



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

Maestría en ciencias (Neurobiología)

**INSTITUTO DE NEUROBIOLOGÍA**

***“La interacción entre los efectos de la retroalimentación retardada del habla y la frecuencia silábica producida por el hablante”***

TESIS QUE PARA OPTAR POR EL GRADO DE:

**Maestra en Ciencias**

Presenta:

**Liliana Sánchez Zepeda**

Tutora principal:

**Dra. M. Florencia Assaneo, INB, UNAM**

Miembros del comité tutorial:

**Dr. Luis Concha Loyola, INB, UNAM**

**Dr. Juan Silva Pereira, Facultad de Estudios Superiores Iztacala, UNAM**

*Santiago de Querétaro, Qro. Agosto de 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.

## RESUMEN

El ciclo de retroalimentación auditivo-motora del habla es un mecanismo cerebral que compara la expectativa sonora y el sonido real de una vocalización para asegurar una correcta producción de habla y hacer correcciones de ser necesarias.

Una de las formas evidenciar al ciclo de retroalimentación auditivo-motora del habla de manera conductual es a partir del experimento de retroalimentación retardada del habla. En este, se añade un retraso temporal a la retroalimentación auditiva de un hablante que como resultado genera errores del habla.

Se ha descrito que al añadir retrasos de alrededor de 200ms se observa una mayor intensidad del efecto. Este resultado sugiere que la ventana temporal en la que el ciclo de retroalimentación auditivo-motora del habla compara el sonido esperado con el percibido tiene un valor fijo de 200 ms. Sin embargo, los trabajos previos no controlaron la velocidad de los hablantes y en condiciones normales de habla la duración media de una sílaba es precisamente 200ms. Esto nos lleva a preguntarnos si la ventana temporal de comparación tiene un valor estable de 200 ms o si varía en correspondencia a la velocidad del habla.

Para dar respuesta a esta pregunta sometimos a un grupo de participantes a un experimento de retroalimentación retardada del habla en el que se añadieron de forma aleatoria distintos valores de retraso temporal en 3 condiciones de habla diferenciadas: velocidad baja (3 sílabas por segundo), media (4 sil/seg) y alta (6 sil/seg).

Se encontró en la muestra de estudio que el retraso temporal más disruptivo es distinto para cada una de las 3 condiciones con una interacción significativa entre el retraso temporal añadido y la velocidad del hablante. Más aún, el retraso más disruptivo se alinea con la duración de las sílabas emitidas por el participante. Este resultado sugiere que la ventana temporal de comparación del ciclo de retroalimentación auditivo-motora del habla no tiene un valor fijo, sino que está parametrizada por la duración silábica.

## ABSTRACT

The auditory-motor feedback cycle of speech is a brain mechanism that compares the expected sound and the actual sound of a vocalization to make necessary corrections.

Behaviorally, the AMFCS can be evaluated through the delayed auditory feedback experiment. In this experiment, a temporal delay is added to the auditory feedback of a speaker, which results in speech errors and a reduction in the speech speed of the participant.

It has been described that adding delays of around 200ms generates the most disfluencies. This suggests that the temporal window in which the the auditory-motor feedback cycle of speech compares the expected sound with the perceived sound has a fixed value of 200 ms. However, previous studies did not control for the speed of the speakers, and under normal speech conditions, the average duration of a syllable is precisely 200 ms. This leads us to wonder whether the comparison temporal window has a stable value of 200 ms or varies with speech speed.

To answer this question, we subjected a group of participants to a delayed auditory feedback experiment in which different values of temporal delay were randomly added in three different speech conditions: slow, medium, and fast.

We found a significant interaction between the most disruptive temporal delay and the speaker's speed. The most disruptive delay aligns with the duration of the syllables emitted by the participant. This result suggests that the comparison temporal window of the AMFCS does not have a fixed value but is parameterized by syllabic duration.

## *Agradecimientos*

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

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

- **Dr. Luis Concha Loyola (INB, UNAM).**
- **Dr. Juan Silva Pereyra (FES Iztacala, UNAM)**

**Dicha investigación fue financiada por el programa PAPIIT - IA2002223 y por la beca otorgada por el CONAHCYT (CVU: 1145895).**

**Agradezco por su apoyo a la comunidad académica del INB y al personal administrativo, particularmente a la coordinación de la Maestría en Ciencias (Neurobiología).**

*Expresar el agradecimiento que siento al concluir este trabajo es, sin duda, un desafío. ¿Cómo describir un sentimiento que desborda el corazón? A pesar de la dificultad, debo intentarlo, comenzando por agradecer a mi familia: a mi madre Lilia, a mi hermano Moisés y a mi padre Mario (aunque ya no esté entre nosotros), por su apoyo incondicional y por enseñarme a superar los obstáculos con determinación. Gracias por su comprensión y amor, incluso en momentos poco convencionales.*

*Un agradecimiento especial a Oz, mi compañero en crimen/novio/mejor por su amor y apoyo incondicional. Agradezco también a su familia por recibirme siempre con amabilidad y preocuparse por mí.*

*Quiero expresar un sincero agradecimiento a todas las personas que conocí durante mi estancia en la maestría. Aunque no pueda mencionarlos a todos, si alguna vez compartiste un momento conmigo, esta gratitud es para ti. Gracias por ser mi red de apoyo, por compartir conocimientos, inquietudes y experiencias. Fue gracias a todos ustedes que enfrenté mi debut como científica con confianza. En particular, agradezco al equipo de Hablab; fue un verdadero placer tenerlos como compañeros de laboratorio.*

*Agradezco a mi asesora, la Dra. Florencia, por ser una guía justa, comprensiva y auténtica. Admiro su valentía y dedicación.*

*A mis amigos y amigas en otros ámbitos de mi vida, les agradezco sinceramente. Nuestras risas y nuestras penas compartidas han sido fundamentales para mantenerme centrada y recordar que hay personas increíbles que me valoran.*



*Finalmente, agradezco a mí misma por haber tenido el coraje de perseguir mi camino como neurocientífica. A pesar de las dudas y dificultades, logré alcanzar metas que no habría creído antes. ¡Qué interesante ha sido este viaje!*

## ÍNDICE

Introducción.....	7
Antecedentes. ....	9
Planteamiento del problema. ....	12
Hipótesis.....	13
Objetivos.....	14
Objetivo general. ....	14
Objetivos específicos.....	14
Métodos. ....	14
a) Participantes.....	14
b) Criterios de exclusión. ....	14
c) Instrumentalización. ....	15
Análisis de datos.....	17
Cálculo de los parámetros de estudio. ....	17
Conteo de disfluencias y/o errores de habla. ....	17
Cálculo de la tasa silábica producida por los participantes. ....	17
Criterio de inclusión construido a partir de la tasa silábica en <i>la condición BASAL</i> . ....	17
Análisis de varianza (ANOVA) de medidas repetidas para comparar el efecto de los diferentes retrasos temporales en la tasa silábica producida.....	18
Ajuste de curvas Weibull a los datos de errores cometidos en función de los retrasos temporales añadidos. ....	18
Análisis estadístico para la disminución de la tasa silábica en <i>DAF</i> . ....	19
Resultados.....	19
Cantidad de disfluencias del habla en relación al retraso temporal añadido. ....	19
Estimación de tasa silábica en condición <i>BASAL</i> y <i>DAF</i> . ....	23
Diseño de experimentos complementarios. ....	25
Participantes. ....	25
I. Interacción entre la tasa silábica y la consciencia del efecto <i>DAF</i> .....	25
II. Control para tasa silábica de 6Hz. ....	26
Análisis de datos complementarios. ....	28
Cálculo del porcentaje de ocasiones en que los sujetos fueron conscientes del efecto <i>DAF</i> en el experimento para evaluar la consciencia del efecto. ....	28
Estimación de la tasa silábica para el control de 6 Hz.....	28
Resultados complementarios. ....	28

Cálculo del porcentaje de ocasiones en las que hay una conciencia del efecto.....	28
Control de la condición de velocidad alta (6Hz) con retraso temporal mínimo. ....	30
Discusión. ....	32
I. El periodo temporal de comparación en el ciclo de retroalimentación auditivo-motora del habla escala con la duración silábica. ....	32
II. Los mecanismos que producen errores en el habla podrían ser distintos a los que generan una disminución en la tasa silábica emitida. ....	33
Conclusiones.....	34
Referencias. ....	35

## **Introducción.**

La interacción de los mecanismos de percepción y producción del habla es un fenómeno ampliamente debatido en la comunidad científica (Pulvermüller & Fadiga, 2010). Diversos grupos de investigación se han preguntado si el funcionamiento de los sistemas auditivos y motores, (asociados con el procesamiento y ejecución del habla, respectivamente) ocurre de manera independiente o si están relacionadas a través de un circuito de acción-percepción.

En estudios recientes se ha explorado en mayor medida la idea de que, para que exista una correcta percepción y producción del habla, debe de existir una interacción interdependiente entre corteza motora y corteza auditiva. Por ejemplo, se ha observado que la escucha pasiva de estímulos verbales que involucren la movilización de los músculos de la lengua y los labios genera una activación en la corteza motora, específicamente en las zonas encargadas de controlar estos músculos. Más aún, esta actividad en la corteza motora es similar a la que se puede observar cuando se está pronunciando el estímulo verbal antes escuchado (Fadiga, et al. 2002; Wilson, et al. 2004). A su vez, la actividad motora producida mientras se escucha un estímulo verbal presenta una intensidad distinta dependiendo de si el estímulo escuchado implica un mayor o menor movimiento de los músculos de lengua y labios (Zhao, et al. 2017).

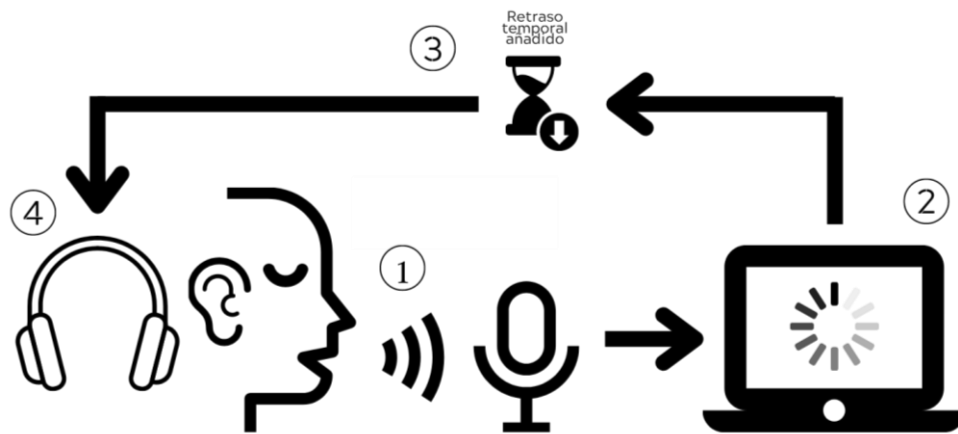
Por otro lado, existe evidencia que fundamenta que la percepción del habla propia es un elemento crucial para una correcta producción del discurso. Un fenómeno que confirma este planteamiento es el aumento de disfluencias en la articulación del habla en personas con sordera postlocutiva (Waldstein, 1990). Además, se ha encontrado que la percepción de un fonema que acabamos de articular, produce una respuesta neurofisiológica en la corteza auditiva menor, en comparación con la actividad generada cuando se escucha pasivamente el mismo fonema (Flinker, et al. 2010).

Uno de los acercamientos que ha permitido investigar las características de este circuito de retroalimentación auditivo-motor a nivel conductual son los protocolos de retroalimentación alterada del habla. Estos paradigmas experimentales implican la manipulación de los parámetros acústicos, como el tono (Chang, et al. 2013; Behroozmand & Larson, 2011) y/o la temporalidad del sonido producido por el habla propia, antes de que esta sea percibida por el hablante. Esta manipulación resulta en una alteración de la expectativa que el sujeto tiene sobre el habla que ha producido y su



percepción de la misma, lo que ocasiona que el o la hablante haga modificaciones en la forma que verbaliza para que la acción producida se ajuste a la expectativa generada (Guenther, 2016, p. 153-172).

Para los propósitos de este trabajo, nos centraremos en la alteración temporal de la señal del habla, es decir, la retroalimentación retardada del habla, también conocida como *DAF* por sus siglas del inglés *delayed auditory feedback*. Esta ocurre cuando se introduce de forma artificial un retraso en tiempo real, dentro del rango de los *ms*, entre la señal acústica producida por el hablante y la escucha de dicha señal generada por ella o él mismo (Lee, 1950). En la **figura 1** se muestra la configuración clásica para el desarrollo de un protocolo *DAF*.



**Figura 1. Diseño general de la estrategia experimental con DAF.** Se ubica al sujeto frente a un monitor y es equipado con auriculares. Se completa la tarea conductual bajo la siguiente secuencia: (1) El o la participante genera una vocalización que es captada por un micrófono. (2) La vocalización del sujeto se registra y procesa a partir de un equipo de cómputo o audio para luego devolver la señal acústica al participante con un retraso temporal agregado (3), el cual será reproducido de vuelta a través de unos auriculares (4).

Como resultado de la exposición al DAF, aumenta el número de errores en el habla, como lo pueden ser prolongaciones, repeticiones de sílabas o sonidos, incremento en la intensidad vocal o bloqueos, además de generar una disminución en la tasa silábica (es decir, el número de sílabas producidas por segundo) que es emitida (Fairbanks, 1955).

En contraste a los efectos producidos por el DAF en personas sanas, se ha demostrado que, en trastornos del habla, como la tartamudez, se presenta un efecto opuesto consistente en una reducción de alrededor del 25% de las disfluencias producidas (Hernández-Jaramillo & Gil-Lozada, 2013). Además, bajo condiciones donde se ha añadido un retraso en la retroalimentación del habla a

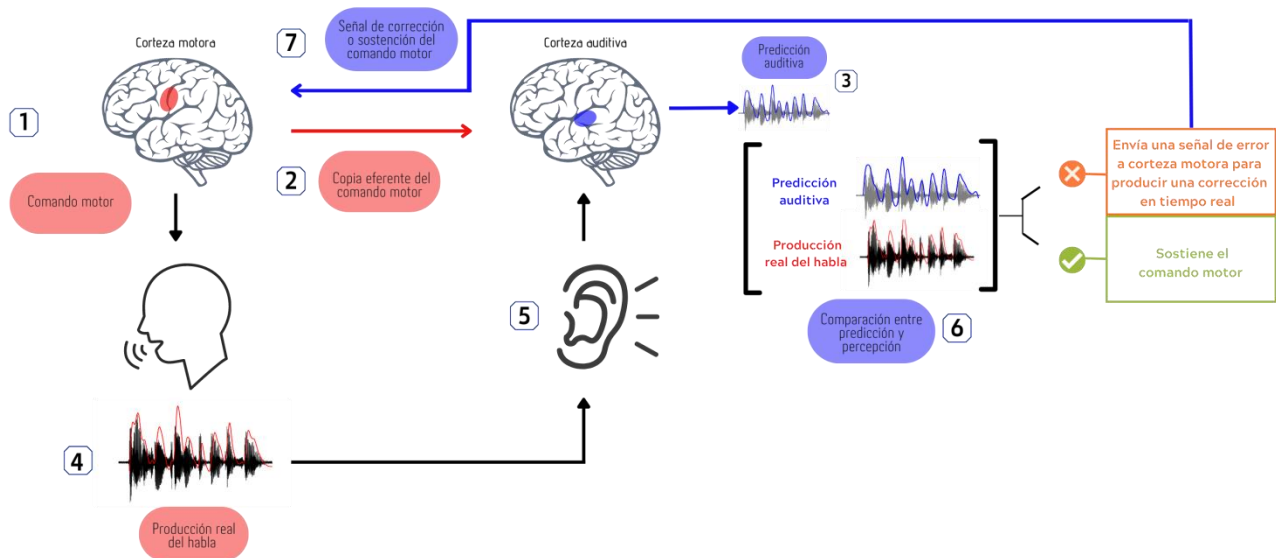
personas con tartamudeo severo, estas muestran una mejora significativamente mayor a la observada en quienes tienen tartamudeo menos grave (Toyomura, et al. 2020). En menor medida, también es posible identificar una mejora en la fluidez de personas que presentan una producción anormal del habla derivada de la enfermedad de Parkinson, al articular bajo los efectos del DAF (Brendel, et al. 2004).

Los hallazgos acerca de cómo opera el DAF en trastornos del habla, particularmente en la tartamudez, ha permitido explorar su potencial terapéutico mediante materiales de apoyo como el dispositivo SpeechEasy con el cual se manipula la temporalidad de la retroalimentación del habla para así aliviar temporalmente el tartamudeo del usuario (Foundas, et al. 2013).

Profundizar en el estudio de la interacción temporal entre los aspectos sensoriales y motores del proceso de habla podría enriquecer nuestra comprensión de los diversos trastornos del habla y proporcionar una base sólida para optimizar las intervenciones clínicas dirigidas a los mismos.

### **Antecedentes.**

Existen diversos modelos neuro computacionales (Houde & Rong, 2011; Tourville & Guenther, 2011) del sistema de producción del habla. Un elemento que dichos modelos tienen en común es un circuito de retroalimentación auditivo-motor, cuya función es el monitoreo en tiempo real de la señal acústica producida y corregir el gesto motor, de ser necesario. Más precisamente, se propone que al activarse un comando motor dirigido a producir habla en la corteza motora, se envía simultáneamente a la corteza auditiva una predicción de la consecuencia acústica de la ejecución motora. De este modo, la corteza auditiva puede hacer una comparación en tiempo real de la consecuencia vocal percibida con la deseada. En caso de que la producción real sea distinta de la que se esperaba escuchar, se genera una respuesta auditiva que codifica el error percibido, la cual le será retransmitida al área motora para que esta pueda hacer las correcciones pertinentes y así obtener la vocalización deseada (Ozker, et al. 2021, véase **Figura 2**).

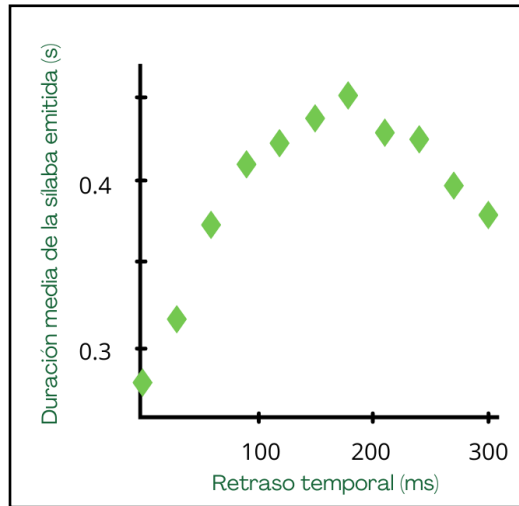


**Figura 2. El ciclo de retroalimentación auditivo-motora del habla.** En el momento en el que la corteza motora genera un comando motor (1) para una producción vocal, se manda una copia eferente de dicho comando a la corteza auditiva (2) a partir de la cual se genera una expectativa de la consecuencia sonora del habla (3). De manera simultánea a la producción motora real del habla y su posterior percepción (4 y 5), se lleva a cabo una comparación entre la vocalización percibida y la expectativa generada por la corteza auditiva (6), en donde se estima si hay una coincidencia. En el caso de que la producción vocal no concuerde con el estimado interno, la corteza auditiva manda una señal de error (7) a la corteza motora para así hacer las correcciones necesarias.

Es importante notar que, si bien la existencia de este circuito de retroalimentación auditivo-motora ha sido ampliamente validada, aún se desconocen sus características temporales. Es decir, cuál es la ventana temporal en la que se lleva a cabo la comparación entre la expectativa sonora y el sonido real del habla.

Una de las maneras en que es posible explorar las características temporales de este proceso de comparación es a partir de la aplicación de paradigmas *DAF*.

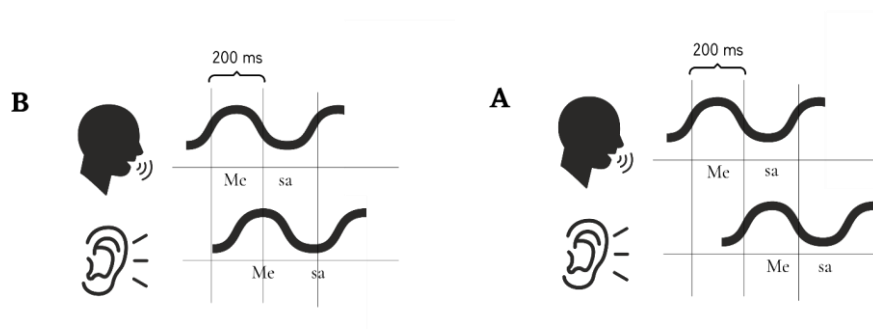
Se ha reportado que los efectos producidos por un experimento *DAF* (errores en el habla y disminución de la tasa silábica) se presentan con una mayor intensidad cuando el retraso temporal agregado a la retroalimentación del habla se encuentra alrededor de 200 ms (Zanini, et al. 1999). Si el retraso tiene valores menores o mayores a los 200 ms los efectos antes mencionados tendrán una menor intensidad. Véase **Figura 3**.



**Figura 3. Relación existente entre el tiempo de enunciación y el retraso temporal aplicado.** Representación de los datos reportados por Black (1951), en donde se detalla la duración media de las sílabas enunciadas por los sujetos en relación al retraso temporal añadido en el experimento. Es posible observar que al introducir retrasos temporales de alrededor de 200 ms los participantes tardaron más tiempo en vocalizar las sílabas que debían repetir. También se señala que una vez que se incluyen retrasos mayores a los 200 ms la duración de las sílabas emitidas disminuyó. Cabe aclarar que la gráfica de errores en función del retraso da lugar a una curva similar con un valor máximo en los 200 ms.

Este fenómeno puede ser explicado a partir de las posturas de autores como Black (1951), quien sugiere la posibilidad de que la duración de las sílabas producidas por un hablante funciona como segmentos críticos del control del habla (en otras palabras, representaría la ventana temporal de comparación-estimación de errores del circuito de retroalimentación auditivo-motor descrito en la **Figura 2**).

De esta forma, al introducir un retraso temporal cercano a los 200 ms, que es una duración que coincide con la de una sílaba producida en condiciones de habla conversacional en distintas lenguas (Kowal, et al., 1983), se genera una perturbación particularmente intrusiva entre lo que se enuncia y lo que se escucha, y esto da lugar a un efecto más pronunciado (Howell, et al. 1983). Por ejemplo, si se le pide a un sujeto que enuncie la palabra “mesa” mientras se ha introducido un retraso auditivo de 200 ms, el cual es cercano a la duración de una sílaba, éste podría estar escuchando /me/ en el momento en que está pronunciando /sa/, en cambio, si el retraso establecido es menor, este no generará una disrupción tan intensa como en el caso antes mencionado (Kaaspar & Rübeling, 2010). Véase la **Figura 4**.



**Figura 4. Disrupci3n en el habla resultante de la aplicaci3n de un retraso auditivo del habla.** Las curvas presentadas representan la enunciaci3n de la palabra “mesa” y su posterior retroalimentaci3n auditiva. En el inciso **A** se puede advertir que ante un retraso de 200 ms (equivalente a la duraci3n promedio de una s3laba durante velocidad de habla conversacional) genera una perturbaci3n considerable entre lo que el sujeto vocaliza y lo que escucha de esa enunciaci3n. En cambio, en el inciso **B**, en el que se representa la introducci3n de un retraso temporal menor la disrupci3n generada es menos acentuada.

### Planteamiento del problema.

A partir del an3lisis de la literatura existente, el hecho de que la incorporaci3n de retrasos temporales de 200 ms produce los efectos DAF en su forma m3s representativa, nos permite sugerir que la ventana en que se hace una comparaci3n de la expectativa auditiva con el sonido real del habla en el Ciclo de Retroalimentaci3n Auditiva-Motora del Habla podr3a tener una duraci3n de 200 ms.

No obstante, tambi3n es posible que este resultado se presente debido a que, como plante3 Black (1951), un retraso temporal de 200 ms tiende a ser m3s disruptivo porque esta medida temporal es similar a la duraci3n media de una s3laba producida en una velocidad de habla conversacional. Si la duraci3n de la s3laba es la que determina qu3 retraso temporal es el m3s disruptivo en un experimento DAF, el modificar la tasa sil3bica causar3a una variaci3n en el retraso temporal necesario para observar el m3ximo efecto.

Cabe destacar que son muy pocos los trabajos de investigaci3n en los que se ha tratado de establecer un control sobre la tasa sil3bica emitida por sus participantes (Stuart, et al. 2002; Hashimoto & Sakai 2003).

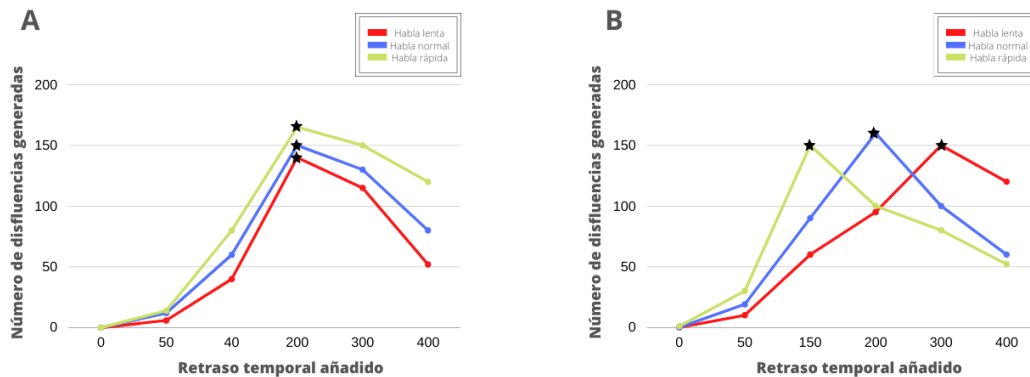
En vista de lo anterior, este trabajo tiene la finalidad de establecer un control m3s preciso de las tasas sil3bicas producidas por los participantes. De esta manera, se podr3 observar con mayor detalle c3mo los efectos de una manipulaci3n *DAF* interact3an con diferentes variaciones en la

duración silábica. El propósito es profundizar en el conocimiento sobre las características temporales del proceso de comparación en el Ciclo de Retroalimentación Auditivo-Motora del habla.

## Hipótesis

**H0.** La ventana de comparación de la retroalimentación auditivo-motora del habla se lleva a cabo en un periodo temporal fijo de 200 ms, independientemente de la duración de las sílabas emitidas por el sujeto, por lo que los efectos de la retroalimentación retardada del habla siempre tendrán su punto más alto al añadirse un retraso temporal de 200 ms. Véase figura 5 A.

**H1.** La ventana de comparación de la retroalimentación auditivo-motora del habla es variable en relación a la duración de las sílabas emitidas por el sujeto, de manera que los efectos más disruptivos



de la retroalimentación retardada del habla pueden presentarse en un retraso temporal mayor o menor en función de si se aumenta o disminuye la duración de las sílabas emitidas. Véase figura 5 B.

**Figura 5. Modelos correspondientes a las hipótesis 0 y 1.** En ambas representaciones se muestra el número de fallos en el habla en relación a los retrasos temporales añadidos. Cada color corresponde distintas tasas silábicas producidas. Las estrellas de color negro representan el punto máximo de errores. **A.** Hipótesis 0 el mayor número de disfluencias se encuentre al añadirse un retraso temporal de 200 ms ya que ese es el tiempo en el que el circuito de retroalimentación auditiva se lleva a cabo independientemente de la tasa silábica producida en cada condición. **B.** En la hipótesis 1 se establece que el número máximo de fallos del habla podría presentarse en puntos distintos en dependencia a la tasa silábica producida por el hablante. Si la duración media de las sílabas es menor, la mayor cantidad de errores se producen al añadir un retraso temporal menor a 200 ms, en cambio, si la duración media de la sílaba es mayor, el retraso temporal en el que se presentan más fallos del habla será mayor a los 200 ms.

## **Objetivos.**

### **Objetivo general.**

Obtener una descripción más detallada de la ventana temporal en la que opera la comparación entre la señal acústica predicha y el sonido real que se lleva a cabo en el ciclo de retroalimentación auditivo-motora del habla, mediante la observación de la interacción entre los efectos conductuales generados por el DAF y la tasa silábica producida por el hablante.

### **Objetivos específicos.**

- a. Puesta a punto de una estación de medición que permita llevar a cabo experimentos conductuales con retrasos temporales del habla. Para cada experimento se desarrollará un programa computacional que registre el habla del participante, la reproduzca posteriormente con un retraso temporal y recolecte grabaciones de las enunciaciones producidas por los participantes.
- b. Colecta y análisis de datos provenientes de los participantes para determinar si el aumento en disfluencias al introducir un retraso temporal guarda una relación directa con la frecuencia silábica producida por el hablante.

## **Métodos.**

### a) **Participantes.**

Se reclutó un total de 15 participantes de entre 20 y 40 años (8 mujeres, edad promedio= 28, DE= 5.014) hablantes nativos de español y quienes no reportaron trastornos cognitivos, del habla, psiquiátricos o de audición. Se excluyeron 5 participantes por no poder mantener la tasa silábica indicada (ver Criterios de exclusión).

### b) **Criterios de exclusión.**

Dado que el objetivo de este estudio fue poder tener un mayor control en la tasa silábica producida por los participantes, se estableció que aquellos o aquellas cuya tasa silábica difiriera

considerablemente de la tasa silábica solicitada serían excluidos. Este criterio se determinó a partir de la estimación de la tasa silábica en la condición donde no se ha añadido un retraso temporal al habla del participante para cada uno de los ensayos (ver más en el apartado de Análisis del experimento 1). Si el participante no lograba mantener una tasa silábica estable y diferenciada para cada una de las condiciones, este era excluido de los análisis posteriores.

### c) **Instrumentalización.**

Los y las participantes se ubicaron frente a una computadora y cerca de un micrófono (SHURE PGA48-XLR) conectado al equipo a través de una interfaz de audio (Focusrite Scarlett 2i2). Se les proporcionaron auriculares especiales para audiometría (ETYMOTIC ER1 Insert Earphones) para minimizar la retroalimentación del habla a través de la vibración ósea (Békésy, 1949). El protocolo DAF utilizado en este estudio se implementó mediante un programa creado en el software *Psychtoolbox* para *MATLAB*.

### **Diseño experimental.**

Cada ensayo del experimento (véase *figura 6*) consiste en tres fases: (1) El o la participante escucha con atención un audio de entrenamiento que consistente en la concatenación de las sílabas /ta/, /te/, /ti/, /to/, /tu/ a una tasa específica (tres, cuatro o seis sílabas por segundo), respetando que se presenten siempre las mismas sílabas y en el mismo orden, mientras observa un punto negro en el centro de la pantalla de la computadora; (2) el punto al centro de la pantalla cambia a color verde y la persona comienza a repetir el tren de silabas a la misma frecuencia mostrada en el entrenamiento mientras recibe retroalimentación de su habla con un retraso de 30 ms. A esta etapa en el experimento se le denomina *BASAL*. Finalmente, (3) el punto verde desaparece por unos segundos en los que el o la participante detiene su vocalización; el punto verde reaparece en la pantalla y el o la participante reanuda inmediatamente la repetición de las sílabas escuchadas mientras se introduce un retraso en la retroalimentación auditiva. Cada ensayo pertenecerá a una de 6 posibles condiciones en donde se añada alguno de los retrasos temporales (30, 80, 160, 240, 320 y 400 ms) de manera aleatoria. A esta fase se le denominó como *DAF*.

La duración de cada una de las fases del experimento está determinada por un número aproximado de sílabas emitidas. Para la fase *BASAL* se espera que las y los sujetos enuncien un

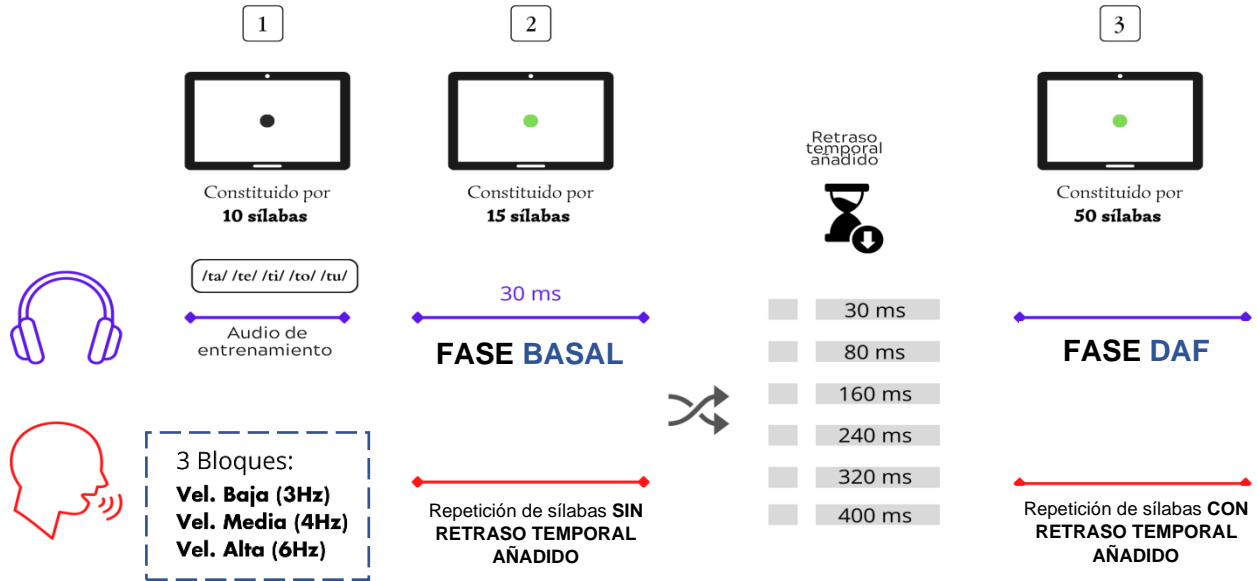


aproximado de 15 sílabas, mientras que la fase *DAF* estará compuesta por alrededor de 50 sílabas. El tiempo de duración de cada ensayo varía dependiendo de la duración de la sílaba correspondiente a cada bloque (0.33 s para 3 Hz, 0.25 s para 4 Hz y 0.16 s para 6 Hz). Cada uno de los bloques se constituye por 72 ensayos, dando un total de 216 ensayos en totalidad del experimento.

La condición *BASAL* será presentada con un retraso de 30 ms, el cual se ha reportado como imperceptible para los participantes y como incapaz de producir cambio alguno en el habla, es decir, no produce errores ni enlentecimiento (Black, 1951).

El experimento en su totalidad se realiza en tres bloques correspondientes a cada una de las tasas silábicas requeridas (tres, cuatro o seis sílabas por segundo), los cuales fueron realizados en días distintos con la intención de contrarrestar los efectos causados por el cansancio o frustración.

Previo a cada ensayo, se realizó un entrenamiento adicional constituyente a tres ensayos idénticos a los que realizarían en el experimento mientras la experimentadora se encontraba presente, con la intención de identificar si las o los participantes necesitaban de retroalimentación para reforzar la indicación de mantener una tasa silábica en particular.



**Figura 6. Representación gráfica del experimento.** El experimento fue dividido en tres bloques distintos correspondientes a cada una de las tasas silábicas (tres, cuatro y seis sílabas por segundo). Cada ensayo estuvo conformado por tres momentos. (1) En primer lugar, la o el participante escucha con atención el audio de entrenamiento que ejemplifica la tasa silábica requerida, mientras este audio se reproduce se podrá ver un punto negro al centro de la pantalla, una vez terminado. (2) Condición *BASAL*, el punto de la pantalla se vuelve verde y la persona repite las sílabas

escuchadas a la misma velocidad mientras escucha la retroalimentación de su voz con 30 ms. (3) Finalmente, el punto verde desaparece por algunos segundos en los que el o la hablante para de hablar; el punto verde reaparece en la pantalla y el o la participante continúa la repetición de sílabas mientras que escucha su voz con uno de los seis retrasos temporales asignados en forma aleatoria para cada ensayo.

## **Análisis de datos.**

### **Cálculo de los parámetros de estudio.**

Se realizaron dos tipos de tratamiento de los datos obtenidos del proceso experimental: (i) el conteo de errores en la articulación producidos por el sujeto y (ii) la estimación de la tasa silábica tanto en la condición *BASAL*, como en la condición *DAF*.

### **Conteo de disfluencias y/o errores de habla.**

Mediante el uso del software *Praat* se revisaron los audios de las vocalizaciones de los sujetos durante la fase *DAF* para marcar de manera manual el número total de errores en el habla producidos, los cuales se clasificaron en cuatro categorías distintas:

- a) Repeticiones.* Reiteración de la sílaba emitida anteriormente.
- b) Intercambios.* Modificación en el orden de las sílabas de la secuencia.
- c) Titubeos.* Indecisión o falta de claridad en la pronunciación.
- d) Errores de articulación.* En donde el sujeto enunció sílabas o sonidos totalmente distintos a los establecidos en la secuencia de sílabas presentadas

### **Cálculo de la tasa silábica producida por los participantes.**

Se desarrolló un programa en MATLAB que utilizó el valor absoluto de la transformada de Hilbert para extraer la envolvente de la señal acústica registrada en las grabaciones de ambas condiciones (*BASAL* y *DAF*). Luego se estimó el espectro de frecuencias de cada envolvente y se consideró el valor de frecuencia con máxima potencia como la tasa silábica del ensayo correspondiente.

### **Criterio de inclusión construido a partir de la tasa silábica en la condición *BASAL*.**

Con el fin de determinar si las y los participantes generaron tasas silábicas diferenciadas en las tres condiciones del experimento, se estableció un criterio de inclusión utilizando los datos de la frecuencia silábica en la condición *BASAL* de cada participante. Para ello, se realizó un cálculo de

medias recortadas, el cual consiste en la exclusión de valores extremos (es decir, los valores que difieren significativamente de la mayoría de los datos en conjunto) para, posteriormente, calcular la media.

Esto permite tener una medida de la tendencia central más precisa y representativa de los datos de cada una de las condiciones, para este estudio en particular los datos extremos excluidos consistieron en el 10% de los datos de los extremos. Aquellos ensayos en los que la frecuencia silábica se alejaron más de un 20% de la media fueron excluidos de los análisis posteriores. Además, aquellos sujetos con más de 15 ensayos excluidos en alguna de las condiciones (3, 4 o 6 Hz) fueron descartados del estudio.

### **Análisis de varianza (ANOVA) de medidas repetidas para comparar el efecto de los diferentes retrasos temporales en la tasa silábica producida.**

Un análisis de varianza de medidas repetidas es una herramienta estadística poderosa utilizada para examinar cómo los cambios en una o más variables independientes afectan los valores observados de una variable dependiente, a medida que los mismos sujetos o elementos son evaluados en diferentes condiciones a lo largo de distintas condiciones. Este análisis se realizó con dos factores intra-sujeto (condición de tasa silábica y retraso temporal). Un factor fueron las tasas silábicas con tres niveles: baja, media y alta (3, 4 y 6 Hz respectivamente). El otro factor fue el retraso temporal añadido con 6 niveles: 30, 80, 160, 240, 320 y 400 ms).

Adicionalmente, cuando se encontraron efectos significativos, estos se investigaron en profundidad mediante análisis post hoc corregidos por el método de Holm.

### **Ajuste de curvas Weibull a los datos de errores cometidos en función de los retrasos temporales añadidos.**

Como complemento para observar el comportamiento de los errores cometidos en cada una de las condiciones de tasa silábica se utilizó una herramienta de ajuste de curvas de MATLAB para ajustar curvas Weibull, ya que la forma de este tipo de funciones presenta cierto parecido a los datos que se suelen obtener en experimentos DAF donde se usan distintos retrasos temporales, ya que el número de errores aumenta de forma progresiva para después alcanzar un punto máximo de errores que posteriormente volverá a disminuir pero con menor pendiente.

El ajuste de las curvas a los datos de cada una de las condiciones está modelado por la siguiente fórmula:

$$a * b * x^{b-1} * \exp^{-a*x^b}$$

En donde  $a$  controla la amplitud de la curva y  $b$  determina la forma de la curva y cómo se distribuyen los valores alrededor de  $a$ , además de la posición en el eje horizontal.

### **Análisis estadístico para la disminución de la tasa silábica en DAF.**

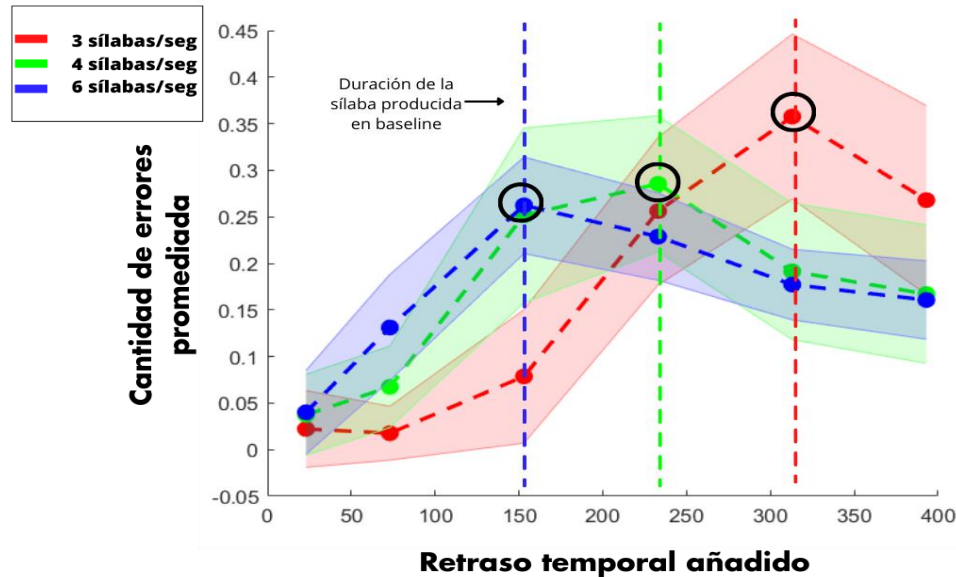
Para cada sujeto y ensayo se calculó la modificación en la tasa silábica inducida por el DAF; para esto, se dividió la tasa silábica durante el DAF por la obtenida durante la condición *BASAL*. Esto nos permite estudiar cuánto se modifica la tasa silábica al incluir un retraso temporal, un valor 1 indicaría que no hay efecto alguno del DAF sobre la tasa silábica. Este valor se promedió por sujeto, retraso temporal y condición de tasa silábica. Para estudiar si existe un enlentecimiento se compararon los valores obtenidos de la tasa silábica normalizada contra 1 por medio de una prueba t de student, corregida por múltiples comparaciones con el método de corrección FDR (False Discovery Rate), técnica estadística que controla la cantidad de falsos positivos al analizar múltiples pruebas o comparaciones (Jafari & Ansari-Pour, 2019).

## **Resultados.**

### **Cantidad de disfluencias del habla en relación al retraso temporal añadido.**

En la *Figura 7* se muestra el promedio de errores cometidos por todas y todos los participantes de la muestra en función de los retrasos temporales añadidos en los tres bloques del experimento. Es posible observar que el retraso temporal necesario para provocar la mayor cantidad de errores varía para cada una de las condiciones de tasas silábicas. En la condición de seis sílabas por segundo el punto máximo de disfluencias se alcanza al añadir un retraso temporal de 160 ms, la cantidad más representativa de errores en la condición de habla de cuatro sílabas por segundo se presenta al añadir un retraso temporal de 240 ms, y en la condición de tres sílabas por segundo la mayor cantidad de errores ocurre al añadir un retraso temporal de 320 ms.

En cada condición el retraso temporal que generó el efecto más pronunciado en cada condición de habla se alinea bien con su correspondiente duración media de la sílaba producida en la condición *BASAL*.



**Figura 7. Cantidad de errores cometidos en relación al retraso añadido.** Se muestra en el eje x retraso temporal añadido y en el eje y el número de errores cometidos por los participantes. Adicionalmente se incluye una línea punteada que representa la duración silábica objetivo para cada una de las condiciones basales. El punto máximo de disfluencias (el cual se denota con un círculo) para cada una de las tasas silábicas se presenta en tres puntos distintos. La única de las condiciones donde se cumple la noción clásica de que al añadir 200 ms de retraso temporal los errores del habla son mayores es en la que se ha pedido al participante producir 4 sílabas por segundo. El área sombreada que se observa en la trayectoria de los datos de cada una de las condiciones corresponde a la desviación estándar de los mismos.

Para corroborar estadísticamente estos hallazgos, se realizó un ANOVA de medidas repetidas con dos factores intra-sujeto:

1. Tasa silábica. Con tres niveles: baja, media y alta (3, 4 y 6 Hz, respectivamente).
2. Retrasos temporales. Con seis niveles: 30, 80, 160, 240, 320 y 400 ms.

Los resultados de este análisis mostraron que, de forma relevante para nuestra hipótesis, existe un efecto significativo de la interacción condición de tasa silábica con retraso temporal ( $F [10, 5.07] = 20.47, p < 0.001, \eta^2 = 0.22$ ).

Para complementar el resultado anterior se realizó un análisis post hoc para observar las diferencias entre los errores cometidos al añadir un retraso temporal de 30 ms (que según la literatura no se produce efectos, además aquí vemos que no distingue entre condiciones de tasa silábica) y los cometidos por retrasos más grandes. Los resultados indican que el punto en donde al agregarse un retraso temporal la cantidad de errores cometidos empieza a ser significativamente distinta es diferente para cada una de las tasas silábicas. En la condición de seis sílabas por segundo, se observó un efecto significativo al agregar un retraso temporal de 80 ms ( $p = 0.029, t = -3.543$ ). Para la condición de cuatro sílabas por segundo se encontró un efecto significativo al añadir un retraso de 160 ms ( $p < 0.001, t = -8.338$ ). Finalmente, en la condición de tres sílabas por segundo se indicó un efecto significativo al agregar un retraso de 240 ms ( $p < 0.001, t = -9.105$ ).

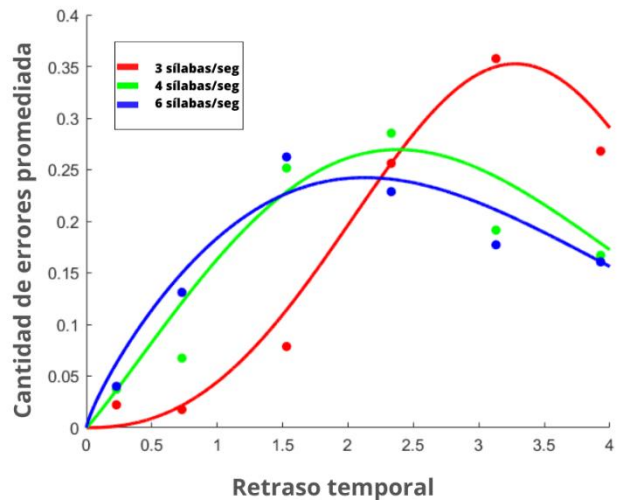
Como complemento a los hallazgos anteriores, se realizó un ajuste de curvas Weibull a los datos antes procesados mediante una herramienta de ajuste de curvas de *MATLAB*, buscando una representación más precisa de las tendencias subyacentes.

En la *figura 8 A* se observan los parámetros obtenidos del ajuste y en la *figura 8 B* la representación gráfica de los mismos. Este ejercicio nos mostró que la curva ajustada a los datos correspondientes a la condición de tasa silábica de 3 Hz es distinta de las curvas ajustadas a las condiciones de 4 y 6 Hz, ya que los valores dentro de su intervalo de confianza no se superponen (véase *figura 8 A*).

A

Condición de velocidad de habla	a	b
3 Hz	0.0135 (0.003658, 0.02336)	3.325 (2.774, 3.874)
4 Hz	0.08565 (0.02747, 0.1438)	2.077 (1.539, 2.616)
6 Hz	0.1139 (0.06373, 0.164)	1.806 (1.468, 2.145)

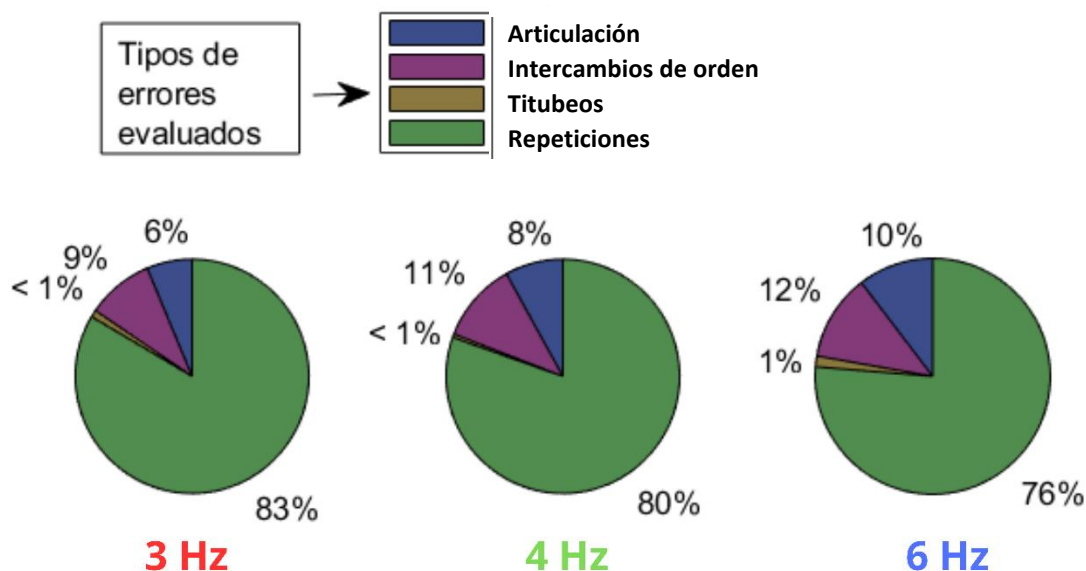
B



**Figura 8. Curvas Weibull ajustadas a los datos de errores producidos.** En el inciso **A** se registran los parámetros obtenidos del ajuste a los datos de cada condición, mostrando en la parte superior el valor que se considera el coeficiente que mejor se ajusta al parámetro y en la parte inferior los límites del intervalo de confianza ajustado, el cual cuenta con una precisión del 95%. Además, en el inciso **B** se muestran representados en puntos del color correspondiente a cada condición la cantidad de errores promediada para cada uno de los retrasos temporales añadidos. Se muestra igualmente graficadas con los colores pertenecientes a cada tasa silábica emitida las curvas que se han ajustado a los datos.

Estos hallazgos implican que la relación entre el retraso temporal y el efecto producido es específica para cada tasa silábica. En otras palabras, la cantidad de retraso temporal necesario para comenzar a observar el efecto *DAF* y el que se necesita para alcanzar la mayor cantidad de errores son distintos en relación a la duración de la sílaba producida.

Adicionalmente, en la **Figura 9** se presenta el porcentaje de cada uno de los tipos de errores cometidos correspondientes al retraso que genera una mayor cantidad de errores para cada condición. En su mayoría, nos encontramos con errores de repetición, los cuales representan más del 80% para las tres condiciones, lo cual nos indica que es el tipo de error más exhibido independientemente de la tasa silábica producida por el hablante. Este hallazgo se alinea con los datos antes reportados por Malloy, et al. (2022). Además, el hecho de que las distribuciones de errores se mantengan iguales a través de las condiciones sugiere que son consecuencia de un mismo mecanismo.



**Figura 9. Porcentaje de cada tipo de errores cometidos.** Se representa el porcentaje de cada uno de los tipos de errores para las distintas condiciones de tasas silábicas (articulación marcada en azul, intercambio marcado con morado, titubeo marcado con color oliva y repetición marcado con verde). En los tres casos la mayoría de los errores corresponden a repeticiones silábicas, representando más del 80% en todos los casos. Los errores de articulación e intercambio constituyen entre el 16 y 18% de los equívocos presentes en todas las condiciones. Los errores de titubeo se presentaron en porcentajes por debajo del 2%.

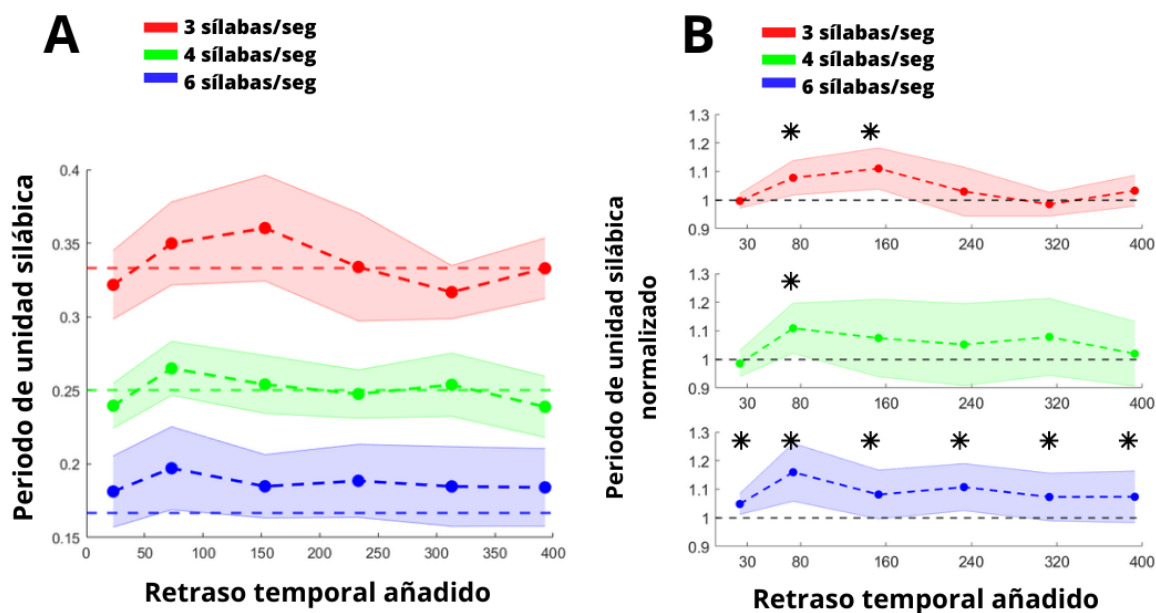
### **Estimación de tasa silábica en condición BASAL y DAF.**

En la *Figura 10 A* se representa gráficamente el promedio de la duración de la unidad silábica producida en la condición *DAF* por los participantes de la muestra, es decir, la duración media de las sílabas producidas por los sujetos en la condición *DAF*. En cambio, en la *Figura 10 B* se presentan los valores de la unidad silábica normalizada, esto se logra dividiendo el valor de la unidad silábica efectuada en la fase *DAF* entre el valor de la unidad silábica producida en la fase *BASAL*. El objetivo es determinar si ha habido un cambio en la tasa silábica entre estas dos condiciones. Para lograrlo, evaluamos cuánto se aleja el resultado obtenido del valor uno.

Para verificar estadísticamente los efectos de los retrasos temporales en la tasa silábica, se compararon los datos obtenidos en cada condición y retraso temporal contra 1. Este análisis muestra que para la condición de 3Hz de los seis retrasos temporales añadidos sólo los valores de 80 y 160 ms produjeron un enlentecimiento significativo (30:  $p_{\text{corr}} = 0.670$ ; 80:  $p_{\text{corr}} p < 0.001$ , 160:  $p_{\text{corr}} p < 0.001$ ; 240:  $p_{\text{corr}} = 0.238$ , 320:  $p_{\text{corr}} = 0.238$ ; 400:  $p_{\text{corr}} = 0.065$ ). Para la condición de 4 Hz sólo el retraso temporal de 80 ms (30:  $p_{\text{corr}} = 0.297$ ; 80:  $p_{\text{corr}} p < 0.001$ , 160:  $p_{\text{corr}} = 0.104$ ; 240:  $p_{\text{corr}} = 0.278$ ,



320:  $p_{\text{corr}} = 0.104$ ; 400:  $p_{\text{corr}} = 0.519$ . Finalmente, para 6Hz todos los retrasos temporales, incluso el de 30 ms reportado en la literatura como imperceptible, modifican significativamente la velocidad del habla (30:  $p_{\text{corr}} p < 0.001$ ; 80:  $p_{\text{corr}} p < 0.001$ , 160:  $p_{\text{corr}} = 0.003$ ; 240:  $p_{\text{corr}} p < 0.001$ , 320:  $p_{\text{corr}} = 0.005$ ; 400:  $p_{\text{corr}} = 0.007$ ). Se observa que para las tres condiciones de tasa silábica existen variaciones en la duración de las unidades silábicas que no concuerdan con lo que se ha reportado antes en la literatura, ya que se esperaría que el retraso temporal que provoca la mayor cantidad de errores en el habla también resultaría en una mayor disminución en la tasa silábica, sin embargo, los resultados de este estudio no respaldan esta observación.



**Figura 10. Duración promedio de la unidad silábica producida en DAF.** En ambos gráficos se presenta en el eje de las  $x$  el retraso temporal añadido y en el eje de las  $y$  la duración de la unidad silábica promedio producido por los participantes en la condición DAF. **A.** Se observa además marcado con una línea punteada con los colores correspondientes a cada condición la duración de una unidad silábica producida para cada una de las condiciones. **B.** Se muestran los datos promedio de la unidad silábica producida normalizados al dividir el valor de la unidad silábica generada en DAF entre la unidad silábica promedio producida en la etapa BASAL para cada uno de los retrasos temporales añadidos los cuales fueron comparados contra el valor 1, representado por una línea punteada.

En resumen, este diseño experimental reveló dos resultados principales:

- a. El retraso temporal necesario para observar el número máximo de errores en el habla se modifica en función de la tasa silábica producida por la o el participante. De forma similar, el retraso temporal requerido para empezar a producirse una cantidad de errores significativamente distinta de cero también cambia según la condición de habla.

- b. Interesantemente, el retraso temporal que generó la mayor cantidad de errores en cada condición no mostró una reducción más pronunciada en la tasa silábica en comparación a la producida en la condición *BASAL*. Esto es especialmente notable en el caso de la condición de tasa silábica alta (6 Hz), donde se observó una disminución en la tasa silábica para todos los retrasos temporales agregados.

Derivado de estos hallazgos, surgieron distintas preguntas que nos motivaron a abordar de manera más detallada los fenómenos observados. Específicamente, estas preguntas se centran en dos aspectos principales:

- a. Indagar si existen otros mecanismos derivados de una manipulación *DAF*, además de la generación de errores y la disminución en la tasa silábica, que se manifiesten de formas distintas al examinarse haciendo variaciones en la duración de las sílabas emitidas
- b. Analizar las razones detrás del comportamiento observado para los datos de disminución de la tasa silábica en la condición de 6 Hz e investigar si existen factores adicionales que pudieran influir en este fenómeno.

Es por ello, que nos ocupamos en diseñar dos experimentos complementarios que no sólo se dirigieran a llenar los vacíos de información identificados, sino también a crear una visión más completa de la dinámica subyacente.

## **Diseño de experimentos complementarios.**

### **Participantes.**

Se reclutó un total de 15 personas adicionales a las que participaron en la ronda anterior de experimentos de entre 24 y 34 años (10 hombres, edad promedio= 27.5, DE= 3.41), hablantes nativos de español y que no reportaron tener trastornos cognitivos, del habla, psiquiátricos o de audición. Se excluyeron a 3 participantes por no poder mantener la tasa silábica indicada (ver Criterios de exclusión).

#### **I. Interacción entre la tasa silábica y la consciencia del efecto DAF.**

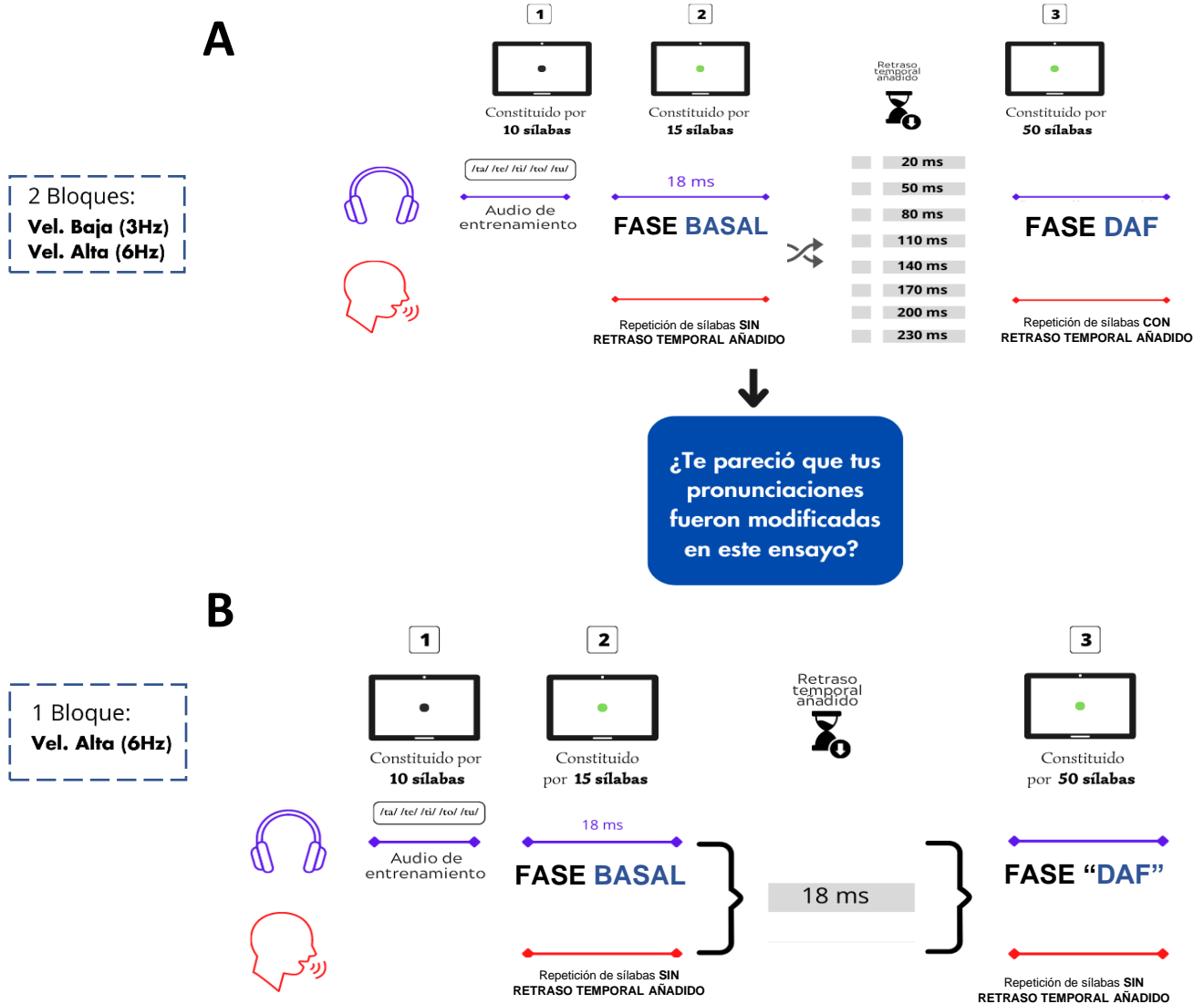
Con la intención de explorar otro tipo de efectos derivados de una manipulación *DAF* y su interacción con la tasa silábica producida, elegimos evaluar el momento en el que los participantes empiezan a tomar consciencia de la alteración en su retroalimentación del habla. Para ello, se tomó el diseño experimental base (véase Figura 11) y se han modificado los valores de retraso temporal que se añaden (20, 50, 80, 110, 140, 170, 200 y 230 ms) y al final de cada uno de los ensayos se le

pide al participante que conteste la pregunta “¿Te pareció que tus pronunciaciones fueron modificadas en este ensayo?” a la que podrá contestar con un sí o no. Constituido por dos bloques correspondientes a las condiciones de tasa silábica baja y alta (3 y 6 sílabas por segundo) con 96 ensayos para un total de 288 ensayos en el experimento completo. (Véase figura 11A).

## **II. Control para tasa silábica de 6Hz.**

Los resultados obtenidos al analizar el efecto del DAF en la tasa silábica para la condición de 6 Hz en el experimento original mostraron una tendencia inesperada. Consideramos la posibilidad de que la disminución en la tasa silábica, incluso al añadir diferentes retrasos temporales que teóricamente no deberían causar dicho enlentecimiento, pueda deberse a otros factores. Estos podrían incluir la dificultad de mantener una tasa silábica constante de 6 sílabas por segundo o la frustración incidida al participar en el experimento DAF (Marshall, 2021).

Para investigar más a detalle se ha utilizado una vez más el diseño experimental base y se le modificaron los valores de retraso temporal añadidos para que en ambas condiciones (*BASAL* y DAF) se añada un solo retraso temporal equivalente a 18 ms, tal como se observa en la **figura 11B**, esto para determinar si la tasa silábica producida por el o la participante se mantiene estable a lo largo de cada ensayo. Se aplicó en un bloque de 12 ensayos por participante.



**Figura 11. Diseños experimentales complementarios.** En el inciso **A** se presenta el diseño experimental creado manteniendo una dinámica similar a la del experimento base, con la diferencia de que se han agregado 8 valores distintos de retraso temporal (20, 50, 80, 110, 140, 170, 200 y 230 ms) y se ha aplicado solamente a las condiciones de tasa silábica baja y alta (3 y 6 Hz). Además, se ha añadido una pantalla al final de cada ensayo, en la cual se le pregunta al sujeto si ha percibido alguna modificación en sus producciones de habla. En el inciso **B** se muestra el experimento para establecer un control más preciso de la condición de tasa silábica alta (6Hz). De igual manera, se ha mantenido la misma estructura del experimento original, pero se ha añadido un valor de retraso temporal mínimo equivalente al presentado en la condición BASAL.

## **Análisis de datos complementarios.**

### **Cálculo del porcentaje de ocasiones en que los sujetos fueron conscientes del efecto**

#### **DAF en el experimento para evaluar la consciencia del efecto.**

De las respuestas proporcionadas por los y las participantes para cada uno de los ensayos de ambos bloques en el experimento, se extrajeron las respuestas *sí*, las cuales representan las ocasiones en las que había consciencia de la manipulación *DAF* aplicada. A continuación, se calculó el porcentaje de estas respuestas para cada uno de los retrasos temporales añadidos.

Posteriormente, se ajustó una curva sigmoidea a los datos de cada uno de los sujetos para determinar el punto de inflexión en las respuestas proporcionadas (tanto para 3, como 6 Hz). De igual manera, se realizó otro ajuste a una curva sigmoidea a los datos promediados de todos los sujetos para determinar el punto de inflexión para cada una de las condiciones.

#### **Estimación de la tasa silábica para el control de 6 Hz.**

De la misma manera en que se realizaron los análisis para estimar la tasa silábica en el experimento original, se extrajo la envolvente de la señal acústica registrada para posteriormente obtener sus espectros de frecuencias y determinar el valor de la tasa silábica producida. Se promediaron los valores obtenidos de cada uno de los ensayos y fases para así contar con un solo valor por participante.

## **Resultados complementarios.**

### **Cálculo del porcentaje de ocasiones en las que hay una consciencia del efecto.**

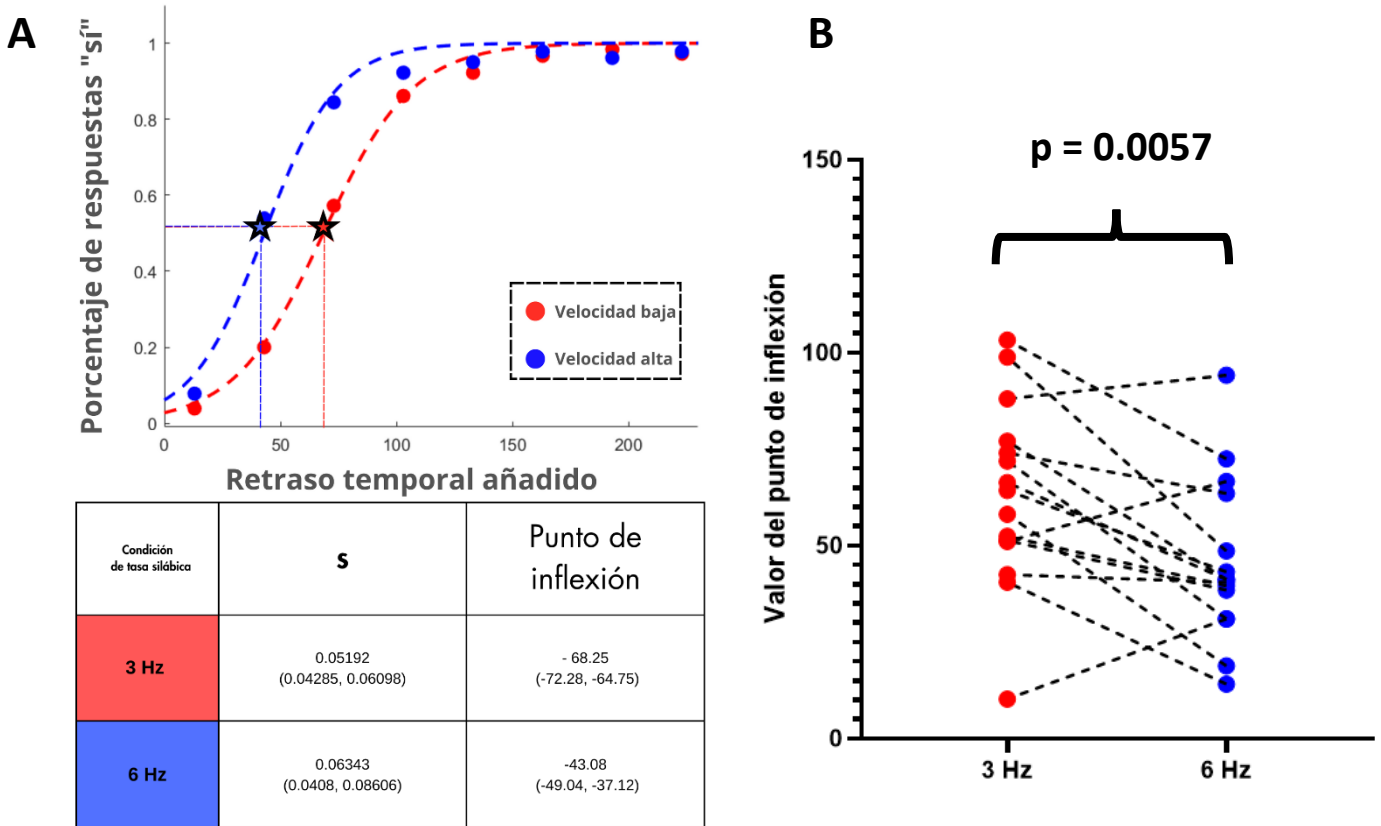
A raíz del hallazgo de que los efectos del *DAF* se manifiestan de forma distinta según la tasa silábica que se esté produciendo, consideramos la posibilidad de explorar otros efectos del *DAF* para observar si también se presenta una variación establecida por las distintas condiciones de tasa silábica. Optamos por evaluar el punto en donde los participantes comienzan a ser conscientes de los efectos generados por el *DAF*.

En la **Figura 12 A** se muestra el promedio del porcentaje de ocasiones en las que las y los participantes fueron conscientes de que se les ha manipulado la retroalimentación del habla en relación a los diferentes retrasos temporales añadidos, esto para ambas condiciones de tasa silábica, así como las curvas que se ajustaron a los datos marcadas con una línea punteada.

También se presenta el registro de los parámetros que se obtuvieron del ajuste correspondientes al punto de inflexión, es decir, el punto donde se produce un cambio perceptible en la respuesta de los participantes, marcando el umbral en el cual la detección o discriminación de un estímulo cambia significativamente. Asimismo, se registran los valores del parámetro  $s$  que controla la pendiente de la curva ajustada lo que determina la rapidez en la que cambia la percepción del participante.

Es posible apreciar que las curvas ajustadas tienen una pendiente similar, pero el punto de inflexión correspondiente a cada tasa silábica se encuentra en valores distintos, siendo el punto de inflexión para la condición de tasa silábica alta (6 Hz) menor que el valor correspondiente a la tasa silábica baja (3 Hz). Esto añade evidencia que respalda al planteamiento de que los efectos generados por el *DAF* pueden variar al modificarse la tasa silábica emitida.

Para verificar estadísticamente si los puntos en los que los sujetos empiezan a ser conscientes del efecto generado son diferentes, se ajustaron curvas sigmoideas a los datos de cada uno de los participantes, de las cuales se obtuvo el punto de inflexión para cada sujeto y condición de tasa silábica. Luego, se compararon los valores obtenidos en cada condición utilizando una prueba  $t$  de student pareada, la cual indicó una diferencia significativa entre condiciones ( $p = 0.0057$ ,  $t = 3.25$ ), tal como se muestra en la **Figura 12 B**.



**Figura 12. Interacción entre la consciencia del efecto DAF y la tasa silábica producida.** En el inciso A podemos observar representados por puntos el porcentaje promedio de respuestas sí tanto para la condición de 3 Hz (marcada en color rojo), como para 6 Hz (marcada en color azul) y las curvas sigmoideas ajustadas a cada una de las condiciones con líneas punteadas, con los valores de sus parámetros registrados en la tabla inferior. Adicionalmente se señalan los puntos de inflexión de cada curva con una estrella. En el inciso B se muestran los valores de los puntos de inflexión para cada uno de los sujetos representados por un punto y siendo comparados entre condiciones, donde se muestra que son significativamente distintos.

**Control de la condición de velocidad alta (6Hz) con retraso temporal mínimo.**

Basándonos en los resultados obtenidos del análisis del experimento original, decidimos explorar si existía algún factor adicional que estuviera generando una disminución en la tasa silábica como resultado de todos los retrasos temporales añadidos, como lo podría ser el cansancio. Con este fin, pedimos a los participantes realizar el experimento sin incorporar ningún retraso temporal, es decir, manteniendo la retroalimentación del habla en 16 ms (el mínimo retraso posible por el hardware utilizado) tanto en la condición *BASAL* como en la de *DAF*. Posteriormente, observamos si se presentaba un enlentecimiento similar al observado en los resultados anteriores.





al valor promedio de cada uno de los sujetos. La línea punteada de color azul representa el valor de 1, frente al que se han comparado los datos de los sujetos.

## Discusión.

La finalidad del presente trabajo fue investigar el valor temporal de la ventana de comparación del Ciclo de Retroalimentación auditivo-motora del habla, mediante un experimento de Retroalimentación Retardada del Habla (DAF). Es posible que, con base en lo que se ha reportado durante años en la literatura, el valor temporal que buscamos sea 200 ms, ya que produce el efecto más significativo en un paradigma DAF. Sin embargo, se conoce que la duración de una sílaba en una velocidad de habla conversacional (entre 4 y 5 sílabas por segundo) se acerca a los 200 ms, y no ha habido un control adecuado de la tasa silábica producida en los trabajos previos de investigación DAF. Por lo tanto, planteamos la posibilidad de que la ventana temporal de comparación en el Ciclo de Retroalimentación auditivo-motora del habla se modifica en relación a la duración de las sílabas emitidas por el hablante. Esto implica que en un experimento DAF aplicando el efecto ante distintas tasas silábicas, el retraso temporal necesario para observar el efecto máximo no siempre será de 200 ms, sino que dependerá de la duración de las sílabas pronunciadas.

Para abordar este punto, empleamos distintas estrategias para examinar cómo los diferentes efectos generados por la manipulación DAF interactúan con tres tasas silábicas distintas. Nuestros resultados revelan principalmente dos hallazgos significativos:

### I. **El periodo temporal de comparación en el ciclo de retroalimentación auditivo-motora del habla escala con la duración silábica.**

Se determinó que la duración del proceso de retroalimentación auditivo-motora en el habla está parametrizada por la duración de la sílaba que se produzca, de manera que se comprueba nuestra hipótesis H1 (Véase *figura 5 B*). Esto se sustenta a partir de las interacciones del número de errores cometidos y los retrasos temporales agregados y en el hecho que el retraso más disruptivo coincide con la duración silábica que llevaba el hablante previo a aplicar el DAF. Estos hallazgos sustentan la hipótesis de Black (1951) en la que se plantea que lo que lleva el control de los mecanismos de compensación durante un experimento DAF es la duración de la sílaba.

Esta conclusión es reforzada por el hecho de que el retraso temporal necesario para que los participantes comiencen a cometer una cantidad significativa de errores y el requerido para que se percaten del cambio en su retroalimentación del habla, tiene valores distintos para cada una de las condiciones de tasa silábica.

En este sentido, podríamos proponer que la duración de la ventana de comparación del Ciclo de Retroalimentación Auditivo-motora del habla es equivalente a la duración de la sílaba emitida. Por consiguiente, es crucial enfatizar que considerar este descubrimiento permitirá crear modelos aún más efectivos que permitan un avance en la comprensión de dicho ciclo.

También es importante recalcar que este puede ser un antecedente importante para sugerir un control más riguroso de las variables evaluadas dentro de un experimento DAF. El conocer que los efectos producidos pueden variar abre un camino de posibilidades para caracterizar más a fondo los mecanismos que dan respuesta a este tipo de manipulaciones.

## II. **Los mecanismos que producen errores en el habla podrían ser distintos a los que generan una disminución en la tasa silábica emitida.**

De manera inesperada, nuestros resultados no replican los mostrados en otros estudios como el de Kalinowski et al (2002) en los que se el retraso temporal donde se produce el mayor número de disfluencias en el habla es el mismo que se requiere para generar el máximo enlentecimiento en la velocidad del habla. En cambio, los datos de nuestra muestra arrojan que para la condición de velocidad baja (3 Hz) solamente los retrasos temporales de 80 y 160 ms generan una disminución en la velocidad. Para la condición de velocidad media (4 Hz) solamente un retraso de 80 ms. Posteriormente, el resto de los retrasos temporales no genera enlentecimiento. Por otro lado, para la condición de velocidad alta (6 Hz) todos los retrasos temporales añadidos mostraron una disminución de velocidad.

Es posible notar que para las condiciones de tasa silábica baja y media (3 y 4 Hz) el retraso temporal a partir del cual ya no se presenta una disminución en la tasa silábica, se alinea con el momento en que los sujetos comienzan a cometer una cantidad de errores significativamente distinta de cero. Esto podría sugerir que los mecanismos de reducción en la tasa silábica y los de producción de errores no ocurren de forma simultánea, sino que aparecen en etapas separadas. Podría ser que el enlentecimiento es una estrategia para disminuir el efecto del DAF y evitar las disfluencias. Una vez

que no se puede compensar porque el retraso es comparable con la duración silábica comienzan las disfluencias y desaparece en enlentecimiento.

Sin embargo, esto no explica el comportamiento de los datos que se observa en la condición de velocidad alta (6 Hz) en donde todos los retrasos temporales causan una disminución en la velocidad. Esto podría deberse a un efecto de fatiga por ser 6Hz una tasa silábica alta para el habla normal. Si bien el experimento control no arroja un enlentecimiento significativo al pedirle a los sujetos que realicen la tarea manteniendo el retraso añadido al mínimo el valor medio se encuentra por arriba de 1, aun cuando realizan solo 16 ensayos, mientras que el experimento principal posee 72. Más experimentos serán necesarios para entender la relación entre las disfluencias y la disminución en la tasa silábica ocasionados por el DAF.

### **Conclusiones.**

Este estudio proporciona evidencia que respalda que la duración de la sílaba emitida controla el periodo temporal en el cual se efectúa la comparación de la expectativa sonora y el sonido real del habla en el Ciclo de Retroalimentación Auditivo-Motora del Habla. Además, los resultados sugieren que la producción de errores en el habla y disminución en la tasa silábica ocasionados por el DAF podrían ser consecuencia de distintos mecanismos.

Considerar estos hallazgos en futuros trabajos podría resultar esencial para elaborar modelos más efectivos del Ciclo de Retroalimentación Auditiva-Motora del Habla. Por otro lado, un conocimiento más acabado de las dinámicas temporales del CRAHM es relevante para mejorar terapias basadas en DAF dirigidas al tratamiento de disfluencias en el habla producidas por trastornos como la tartamudez o derivadas de sintomatología de patologías como el síndrome de Parkinson y algunos tipos de afasias.

## Referencias.

**Behroozmand, R., & Larson, C. R.** (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, *12*(1).

<https://doi.org/10.1186/1471-2202-12-54>

**Békésy, G. V.** (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America*, *21*, 217–232.

<https://doi.org/10.1121/1.1906501>

**Black, J. W.** (1951). The Effect of Delayed Side-Tone Upon Vocal Rate and Intensity. *Journal of Speech and Hearing Disorders*, *16*(1), 56–60. <https://doi.org/10.1044/jshd.1601.56>

**Brendel, B., Lowit, A., & Howell, P.** (2004). The effects of delayed and frequency shifted feedback on speakers with Parkinson's Disease. *Med Speech Lang Pathol*, *1*(December), 131–138.

**Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F.** (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences*, *110*(7), 2653–2658. <https://doi.org/10.1073/pnas.1216827110>

**Chapin, C., Blumstein, S., Meissner, B., & Boller, F.** (1981). Speech production mechanisms in aphasia: A delayed auditory feedback study. *Brain and Language*, *14*(1), 106–113.

[https://doi.org/10.1016/0093-934x\(81\)90068-7](https://doi.org/10.1016/0093-934x(81)90068-7)

**Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G.** (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, *15*(2), 399–402. <https://doi.org/10.1046/j.0953-816x.2001.01874.x>

**Fairbanks, G.** (1955). Selective Vocal Effects of Delayed Auditory Feedback. *Journal of Speech and Hearing Disorders*, 20(4), 333–346. <https://doi.org/10.1044/jshd.2004.333>

**Flinker, A., Chang, E. F., Kirsch, H. E., Barbaro, N. M., Crone, N. E., & Knight, R. T.** (2010). Single-Trial Speech Suppression of Auditory Cortex Activity in Humans. *Journal of Neuroscience*, 30(49), 16643–16650. <https://doi.org/10.1523/jneurosci.1809-10.2010>

**Foundas, A. L., Mock, J. R., Corey, D. M., Golob, E. J., & Conture, E. G.** (2013). The SpeechEasy device in stuttering and nonstuttering adults: Fluency effects while speaking and reading. *Brain and Language*, 126(2), 141–150.

**Guenther, F. H.** (2016). *Neural Control of Speech (The MIT Press)* (Ed. ilustrada). The MIT Press.

**Hernández-Jaramillo J, J., & Gil-Lozada Y, Y.** (2014). Efectos de la retroalimentación auditiva retardada en los patrones de tartamudez. *Ciencias de La Salud*, 12(2), 243–251. <https://doi.org/10.12804/revsalud12.2.2014.09>

**Hickok, G., Houde, J., & Rong, F.** (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, 69(3), 407–422. <https://doi.org/10.1016/j.neuron.2011.01.019>

**Howell, P., Powell, D. J., & Khan, I.** (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 9(5), 772–784. <https://doi.org/10.1037/0096-1523.9.5.772>

**Jafari, M., & Ansari-Pour, N.** (2019). Why, when and how to adjust your P values? *PubMed*, 20(4), 604–607. <https://doi.org/10.22074/cellj.2019.5992>

**Kaspar, K., & Rübeling, H.** (2011). Rhythmic Versus Phonemic Interference in Delayed Auditory Feedback. *Journal of Speech, Language, and Hearing Research*, 54(3), 932–943.

[https://doi.org/10.1044/1092-4388\(2010/10-0109\)](https://doi.org/10.1044/1092-4388(2010/10-0109))

**Kowal, S., Wiese, R., & O’Connell, D. C.** (1983). The use of time in storytelling. *Language and Speech*, 26(4), 377–392. <https://doi.org/10.1177/002383098302600405>

**Lee, B. S.** (1950). Effects of Delayed Speech Feedback. *The Journal of the Acoustical Society of America*, 22(6), 824–826. <https://doi.org/10.1121/1.1906696>

**Malloy, J. R., Nistal, D. A., Heyne, M., Tardif, M., & Bohland, J. W.** (2022). Delayed auditory feedback elicits specific patterns of serial order errors in a paced syllable sequence production task. *Journal of Speech Language and Hearing Research*, 65(5), 1800–1821. [https://doi.org/10.1044/2022\\_jslhr-21-00427](https://doi.org/10.1044/2022_jslhr-21-00427)

**Marshall, O.** (2021). The Maniac-Making Machine: A Media History of Delayed Auditory Feedback. *Technology and Culture*, 62(3), 839-860. <https://doi:10.1353/tech.2021.0109>.

**Pulvermüller, F., & Fadiga, L.** (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Perspectives*, 11(May), 351–360. <https://doi.org/10.1038/nrn2811>

**Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K.** (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, 111(5), 2237. <https://doi.org/10.1121/1.1466868>

**Tourville, J. A., & Guenther, F. H.** (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981.

<https://doi.org/10.1080/01690960903498424>

**Toyomura, A., Miyashiro, D., Kuriki, S., & Sowman, P. F.** (2020). Speech-Induced Suppression for Delayed Auditory Feedback in Adults Who Do and Do Not Stutter. *Frontiers in Human Neuroscience*, 14. <https://doi.org/10.3389/fnhum.2020.00150>

**Waldstein, R. S.** (1990). Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *The Journal of the Acoustical Society of America*, 88(5), 2099–2114. <https://doi.org/10.1121/1.400107>

**Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M.** (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701–702. <https://doi.org/10.1038/nn1263>

**Zanini, S., Clarici, A., Fabbro, F., & Bava, A.** (1999). Speaking Speed Effects on Delayed Auditory Feedback Disruption of Speech Fluency. *Perceptual and Motor Skills*, 89(3\_suppl), 1095–1109. <https://doi.org/10.2466/pms.1999.89.3f.1095>

**Zhao, B., Dang, J., & Zhang, G.** (2017). A Neuro-Experimental Evidence for the Motor Theory of Speech Perception. *Interspeech 2017*, 2441–2445. <https://doi.org/10.21437/interspeech.2017-1741>