finches (*Lonchura striata* var. *domestica*). *Ethology*, 116(6), 481-489.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34.

Toro JM, Trobalón JB. 2005 Statistical computations over a speech stream in a rodent. *Percept. Psychophys.* 67, 867– 875. (doi:10.3758/ BF03193539)

Voeten, C. C., & Voeten, M. C. C. (2021). Package ‘buildmer’.

Bibliografía

1. Lizcano-Cortés, F., Gómez-Varela, I., Mares, C., Wallisch, P., Orpella, J., Poeppel, D., ... & Assaneo, M. F. (2022). Speech-to-Speech Synchronization protocol to classify human participants as high or low auditory-motor synchronizers. *STAR protocols*, 3(2), 101248.

Lista de figuras y tablas

Figuras

Figura 1. Representación esquemática de un bloque de aprendizaje estadístico de palabras.	11
Figura 2. Evolución de la onda de amplitud a través del tiempo, o la envolvente del habla.....	12
Figura 3. Diseño experimental.....	16
Figura 4. Test de Sincronización Espontánea del Habla (SSS test).....	17
Figura 5. Frecuencias silábicas token para distintas posiciones.....	19
Figura 6. Comparación de los estímulos acústicos.....	22
Figura 7. Distribuciones de frecuencias y duraciones para la condición no isócrona.....	23
Figura 8. Aprendizaje para los distintos pseudolenguajes.....	27
Figura 9. Nivel de aprendizaje entre métodos de síntesis de estímulos.....	28
Figura 10. Nivel de aprendizaje para la condición sin fonotáctica.....	31
Figura 11. Histograma de PLV en el SSS test (media a través de bloques).....	33
Figura 12. Aprendizaje mediado por la fonotáctica para dos grupos de sincronización.....	36
Figura 13. Aprendizaje para los grupos de sincronización en función de la fonotáctica para el experimento de duraciones planas.....	39

Tablas

Tabla 1. Modelos estadísticos seleccionados.....	34
Tabla 2. Análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) para el modelo estadístico seleccionado en Frecuencias Planas.....	35
Tabla 3. Análisis de desviación (pruebas de chi-cuadrado de Wald tipo III) para el modelo estadístico seleccionado en Duraciones Planas.....	37

Apéndices

Apéndice A. Valor de fase de bloqueo o Phase Locking Value

El nivel de sincronización entre dos señales acústicas (los estímulos auditivos y la producción de habla de los participantes), se calculó midiendo el Phase-locking value de las envolventes del habla y la envolvente coclear de los estímulos. A su vez, las envolventes se obtuvieron tomando el valor absoluto de los espectros de ambas señales a través de la transformada de Hilbert.

La fórmula para el Phase-locking value se define de la siguiente manera:

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

donde t es el tiempo discretizado, T es el número total de puntos en el tiempo, y θ_1 es la fase de la primera y θ_2 de la segunda señal.

Apéndice B. Lista de pseudopalabras y palabras falsas de cada pseudolenguaje

Pseudolenguaje 1		Pseudolenguaje 2		Pseudolenguaje 3		Pseudolenguaje 4	
Pseudo palabras	Palabras falsas	Pseudo palabras	Palabras falsas	Pseudo palabras	Palabras falsas	Pseudo palabras	Palabras falsas
<p>sipota nafemo mecabi fabani</p>	<p>bisipo mosipo nisipo binafe ninafe tanafe momeca nimeca tameca bifaba mofaba tafaba</p>	<p>dideti miñolo tocoña mugullo</p>	<p>lodide ñadide lloidide ñamiño timiño llomiño lotoco titoco llotoco lomugu ñamugu timugu</p>	<p>bogoli lurrite nerroda lenoga</p>	<p>dabogo gabogo tebogo dalurri galurri lilurri ganerro linerro tenerro daleno lileno teleno</p>	<p>petuma redola sasojo sekerra</p>	<p>jopetu lapetu rapetu joreda maredo raredo joseke laseke maseke lasaso masaso rasaso</p>