

UNIVERSIDAD DE COSTA RICA

SISTEMA DE ESTUDIOS DE POSGRADO

**MODULACIÓN DE LA MEMORIA DE TRABAJO SEGÚN LA OPACIDAD DEL
LENGUAJE DURANTE LA LECTURA EN VOZ ALTA DE PALABRAS: UN
ACERCAMIENTO EXPERIMENTAL CON RASTREO OCULOMOTOR**

Tesis sometida a la consideración de la Comisión del Programa de Estudios de Posgrado en
Ciencias Cognoscitivas para optar al grado y título de Maestría Académica en Ciencias
Cognoscitivas

ALEJANDRO CAMBRONERO DELGADILLO

Ciudad Universitaria Rodrigo Facio, Costa Rica

2020

Dedicatoria

À mon ami Thomas

Gracias por ser mi mentor durante el proceso de desarrollo de tesis en el apasionante campo de la psicolingüística. Ya sea en las largas y extenuantes sesiones en el laboratorio o disfrutando de un par de cervezas, recordaré con gran aprecio los momentos que compartimos juntos. Cuando comencé este proyecto tan solo era un asistente con grandes deseos de aprender a hacer ciencia en el ámbito de los estudios experimentales con bilingües.

A lo largo del proceso encontré escollos que me hicieron dudar y perder la motivación, no obstante, los consejos oportunos y las palabras de apoyo que recibí siempre me hicieron levantarme y seguir adelante. Ahora, al término de una etapa crucial en mi vida, miro hacia atrás y entiendo la importancia y la sabiduría que encerraban cada una de las lecciones aprendidas, las cuales avivaron en mí la curiosidad científica y me ayudaron a nacer como investigador.

Agradecimientos

Agradezco, en primer lugar, a la Vicerrectoría de Investigación de la Universidad de Costa Rica por el apoyo brindado mediante el Fondo de Apoyo a Tesis de Posgrado 2018.

A mis lectores, Ximena del Río y Adrián Vergara, quienes con su mirada crítica aportaron valiosas contribuciones para el desarrollo de la investigación.

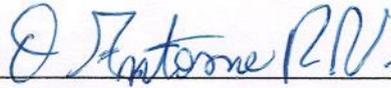
A Mercedes Villalobos por su apoyo durante todo el proceso de desarrollo de mi tesis.

A Valentina de Rooy por su acompañamiento durante la maestría.

A Antonio Leoni de León por su ayuda y disposición en aras del buen desarrollo del proyecto del cual se desprende esta tesis.

A Miao Chen, Catherine Corrales, Verónica Marín, Karen Monge y Allan Vargas por su asistencia en las diversas tareas logísticas para la consecución del proyecto.

“Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Ciencias Cognoscitivas de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Ciencias Cognoscitivas”



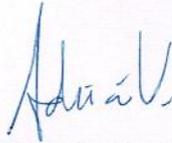
Dr. Odir Rodríguez Villagra
Representante del Decano
Sistema de Estudios de Posgrado



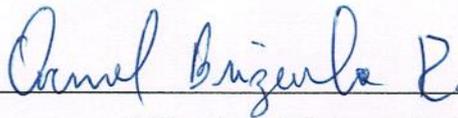
Dr. Thomas Castelain
Director de Tesis



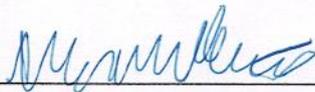
M.A. Ximena del Río Urrutia
Asesora



Dr. Adrián Vergara
Asesor



MSc. Armel Brizuela Rodríguez
Representante del Director
Programa de Posgrado en Ciencias Cognoscitivas



Alejandro Cambronero Delgadillo
Candidato

Tabla de contenido

Dedicatoria.....	ii
Agradecimientos.....	iii
Hoja de aprobación.....	iv
Tabla de contenido.....	v
Resumen.....	vi
Lista de tablas.....	viii
Lista de figuras.....	viii
Lista de abreviaturas.....	ix
Justificación.....	2
Estado de la Cuestión.....	3
Objetivos.....	11
Metodología.....	12
Resultados.....	19
Discusión.....	27
Conclusiones.....	29
Referencias.....	30
<i>Apéndice 1: Paradigmas psicolingüísticos para el estudio del bilingüismo: propuesta para el control y la selección asistida de estímulos.....</i>	<i>37</i>

Resumen

La hipótesis de profundidad ortográfica (Katz y Feldman, 1983; Katz y Frost, 1992) plantea que las lenguas pueden clasificarse en un continuo de transparentes a opacas según el grado de correspondencia más o menos directo que existe entre los grafemas y los fonemas. Diversos estudios conductuales y de neuroimagen (Buetler et al., 2014) ofrecen evidencia para sostener que los lectores adaptan su procesamiento del lenguaje de acuerdo con la profundidad ortográfica de cada lengua. A pesar de que la hipótesis de profundidad ortográfica posee gran relevancia para la comprensión de los procesos de lectura, los estudios sobre el tema son escasos (Kochva y Breznitz, 2012). El procesamiento de información en tiempo real ejerce una carga cognitiva en la memoria de trabajo. Durante el proceso de lectura en voz alta, el lapso entre la primera fijación de la palabra y el inicio de su pronunciación (lapso ojo-voz) ha sido asociado a la actualización del *buffer* de la memoria de trabajo (Laubrock y Kliegl, 2015). Con la intención de evaluar cómo el contexto lingüístico modula el procesamiento cognitivo, se llevó a cabo un experimento de lectura con bilingües cuyas lenguas difieren en profundidad ortográfica: español (transparente) e inglés (opaca). Se partió de la hipótesis de un procesamiento de palabras más rápido (lapso ojo-voz menor) en la lengua opaca en contraste con la lengua transparente. Además, se esperaba un procesamiento de pseudopalabras más rápido (lapso ojo-voz menor) en la lengua transparente con contraste con lengua opaca. El análisis de la conducta oculomotora y del lapso ojo-voz de los participantes sugiere que la lectura de los bilingües es sensible tanto a modulaciones del contexto lingüístico como a factores intraindividuales. Estos hallazgos poseen implicaciones sustanciales para el desarrollo de modelos cognitivos de lectura y para el mejoramiento del diseño de metodologías para la alfabetización

Palabras clave: Bilingüismo, profundidad ortográfica, memoria de trabajo, lapso ojo-voz.

Abstract

The orthographic depth hypothesis (Katz & Feldman, 1983; Katz & Frost, 1992) states that languages can be classified on a continuum from transparent to opaque according to the degree of correspondence between graphemes and phonemes. Various behavioral and neuroimaging studies (Buetler et al., 2014) have provided evidence to support the claim that readers adapt their language processing strategies according to the orthographic depth of each language. Although the orthographic depth hypothesis has great relevance for the understanding of reading processes, studies on the subject are scarce (Kochva & Breznitz, 2012). Real-time information processing exerts a cognitive load on working memory. During reading aloud processes, the time between the first fixation of a word and the beginning of its pronunciation (eye-voice span) has been associated with the working memory buffer update (Laubrock & Kliegl, 2015). In order to evaluate how the linguistic context modulates cognitive processing, a reading experiment was carried out with bilinguals whose languages differ in orthographic depth: Spanish (transparent) and English (opaque). We hypothesized faster word processing (shorter eye-voice span) in the opaque language in contrast to the transparent language. In addition, faster pseudo-word processing (shorter eye-voice span) was expected in the transparent language in contrast to the opaque language. The analysis of the participants' oculomotor behavior and eye-voice span suggests that bilinguals' reading strategies are sensitive to linguistic context's modulations and to intra-individual factors. These findings have substantial implications for the development of cognitive models of reading and for improving the design of literacy methodologies.

Keywords: Bilingualism, orthographic depth, working memory, eye-voice span.

Lista de tablas

Tabla 1. Nivel de proficiencia en la L1 y la L2 de los participantes.....	18
Tabla 2. Promedios (y desviaciones estándares) de la Localización de la Primera Fijación, el Número de Fijaciones y el Lapso Ojo-Voz en función del L1 de los participantes, el Idioma y el Tipo de palabras.....	20
Tabla 3. Estimaciones de los parámetros del modelo Localización de la Primera Fijación en función de la L1 de los participantes, el Idioma y el Tipo de palabra.....	24
Tabla 4. Estimaciones de los parámetros del modelo Número de Fijaciones en función de la L1 de los participantes, el Idioma y el Tipo de palabra.....	25
Tabla 5. Estimaciones de los parámetros del modelo Lapso Ojo-Voz en función de la L1 de los participantes, el Idioma y el Tipo de palabra.....	26

Lista de figuras

Figura 1. Representación esquemática del modelo de doble ruta.....	4
Figura 2. Representación gráfica del lapso ojo-voz.....	9
Figura 3. Representación esquemática de la manera en la cual fueron presentados los estímulos.....	17
Figura 4. Número de fijaciones para Palabras en función del Idioma y la L1 de los participantes.....	21
Figura 5. Número de fijaciones para Pseudo-palabras en función del Idioma y la L1 de los participantes.....	22
Figura 6. Lapso Ojo-Voz para Palabras en función del Idioma y la L1 de los participantes.....	23

Figura 7. Lapsos Ojo-Voz para Pseudo-palabras en función del Idioma y la L1 de los participantes.....23

Lista de abreviaturas

CGF Conversión grafo-fonológica

HPO Hipótesis de la profundidad ortográfica

LOV Lapsos ojo-voz

LPF Localización de la primera fijación

L1 Primera lengua

L2 Segunda lengua

MDR Modelo de doble ruta

MLM Modelos lineales de efectos mixtos

NF Número de fijaciones



UNIVERSIDAD DE
COSTA RICA

SEP Sistema de
Estudios de Posgrado

Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.

Yo, Alejandro Javier Cambrono Delgadillo, con cédula de identidad 1-1387-0359, en mi condición de autor del TFG titulado Modulación de la Memoria de Trabajo Según la Opacidad del Lenguaje Durante la Lectura en Voz Alta de Palabras: un Acercamiento Experimental con Rastreo Oculomotor

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI NO *

*En caso de la negativa favor indicar el tiempo de restricción: _____ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

INFORMACIÓN DEL ESTUDIANTE:

Nombre Completo: Alejandro Javier Cambrono Delgadillo

Número de Carné: A71313 Número de cédula: 1-1387-0359

Correo Electrónico: alejandro.cambrono@ucr.ac.cr

Fecha: 7/7/2020 Número de teléfono: 22793170

Nombre del Director (a) de Tesis o Tutor (a): Thomas Castelain

FIRMA ESTUDIANTE

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

Justificación

Debido a cambios en las dinámicas del mundo globalizado, la adquisición de una segunda lengua (L2) se ha convertido en una necesidad para mantenerse competitivo en el mercado laboral costarricense (Barquero, 2019). Lo anterior ha impulsado la creación y promoción de políticas estatales para fomentar el aprendizaje de una L2 (Barquero, 2020) e incrementar el porcentaje de población bilingüe en el país (Alvarado, 2020). Junto con su utilidad como una herramienta para la obtención de empleo, el aprendizaje de una L2 reconfigura el sistema cognitivo de las personas en términos de estructura, posibilidades de enunciación y contexto lingüístico con el objetivo de adquirir las competencias necesarias para procesar información y comunicarse en otra lengua (Costa y Sebastián-Gallés, 2014).

Siguiendo la hipótesis de profundidad ortográfica (Katz y Feldman, 1983; Katz y Frost, 1992), las lenguas pueden clasificarse en un continuo de transparentes a opacas según el grado de correspondencia más o menos directo que existe entre los grafemas y los fonemas. Actualmente, un creciente número de estudios conductuales y de neuroimagen (Buetler et al., 2014) ha aportado evidencia para sostener que los procesos de lectura en voz alta de encuentran modulados por la profundidad ortográfica del lenguaje leído. Aun cuando la hipótesis de la profundidad de la profundidad ortográfica constituye un marco sólido para comprensión de los procesos de lectura, los estudios sobre el tópico son escasos (Bar-Kochva y Bresnitz, 2012).

Con la intención de explorar de qué manera la opacidad del lenguaje modula el procesamiento cognitivo durante los procesos de lectura en voz alta, llevamos a cabo un estudio con bilingües cuyas lenguas difieren en profundidad ortográfica (español como lengua transparente e inglés como lengua opaca). En el estudio evaluamos el lapso entre la primera fijación de palabras y pseudo-palabras y el inicio de su pronunciación, lapso que ha sido asociado con la actualización del *buffer* de la memoria de trabajo (Laubrock y Kliegl, 2015). Los resultados que obtuvimos arrojan luz hacia la comprensión de los procesos de lectura en bilingües cuyas lenguas difieren en profundidad ortográfica y ofrecen nuevas perspectivas a la discusión de la “guerra de lecturas” (Castles, Rastle, y Nation, 2018). Estas perspectivas poseen implicaciones sustanciales hacia el desarrollo de métodos de

alfabetización más integrales y personalizados, los cuales comprendan factores tanto del contexto lingüístico (lengua enseñada) como individuales (lengua del aprendiente). Por último, los hallazgos encontrados representan importantes insumos para el avance de los modelos cognitivos de lectura bilingüe.

Estado de la cuestión

Hipótesis de profundidad ortográfica

El proceso de lectura en voz alta implica la adquisición de información lingüística mediante la decodificación de grafemas en fonemas. Dependiendo de la lengua, la consistencia de las correspondencias de grafemas a fonemas varía en gran medida (Ellis et al., 2004). Esta variabilidad constituye la base de la hipótesis de la profundidad ortográfica (HPO; Katz y Feldman, 1983), según la cual las lenguas pueden clasificarse en un continuo de transparentes a opacas de acuerdo con el grado más o menos directo de estas correspondencias.

Al pronunciar las palabras escritas, los hablantes recuperan las correspondencias grafo-fonológicas mediante la utilización de reglas almacenadas en la memoria (Frost, 1998). Dependiendo de la opacidad de la lengua leída, su procesamiento variará: por un lado, la lectura de lenguas opacas —como el inglés— posee reglas de conversión ambiguas y altamente variables, en las cuales un grafema puede representar distintos fonemas. Por otro lado, la lectura de lenguas transparentes —como el español— sigue reglas de conversión grafo-fonológicas (CGF) simples, las cuales son usualmente unívocas (un grafema corresponde a un fonema) o presentan poca variación. Esta distinción ha llevado a postular la existencia de rutas diferenciadas de lectura en voz alta según la naturaleza transparente u opaca de la lengua (Katz y Frost, 1992).

De acuerdo con el modelo de lectura doble ruta (MDR; Coltheart, 2005) la lectura en voz alta es un proceso que es posible gracias al uso de dos vías: por un lado, la vía directa (léxica), la cual se basa en el acceso léxico directo y el uso del lexicón mental para recuperar la información fonológica de las palabras. Se ha postulado que el uso de esta vía es más prevalente en la lectura de lenguas opacas y en lectores proficientes, ya que implica la existencia de un vínculo semántico fuerte con la forma de las palabras (Nation y Castles, 2017).

Por otro lado, la lectura se da mediante la vía indirecta (no-léxica), la cual favorece la fonología y las correspondencias símbolo-sonido uno a uno para la traducción del código escrito a la oralidad. Esta vía se considera la vía por defecto para la lectura en lenguas transparentes y en lectores iniciantes, dada su dependencia en el procesamiento de unidades más pequeñas para la recuperación de la pronunciación de las palabras (Nation y Castles, 2017).

Para cada una de las rutas, el modelo de doble ruta especifica estrategias de lectura diferenciadas para ensamblar la fonología de las palabras a partir de sus grafemas: mientras que en la ruta directa el procesamiento de palabras ocurre de manera global y en paralelo, en la ruta indirecta se da de manera local y en serie (de izquierda a derecha, en el sentido de la lectura). Tal y como se observa en la Figura 1, la ruta mediante la cual se llevará a cabo el proceso de lectura en voz alta es activada luego de la identificación de las letras que componen la palabra que está siendo leída.

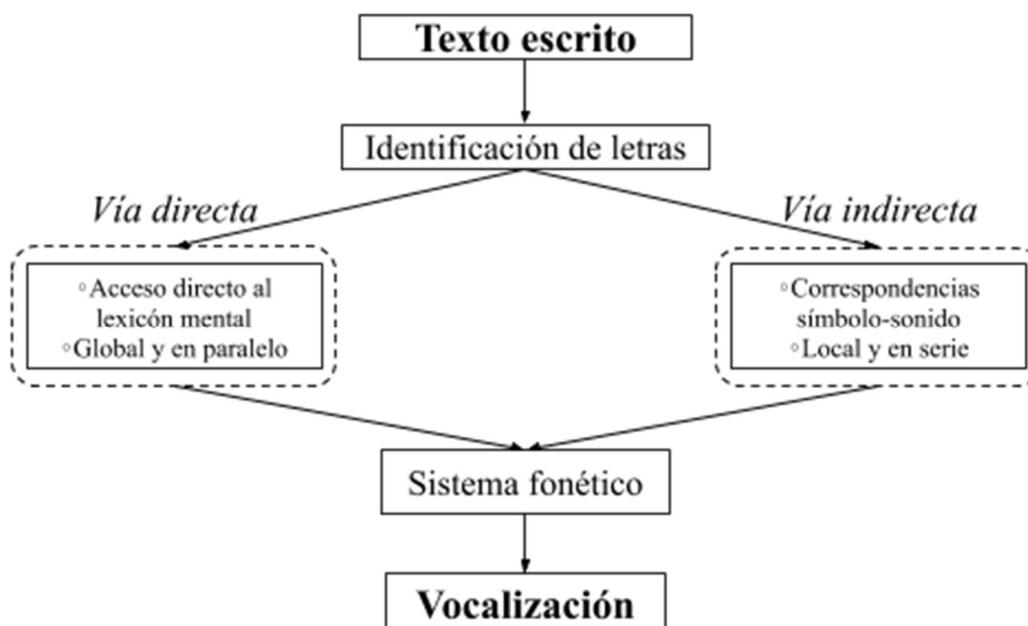


Figura 1. Representación esquemática del modelo de doble ruta. Adaptado de Coltheart, Rastle, Perry, Langdon, y Ziegler (2001).

Existe un conjunto de estudios comportamentales y de neuroimagen que han puesto a prueba estos supuestos teóricos (Fukuda y Fukuda, 2009; Joshi, Tao, Aaron, y Quiroz, 2012; Seymour et al., 2003; Simon, Bernard, Lalonde, y Rebaï, 2006; Ziegler, Perry,

Jacobs, y Braun, 2001); ; no obstante, las investigaciones sobre el tema continúan siendo escasas debido a la complejidad en el control de variables de confusión (Bar-Kochva y Breznitz, 2012).

Recientemente, esta limitación ha sido abordada mediante estudios con bilingües (Buetler et al., 2014; de León Rodríguez et al., 2016a; de León Rodríguez et al., 2016; Rau, Moll, Snowling, y Landerl, 2015) quienes parecen ser un buen modelo para evaluar el impacto de la profundidad ortográfica en el procesamiento del lenguaje. En el siguiente apartado se discutirán las principales implicaciones derivadas de estos estudios, los cuales exploran relación entre opacidad de lenguaje y procesos de lectura en individuos cuyas lenguas difieren en términos de profundidad ortográfica.

Opacidad del lenguaje y bilingüismo

El bilingüismo es la habilidad para utilizar una segunda lengua (L2) de forma cotidiana (Halsband, 2006). Si bien esta definición cubre todo el espectro de bilingües, existen diversos factores que inciden en la manera en cómo el bilingüismo se caracteriza y se expresa (Grosjean, 1992). Para fines de la presente discusión, es necesario precisar dos tipologías de bilingües: en primer lugar, según el orden de adquisición de las lenguas, se clasifican como bilingües simultáneos a quienes adquieren una L2 de manera paralela con su L1, mientras que se conoce como bilingües sucesivos a las personas que adquieren una L2 de manera asincrónica con respecto a su L1 (Mitchell, Myles y Marsden, 2019). En segundo lugar, según la diferencia de proficiencia entre las lenguas, se denomina bilingües dominantes a quienes poseen mayor dominio en una de sus lenguas con respecto a la otra (normalmente la L1), y bilingües balanceados a quienes poseen un nivel equivalente entre ambas lenguas (Peal y Lambert, 1962).

Recientemente, ha surgido una línea de investigación interdisciplinaria para el estudio de la hipótesis de la profundidad ortográfica en los procesos de lectura de bilingües con distintos perfiles (Buetler et al., 2014; de León et al, 2016a, 2016b; Rau et al., 2015). Los hallazgos de estos estudios han aportado evidencia proveniente tanto del estudio de la conducta oculomotora como de la actividad cerebral para respaldar la hipótesis de patrones de lectura diferenciados según la profundidad ortográfica. Además, mostraron que en

conjunción con factores *bottom-up* (el contexto de lenguaje), los factores *top-down* (proficiencia del bilingüe) tienen un rol importante en el procesamiento del lenguaje escrito.

En el estudio de Butler et al. (2014), se pidió a bilingües simultáneos que leyeran palabras y pseudo-palabras en voz alta en su idioma transparente (alemán) o en su idioma opaco (francés) mientras su actividad cerebral se registraba con la técnica de electroencefalografía de alta densidad. Los resultados obtenidos ofrecieron evidencia para sostener que durante los procesos de lectura en voz alta ocurre una activación de redes neuronales diferenciadas, la cual es modulada por la profundidad ortográfica del idioma leído.

Los estudios en los cuales se utilizó la técnica del rastreo oculomotor evaluaron la hipótesis de la profundidad ortográfica en bilingües con distintos perfiles. Rau et al. (2015) trabajaron con adultos y niños ingleses (lengua opaca) y alemanes (lengua transparente) en un paradigma de lectura en voz alta de oraciones en ambas lenguas. Dichas oraciones contenían un estímulo clave, el cual variaba en estatus léxico (palabras de baja y alta frecuencia léxica, y pseudo-palabras) y largo de las palabras. Los resultados del estudio aportaron evidencia para decir que hubo una diferencia entre los niños ingleses y los alemanes en términos del tiempo que duraron procesando las pseudo-palabras, ya que los últimos tomaron menos tiempo que los primeros. Aun cuando estos grupos no mostraron diferencias temporales en la lectura de palabras, los niños alemanes utilizaron una estrategia de decodificación con unidades más pequeñas (mayor cantidad de fijaciones) en contraste con los niños ingleses.

Adicionalmente, los adultos tomaron menos tiempo en procesar pseudo-palabras cortas en contraste con pseudo-palabras largas, mas este efecto fue de mayor magnitud en los adultos alemanes que en los ingleses. En síntesis, los autores mencionan que los grupos de adultos utilizaron el mismo patrón de lectura para procesar las pseudo-palabras (unidades más pequeñas), pero los lectores alemanes fueron más eficaces que los ingleses debido a su experiencia en el uso de este patrón, en contraste con los ingleses quienes se

crea que poseen un patrón de lectura basado en el procesamiento de unidades más grandes (Ziegler y Goswami, 2005).

En el primer estudio llevado a cabo por de León et al. (2016a), se le pidió a un grupo de bilingües altamente proficientes que leyeran palabras y pseudo-palabras en su lengua transparente (alemán) y en su lengua opaca (francés). Los resultados obtenidos apoyan la preferencia de conversiones grafo-fonológicas directas tanto en la lectura de palabras en lenguas transparentes como en la lectura de pseudo-palabras, lo cual se refleja en una localización de la primera fijación más cercana al inicio en las palabras del idioma transparente y en las pseudo-palabras en contraste con la lectura de palabras en el idioma opaco. Lo anterior sugiere que, dependiendo de la opacidad del lenguaje, los mismos sujetos utilizan diferentes patrones de lectura (factores *bottom-up*).

En el segundo estudio realizado por de León et al. (2016b), el mismo procedimiento fue llevado a cabo, pero con bilingües sucesivos (alemán-francés y francés-alemán) con una diferencia de proficiencia entre su primera lengua (nativa) y su segunda lengua (baja proficiencia). Este grupo fue escogido con el fin de ahondar en los resultados del primer estudio (de León et al., 2016a), en concreto, con el hallazgo de las diferencias en la localización de la primera fijación en las pseudo-palabras, el cual sugería que el vínculo semántico existente con la forma de las palabras —ausente en las pseudo-palabras— podría tener injerencia en la selección de la ruta de lectura.

De manera conjunta con la localización de la primera fijación, en este segundo estudio se registró el número de fijaciones por estímulo con el fin de evaluar la hipótesis de una estrategia de procesamiento serial de lectura en lenguas transparentes y en pseudo-palabras (Coltheart et al., 2001; Taylor, Rastle, y Davis, 2013). También, un aumento en el número de fijaciones durante la lectura era esperado en la L1 transparente en contraste con la L1 opaca de los bilingües. Para el caso de los bilingües con una L1 opaca y leyendo en su L1, se esperaba un aumento de fijaciones en pseudo-palabras en contraste con palabras. Las predicciones con respecto a la localización de la localización de la primera fijación eran las mismas que en el primer estudio (de León et al., 2016a) pero para la L1 de ambos grupos. Para explorar la hipótesis del vínculo semántico, una localización de la primera

fijación más cercana al inicio de los estímulos era esperada en la L2 en contraste con la L1, pero únicamente en el grupo de bilingües con L1 opaca.

Los resultados obtenidos se alinearon con las predicciones con una excepción: la localización de la primera fijación no difirió en el caso de bilingües que leían en su L1 opaca. El número de fijaciones sí presentó variaciones, lo cual soporta la hipótesis de un procesamiento serial de las pseudo-palabras y también indica que los bilingües sucesivos poseen una estrategia de lectura por defecto en su L1, la cual cambia luego de haber realizado la primera fijación. Los resultados de este segundo estudio apoyan los resultados del primer estudio (de León et al., 2016a), ya que los bilingües sucesivos que leen en una L1 opaca procesan las palabras de una manera global, mientras que la lectura en una L2 exhibe vínculos semánticos más débiles, lo cual da paso a una estrategia de lectura local y en serie.

En resumen, estos estudios han aportado evidencia para sostener que existe una activación diferenciada de rutas de lectura en voz alta, diferenciación que se halla modulada por la profundidad ortográfica de la lengua. Además, aun cuando los bilingües son capaces de cambiar su estrategia de decodificación según el contexto lingüístico, la opacidad de su lengua dominante dicta su estrategia de lectura por defecto, mientras que el patrón serial-local es preferido en la lengua menos dominante, independientemente de su opacidad.

En conjunto, estos resultados recalcan la importancia de incluir factores tanto individuales como contextuales en los estudios de procesos de lectura. También, abren nuevas direcciones en el ámbito del estudio del bilingüismo; por ejemplo, dado que se han establecido rutas claramente diferenciadas para la conversión grafo-fonológica en lenguas que varían en profundidad ortográfica, y dado que los bilingües dominantes favorecen una estrategia de lectura correspondiente a su L1, es válido pensar que el procesamiento cognitivo de los bilingües se verá afectado por el contexto lingüístico. En el siguiente apartado discutiremos las posibles relaciones entre la profundidad ortográfica y la modulación de los procesos cognitivos durante la lectura en voz alta. Además, discutiremos cómo estos supuestos pueden ser evaluados mediante el estudio de la lectura en bilingües.

Procesos de lectura en voz alta y modulación de procesos cognitivos

Dada la hipótesis de que el uso de códigos fonológicos y léxicos varía entre lenguas transparentes y opacas, lo cual se refleja en el procesamiento de unidades cortas y largas respectivamente, se espera un procesamiento cognitivo de palabras más rápido en lenguas opacas en contraste con lenguas transparentes. Esta afirmación puede ser evaluada mediante la medición de la carga de memoria de trabajo durante el proceso de lectura en voz alta, en concreto, estimando los costes de *buffering* en el bucle fonológico durante los procesos de conversión grafo-fonológica.

Durante la lectura en voz alta, la mirada usualmente se adelanta a la voz (Buswell, 1921). Este lapso entre la primera fijación de la palabra y el inicio de su pronunciación se denomina lapso ojo-voz (Inhoff, Solomon, Radach, y Seymour, 2011). La representación gráfica de esta medida puede observarse en la Figura 2.

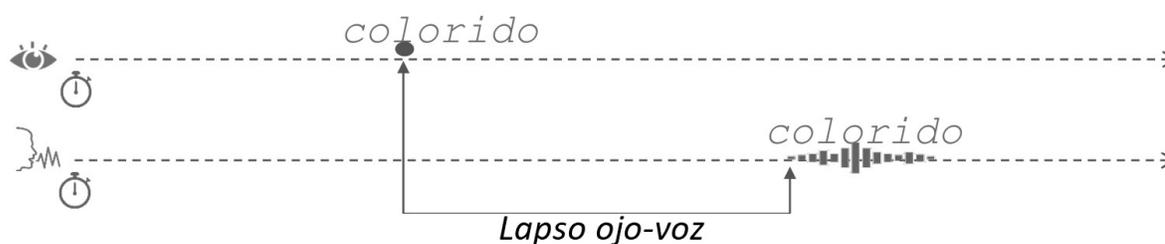


Figura 2. Representación gráfica del lapso ojo-voz. Durante la lectura en voz alta, la pronunciación de las palabras usualmente ocurre luego de su fijación con la mirada. La línea punteada simboliza la temporalidad. Adaptado de Costa, Falé y Luegi, 2016.

El lapso ojo-voz se ha propuesto como un índice que permitiría informar acerca de la naturaleza de los factores cognitivos involucrados en el procesamiento del lenguaje. En un estudio reciente (Laubrock y Kliegl, 2015) se registró la conducta oculomotora de participantes mientras leían oraciones en alemán en voz alta. Basándose en los resultados obtenidos, los autores proponen el lapso ojo-voz como una medida indirecta de la carga cognitiva asociada al *buffering* del código fonológico traducido de las palabras escritas, ya que observan que esta medida es mejor predictora de la duración de las fijaciones en contraste con otros predictores usados previamente, tales como la frecuencia léxica y el largo en caracteres (Kliegl, Grabner, Rolfs y Engbert, 2004).

Esta observación puede explicarse en vista de que la respuesta solicitada a los participantes (la pronunciación de la palabra escrita) requiere de una integración en serie de la información fonológica en el bucle fonológico de la memoria de trabajo. Aun cuando es esperable que la información fonológica sea integrada en serie, esto no puede ser asumido para el caso de la activación de las correspondientes unidades léxicas en la memoria.

Siguiendo los estudios llevados a cabo por de León Rodríguez et al. (2016a, 2016b), al incluir el lapso ojo-voz en estudios sobre bilingüismo y opacidad del lenguaje se podría profundizar en la comprensión de los procesos de lectura en bilingües. Dado que los bilingües sucesivos con una L1 opaca serían más sensibles a modulaciones léxicas, se esperaría que estos presenten un lapso ojo-voz más reducido durante la lectura de palabras con respecto a bilingües con una L1 transparente leyendo palabras en su L1. Adicionalmente, y en línea con los hallazgos de Rau et al. (2015), si el lapso ojo-voz fuese incluido como una medida durante la lectura de pseudo-palabras, los bilingües sucesivos con una L1 opaca y menor proficiencia en su L2 deberían presentar un lapso ojo-voz más largo en contraste con bilingües con una L1 transparente, dado que estos últimos poseen mayor experiencia en el procesamiento de la lectura local y en serie en comparación con los primeros, quienes exhiben patrones de lectura de naturaleza más global y paralela.

Con el objetivo de determinar de qué manera el contexto lingüístico modula el procesamiento cognitivo en bilingües, en el presente estudio se evaluó el lapso ojo-voz en dos grupos de bilingües dominantes (bilingües con un mayor dominio en una de sus lenguas con respecto a la otra) cuyas lenguas difieren en el grado de profundidad ortográfica (español como lengua transparente e inglés como lengua opaca). Así, palabras y pseudo-palabras en un contexto en español y en un contexto en inglés fueron presentadas a ambos grupos y se les pidió que las leyeran en voz alta y a un ritmo natural. Las medidas oculomotoras estándar en este tipo de paradigmas fueron registradas: localización de la primera fijación y número de fijaciones. Conjuntamente, fueron registradas las respuestas vocales de los participantes, en concreto, el inicio de la pronunciación de las palabras.

En vista de que se ha propuesto que el uso de códigos fonológicos y léxicos varía de lenguas transparentes a lenguas opacas (de León Rodríguez et al., 2016a; 2016b; Rau et al.,

2015), el lapso ojo-voz también debería ser modulado por la profundidad ortográfica. Dado que se considera que la lectura en voz alta en lenguas opacas procesa menos unidades en contraste con la lectura en voz alta en lenguas transparentes, el lapso ojo-voz debería ser más corto en la lectura de palabras en lenguas opacas que en lenguas transparentes. Así, se espera un procesamiento de palabras más rápido en la lengua opaca que en la lengua transparente. En cambio, dado el supuesto de una preferencia por el uso de la vía indirecta en la lectura de pseudo-palabras, estas deberían ser procesadas más rápido en las lenguas transparentes que en las lenguas opacas, lo cual será reflejado en un lapso ojo-voz más reducido en la lectura de pseudo-palabras en lenguas transparentes que en lenguas opacas.

Adicionalmente, dado a que se asume que el vínculo léxico con la forma de las palabras es más fuerte en la lengua dominante (Nation y Castles, 2017), se espera que las palabras sean procesadas de manera más eficiente en la L1 con respecto a la L2, lo cual será reflejado con un número de fijaciones reducido y un lapso ojo-voz (LOV) menor en la lectura de palabras en la L1 con respecto a la lectura de palabras en la L2. Puesto que las pseudo-palabras son carentes de lexicalidad (es decir, que no poseen un vínculo léxico asociado a su forma escrita), ambos grupos de bilingües deberían procesarlas de manera menos eficiente que las palabras en los dos contextos lingüísticos. Así, se espera que durante la lectura de pseudo-palabras en ambos contextos (español e inglés) se observe un número de fijaciones (NF) mayor, un lapso ojo-voz más largo y una localización de la primera fijación (LPF) más cercana al inicio en contraste con la lectura de palabras.

Objetivos

Objetivo General

Determinar de qué manera la opacidad del lenguaje modula el procesamiento cognitivo en sujetos bilingües.

Objetivos Específicos

1. Comparar el lapso ojo-voz entre dos lenguas que varían en su profundidad ortográfica.

2. Discutir cómo la opacidad del lenguaje modula la memoria de trabajo durante los procesos de lectura en voz alta.

Metodología

Participantes

Un total de 25 bilingües dominantes (17 mujeres, $M_{edad} = 26.12$ años, $SD_{edad} = 8.87$, [18-52 años]) participaron en el estudio.¹ Doce de ellos tenían español como su primera lengua (L1) e inglés como su segunda lengua (L2). Los 13 restantes presentaban el perfil inverso (inglés como L1 y español como L2). Todos los participantes tenían visión normal o corregida a normal y no conocían las hipótesis experimentales. Luego de su participación en el estudio, los participantes recibieron un pequeño obsequio de agradecimiento. El procedimiento experimental fue aprobado por el Comité Ético Científico de la Universidad de Costa Rica.

Evaluación del nivel de bilingüismo

Con el fin de evaluar el nivel de bilingüismo en ambas lenguas, se aplicaron las versiones extendidas de las tareas de decisión léxica LexTALE y LexTALE-Esp (Hoversten, Brothers, Swaab, y Traxler, 2017; Izura, Cuetos, y Brysbaert, 2014; Lemhöfer y Broersma, 2012). LexTALE y LexTALE-Esp son pruebas en inglés y en español respectivamente, las cuales han sido validadas y son consideradas fiables para estimar el conocimiento léxico del idioma de los sujetos. La puntuación de este conocimiento se representa en una escala de 0 a 100: el rango 0-59 equivale a proficiencia baja y baja intermedia; el rango 60-80 equivale a proficiencia intermedia superior; el rango 80-100 equivale a proficiencia avanzada baja y avanzada alta. Ambas pruebas fueron adaptadas

¹ Los datos analizados para efectos del presente Trabajo Final de Graduación provienen de la cohorte de bilingües que participaron en el estudio previo a la pandemia global a causa de la COVID-19, la cual impidió continuar con la recolección de datos. Es importante notar que estos participantes no constituyen la muestra definitiva que será utilizada para la publicación de los resultados en un artículo científico, sino que este contará con un mayor número de participantes.

para su aplicación mediante el software E-Prime 2.0 (Schneider, Eschman y Zuccolotto, 2002).

Materiales

La fase de selección y control de estímulos exigió una labor sumamente meticulosa con el fin de contar con conjuntos de palabras en dos lenguas diferentes y pseudo-palabras comunes que fueran equivalentes entre lenguas según una serie de factores que han sido ampliamente descritos como moduladores del procesamiento de la información escrita en bilingües. La lógica tras la escogencia de estos factores y el detalle completo de cómo fueron controlados pueden consultarse en el Anexo 1 - *Paradigmas psicolingüísticos para el estudio del bilingüismo: propuesta para el control y la selección asistida de estímulos*, el cual constituye un manuscrito (sometido a publicación) derivado del presente Trabajo Final de Graduación. En este manuscrito ofrecemos una guía metodológica y la implementación de una herramienta para el control y la selección asistida de estímulos en paradigmas psicolingüísticos para el estudio del bilingüismo. Esta propuesta fue diseñada a partir del robustecimiento de la rigurosidad de estudios de corte similar (de León Rodríguez et al., 2016a, 2016b). A continuación, reseñaremos la versión reducida del procedimiento para la construcción de los estímulos.

El conjunto de estímulos se compone por 160 palabras por lengua (40 sustantivos de cinco letras, 40 sustantivos de seis letras, 40 sustantivos de siete letras y 40 sustantivos de ocho letras) y 160 pseudo-palabras (40 pseudo-palabras de cinco letras, 40 pseudo-palabras de seis letras, 40 pseudo-palabras de siete letras y 40 pseudo-palabras de ocho letras). Todos los estímulos fueron presentados en mayúsculas, en fuente *Courier News 72 pt* y en negrita para asegurar contextos equivalentes en cada condición de lengua. Las palabras fueron equiparadas entre lenguas en términos de regularidad, frecuencia léxica, número de sílabas, frecuencia de bigramas en posición específica, valencia y excitación emocional y tamaño de vecindario ortográfico. Además, se excluyeron los homógrafos y los homófonos entre lenguas, y palabras en español con acentos ortográficos o con la letra “ñ”. Las pseudo-palabras eran ortográfica y fonológicamente válidas entre lenguas, y fueron

controladas por frecuencia de bigramas en posición específica y tamaño de vecindario ortográfico.

Las palabras fueron seleccionadas de bases en inglés (Warriner, Kuperman y Brysbaert, 2013) y en español (Stadthagen-Gonzalez, Imbault, Pérez, y Brysbaert, 2017) que incluían normas para valencia y excitación emocionales. Para calcular la frecuencia léxica, el tamaño de vecindario ortográfico y la frecuencia de bigramas en posición específica, las palabras fueron sometidas a *Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities* (Marian, Bartolotti, Chabal, Shook y White, 2012), la cual es una herramienta multilingüe en línea para la recuperación de información fonológica, ortográfica y de vecindario léxico. Para determinar la regularidad de las palabras en el idioma opaco (inglés), se utilizó el *Dual Route Cascaded Model* (Coltheart et al., 2001), el cual es un modelo computacional para la identificación y lectura en voz alta de palabras.

Para crear las pseudo-palabras, se tomaron las palabras seleccionadas y se ingresaron en Wuggy (Keuleers y Brysbaert, 2010), el cual es un generador de pseudo-palabras fonotácticamente válidas en múltiples idiomas. Dado que Wuggy utiliza palabras como base para generar pseudo-palabras, los conjuntos de palabras en inglés y en español fueron usados para crear las pseudo-palabras en inglés y en español respectivamente. El generador fue configurado para producir diez pseudo-palabras por cada palabra en un periodo de búsqueda de 30s, lo cual generó un total de 4800 pseudo-palabras candidatas por lengua. Luego, estas candidatas fueron sometidas a un procedimiento de selección en tres etapas.

En primer lugar, dos jueces hispanohablantes nativos calificaron las pseudo-palabras candidatas provenientes de las palabras en inglés de acuerdo con su similitud a palabras en español. Un procedimiento similar fue llevado a cabo por dos jueces angloparlantes nativos con las pseudo-palabras provenientes de las palabras en español. En segundo lugar, los bigramas en posición específica fueron calculados para todas las candidatas en ambos idiomas; si algún bigrama era inválido (es decir, que no ocurría en alguna de las dos lenguas), la pseudo-palabra se descartaba. Por último, el conjunto final de

pseudo-palabras estaba constituido por 80 pseudo-palabras provenientes de las palabras inglés y 80 pseudo-palabras provenientes de las palabras en español, distribuidas de manera igualitaria en las cuatro categorías de largo (cinco, seis, siete y ocho letras).

Dispositivos

El experimento fue diseñado y ejecutado mediante el software E-prime 2.0 (Schneider, Eschman y Zuccolotto, 2002). La conducta oculomotora fue grabada con un sistema de rastreo de oscuro-pupila basado en video (SMI iView X™ RED 500), el cual posee una frecuencia de muestreo de 500 Hz y una resolución espacial de 0.03°. El procedimiento de calibración utilizado fue de 9 puntos. Las respuestas vocales de los participantes fueron grabadas con un micrófono ACM-51B (Cyber Acoustics, Vancouver, WA) acoplado a una caja de respuesta Chronos® (Psychology Software Tools, Pittsburgh, PA). Esta configuración permitió sincronizar y marcar temporalmente las respuestas vocales con la respuesta oculomotora. Los estímulos se presentaron en una pantalla LCD de 22 pulgadas, con una resolución de 1680 x 1050 píxeles y una frecuencia de actualización de 60 Hz. Para el procesamiento de la conducta oculomotora se utilizó el programa SMI Experiment Suite™ (SensoMotoric Instruments GmbH, Teltow, Alemania).

Procedimiento

El procedimiento se llevó a cabo en una habitación silenciosa. De acuerdo con las especificaciones del sistema de rastreo oculomotor, los participantes se sentaron a una distancia de 65cm frente a la pantalla de presentación de los estímulos. Con el objetivo de minimizar el movimiento de los participantes, un reposacabezas fue utilizado y se les solicitó mantenerse quietos durante las fases de grabación y evitar parpadear durante la pronunciación de los estímulos. Los participantes llegaron al sitio de evaluación y se les brindó el consentimiento informado para que lo leyeran y lo firmaran. Las evaluaciones en ambos idiomas (español e inglés) fueron llevadas a cabo en la misma sesión, con un tiempo de descanso de 10 minutos entre cada una. El procedimiento dio inicio con las tareas de decisión léxica, las cuales fueron contrabalanceadas de acuerdo con la L1 del participante. Luego de completar ambas tareas, se prosiguió a la tarea principal. El orden de evaluación

de lengua de la tarea principal (español-inglés o inglés-español) se determinó de acuerdo con el orden de la tarea de decisión léxica, empezando con la última lengua evaluada.

La tarea principal estuvo compuesta por tres fases: la fase de activación lingüística, la fase de instrucciones y entrenamiento y la fase de prueba. Aparte de la lengua de presentación, la tarea principal era equivalente en ambos contextos lingüísticos (español o inglés). La fase de prueba en cada contexto fue dirigida por un experimentador entrenado, utilizando el mismo procedimiento y hablando ya sea en español o en inglés dependiendo de la lengua en evaluación. Cada procedimiento en cada idioma duró aproximadamente 40 minutos (80 minutos en total).

Durante la fase de activación lingüística, los participantes debían leer un texto de una dificultad intermedia-alta en voz alta. Los textos (The End of the Flu Virus?, 2020; Google y el derecho al olvido, 2020) y sus respectivas traducciones fueron seleccionados de *Dual Text*, una revista en línea de textos paralelos (español e inglés) dirigida a bilingües. La versión original o la traducción de cada texto fue presentada una vez en cada contexto de lengua de manera contrabalanceada. Las traducciones fueron adaptadas para controlar por frecuencia léxica, largo de las palabras, largo de los párrafos y tamaño de vecindario léxico con respecto a su original. Luego de leer el texto, se proseguía con la fase de instrucciones.

La segunda fase iniciaba con la presentación de las instrucciones y la etapa de entrenamiento. En este momento, los participantes fueron instruidos sobre cómo llevar a cabo la tarea, además, se les indicó que todos los estímulos se presentarían en mayúscula. A cada participante se le dijo que debía leer en voz alta palabras o palabras inventadas, y que estas se presentarían en bloques separados. Para el caso de las pseudo-palabras, se les pidió que las leyeran como si se tratase de palabras reales en español o inglés dependiendo del idioma evaluado. La existencia de pseudo-palabras comunes entre lenguas fue omitida con el fin de que los participantes estuviesen completamente inmersos en cada contexto de lengua. En caso de ser necesario, se brindó información adicional en el idioma evaluado. Previo a la tarea principal, los participantes debían leer cuatro palabras y cuatro pseudo-palabras de práctica como entrenamiento. Las palabras y las pseudo-palabras de

entrenamiento diferían de las utilizadas en la tarea principal y en las pruebas de decisión léxica.

La fase de prueba estaba compuesta por un total de ocho bloques con 40 estímulos cada uno. Las categorías de estímulos (palabras y pseudo-palabras) fueron presentados en bloques separados: cuatro bloques continuos de palabras y cuatro bloques continuos de pseudo-palabras. El orden (palabras/pseudo-palabras o pseudo-palabras/palabras) fue contrabalanceado entre participantes. Los estímulos dentro de cada bloque fueron presentados de manera aleatoria tal y como se muestra en la Figura 3.

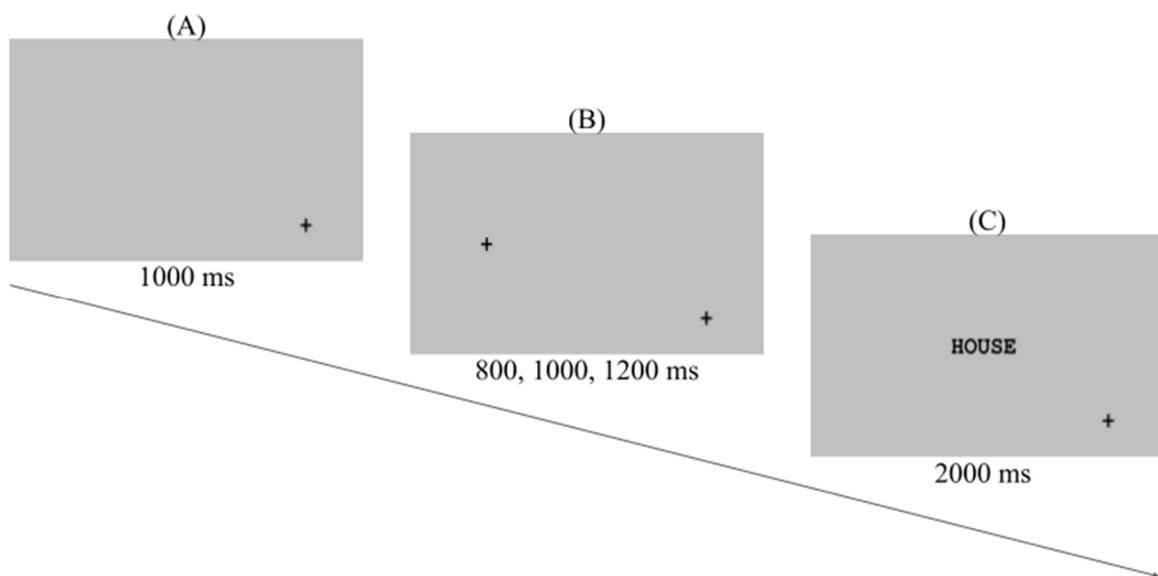


Figura 3. Representación esquemática de la manera en la cual fueron presentados los estímulos. (A): los participantes debían fijar con la mirada la cruz inferior-derecha por 1000 ms para activar la aparición de la cruz izquierda. (B): los participantes debían fijar la cruz de la izquierda, la cual era presentada aleatoriamente ya sea por 800 ms, 1000 ms o 1200 ms. Luego de fijarla por el tiempo requerido, se activaba la aparición del estímulo. (C): los participantes debían leer en voz alta el estímulo, el cual era presentado por 2000 ms. Luego, debían dirigir su mirada a la cruz de fijación en la parte inferior-derecha. Todas las cruces estaban programadas con un área de interés para detectar la mirada y activar cada aparición. La flecha simboliza la línea de tiempo.

Al inicio de cada bloque, los participantes leían una instrucción que indicaba la categoría del bloque que seguía a continuación (palabras o pseudo-palabras). Al final de cada bloque, había un periodo de descanso de 30s en el cual se le solicitaba al participante que cerrara los ojos previo a una nueva calibración. Durante el procedimiento, el experimentador interactuó con el participante únicamente en los periodos de descanso y durante las calibraciones. En total, se realizaron 18 calibraciones por participante: una antes

de la fase de activación lingüística y ocho durante la fase de prueba (9 en cada contexto de lengua). Solo fueron aceptadas aquellas calibraciones con una desviación inferior a 1°. Luego de evaluar ambos idiomas, la existencia de pseudo-palabras comunes fue revelada.

Resultados

Procesamiento de datos

Los participantes incluidos en este Trabajo de Final de Graduación fueron seleccionados teniendo en consideración la equiparación de los niveles de la L1 y la L2 entre grupos, tal y como se muestra en la Tabla 1.

Tabla 1

Nivel de proficiencia en la L1 y la L2 de los participantes.

L1 de Participante	Nivel L1	Nivel L2
Español	91.42 (3.38)	72.50 (11.59)
Inglés	91.15 (6.00)	69.69 (8.42)
<i>t-test</i>	0.90	0.52

Nota. L1 = primera lengua, L2 = segunda lengua. Se muestra la puntuación promedio para el nivel de la L1 y la L2 de los participantes. Entre paréntesis las desviaciones estándares correspondientes. *T-test* reporta entre el nivel de las L1 y el nivel de las L2.

Para la limpieza de los datos se llevó a cabo un procedimiento en múltiples etapas. En primer lugar, se eliminaron todos aquellos ensayos en los cuales el participante registró una conducta oculomotora diferente a la instruida, es decir, que su mirada no siguió el orden: 1) cruz de fijación izquierda, 2) estímulo y 3) cruz de fijación inferior-derecha. Lo anterior con el fin de asegurar que la sacada inicial hacia el estímulo siempre fuese lanzada desde el mismo punto.

En segundo lugar, se eliminaron los ensayos en los cuales no hubo respuesta vocal por parte del participante y aquellos en los cuales hubo titubeo antes o durante la pronunciación de los estímulos. Además, se excluyeron los ensayos para los cuales no fue posible marcar el inicio de la pronunciación debido a factores como ruido ambiental al momento de la grabación. Es importante mencionar que, aunque se implementó un programa computacional para el marcaje automático del inicio de la pronunciación, los

valores utilizados para esta medida fueron marcados manualmente por asistentes entrenados siguiendo un protocolo preestablecido, ya que el marcaje manual del inicio de la pronunciación es considerado el estándar de referencia en este tipo de paradigmas (Roux, Armstrong y Carreiras, 2017).

Análisis de datos

El procesamiento de los datos, los análisis y los gráficos se realizaron mediante R en Rstudio (R Development Core Team, 2020; RStudio Team, 2020) utilizando los siguientes paquetes: tidyverse (Wickham, 2017), lme4 (Bates, Maechler, Bolker, & Walker, 2015), lmerTest (Kuznetsova, Brockhoff, y Christensen, 2017), dplyr (Wickham, François, Henry, y Müller, 2020), ggplot2 (Wickham, 2016) y lsmeans (Lenth, 2016).

Adaptamos modelos lineales de efectos mixtos (MLM) para evaluar los efectos del Idioma (Inglés codificado como -1 y Español codificado como 1) del Tipo de palabra (Palabras codificado como -1 y Pseudo-palabras codificado como 1) como factor intrasujetos, la L1 de los participantes (InglésL1 codificado como -1 y EspañolL1 codificado como 1) como factor intersujetos en la Localización de la Primera Fijación (LPF), el Número de Fijaciones (NF) y el Lapso Ojo-Voz (LOV).

Estadísticas descriptivas

En la Tabla 2 se resumen las estadísticas descriptivas de las tres variables de interés: Localización de la Primera Fijación (LPF), Número de Fijaciones (NF) y Lapso Ojo-Voz (LOV) en función de la L1 de los participantes (Inglés vs. Español), el Idioma (Inglés vs. Español) y el Tipo de palabras (Palabras vs. Pseudo-palabras).

Localización de la Primera Fijación

La Tabla 3 muestra los resultados de los análisis de modelos lineales de efectos mixtos para la variable Localización de la Primera Fijación (LPF). La inspección de los parámetros estimados del modelo indica un efecto principal del Tipo de palabra. El valor del coeficiente estimado de regresión sugiere que la localización de la primera fijación es más cercana del inicio de los estímulos en el caso de las Pseudo-palabras en comparación con las Palabras.

Tabla 2

Promedios (y desviaciones estándares) de la Localización de la Primera Fijación, el Número de Fijaciones y el Lapso Ojo-Voz en función del L1 de los participantes, el Idioma y el Tipo de palabras.

L1 Participante	Idioma	Palabras	Pseudo-palabras
<i>Localización de la Primera Fijación (LPF)</i>			
L1 Español	Español	28.14 (10.62)	27.38 (11.16)
L1 Español	Inglés	28.33 (10.83)	26.80 (10.52)
L1 Inglés	Inglés	30.97 (9.27)	30.02 (9.44)
L1 Inglés	Español	29.25 (9.95)	29.26 (9.34)
<i>Número de Fijaciones (NF)</i>			
L1 Español	Español	4.40 (1.49)	5.30 (1.56)
L1 Español	Inglés	4.89 (1.57)	5.10 (1.49)
L1 Inglés	Inglés	3.90 (1.58)	4.77 (1.61)
L1 Inglés	Español	4.47 (1.68)	5.02 (1.62)
<i>Lapso Ojo-Voz (LOV)</i>			
L1 Español	Español	588.35 (216.78)	761.57 (248.49)
L1 Español	Inglés	745.80 (231.27)	828.39 (244.86)
L1 Inglés	Inglés	555.57 (158.85)	758.65 (245.18)
L1 Inglés	Español	653.71 (211.20)	750.64 (249.23)

Nota. Los resultados de la LPF se expresan en porcentaje del largo horizontal del estímulo (0% corresponde al inicio del estímulo y 100% al final de este); los resultados del NF se reportan en cantidad de fijaciones por estímulo; los resultados del LOV en milisegundos.

Número de Fijaciones

La Tabla 4 muestra los resultados de los análisis de modelos lineales de efectos mixtos para la variable Número de Fijaciones (NF). La inspección de los parámetros estimados del modelo indica un efecto principal del Tipo de palabra, un efecto principal del Idioma y una interacción de tres vías entre L1 Participante * Idioma * Tipo de palabra. Con la intención de ofrecer una representación gráfica que permita visualizar de manera más clara la modulación de la conducta oculomotora y el lapso ojo-voz de los participantes según las variables de interés, se muestran estos resultados en los 4 diagramas de violín que se presentan a continuación (Figura 4, Figura 5, Figura 6 y Figura 7). Para interpretar estos diagramas, se debe tener en consideración lo siguiente: en primer lugar, la línea central vertical se extiende desde el valor mínimo hasta el valor máximo adyacentes. En caso de haber valores atípicos (muy bajos o altos), estos se representan con puntos negros bajo o sobre la línea central respectivamente. En segundo lugar, la caja ubicada sobre esta línea

central vertical representa el rango intercuartil. Además, la raya horizontal que se halla sobre esta caja representa la mediana de los datos. Por último, el ancho de los diagramas representa la distribución de los valores: cuanto más gruesa la sección, más valores se hallan agrupados en esta

Sin embargo, para poder descifrarla, se recurrió a contrastes post-hoc de comparaciones múltiples de idioma por nivel de L1 y tipo de palabra mediante el uso de Tests HSD de Tukey. Estos indican que, cuando están leyendo pseudopalabras, tanto los participantes nativos del inglés ($\beta = -0.191$, $SE = 0.106$, $z = -1.800$, $p = 0.07$) como los participantes nativos del español ($\beta = -0.207$, $SE = 0.111$, $z = -1.854$, $p = 0.06$) no muestran un efecto significativo de idioma (el inglés comparado con el español) en el número de fijaciones. Adicionalmente, se observó que, tanto los participantes nativos del inglés ($\beta = -0.646$, $SE = 0.150$, $z = -4.301$, $p < .001$) como los participantes nativos del español ($\beta = 0.487$, $SE = 0.155$, $z = 3.132$, $p < .01$) hacen significativamente más fijaciones durante la lectura de palabras en su L2 en contraste con la lectura de palabras en su L1.

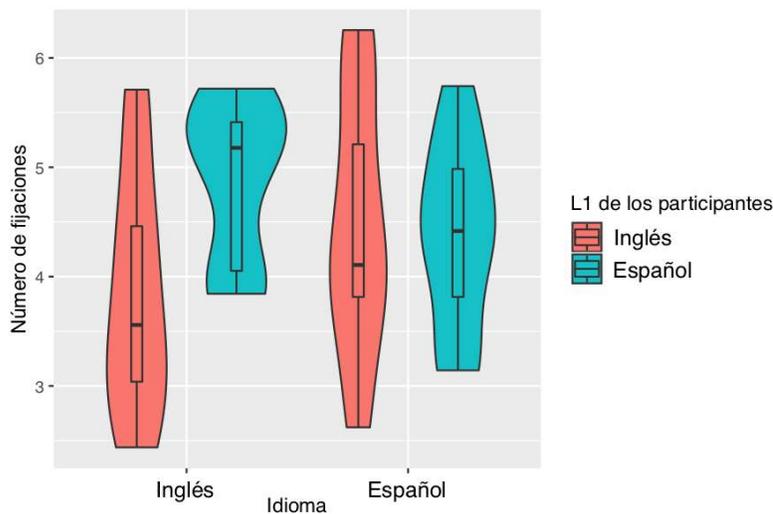


Figura 4. Número de fijaciones para Palabras en función del Idioma y la L1 de los participantes.

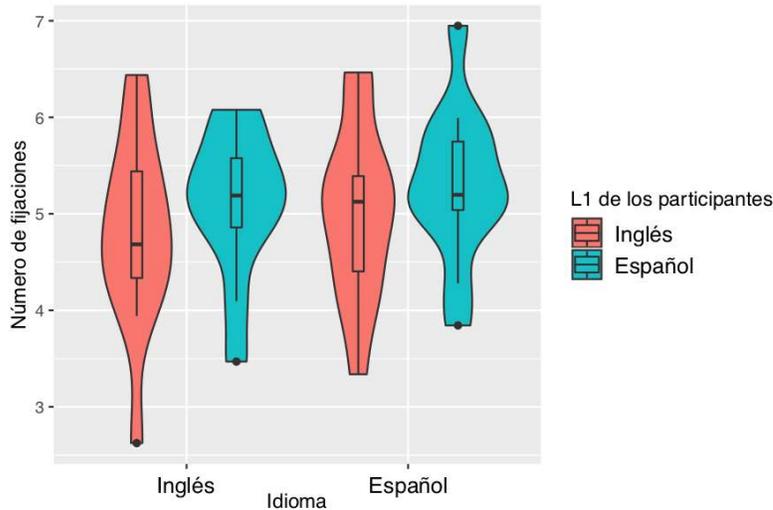


Figura 5. Número de fijaciones para Pseudo-palabras en función del Idioma y la L1 de los participantes.

Lapso Ojo-Voz

La Tabla 5 muestra los resultados de los análisis de modelos lineales de efectos mixtos para la variable Lapso Ojo-Voz (LOV). La inspección de los parámetros estimados del modelo indica un efecto principal del Tipo de palabra, un efecto principal del Idioma y una interacción de tres vías entre L1 Participante * Idioma * Tipo de palabra. Las figuras 6 y 7 permiten apreciar las respuestas observadas en función de las variables de interés y provee unas claves para interpretar esa interacción.

Sin embargo, para poder descifrarla, se recurrió a contrastes post-hoc de comparaciones múltiples de idioma por nivel de L1 y tipo de palabra mediante el uso de Tests HSD de Tukey. Estos indican que, cuando están leyendo pseudopalabras, los participantes nativos del inglés ($\beta = 5.74$, $SE = 22.59$, $z = 0.254$, $p = 0.79$) no muestran un efecto significativo de idioma (el inglés comparado con el español) en el LOV. En el caso de los participantes nativos del español, se observó que su LOV es mayor durante la lectura de pseudo-palabras en inglés en comparación con la lectura de pseudo-palabras en español ($\beta = 72.93$, $SE = 23.60$, $z = 3.091$, $p < .01$). Adicionalmente, se observó que, tanto los participantes nativos del inglés ($\beta = -103.10$, $SE = 25.98$, $z = -3.968$, $p < .001$) como los participantes nativos del español ($\beta = 152.15$, $SE = 26.90$, $z = 5.655$, $p < .001$) mostraron un LOV significativamente más corto durante la lectura de palabras en su L1 en comparación con la lectura de palabras en su L2.

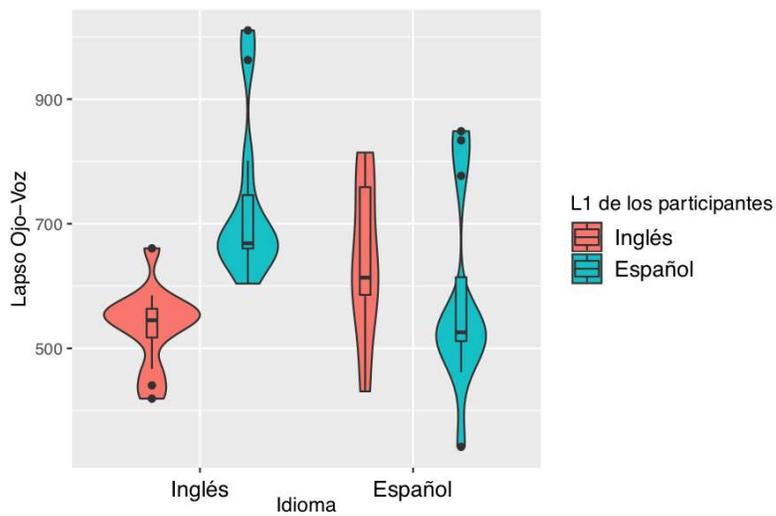


Figura 6. Lapso Ojo-Voz para Palabras en función del Idioma y la L1 de los participantes.

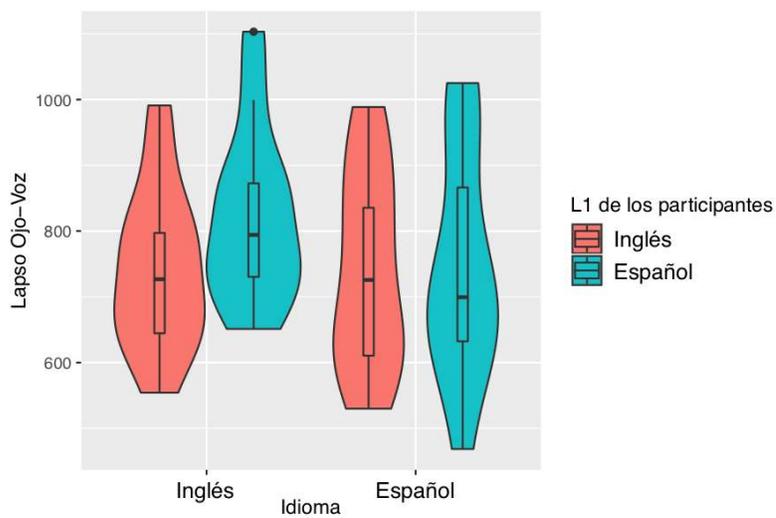


Figura 7. Lapso Ojo-Voz para Pseudo-palabras en función del Idioma y la L1 de los participantes.

Tabla 3

Estimaciones de los parámetros del modelo Localización de la Primera Fijación en función de la L1 de los participantes, el Idioma y el Tipo de palabra.

Efecto	Parámetro	Estimado	Error estándar	df	Valor <i>t</i>	Valor <i>p</i>
Intercepto	<i>Intercepto</i>	27.980	1.069	24	26.158	<.001
Pendiente de <i>Tipo de palabra</i>	<i>P-PP</i>	1.074	0.466	39	2.307	<.05
Pendiente de <i>Idioma</i>	<i>ESP-IN</i>	0.104	0.805	23	0.129	0.898
Pendiente de <i>L1 Participante</i>	<i>ESPL1-INL1</i>	-2.562	2.107	23	-1.126	0.236
Interacción de <i>Tipo de palabra * Idioma</i>	<i>P-PP * ESP-IN</i>	-1.172	0.781	29	-1.500	0.144
Interacción de <i>Tipo de palabra * L1 Participante</i>	<i>P-PP * ESPL1-INL1</i>	0.510	0.812	23	0.627	0.536
Interacción de <i>Idioma * L1 Participante</i>	<i>ESP-IN * ESPL1-INL1</i>	1.628	1.610	23	1.011	0.323
Interacción de <i>Tipo de palabra * Idioma * L1 Participante</i>	<i>P-PP * ESP-IN * ESPL1-INL1</i>	0.291	1.471	23	0.198	0.845

Nota. *Tipo de palabra* = Palabras (P) y Pseudo-palabras (PP), *Idioma*= Español (ESP) e Inglés (IN), *L1 Participante* = L1 Español (ESPL1) y L1 Inglés (INL1), y *P-PP* = contrastes (i.e., diferencias) entre Palabras y Pseudo-palabras, *ESP-IN* = contrastes (i.e., diferencias) entre Español e Inglés, *ESPL1- INL1* = contrastes (i.e., diferencias) entre L1 Español y L1 Inglés.

Tabla 4

Estimaciones de los parámetros del modelo Número de Fijaciones en función de la L1 de los participantes, el Idioma y el Tipo de palabra.

Efecto	Parámetro	Estimado	Error estándar	df	Valor <i>t</i>	Valor <i>p</i>
Intercepto	<i>Intercepto</i>	5.067	0.175	25	28.904	<.001
Pendiente de <i>Tipo de palabra</i>	<i>P-PP</i>	-0.675	0.097	38	-6.945	<.001
Pendiente de <i>Idioma</i>	<i>ESP-IN</i>	0.199	0.077	23	2.582	<.05
Pendiente de <i>L1 Participante</i>	<i>ESPL1-INL1</i>	0.304	0.342	22	0.889	0.383
Interacción de <i>Tipo de palabra * Idioma</i>	<i>P-PP * ESP-IN</i>	-0.120	0.134	33	-0.892	0.379
Interacción de <i>Tipo de palabra * L1 Participante</i>	<i>P-PP * ESPL1-INL1</i>	0.174	0.171	22	1.016	0.320
Interacción de <i>Idioma * L1 Participante</i>	<i>ESP-IN * ESPL1-INL1</i>	0.016	0.154	23	0.104	0.918
Interacción de <i>Tipo de palabra * Idioma * L1 Participante</i>	<i>P-PP * ESP-IN * ESPL1-INL1</i>	-1.149	0.246	23	-4.665	<.001

Nota. *Tipo de palabra* = Palabras (P) y Pseudo-palabras (PP), *Idioma*= Español (ESP) e Inglés (IN), *L1 Participante* = L1 Español (ESPL1) y L1 Inglés (INL1), y *P-PP* = contrastes (i.e., diferencias) entre Palabras y Pseudo-palabras, *ESP-IN* = contrastes (i.e., diferencias) entre Español e Inglés, *ESPL1- INL1* = contrastes (i.e., diferencias) entre L1 Español y L1 Inglés.

Tabla 5

Estimaciones de los parámetros del modelo Lapso Ojo-Voz en función de la L1 de los participantes, el Idioma y el Tipo de palabra.

Efecto	Parámetro	Estimado	Error estándar	df	Valor <i>t</i>	Valor <i>p</i>
Intercepto	<i>Intercepto</i>	756.94	28.63	25	26.755	<.001
Pendiente de <i>Tipo de palabra</i>	<i>P-PP</i>	-142.82	16.60	32	-8.605	<.001
Pendiente de <i>Idioma</i>	<i>ESP-IN</i>	-39.34	16.33	23	-2.409	<.05
Pendiente de <i>L1 Participante</i>	<i>ESPL1-INL1</i>	43.95	56.30	22	0.786	0.443
Interacción de <i>Tipo de palabra * Idioma</i>	<i>P-PP * ESP-IN</i>	14.81	20.84	30	0.711	0.483
Interacción de <i>Tipo de palabra * L1 Participante</i>	<i>P-PP * ESPL1-INL1</i>	23.61	30.64	23	0.771	0.449
Interacción de <i>Idioma * L1 Participante</i>	<i>ESP-IN * ESPL1-INL1</i>	-67.19	32.66	23	-2.057	0.05
Interacción de <i>Tipo de palabra * Idioma * L1 Participante</i>	<i>P-PP * ESP-IN * ESPL1-INL1</i>	-188.06	38.98	23	-4.825	<.001

Nota. *Tipo de palabra* = Palabras (P) y Pseudo-palabras (PP), *Idioma*= Español (ESP) e Inglés (IN), *L1 Participante* = L1 Español (ESPL1) y L1 Inglés (INL1), y *P-PP* = contrastes (i.e., diferencias) entre Palabras y Pseudo-palabras, *ESP-IN* = contrastes (i.e., diferencias) entre Español e Inglés, *ESPL1- INL1* = contrastes (i.e., diferencias) entre L1 Español y L1 Inglés.

Discusión

El objetivo de este Trabajo Final de Graduación fue determinar de qué manera la opacidad del lenguaje modula el procesamiento cognitivo en bilingües. Para ello, se evaluaron las estrategias de lectura en voz alta en dos grupos de bilingües que poseen lenguas que varían en términos de profundidad ortográfica: el primer grupo presentaba una L1 transparente (español) y una L2 opaca (inglés) mientras que el segundo grupo presentaba el patrón opuesto. Ambos grupos contaban con un nivel de proficiencia menor en su L2 con respecto a su L1. La conducta oculomotora y las respuestas vocales de los bilingües fueron registradas mientras leían palabras y pseudo-palabras en un contexto lingüístico monolingüe (español o inglés).

Estrategias de lectura en bilingües dominantes

En la lectura de palabras, ambos grupos de bilingües llevaron a cabo más fijaciones en su L2 con respecto a su L1. En línea con la predicción de un vínculo léxico más débil en la lengua no dominante (Castles et al., 2018), este aumento en el número de fijaciones era esperable ya que, independientemente de la profundidad ortográfica, al leer en su L2 los bilingües dominantes parecen adoptar una vía de conversión grafo-fonológica indirecta, la cual se basa en el procesamiento serial y local de unidades ortográficas pequeñas (Cop, Drieghe y Duyck, 2015; de León et al., 2016a; 2016b).

En cuanto a la localización de la primera fijación, los dos grupos de bilingües exhibieron una localización de la primera fijación más cercana al inicio de las pseudo-palabras en contraste con las palabras. Esto fue observado para ambos idiomas (L1 y L2) y para ambos contextos lingüísticos (español e inglés). Este resultado se alinea con la hipótesis de la modulación de las estrategias de lectura según la lexicalidad: durante la lectura de pseudo-palabras —las cuales carecen vínculo semántico— se ha observado que los lectores activan una vía de conversión grafo-fonológica indirecta (Buetler et al, 2014). Según el modelo de doble ruta (Coltheart et al., 2001), la traducción del código escrito a la oralidad en esta vía ocurre de manera local y en serie (de izquierda a derecha en el sentido de lectura). Así, la localización de la primera fijación más cercana al inicio en las pseudo-

palabras en comparación con las palabras parece indicar que la lexicalidad posee efectos muy tempranos que modulan la programación de la primera sacada hacia el inicio del estímulo (Rayner, 2009). Lo anterior es respaldado por investigaciones previas que vinculan una localización de la primera fijación alejada del inicio y más cercana al medio de las palabras con estrategias de lectura globales y en paralelo, las cuales son características de la lectura de palabras con un vínculo semántico fuerte (Vitu, O’regan, y Mittau, 1990).

Modulación del procesamiento cognitivo según opacidad del lenguaje

Durante la lectura de palabras, ambos grupos de bilingües presentaron un lapso ojo-voz más corto durante la lectura de palabras en su L1 con respecto a la lectura de palabras L2. Este hallazgo provee evidencia para sostener la hipótesis de la existencia de un vínculo léxico más robusto con las formas de las palabras en la L1 en contraste con la L2 (Nation y Castles, 2017), lo cual da paso a un reconocimiento más rápido de las palabras escritas y un subsecuente acceso léxico más directo en la L1. Como se mencionó previamente, esta hipótesis también es soportada por las diferencias discutidas previamente en el número de fijaciones entre la L1 y la L2 de los participantes durante la lectura de palabras.

En la lectura de pseudo-palabras, los bilingües con una L1 opaca (inglés) no presentaron diferencias estadísticas en la duración del lapso ojo-voz entre contextos lingüísticos (español vs. inglés). Lo anterior puede explicarse de la siguiente manera: dada la ausencia de lexicalidad de las pseudo-palabras, su lectura exigiría una conversión grafofonológica mediante una vía indirecta (Coltheart et al., 2001). Para el caso de los bilingües dominantes con una L1 opaca, el uso cotidiano de esta vía es raro en contraste con el uso de la vía directa (Ziegler y Goswami, 2005). Así, es posible argumentar que la lectura de pseudo-palabras para bilingües con L1 opaca y menor proficiencia en su L2 supondría una estrategia de lectura similar, independientemente del contexto lingüístico en el cual son presentadas (español o inglés en este caso).

Los bilingües con una L1 transparente (español) exhibieron un lapso ojo-voz menor durante la lectura de pseudo-palabras en español en comparación con la lectura de pseudo-palabras en inglés. En este caso, podría discutirse que el contexto lingüístico posee un

efecto sobre el procesamiento del lenguaje escrito: dado que se propone que la lectura de pseudo-palabras se lleva a cabo mediante la vía indirecta (Coltheart et al., 2001), y dado que se ha observado la prevalencia de esta vía durante la lectura en lenguas transparentes (de León Rodríguez et al., 2016a), es razonable pensar que, dada la familiaridad con el uso de estrategias de lectura locales y en serie, los bilingües dominantes con una L1 transparente presentarían costes de procesamiento de pseudo-palabras menores en un contexto transparente en comparación con un contexto opaco. Lo anterior en vista de que el contexto opaco induciría al uso de una vía directa, la cual es menos usual en la lectura en lenguas transparentes (de León Rodríguez et al., 2016a).

Conclusiones

El estudio experimental de los procesos de lectura ha generado una línea de investigación que pretende explicar la modulación del procesamiento del lenguaje escrito según diversos factores. Mediante el estudio de estos procesos en bilingües cuyas lenguas varían en su grado de profundidad ortográfica (español como lengua transparente e inglés como lengua opaca), el presente Trabajo Final de Graduación viene a enriquecer el debate en curso sobre los mecanismos y recursos que utilizan los lectores para llevar a cabo la traducción del código escrito a la oralidad.

Los hallazgos encontrados permiten argumentar que la lectura en voz alta de palabras por parte de bilingües dominantes es sensible a modulaciones del contexto lingüístico, a saber, la profundidad ortográfica de la lengua leída. Igualmente, a partir de los resultados obtenidos es válido sostener que los factores intraindividuales (primera y segunda lengua, proficiencia en la lengua) juegan un papel cardinal en las estrategias de conversión grafo-fonológica. En conjunto, los aportes derivados de este estudio ponen en evidencia la complejidad inherente a los procesos cognitivos subyacentes a los procesos de lectura de palabras en voz alta.

Paralelamente con los insumos para el robustecimiento de perspectivas teóricas sobre procesos de conversión grafo-fonológica y de modelos cognitivos de lectura de palabras en bilingües, los resultados de esta investigación apuntan la necesidad de estudios que exploren los efectos de la profundidad ortográfica en diversos contextos lingüísticos. Si

bien se puso en evidencia la transversalidad de factores contextuales durante los procesos de lectura de palabras aisladas, no es posible generalizar estos hallazgos a procesos de lectura en niveles superiores. Por esta razón, es preciso ahondar en el tema mediante la implementación de paradigmas experimentales de lectura de oraciones, párrafos y textos.

Asimismo, aun cuando fue posible observar la influencia de factores individuales en el procesamiento de palabras escritas, ha de tenerse en consideración que la población estudiada poseía un perfil de bilingüismo específico (bilingües sucesivos con una L1 dominante y cuyas lenguas contrastan en términos de profundidad ortográfica). Por ello, es importante llevar a cabo estudios complementarios con poblaciones que posean distintos perfiles de bilingüismo, lo cual permitiría desagregar de manera más fina los efectos que fueron observados.

Referencias

- Alvarado, C. (2020, 15 de febrero). Página quince: Costa Rica será bilingüe en el 2040. *La Nación*. Recuperado de <https://www.nacion.com/opinion/columnistas/pagina-quince-costa-rica-sera-bilingue-en-el-2040/7PDLHZBAWVHOBMXF3YUC3YRZP4/story/>
- Barquero, K. (2019, 28 de octubre). Aumenta oferta del inglés que el sector productivo requiere. *La República*. Recuperado de <https://www.larepublica.net/noticia/aumenta-oferta-del-ingles-que-el-sector-productivo-requiere>
- Barquero, K. (2020, 13 de febrero). En 20 años, todos los estudiantes del país serán bilingües gracias a política educativa. *La República*. Recuperado de <https://www.larepublica.net/noticia/en-20-anos-todos-los-estudiantes-del-pais-seran-bilingues-gracias-a-politica-educativa>

- Bar-Kochva, I., y Breznitz, Z. (2012). Does the Reading of Different Orthographies Produce Distinct Brain Activity Patterns? An ERP Study. *PLoS ONE*, 7(5), e36030. <https://doi.org/10.1371/journal.pone.0036030>
- Buetler, K. A., de León Rodríguez, D., Laganaro, M., Müri, R., Spierer, L., y Annoni, J.-M. (2014). Language context modulates reading route: An electrical neuroimaging study. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00083>
- Buswell, G. T. (1921). The relationship between eye-perception and voice-response in reading. *Journal of Educational Psychology*, 12(4), 217-227. <https://doi.org/10.1037/h0070548>
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., y Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach. En M. J. Snowling y C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 6-23). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470757642.ch1>
- Cop, U., Drieghe, D., & Duyck, W. (2015). Eye Movement Patterns in Natural Reading: A Comparison of Monolingual and Bilingual Reading of a Novel. *PLOS ONE*, 10(8), e0134008. <https://doi.org/10.1371/journal.pone.0134008>
- Costa, A., Falé, I., y Luegi, P. (2016). Searching for behavioral correlates of text complexity in reading. En *Twenty-Third Annual Meeting-Society for the Scientific Study of Reading*.
- de León Rodríguez, D., Buetler, K. A., Eggenberger, N., Preisig, B. C., Schumacher, R., Laganaro, M., Nyffeler, T., Annoni, J.-M., y Müri, R. M. (2016a). The modulation

of reading strategies by language opacity in early bilinguals: An eye movement study. *Bilingualism: Language and Cognition*, 19(3), 567-577.

<https://doi.org/10.1017/S1366728915000310>

de León Rodríguez, D., Buetler, K. A., Eggenberger, N., Laganaro, M., Nyffeler, T., Annoni, J.-M., y Müri, R. M. (2016b). The Impact of Language Opacity and Proficiency on Reading Strategies in Bilinguals: An Eye Movement Study. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00649>

Ellis, N. C., Natsume, M., Stavropoulou, K., Hoxhallari, L., Daal, V. H. P., Polyzoe, N., Tsipa, M.-L., y Petalas, M. (2004). The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading Research Quarterly*, 39(4), 438-468. <https://doi.org/10.1598/RRQ.39.4.5>

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, 123(1), 71-99.

<https://doi.org/10.1037/0033-2909.123.1.71>

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 35(5), 263-279. <https://doi.org/10.1017/S0140525X11001841>

Fukuda, R., y Fukuda, T. (2009). Comparison of Reading Capacity for Japanese, German, and English. *Perceptual and Motor Skills*, 108(1), 281-296.

<https://doi.org/10.2466/pms.108.1.281-296>

Google y "el derecho al olvido". (2020). Dual Texts: Bilingual Magazine for Advanced Learners. Recuperado de <http://www.dualtexts.com/parallel/25-the-end-of-the-flu-virus?highlight=WyJmbHUjXQ==>

Grosjean, F. (1992). Another View of Bilingualism. *Cognitive Processing in Bilinguals Advances in Psychology*, 51-62. doi:10.1016/s0166-4115(08)61487-9

Halsband, U. (2006). Bilingual and multilingual language processing. *Journal of Physiology-Paris*, 99(4-6), 355-369.

<https://doi.org/10.1016/j.jphysparis.2006.03.016>

- Hoversten, L. J., Brothers, T., Swaab, T. Y., y Traxler, M. J. (2017). Early processing of orthographic language membership information in bilingual visual word recognition: Evidence from ERPs. *Neuropsychologia*, *103*, 183-190.
<https://doi.org/10.1016/j.neuropsychologia.2017.07.026>
- Inhoff, A. W., Solomon, M., Radach, R., y Seymour, B. A. (2011). Temporal dynamics of the eye-voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*(5), 543-558.
<https://doi.org/10.1080/20445911.2011.546782>
- Izura, C., Cuetos, F., y Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, *35*(1), 49-66.
- Joshi, R. M., Tao, S., Aaron, P. G., y Quiroz, B. (2012). Cognitive Component of Componential Model of Reading Applied to Different Orthographies. *Journal of Learning Disabilities*, *45*(5), 480-486. <https://doi.org/10.1177/0022219411432690>
- Katz, L., y Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 157-166.
<https://doi.org/10.1037/0278-7393.9.1.157>
- Katz, L., y Frost, R. (1992). Chapter 4 The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. En *Advances in Psychology* (Vol. 94, pp. 67-84). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)
- Keuleers, E., y Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633. <https://doi.org/10.3758/BRM.42.3.627>
- Kliegl, R., Grabner, E., Rolfs, M., y Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 262-284.
<https://doi.org/10.1080/09541440340000213>

- Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). “lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, 82(13), 1–26. doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13).
- Laubrock, J., y Kliegl, R. (2015). The eye-voice span during reading aloud. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01432>
- Lemhöfer, K., y Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44(2), 325-343. <https://doi.org/10.3758/s13428-011-0146-0>
- Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1-33. doi:[10.18637/jss.v069.i01](https://doi.org/10.18637/jss.v069.i01)
- Marian, V., Blumenfeld, H. K., y Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940-967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Mitchell, R., Myles, F., y Marsden, E. (2019). *Second language learning theories* (4ta ed.). Routledge.
- Nation, K., y Castles, A. (2017). Putting the learning in orthographic learning. En Cain, K., Compton, D., Parrila, R. (Eds.), *Theories of reading development* (pp. 147–168). Amsterdam, The Netherlands: John Benjamins.
- Peal, E., y Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, 76(27), 1-23. <https://doi.org/10.1037/h0093840>
- Rau, A. K., Moll, K., Snowling, M. J., y Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, 130, 92-105. <https://doi.org/10.1016/j.jecp.2014.09.012>

- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506. <https://doi.org/10.1080/17470210902816461>
- Roux, F., Armstrong, B. C., y Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, 49(5), 1864-1881. <https://doi.org/10.3758/s13428-016-0830-1>
- RStudio Team (2020). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA <http://www.rstudio.com/>
- Schneider, W., Eschman, A., y Zuccolotto, A. (2002). E-prime (version 2.0). *Computer software and manual*. Pittsburgh, PA: Psychology Software Tools Inc.
- Seymour, P. H. K., Aro, M., Erskine, J. M., y collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143-174. <https://doi.org/10.1348/000712603321661859>
- Simon, G., Bernard, C., Lalonde, R., y Rebaï, M. (2006). Orthographic transparency and grapheme–phoneme conversion: An ERP study in Arabic and French readers. *Brain Research*, 1104(1), 141-152. <https://doi.org/10.1016/j.brainres.2006.05.071>
- Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., y Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 49(1), 111-123. <https://doi.org/10.3758/s13428-015-0700-2>
- Taylor, J. S. H., Rastle, K., y Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, 139(4), 766-791. <https://doi.org/10.1037/a0030266>
- The end of the flu virus? (2020). Dual Texts: Bilingual Magazine for Advanced Learners. Retrieved from <http://www.dualtexts.com/parallel/41-the-right-to-be-forgotten?highlight=WyJnb29nbGUiLCJnb29nbGUncyJd>

- Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, 47(6), 583-600.
<https://doi.org/10.3758/BF03203111>
- Warriner, A. B., Kuperman, V., y Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191-1207. <https://doi.org/10.3758/s13428-012-0314-x>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. <https://ggplot2.tidyverse.org>.
- Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation. R package version 0.8.4. <https://CRAN.R-project.org/package=dplyr>
- Ziegler, J. C., Perry, C., Jacobs, A. M., y Braun, M. (2001). Identical Words are Read Differently in Different Languages. *Psychological Science*, 12(5), 379-384.
<https://doi.org/10.1111/1467-9280.00370>
- Ziegler, J. C., y Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, 131(1), 3-29. <https://doi.org/10.1037/0033-2909.131.1.3>

Apéndice 1: Paradigmas psicolingüísticos para el estudio del bilingüismo: propuesta para el control y la selección asistida de estímulos

Paradigmas psicolingüísticos para el estudio del bilingüismo: propuesta para el control y la selección asistida de estímulos

Psycholinguistic paradigms for the study of bilingualism: a proposal for an assisted control and selection of stimulus

Alejandro Cambroner Delgadillo ^a, Diego de León Rodríguez ^b, y Thomas Castelain ^{a, c *}

^a Instituto de Investigaciones Lingüísticas, Universidad de Costa Rica, San José, Costa Rica

^b École Polytechnique Fédérale de Lausanne, Lausanne, Suisse

^c Instituto de Investigaciones Psicológicas, Universidad de Costa Rica, San José, Costa Rica

Agradecimientos. Los autores agradecen a Antonio Leoni de León por su apoyo durante el proceso investigativo.

Financiamiento. Este trabajo contó con el apoyo de la Vicerrectoría de Investigación de la Universidad de Costa Rica [Proyecto número 745-B8-311] y una beca *Fondo de Apoyo a Tesis de Posgrado 2018* (de la misma institución) otorgada a Alejandro Cambroner Delgadillo.

* **Persona de contacto.** Thomas Castelain, Universidad de Costa Rica, Instituto de Investigaciones Psicológicas, Instituto de Investigaciones Lingüísticas, Costa Rica; correo electrónico: thomas.castelain@ucr.ac.cr

Resumen

La selección y el control de estímulos es una tarea crítica para los estudios en psicolingüística. Esta etapa suele ser demandante para lo(a)s investigadore(a)s, más aún cuando se pretende comparar distintas lenguas. Para facilitar este proceso, brindamos una propuesta metodológica para la selección de estímulos en paradigmas psicolingüísticos aplicados al bilingüismo, específicamente adaptada al estudio de los procesos de lectura. Después de discutir brevemente los principales factores que se han visto asociados a la modulación del procesamiento del lenguaje escrito, se propone un algoritmo práctico diseñado privilegiando el control automatizado y/o asistido mediante herramientas informáticas con el fin de aumentar la rigurosidad del control y minimizar el error humano. Finalmente, se ofrece una implementación del algoritmo la cual puede adaptarse a distintos lenguajes y paradigmas de investigación.

Palabras clave: Psicolingüística, bilingüismo, selección de estímulos, metodología

Abstract

The control and selection of stimuli is a critical task in psycholinguistics' studies. This stage is usually demanding for researchers, especially when comparing different languages. To alleviate this process, we offer a methodological proposal for the selection of stimuli in psycholinguistic paradigms applied to bilingualism, specifically adapted to the study of reading processes. After briefly discussing the main factors that have been associated with the modulation of written language processing, we propose a practical algorithm designed to favor automated and/or computer-assisted control in order to increase the rigor of control and minimize human error. Finally, an implementation of the algorithm is offered which can be adapted to different languages and research paradigms.

Keywords: Psycholinguistics, bilingualism, stimuli selection, methodology

Introducción

El bilingüismo es la habilidad para utilizar una segunda lengua (L2) de forma cotidiana (Halsband, 2006). A partir de la década de 1980 se ha dado un aumento sostenido de la cantidad de bilingües a nivel global (Drazer, 2018), y actualmente más de la mitad de la población mundial utiliza una L2 con una frecuencia suficiente para ser considerada bilingüe (Ansaldo, Marcotte, Scherer, & Raboyeau, 2008).

Este aumento en la cantidad de personas que hablan dos lenguas ha surgido debido a una multiplicidad de factores tales como las grandes migraciones humanas, políticas estatales que promueven la adquisición de una segunda lengua y cambios en las dinámicas del mercado, los cuales exigen a los individuos el dominio de una L2 para desenvolverse en su entorno (Carlson & McHenry, 2006).

Más allá de su faceta como una competencia útil, la adquisición de una segunda lengua reconfigura el sistema cognitivo en términos de estructura, posibilidades de enunciación y contexto lingüístico con el objetivo de que el hablante sea capaz de procesar la información y comunicarse en otro idioma (Costa & Sebastián-Gallés, 2014). Actualmente se ha desarrollado una vertiente de investigación la cual tiene como objetivo avanzar el conocimiento del bilingüismo y su relación más amplia con otros procesos mentales (Sönmez, 2019).

Si bien se ha mostrado que los bilingües poseen un mejor rendimiento que los monolingües en tareas ejecutivas (Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés 2010), en procesos atencionales (Brito, Murphy, Vaidya, & Barr, 2016) y en memoria de trabajo (Morales, Calvo, & Bialystok, 2013), los beneficios asociados al manejo de una L2 han sido fuertemente cuestionados (Lehtonen et al., 2018) e inclusive refutados (Nichols, Wild, Stojanoski, Battista, & Owen, 2020).

Así, los resultados de estas investigaciones y el debate suscitado han despertado un renovado interés hacia la comprensión del bilingüismo y sus implicaciones en la mente humana, de manera que el estudio del fenómeno se presenta como un área promisoría y atractiva para investigadores en el ámbito de los estudios del lenguaje y en las ciencias cognoscitivas en general.

Ahora bien, la investigación sobre bilingüismo presenta un gran reto metodológico puesto que el lenguaje es un proceso cognitivo complejo, el cual normalmente acaece en contextos cargados de información proveniente tanto de los estados internos de los interlocutores como de los referentes del mundo externo (Mufwene, Coupe, & Pellegrino, 2017). Así, el lenguaje se halla atravesado por una multiplicidad de factores que poseen efectos concretos sobre cómo este es procesado y producido.

Con el objetivo de poder discernir de manera precisa los efectos que posee cierto factor de interés, es preciso establecer condiciones en las cuales se controle la influencia de aquellos otros factores que no nos interesan y que pueden estar causando ruido sobre el fenómeno que pretendemos observar.

El procedimiento para desagregar estos efectos no es trivial, ya que exige al(la) investigador(a) un protocolo con un control de variables minucioso, el cual dé paso a una metodología fiable para la obtención de resultados válidos e informativos. Aunado al reto que supone trabajar en paradigmas experimentales con lenguaje, los estudios con bilingüismo agregan otra dimensión de complejidad, ya que debe hallarse una manera adecuada para asegurar una equivalencia de factores entre dos lenguas diferentes. Al referirse a la literatura especializada, es posible observar gran diversidad y variabilidad en el control de factores utilizados. Ante esta situación, quien desee llevar a cabo estudios sobre bilingüismo puede sentirse abrumado y sin un norte claro con respecto a las decisiones metodológicas que debe tomar para dar una respuesta válida a su pregunta de investigación.

Con el objetivo de aportar una guía clara y práctica para quienes investigan en el ámbito de la psicolingüística y la lectura en bilingües, a continuación, ofrecemos una propuesta metodológica para el control y la selección de estímulos en este tipo de paradigmas. En primer lugar, repasaremos un conjunto de factores que han sido ampliamente descritos como moduladores del lenguaje escrito. Luego, presentaremos un algoritmo para el control de estos factores que ha sido estructurado privilegiando el control automatizado y/o asistido mediante herramientas informáticas con la intención de aumentar la rigurosidad de dicho control y minimizar el error humano. Por último, pondremos a disposición una implementación del algoritmo (<https://tinyurl.com/StimDatabase>), la cual es adaptable según los estímulos, los idiomas y el tipo de estudio.

Dificultades metodológicas en el estudio del bilingüismo y la lectura

Debido a su carácter interdisciplinario y su abordaje empírico al estudio del lenguaje, la psicolingüística se ha posicionado como una alternativa promisoriosa para la comprensión del bilingüismo (Grosjean, Li, & Bialystok, 2013). En cuanto al estudio de la lectura, es importante recalcar que el registro escrito posee especificidades propias que lo diferencian del registro oral (Luiselli et al., 2013). ; por ejemplo, contamos con evidencia tanto conductual como neurofisiológica que nos indica que hay factores que poseen efectos específicos sobre los procesos de lectura en bilingües: la edad de adquisición de una L2 (Larraza, Samuel, & Oñederra, 2016), la influencia de la L1 sobre la L2 (Pavlenko, 2010), el grado de exposición a una L2 (Unsworth, Chondrogianni, & Skarabela, 2018), la diferencia de proficiencia entre la L1 y la L2 (Cieślicka, Heredia, & Olivares, 2014) y el contexto de adquisición de una L2 (Hayakawa, Tannenbaum, Costa, Corey, & Keysar, 2017).

Así, cuando se desea plantear una investigación sobre lectura en bilingües, es necesario el control adecuado de un cúmulo de factores vinculados al procesamiento del lenguaje escrito que permita, por un lado, tomar en consideración estos moduladores de la lectura y, por otro lado, establecer un contexto de equivalencia entre lenguas distintas (Kaushanskaya, Blumenfeld, & Marian, 2019).

Este ejercicio de control y equiparación es laborioso y requiere una inversión de tiempo considerable para ser llevado a cabo adecuadamente. Aunado a lo anterior, ha de tenerse en consideración que a medida que se aumenta la cantidad de factores y el rigor para controlarlos, ocurren dos efectos colaterales. Por un lado, se va reduciendo el número de elementos disponibles que son candidatos para formar parte de nuestra selección de estímulos definitiva. Por otro lado, la búsqueda perfeccionista en aras de equivalencia puede conducir a estímulos rebuscados que no corresponden con los usos del lenguaje que ocurren en contextos cotidianos (Bernardy, Lappin, & Lau, 2018).

De esta manera, en ausencia de una metodología correcta para concretar esta tarea es fácil caer en dificultades o sesgos que vendrán en detrimento de la calidad de los resultados investigativos (Guasch, Boada, Ferré, & Sánchez-Casas, 2013). Por ello, quienes

se enfrenten a este proceso sin una guía adecuada pueden hallarse perdidos ante la complejidad y las particularidades de la tarea.

Ante esta situación, es primordial tener claro qué particularidades sobre el bilingüismo deseamos estudiar con el fin de optar por una metodología pertinente, que se ajuste a nuestros requerimientos y con un control de factores que asegure que la modulación del fenómeno que se espera observar sea producto de las manipulaciones experimentales y no del ruido proveniente de factores no controlados.

Con la intención de ofrecer un esquema general del tratamiento del lenguaje escrito en el bilingüismo, a continuación, repasaremos un modelo cognitivo que explica este proceso en sus distintos niveles.

Niveles de procesamiento del lenguaje en la lectura bilingüe: el modelo BIA+

Con el fin de explicar el procesamiento del lenguaje en los bilingües se han desarrollado distintos modelos cognitivos que ilustran los niveles (subléxico, léxico, semántico) de tratamiento de la información lingüística (Li, 2013). Si bien la oferta de modelos es variada (ver Basnight-Brown, 2014), en esta sección nos remitiremos al *Bilingual interactive activation plus* (BIA+; Dijkstra & van Heuven, 2002) con el fin de contar con un marco referencial para situar los distintos factores que discutiremos en la siguiente sección.

El BIA+ es un modelo conexionista (Figura 1) para el reconocimiento de palabras escritas en bilingües. De forma breve, el modelo concibe este proceso de la siguiente manera: al leer una palabra, sus rasgos visuales activan las unidades ortográficas subléxicas. En seguida, estas unidades activan de manera paralela las unidades ortográficas léxicas y las unidades fonológicas subléxicas. Luego, las unidades fonológicas subléxicas activan las unidades fonológicas léxicas, las cuales junto con las unidades ortográficas léxicas coactivan los nodos del lenguaje y el nivel semántico. Una vez activados los nodos del lenguaje y el nivel semántico, el bilingüe es capaz de asignar la palabra a uno u otro idioma y recuperar su significado. Una vez identificada la palabra, el sistema de decisión se activa, el cual se encarga de especificar las acciones a seguir según la tarea que realice el lector (traducción, tarea de decisión léxica, lectura en voz alta, etcétera).

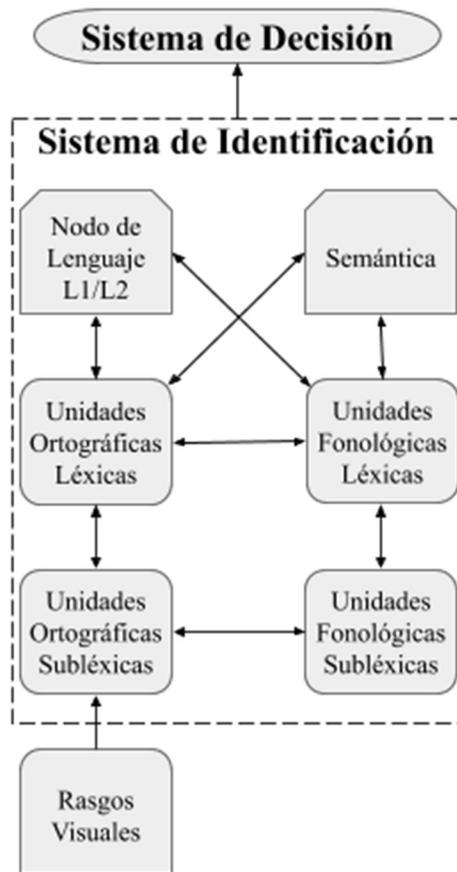


Figura 3. Modelo BIA+ para la lectura en bilingües. Adaptado de van Kesteren, Dijkstra, y de Smedt (2012).

Como se puede apreciar en este modelo, la activación de los distintos niveles de procesamiento de la información lingüística ocurre tanto de forma paralela como en serie, iniciando en los niveles subléxicos hasta culminar en el sistema de decisión. Al tener en cuenta este principio de activación interactivo, podemos entender la razón y la importancia del control de factores en distintos niveles. Previo a la siguiente sección es importante mencionar que existen diversas posturas con respecto a la categorización del nivel al cual pertenecen los distintos factores que modulan el procesamiento del lenguaje escrito, mas esta discusión rebasa el alcance de este artículo (para una revisión del tema, ver Goldrick et al., 2014).

Principales factores asociados a la modulación de los procesos de lectura en bilingües

Como mencionamos previamente, el lenguaje es un fenómeno complejo que se encuentra atravesado por una multitud de factores que modulan su procesamiento y producción. Con el objetivo de reducir al mínimo la influencia de otros factores que puedan causar ruido, es preciso llevar a cabo un control que sustente y dé validez a nuestro protocolo experimental. Claramente, este control puede variar en menor o mayor grado de un estudio a otro, mas, es importante considerar que las decisiones metodológicas tomadas sean consecuentes con nuestras hipótesis experimentales. Al trabajar con dos lenguas, se debe tener la precaución de incluir factores que son específicos al tratamiento de información lingüística en bilingües (Kroll & Bialystok, 2013). En seguida, discutiremos una serie de factores los cuales han sido asociados a la modulación de los procesos de lectura en bilingües. El compendio que presentamos no tiene la pretensión de ser exhaustivo; no obstante, constituye un buen punto de partida para el desarrollo de paradigmas experimentales con población bilingüe.

Perfil de bilingüismo. En la investigación con bilingües es fundamental incluir un instrumento que permita evaluar de manera fiable el perfil del bilingüe. El instrumento debe estar compuesto por ítems para sondear elementos como la edad y el orden de adquisición de las lenguas, el contexto en que se aprendieron, la exposición a cada una y la proficiencia en ambas (Bedore et al., 2012). La importancia de lo anterior radica en que el perfil bilingüe permite tener en cuenta un cúmulo de factores que se han visto asociados con la modulación en el procesamiento y producción del lenguaje (Kroll, Bobb, & Hoshino, 2014).

La oferta de instrumentos para evaluar el perfil bilingüe es vasta (Kunnan, 2013). A grandes rasgos, existen tres tipos: pruebas objetivas, medidas subjetivas y exámenes de lengua. En primer lugar, las pruebas objetivas son evaluaciones que permiten cuantificar el rendimiento del individuo en el tratamiento de la información del lenguaje. Entre ellas se halla la cantidad de palabras leídas por minuto (Segalowitz & Hébert, 1990), las evaluaciones de gramaticalidad y aceptabilidad (Langsford, Perfors, Hendrickson, Kennedy, & Navarro, 2018) y las tareas de decisión léxica (Zhang & Thompson, 2004).

Dada la facilidad de implementación de esta última, su disponibilidad en diversas lenguas y su fiabilidad (Whitford & Titone, 2016), es considerada una prueba de referencia en la investigación con bilingües.

En segundo lugar, las medidas subjetivas se presentan como cuestionarios con ítems concernientes al dominio, proficiencia y cronología de adquisición de las lenguas (Kaushanskaya et al., 2019). Si bien son fiables para estimar el grado de bilingüismo de las personas (Sheng, Lu, & Gollan, 2014), se recomienda utilizarlos en complemento con una medida objetiva que nos dé un indicador estandarizado que facilite el contraste entre distintos participantes.

Por último, los exámenes de lengua son pruebas sistemáticas que permiten obtener una estimación fiable sobre la proficiencia y el conocimiento de una lengua mediante la evaluación de las habilidades escritas, orales y de escucha en una lengua (Educational Testing Service, 2017). Estas proveen el diagnóstico más exhaustivo sobre el nivel de bilingüismo de los sujetos; no obstante, suponen una cantidad considerable de tiempo para ser completados.

Categoría gramatical. La categoría gramatical refiere a la clasificación de las palabras de acuerdo con sus propiedades morfológicas, semánticas o sintácticas (Bosque, 2007). Diversos estudios han aportado evidencia para sostener que el procesamiento del lenguaje está mediado por información relativa a la categoría gramatical de las palabras (Hillis & Caramazza, 1995). ; por ejemplo, se ha observado que en el contexto de lectura de palabras aisladas, la corteza frontal dorsolateral se activa durante la lectura de verbos mas no al leer sustantivos (Perani et al., 1999). Interesantemente, este efecto prevalece en la lectura de homónimos que pueden pertenecer a más de una categoría gramatical ("carga" como verbo o sustantivo), ya que existe una mediación diferenciada de rutas neuronales según la categoría a la cual se adscribe la palabra en el contexto de presentación (Tranel, Martin, Damasio, Grabowski, & Hichwa, 2005).

Homógrafos. Aquellas palabras que comparten la misma representación ortográfica pero cuyo significado difiere son llamadas homógrafos. La existencia de homógrafos entre lenguas es común (; por ejemplo, la palabra española "pan" y la palabra inglesa "pan") y su procesamiento en lectura bilingüe conlleva procesos específicos (Hoshino & Thierry,

2012). Así, se ha postulado que al leer homógrafos interlengua los bilingües coactivan ambas representaciones semánticas de la palabra, lo que deviene en costes de procesamiento más elevados asociados con la inhibición de la palabra en la otra lengua (Durlík, Szewczyk, Muszyński, & Wodniecka, 2016).

Homófonos. Las palabras cuya representaciones ortográficas son diferentes pero su pronunciación es la misma son llamadas homófonos. Durante la lectura de homófonos interlengua, se ha observado que los bilingües muestran costes de procesamiento reducidos en contraste con la lectura de palabras control (Carrasco-Ortiz, Midgley, & Frenck-Mestre, 2012). Se ha postulado que esta facilitación es resultado de la coactivación de las representaciones fonológicas de la palabra en ambos idiomas, lo cual facilitaría su reconocimiento.

Frecuencia léxica. La frecuencia léxica refiere a la proporción de ocurrencia de una palabra en una lengua, y es considerada una de las variables más críticas que deben ser controladas en estudios de psicolingüística (Murray & Forster, 2004). Uno de los efectos más explorados en estudios del lenguaje es el “efecto de la frecuencia léxica”, el cual refiere a la facilitación en el procesamiento de las palabras a medida que su frecuencia aumenta (Brysbaert, Mandera, & Keuleers, 2018). La robustez de este efecto tiene larga data y ha sido ampliamente replicado en diferentes contextos (para una revisión, ver Bontrager, 1991).

Vecinos léxicos. Un vecino léxico es toda aquella palabra que difiere de otra por una letra mientras se respeta el orden de sus letras restantes (Coltheart, Davelaar, Jonasson & Besner, 1977); por ejemplo, “carro” y “cargo”. Se han descrito dos fenómenos principales relacionados con los vecinos ortográficos y la lectura: el efecto de la frecuencia de vecinos y el efecto de la densidad de vecinos (Perea & Rosa, 2000). Por un lado, el primer efecto refiere al aumento en el tiempo de reconocimiento de una palabra a medida que la frecuencia léxica de sus vecinos sea mayor. Por otro lado, el efecto de la densidad de vecinos corresponde a la facilitación en el reconocimiento de la palabra a medida que posea más vecinos. La magnitud de este segundo efecto es mayor cuando las palabras vecinas poseen una frecuencia léxica baja.

Bigramas ortográficos. Un bigrama ortográfico corresponde a la(s) secuencia(s) de 2 letras adyacentes que componen una palabra (la palabra “casa” está compuesta por 3 bigramas: “ca-”, “-as-” y “-sa”). A medida que la frecuencia de bigramas que compone una palabra es mayor, su reconocimiento se facilita (Oganian et al., 2016). Adicionalmente, la posición específica del bigrama dentro la palabra también modula la magnitud de este efecto (Muncer, Knight, & Adams, 2014). En bilingüismo es crucial controlar este factor debido a que la frecuencia de bigramas varía considerablemente entre lenguas (Frauenfelder et al., 1993).

Sílabas. Una sílaba es cada una de las unidades de articulación fonológicas que conforman una palabra (“Syllable”, 2020). En la lectura, la cantidad de sílabas de una palabra modula cómo esta es procesada, ya que a medida que aumenta el número de sílabas, el reconocimiento de la palabra se torna menos eficiente (Ferrand & New, 2003). Debido a la diversidad en las correspondencias grafofonológicas que existe entre las lenguas y su impacto en el número de sílabas (Katz & Frost, 1992), el control interlengua de este factor en el diseño experimental es crucial.

Ahora expuestos los principales factores que deben ser controlados al montar un protocolo experimental en psicolingüística, expondremos el algoritmo que desarrollamos para un control y una selección asistida de los estímulos en estudios sobre el bilingüismo. Adicionalmente, la implementación de este algoritmo puede consultarse en <https://tinyurl.com/StimDatabase>, la cual contiene una matriz que puede ser alimentada con estímulos propios y factores complementarios para así adaptarla a cualquier idioma o tema de investigación..

Propuesta de algoritmo para el control de estímulos en paradigmas psicolingüísticos

A continuación, se ofrece una metodología para la selección de estímulos en paradigmas psicolingüísticos aplicados al bilingüismo (ver Tabla 1). El algoritmo aquí propuesto fue creado para ser utilizado en paradigmas de lectura bilingüe, ; no obstante, está diseñado de manera que puede adaptarse e implementarse fácilmente según las necesidades investigativas del fenómeno en estudio.

Con la intención de minimizar el error humano, el orden de las etapas del algoritmo favorece el control automatizado y/o asistido mediante herramientas informáticas y se relega el control manual de estímulos a una fase final de verificación. A continuación, se describe en detalle cada etapa del algoritmo y su implementación.

Tabla 1

Etapas de control de estímulos en paradigmas en psicolingüística y bilingüismo

Etapas
1. Selección inicial de las palabras por utilizar
2. Revisión de errores en las bases de palabras
3. Elección según categoría gramatical
4. Cotejo de homógrafos entre lenguas
5. Cotejo de homófonos entre lenguas
6. Frecuencia léxica
7. Número de vecinos
8. Bigramas
9. Número de sílabas
10. Equiparación asistida de factores entre lenguas
11. Revisión manual del conjunto restante de palabras

1. Selección inicial de las palabras por utilizar

En la literatura especializada se encuentran disponibles bases de datos en distintos idiomas, las cuales contienen palabras que han sido controladas por diversos factores como la prevalencia (Brysbaert, Mandera, McCormick, & Keuleers, 2019) y uso gramatical (Friendly, Franklin, Hoffman, & Rubin, 1982). Dependiendo del objetivo de nuestro estudio, dichas bases constituyen un buen punto de partida ya que ofrecen un corpus inicial que nos permite seleccionar palabras que se ajusten a determinado criterio.

; por ejemplo, en la implementación del algoritmo que ofrecemos (<https://tinyurl.com/StimDatabase>), las palabras seleccionadas fueron tomadas de una base en español (Stadthagen-Gonzalez, Imbault, Pérez Sánchez, & Brysbaert, 2017) y una base en inglés (Warriner, Kuperman, & Brysbaert, 2013) las cuales controlan por la valencia emocional (negativa, neutra, positiva) y la excitación (alta, neutra, baja) que evocan las palabras.

2. Revisión de errores en las bases de palabras

Dada la gran cantidad de entradas que presentan las bases previamente mencionadas, es posible que inadvertidamente contengan palabras con faltas ortográficas o con errores tipográficos. Con el fin de excluir estas faltas, una estrategia consiste en someter a consulta la lista completa de palabras a un diccionario digital (; por ejemplo, Real Academia Española, 2014) y, luego, mantener únicamente las coincidencias. Al hacer uso de esta estrategia debe tenerse en cuenta que las bases de datos pueden contener palabras que no se encuentran registradas en el diccionario utilizado (tal como los neologismos), de manera que es preciso revisar manualmente las palabras que no coinciden para no descartarlas por omisión.

3. Elección según categoría gramatical

Para determinar la categoría gramatical de las palabras es preciso hacer uso de un analizador sintáctico, el cual es un programa computacional que determina el género, el número y la(s) parte(s) del discurso a la(s) cual(es) pertenece una palabra, ya sea cuando esta se halla aislada o en el contexto de una oración. Estas herramientas cuentan con la ventaja de que cuando una palabra pertenece a más de una categoría gramatical, es posible observar la proporción de veces que es categorizada en cada una de las posibles partes de la oración y, con base en esta proporción, se pueden tomar decisiones con respecto a su inclusión o no en el conjunto final de estímulos.

Actualmente, existe una oferta de analizadores sintácticos de acceso libre y en diversas lenguas (en inglés, Rayson & Garside, 1998; en español, Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013) los cuales son capaces de analizar y etiquetar automáticamente miles de palabras en cuestión de segundos. El análisis automatizado es útil mas debe tomarse únicamente como referencia y ser corroborado por un experto

humano, ya que los analizadores son propensos a errores en ciertos contextos lingüísticos (Keselman et al., 2008).

4. Cotejo de homógrafos

Para excluir los homógrafos interlengua, las listas de palabras preseleccionadas de ambos idiomas deben someterse a un diccionario de la otra lengua (; por ejemplo, Cambridge University Press, 2019). Si se obtienen coincidencias, estas palabras deben eliminarse del conjunto final de estímulos. Luego, deben compararse las listas para tener certeza de que no presenten palabras en común las cuales no fueron detectadas con este método.

5. Cotejo de homófonos

Para eliminar los homófonos entre lenguas, nos podemos valer de transcriptores fonéticos automáticos (Weide, 1998) para obtener la representación fonética de las palabras. Una vez que se cuenta con estas representaciones para las palabras preseleccionadas, el primer paso es descartar los homófonos intralengua (p.e. “ha” y “a” en español).

Luego, se procede a comparar las representaciones de ambas lenguas entre sí para eliminar los homófonos interlengua. Una consideración válida al momento de escoger el transcriptor es preferir aquellos que cuenten con la opción de utilizar fonologías específicas (; por ejemplo, Duchon et al., 2013), y así, poder obtener la representación fonética de las palabras más cercana a la variedad dialectal utilizada por la población de bilingües con quien se pretende trabajar.

6. Frecuencia léxica

Actualmente disponemos de una gran cantidad de bases (ver New, Pallier, Brysbaert, & Ferrand, 2004) para calcular la frecuencia léxica. Al escoger esta base, se debe tomar en consideración el corpus utilizado para la estimación de frecuencias, ya que existen fuentes y registros que han demostrado ser más apropiados y más representativos según el grupo en estudio y los usos del lenguaje (Gimenes & New, 2016). Aunado a esto, en la medida de lo posible se debe optar por bases que posean la opción de discriminar la frecuencia léxica según la variable dialectal utilizada por la población en estudio (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015).

7. Número de vecinos

La cantidad de vecinos ortográficos de una palabra se estima mediante el uso de programas que recuperan el número y la lista de palabras que difieren de la palabra clave por una letra. Entre estos, CLEARPOND (Marian, Bartolotti, Chabal, & Shook, 2012) cuenta con la ventaja de que permite recuperar distintos tipos de vecinos ortográficos (delección, adición o sustitución de una letra).

Dado que se cree que en el reconocimiento de palabras los bilingües presentan una activación paralela de sus dos lenguas (van Heuven, Dijkstra, & Grainger, 1998), es aconsejable controlar los estímulos por vecinos en la otra lengua, en otras palabras, calcular el número de vecinos en x lengua que posee una palabra en y lengua.

8. Bigramas

Para contar con un índice que nos permita controlar este factor, es preciso sumar la proporción de ocurrencia de cada uno de los bigramas que componen una palabra en determinada lengua. Afortunadamente, el cálculo de este factor puede llevarse a cabo de manera expedita mediante el uso de herramientas disponibles en línea como EsPal (Duchon et al., 2013) y las diversas versiones de Subtlex (van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Preferiblemente, se debe optar por utilizar la medida de bigramas en posición específica, ya que como se mencionó previamente, hay evidencia que indica que la posición del bigrama dentro de la palabra tiene efectos específicos sobre los procesos de lectura (Muncer et al., 2014).

9. Número de sílabas

El número de sílabas de las palabras puede ser obtenido a través de su silabificación, la cual puede realizarse ya sea de forma automatizada o mediante el criterio humano. Para el primer método, existen programas computacionales como Wuggy (Keuleers & Brysbaert, 2010), el cual permite silabificar listas de palabras en diversos idiomas. Para el segundo método, se pueden utilizar jueces nativos del idioma que silabifiquen cada palabra y anoten el número de sílabas para cada una de ellas. Idealmente, ambos métodos deben llevarse a cabo de manera complementaria para asegurar la calidad del resultado.

10. Equiparación asistida de factores entre lenguas

Para poder lograr un control simultáneo de las dimensiones expuestas anteriormente, es necesario tabular las palabras con los valores numéricos de los factores de interés en una planilla de cálculo, y así pasar a la etapa de equiparación estadística entre lenguas. La lógica de esta fase consiste en seleccionar estímulos en ambas lenguas para formar dos conjuntos —uno por lengua— que sean equivalentes entre sí al aplicarse una prueba de distribución estadística.

Para conseguir esta equivalencia, puede optarse por un abordaje de escogencia de estímulos por adición uno a uno, siempre tratando de que cada nuevo estímulo que se agregue a un conjunto tenga su contraparte equivalente en la otra lengua. Dado que es raro encontrar palabras en idiomas distintos que sean equivalentes entre sí para todos sus factores, una estrategia que facilita esta fase es iniciar mediante la construcción de conjuntos de palabras que se hallen controlados por el factor numérico que posea más varianza (usualmente la frecuencia léxica).

Una vez que se cuenta con estos conjuntos de base, se prosigue a controlar el segundo factor con más varianza mediante la substitución de estímulos con otros que permitan equilibrar este segundo factor. Este proceso puede repetirse n veces hasta controlar todos los factores de nuestro interés. Como es de esperarse, cada uno de los factores adicionales que se desee equilibrar supone una dificultad mayor para la equiparación en ambas lenguas, dado que a medida que se reemplazan estímulos para equilibrar uno de los factores, paralelamente debe velarse por que el resto de los factores mantenga su equilibrio.

Con el objetivo de brindar una implementación del algoritmo que permita ejemplificar la fase de equiparación de factores y servir como una herramienta auxiliar en el proceso de selección y construcción de estímulos, hemos construido una base bilingüe (español e inglés) con 320 estímulos (160 por lengua) los cuales se encuentran controlados por los distintos factores reseñados en el artículo. En la tabla 2, mostramos un extracto de la implementación del algoritmo disponible en <https://tinyurl.com/StimDatabase>.

Como se puede observar en la Tabla 2, cada estímulo (Palabra) se encuentra tabulado junto con cada uno de sus respectivos valores para los factores de frecuencia léxica (por millón), número de vecinos ortográficos (por substitución), suma de bigramas en

posición específica y número de sílabas. En el medio de la tabla se encuentran diferentes pruebas *t* de Student para comprobar la equivalencia de factores entre ambos conjuntos. Retomando lo explicado anteriormente, la lógica del trabajo de equiparación se basa en la búsqueda de equivalencia de factores para conjuntos de estímulos en idiomas diferentes.

Dada la adaptabilidad de la implementación del algoritmo que ofrecemos, es posible construir conjuntos equivalentes de estímulos en dos (o más) lenguas, los cuales se hallen controlados por n cantidad de factores de interés.

Como nota adicional, es importante advertir que en la tabla 2 mantuvimos únicamente aquellos factores discutidos en el artículo; no obstante, en la versión en línea las palabras también se hallan controladas por número de caracteres y valencia y excitación emocional, factores que tienen efectos específicos en el tratamiento de la información escrita (Kramer & McLean, 2019; Rodríguez-Ferreiro & Davies, 2019). Además, la versión en línea contiene la distribución de los factores por cuartil y por lengua, lo que permite un control más fino y una guía más clara durante la labor de sustitución y equiparación de estímulos

Tabla 2

Extracto de la tabla de selección de estímulos disponible en <https://tinyurl.com/StimDatabase>.

<i>Inglés</i>					Prueba t de Student	<i>Español</i>				
Palabra	Frecuencia	Vecinos	Bigramas	Sílabas		Palabra	Frecuencia	Vecinos	Bigramas	Sílabas
anecdote	1,0784	0	0,0469	3	Frecuencia 0,935904052	cilindro	1,0817	0	0,0499	3
asteroid	1,6863	0	0,0596	3		diagrama	1,6587	0	0,0699	3
bouncer	1,5882	2	0,0776	2	Vecinos 1	secador	1,7788	2	0,0823	3
anomaly	1,2353	0	0,0318	4	Bigramas 0,960907926	alfiler	1,2019	0	0,0431	3
iodine	1,2549	0	0,0369	3		frijol	1,2019	0	0,0156	2
anthem	1,8824	0	0,0325	2	Sílabas 1	hamaca	1,7788	0	0,0389	3
peril	1,9216	0	0,0405	2		bucle	1,851	0	0,0184	2
lotto	1,6667	1	0,0288	2		farol	1,6587	1	0,0405	2

Nota. A la izquierda, se presenta una serie de palabras en inglés con los valores numéricos de cuatro factores de interés: Frecuencia léxica, Vecinos ortográficos, Bigramas, Número de sílabas. A la derecha se encuentra, una serie de palabras en español con los valores numéricos de los mismos factores. En el centro, se reportan los niveles de significancia de los T-tests comparando los valores respectivos de los cuatro factores de interés (Frecuencia léxica, Vecinos ortográficos, Bigramas, Número de sílabas) entre los dos idiomas presentados aquí (inglés y español); el objetivo siendo intentar mantener un nivel de significancia lo más cercano a 1

11. Revisión manual del conjunto restante de palabras

Una vez que se hayan completado las fases previas, es necesario realizar una minuciosa revisión manual de las palabras y valores resultantes para asegurarse de que cumplen con las características deseadas. Básicamente, esta revisión consiste en que una persona externa que no haya manipulado los estímulos tome el conjunto resultante de las fases anteriores y calcule cada uno de los factores por cuenta propia. Luego, debe tabular los estímulos junto con los factores y contrastar sus resultados con los obtenidos originalmente. Idealmente, esta revisión final debe llevarse a cabo más de una vez y por distintas personas, ya que es una garantía final de la calidad del trabajo de las fases previas.

Conclusión

Al constituir la base del planteamiento del protocolo experimental, el control de factores en paradigmas psicolingüísticos aplicados al bilingüismo es una de las etapas más críticas del proceso investigativo, en especial dado que el rigor de este control incide directamente en la calidad del estudio (Guasch et al., 2013). Por ello, no es extraño que para finiquitar esta etapa de manera adecuada sea necesario invertir una cantidad de tiempo y esfuerzo considerable para asegurar la calidad de los resultados.

Con el fin de ofrecer una guía clara y concisa para facilitar este proceso, hemos detallado una metodología de control y selección estímulos la cual consideramos puede ser de gran utilidad para quienes realizan estudios en el ámbito de la psicolingüística ya que permite adaptarse e implementarse fácilmente según las necesidades investigativas del fenómeno en estudio. Hemos de advertir que nuestra propuesta no invalida las metodologías tradicionalmente utilizadas para el control de estímulos, ; no obstante, consideramos que posibilita que los investigadores tengan en consideración un conglomerado de factores cuyos efectos sobre el procesamiento del lenguaje escrito son ampliamente conocidos y robustos.

Así, el aumentar el rigor en la creación de estímulos brinda mayor validez a los resultados experimentales, lo que a su vez da mayor fiabilidad a los datos que servirán para

sostener las hipótesis planteadas. Por supuesto, no es posible obviar el hecho de que existen controles complementarios para robustecer la metodología aquí planteada, ; por ejemplo, que el cálculo de cierta norma de nuestro interés (como el caso de la valencia emocional) se realice a partir de una evaluación llevada a cabo por jueces locales del contexto en el cual se desarrollará el estudio. Esta consideración deja abiertas las puertas a futuros trabajos que deseen retomar nuestra metodología con el objetivo de mejorarla.

Referencias

- Ansaldo, A. I., Marcotte, K., Scherer, L., & Raboyeau, G. (2008). Language therapy and bilingual aphasia: Clinical implications of psycholinguistic and neuroimaging research. *Journal of Neurolinguistics*, 21(6), 539–557.
<https://doi.org/10.1016/j.jneuroling.2008.02.001>
- Basnight-Brown, D. M. (2014). Models of Lexical Access and Bilingualism. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of Bilingual Memory* (pp. 85–107). New York: Springer. https://doi.org/10.1007/978-1-4614-9218-4_5
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., ... Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 15(3), 616–629. <https://doi.org/10.1017/S1366728912000090>
- Bernardy, J.-P., Lappin, S., & Lau, J. H. (2018). The Influence of Context on Sentence Acceptability Judgements. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 456–461.
<https://doi.org/10.18653/v1/P18-2073>
- Bontrager, T. (1991). The development of word frequency lists prior to the 1944 Thorndike-Lorge list. *Reading Psychology*, 12(2), 91–116.
<https://doi.org/10.1080/0270271910120201>
- Bosque, I. (2007). *Las Categorías Gramaticales. Relaciones & Diferencias*. Madrid:

Editorial Síntesis.

- Brito, N. H., Murphy, E. R., Vaidya, C., & Barr, R. (2016). Do bilingual advantages in attentional control influence memory encoding during a divided attention task? *Bilingualism: Language and Cognition*, *19*(3), 621–629. <https://doi.org/10.1017/S1366728915000851>
- Brybaert, M., Mander, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, *27*(1), 45–50. <https://doi.org/10.1177/0963721417727521>
- Brybaert, M., Mander, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, *51*(2), 467–479. <https://doi.org/10.3758/s13428-018-1077-9>
- Cambridge University Press. (2019). *Cambridge Dictionary*. Consultado en <https://dictionary.cambridge.org/es/diccionario/ingles-espanol/>
- Carrasco-Ortiz, H., Midgley, K. J., & Frenck-Mestre, C. (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones: Phonological representations in bilingual word recognition. *Psychophysiology*, *49*(4), 531–543. <https://doi.org/10.1111/j.1469-8986.2011.01333.x>
- Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling*, *43*(2), 70–83. <https://doi.org/10.1002/j.2161-1920.2006.tb00008.x>
- Cieślicka, A. B., Heredia, R. R., & Olivares, M. (2014). It's All in the Eyes: How Language Dominance, Salience, and Context Affect Eye Movements During Idiomatic Language Processing. In M. Pawlak & L. Aronin (Eds.), *Essential Topics in Applied Linguistics and Multilingualism* (pp. 21–41). New York: Springer International Publishing. https://doi.org/10.1007/978-3-319-01414-2_2
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal

- lexicon. En S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). New York: Academic Press.
- Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, *15*(5), 336–345.
<https://doi.org/10.1038/nrn3709>
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. <https://doi.org/10.1017/S1366728902003012>
- Drazer, M. (2018, 23 de abril). Bilingüismo: el enorme potencial de hablar más de una lengua. *Deutsche Welle*. Recuperado de <https://www.dw.com/es/biling%C3%BCismo-el-enorme-potencial-de-hablar-m%C3%A1s-de-una-lengua/a-43499721>
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246–1258. <https://doi.org/10.3758/s13428-013-0326-1>
- Durlik, J., Szewczyk, J., Muszyński, M., & Wodniecka, Z. (2016). Interference and Inhibition in Bilingual Language Comprehension: Evidence from Polish-English Interlingual Homographs. *PLOS ONE*, *11*(3), e0151430.
<https://doi.org/10.1371/journal.pone.0151430>
- Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*(4), 375–399. <https://doi.org/10.3758/BF03203275>
- Gimenes, M., & New, B. (2016). Worldlex: Twitter and blog word frequencies for 66 languages. *Behavior Research Methods*, *48*(3), 963–972.
<https://doi.org/10.3758/s13428-015-0621-0>
- Goldrick, M., Ferreira, V. S., Miozzo, M., Miceli, G., & Costa, V. (2014). The Role of

Lexical and Sublexical Orthography in Writing. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production*. Oxford: Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199735471.013.024>

Grosjean, F., Li, P., & Bialystok, E. (2013). *The psycholinguistics of bilingualism*. New Jersey: Wiley-Blackwell/John Wiley & Sons.

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, 45(3), 765–771. <https://doi.org/10.3758/s13428-012-0296-8>

Halsband, U. (2006). Bilingual and multilingual language processing. *Journal of Physiology-Paris*, 99(4–6), 355–369.

<https://doi.org/10.1016/j.jphysparis.2006.03.016>

Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking More or Feeling Less? Explaining the Foreign-Language Effect on Moral Judgment. *Psychological Science*, 28(10), 1387–1397.

<https://doi.org/10.1177/0956797617720944>

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition*, 13(3), 315–325.

<https://doi.org/10.1017/S1366728909990010>

Hillis, A. E., & Caramazza, A. (1995). Representation of Grammatical Categories of Words in the Brain. *Journal of Cognitive Neuroscience*, 7(3), 396–407.

<https://doi.org/10.1162/jocn.1995.7.3.396>

Hoshino, N., & Thierry, G. (2012). Do Spanish–English bilinguals have their fingers in two pies – or is it their toes? An electrophysiological investigation of semantic access in bilinguals. *Frontiers in Psychology*, 3.

<https://doi.org/10.3389/fpsyg.2012.00009>

- Katz, L., & Frost, R. (1992). Chapter 4 The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. *Advances in Psychology*, 94, 67–84. [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 1–6. <https://doi.org/10.1017/S1366728919000038>
- Keselman, A., Smith, C. A., Divita, G., Kim, H., Browne, A. C., Leroy, G., & Zeng-Treitler, Q. (2008). Consumer Health Concepts That Do Not Map to the UMLS: Where Do They Fit? *Journal of the American Medical Informatics Association*, 15(4), 496–505. <https://doi.org/10.1197/jamia.M2599>
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627–633. <https://doi.org/10.3758/BRM.42.3.627>
- Kramer, B., & McLean, S. (2019). L2 reading rate and word length: The necessity of character-based measurement. *Reading in a Foreign Language*, 31(2), 201–225
- Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25(5), 497–514. <https://doi.org/10.1080/20445911.2013.799170>
- Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two Languages in Mind: Bilingualism as a Tool to Investigate Language, Cognition, and the Brain. *Current Directions in Psychological Science*, 23(3), 159–163. <https://doi.org/10.1177/0963721414528511>
- Kunnan, A. J. (Ed.). (2013). *The Companion to Language Assessment*. <https://doi.org/10.1002/9781118411360>
- Langsford, S., Perfors, A., Hendrickson, A. T., Kennedy, L. A., & Navarro, D. J. (2018). Quantifying sentence acceptability measures: Reliability, bias, and variability. *Glossa: A Journal of General Linguistics*, 3(1), 37.

<https://doi.org/10.5334/gigl.396>

- Larraza, S., Samuel, A. G., & Oñederra, M. L. (2016). Listening to accented speech in a second language: First language and age of acquisition effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(11), 1774–1797. <https://doi.org/10.1037/xlm0000252>
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144(4), 394–425. <https://doi.org/10.1037/bul0000142>
- Li, P. (2013). Computational modeling of bilingualism: How can models tell us more about the bilingual mind? *Bilingualism: Language and Cognition*, 16(2), 241–245. <https://doi.org/10.1017/S1366728913000059>
- Luiselli, J., Happé, F., Hurst, H., Freeman, S., Goldstein, G., Mazefsky, C., Carter, A. S., Kaufman, A. S., Simmons, E. S., Eernisse, E. R., Tsatsanis, K., Fluit, F., Klein-Tasman, B., Feinstein, A., Kahl, A., Poyau, S., Jones, C. R. G., Volkmar, F. R., Wolf, J. M., ... Newman, D. B. (2013). Written Language. In F. R. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders* (pp. 3408–3409). New York: Springer. https://doi.org/10.1007/978-1-4419-1698-3_1125
- Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, 7(8), e43230. <https://doi.org/10.1371/journal.pone.0043230>
- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114(2), 187–202. <https://doi.org/10.1016/j.jecp.2012.09.002>
- Mufwene, S., Coupe, C., & Pellegrino, F. (Eds.). (2017). *Complexity in Language: Developmental and evolutionary perspectives*. Cambridge University Press. <https://doi.org/10.1017/9781107294264>

- Muncer, S. J., Knight, D., & Adams, J. W. (2014). Bigram Frequency, Number of Syllables and Morphemes and Their Effects on Lexical Decision and Word Naming. *Journal of Psycholinguistic Research*, 43(3), 241–254. <https://doi.org/10.1007/s10936-013-9252-8>
- Murray, W. S., & Forster, K. I. (2004). Serial Mechanisms in Lexical Access: The Rank Hypothesis. *Psychological Review*, 111(3), 721–756. <https://doi.org/10.1037/0033-295X.111.3.721>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516–524. <https://doi.org/10.3758/BF03195598>
- Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism Affords No General Cognitive Advantages: A Population Study of Executive Function in 11,000 People. *Psychological Science*, 31(5), 548–567. <https://doi.org/10.1177/0956797620903113>
- Pavlenko, A. (2010). Verbs of motion in L1 Russian of Russian–English bilinguals. *Bilingualism: Language and Cognition*, 13(1), 49–62. <https://doi.org/10.1017/S1366728909990198>
- Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., & Fazio, F. (1999). The neural correlates of verb and noun processing. *Brain*, 122(12), 2337–2344. <https://doi.org/10.1093/brain/122.12.2337>
- Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicológica*, 21(3), 327–340.
- Rayson, P., & Garside, R. (1998). The claws web tagger. *ICAME Journal*, 22, 121–123.
- Real Academia Española. (2014). *Diccionario de la lengua española* (23.a ed.). Consultado en <https://dle.rae.es/>
- Rodríguez-Ferreiro, J., & Davies, R. (2019). The graded effect of valence on word recognition in Spanish. *Journal of Experimental Psychology: Learning, Memory,*

- and Cognition, 45(5), 851–868. <https://doi.org/10.1037/xlm0000616>
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*, 47(4), 1085–1094. <https://doi.org/10.3758/s13428-014-0528-1>
- Segalowitz, N., & Hébert, M. (1990). Phonological Recoding in the First and Second Language Reading of Skilled Bilinguals. *Language Learning*, 40(4), 503–538. <https://doi.org/10.1111/j.1467-1770.1990.tb00604.x>
- Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin–English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition*, 17(2), 364–383. <https://doi.org/10.1017/S1366728913000424>
- Sönmez, H. (2019). Review of studies focused on bilingualism. *Journal of Language and Linguistic Studies*, 15(3), 1045–1068. <https://doi.org/10.17263/jlls.631548>
- Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*, 49(1), 111–123. <https://doi.org/10.3758/s13428-015-0700-2>
- Syllable. (2020). En *Apa Dictionary of Psychology*. Recuperado de <https://dictionary.apa.org/syllable>
- Tranel, D., Martin, C., Damasio, H., Grabowski, T., & Hichwa, R. (2005). Effects of noun-verb homonymy on the neural correlates of naming concrete entities and actions. *Brain and Language*, 92(3), 288–299. <https://doi.org/10.1016/j.bandl.2004.01.011>
- Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential Measures Can Be Used as a Proxy for Language Dominance in Bilingual Language Acquisition Research. *Frontiers in Psychology*, 9, 1809. <https://doi.org/10.3389/fpsyg.2018.01809>
- van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, 39(3),

458–483. <https://doi.org/10.1006/jmla.1998.2584>

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

<https://doi.org/10.1080/17470218.2013.850521>

van Kesteren, R., Dijkstra, T., & de Smedt, K. (2012). Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *Quarterly Journal of Experimental Psychology*, *65*(11), 2129–2154.

<https://doi.org/10.1080/17470218.2012.679946>

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. <https://doi.org/10.3758/s13428-012-0314-x>

Weide, R. L. (1998). The CMU pronouncing dictionary. URL:

<http://www.speech.cs.cmu.edu/cgi-bin/cmudict>

Whitford, V., & Titone, D. (2016). Eye movements and the perceptual span during first- and second-language sentence reading in bilingual older adults. *Psychology and Aging*, *31*(1), 58–70. <https://doi.org/10.1037/a0039971>

Zhang, S., & Thompson, N. (2004). DIALANG: A Diagnostic Language Assessment System (review). *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, *61*(2), 290–293. <https://doi.org/10.1353/cml.2005.0011>