

Año 41, Número 77, julio de 2023

ISSN 0185-2647



ELA ESTUDIOS DE LINGÜÍSTICA APLICADA



Universidad Nacional Autónoma de México

ELA ESTUDIOS DE
LINGÜÍSTICA APLICADA

Año 41 ▶ Número 77 ▶ julio de 2023

Universidad Nacional Autónoma de México
Escuela Nacional de Lenguas, Lingüística y Traducción

ARTÍCULOS ►

- Aproximación al análisis prosódico del foco contrastivo en hablantes preescolares y adultos del español mexicano: datos de producción y percepción 7
An approach to the prosodic analysis of contrastive focus of preschool and adult Mexican Spanish speakers: data from production and perception
Brenda Muñoz Martínez, Natalia Arias Trejo & Erika Mendoza Vázquez
- La activación de la información fonológica en aprendientes de segundas lenguas por medio de imágenes 49
Activation of the phonological information through images in second language learners
Alma Luz Rodríguez Lázaro & Natalia Arias Trejo
- Can foreign language learners discover the rules of non-salient forms? 85
¿Pueden los aprendices de una lengua extranjera descubrir las reglas de formas lingüísticas no sobresalientes?
José Luis Moreno-Vega & Ana Mónica Preciado-Sánchez
- Metáfora enactiva en la música didáctica virtual: una mirada crítica decolonial 119
Enactive metaphor in a virtual didactic musical performance: A decolonial critical look
Luis Escobar L.-Dellamary & Luis Quintero Castro

Formas apelativas de segunda persona en
entrevistas político-presidenciales mexicanas 157
Address forms in Mexican political-presidential
interviews
Melanie del Carmen Salgado López

Lineamientos para publicar en *Estudios de*
Lingüística Aplicada 189



Artículos

Aproximación al análisis prosódico del foco contrastivo en hablantes preescolares y adultos del español mexicano: datos de producción y percepción

An approach to the prosodic analysis of contrastive focus of preschool and adult Mexican Spanish speakers: data from production and perception

Brenda Muñoz Martínez
Universidad Nacional Autónoma de México, Facultad de Psicología,
Laboratorio de Psicolingüística
brenda.mumar@comunidad.unam.mx

Natalia Arias Trejo
Universidad Nacional Autónoma de México, Facultad de Psicología,
Laboratorio de Psicolingüística
nariast@unam.mx

Erika Mendoza Vázquez
Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas
emvazquez03@gmail.com

Recepción: 3 de agosto del 2022
Aceptación: 21 de marzo del 2023
doi: 10.22201/enallt.01852647p.2023.77.1038

Resumen

El foco suele marcarse al final de los enunciados en español, en especial en su función informativa; sin embargo, puede colocarse en otras posiciones con la función de contrastar o corregir información. Estudios de producción han señalado que niños hablantes del español mexicano en ocasiones marcan el foco al inicio de oraciones usando pistas prosódicas distintas a los adultos. Con base en datos de producción, el presente estudio evalúa la comprensión del foco al inicio y final de oraciones sintácticamente idénticas por parte de niños preescolares hablantes del español mexicano, en comparación con hablantes adultos. Para este fin, se analizan acústicamente las respuestas a enunciados con énfasis prosódico en el sujeto (inicio) y atributo locativo (final), en las cuales los participantes corrigen el elemento focalizado al escuchar descripciones erróneas sobre una serie de imágenes. Se espera que ambas poblaciones usen rasgos acústicos diferentes para marcar el foco dependiendo de su percepción de la posición de este en las oraciones. Los resultados mostraron que tanto niños como adultos percibían el foco en ambas posiciones, pero solo los niños enfatizaban más el sujeto.

Palabras clave: prosodia; adquisición del lenguaje; comprensión; focalización; análisis acústico

Abstract

Focus is usually marked at the end of phrases in Spanish, especially in its informative function. However, it can be placed in other positions of the utterance for contrasting or correcting information. Language production studies have shown that Mexican Spanish-speaking children sometimes set the prosodic focus at the beginning of sentences using prosodic cues different from those employed by adults. Based on production data, the present study evaluates the comprehension of the prosodic focus at the beginning and at the end of syntactically identical sentences by preschool Mexican Spanish-speaking children compared with adult speakers. Responses to utterances with prosodic prominence in the subject (at the beginning) and in the prepositional predicative (at the end) are acoustically analyzed, where participants correct the focalized element upon hearing wrong descriptions of a series of images. Both adults and children are expected to use different acoustic features to set the focus, depending on their perception of its position in sentences. The results showed that children and adults perceived the prosodic focus in both positions, but only children give more emphasis on the subject.

Keywords: prosody; language acquisition; comprehension; focalization; acoustic analysis

1. Introducción

La prosodia se refiere a elementos como la entonación, las pausas, el ritmo y el volumen de las frases habladas (Pynte, 1998: 79). Estudios diversos han demostrado que los elementos prosódicos son la base de la adquisición del lenguaje, ya que los infantes empiezan a distinguirlos antes de nacer (Shahidullah & Hepper, 1994; Fifer & Moon, 1994; Nazzi, Bertoncini & Mehler, 1998; Kisilevsky, Hains, Lee, Xie, Huang, Ye, Zhang & Wang, 2003; Kisilevsky, Hains, Brown, Lee, Cowperthwaite, Stutzman, Swansburg, Lee, Xie, Huang, Ye, Zhang & Wang, 2009) y, en sus primeros años, les ayudan a identificar las palabras del flujo del habla y a mapear sus significados (Johnson & Jusczyk, 2001; Goswami, 2008).

Un elemento importante para el proceso de adquisición del lenguaje es el foco, usualmente definido como la información nueva o relevante en un contexto determinado (Vallduví & Engdahl, 1996: 461). Esta definición alude a una conversación en la que el hablante tiene una idea de lo que el oyente sabe y, por lo tanto, enfatiza la información que el oyente no conoce o no toma en cuenta (Dorta Luis, 2008: 109). Un ejemplo son las siguientes oraciones, obtenidas de Mora Bustos (2010: 220):

- (1) a. Pedro compró el PERIÓDICO.
- b. El periódico fue comprado por PEDRO.

En este caso, (1a) también podría interpretarse como un foco amplio; es decir, toda la oración sería el foco, y podría responder a la pregunta, real o hipotética, “¿qué pasó?”; en contraste con su interpretación como foco estrecho, que funcionaría si, en el contexto de la conversación, el foco se centra en un elemento del enunciado

* Agradecemos al Consejo Nacional de Ciencia y Tecnología (Conacyt) por la beca (clave 2019-000037-02NACF) otorgada durante la realización de este proyecto (CVU 744095).

respondiendo a una pregunta específica, como fue ejemplificado al principio de este párrafo (Lambrecht, 1994: 17).

En contraste, diversos autores han definido al foco en términos de la estructura de la información. Así, Chafe (1974) lo caracteriza como un elemento que el hablante usa para modificar la información que el oyente conoce. Otras definiciones hacen referencia a la porción del enunciado que no se presupone por el hablante (Lambrecht, 1994: 207); la parte más informativa de la oración que relaciona el enunciado con el contexto y con condiciones de verdad (Reinhart, 2006: 1); o lo que se predica del tópico o tema de la oración (Gundel & Fretheim, 2004: 176). Asimismo, Krifka (2008: 247) lo define como el elemento del enunciado que señala la presencia de alternativas relevantes para interpretarlo y Dik (1997: 326), como la información más importante, según el hablante, de una expresión lingüística en un contexto comunicativo que se desea que el oyente integre a su información pragmática añadiendo o sustituyendo información. Esta última definición es la que se retomará en este estudio.

De acuerdo con la clasificación de tipos de foco de Dik (1997: 332–334), el foco en (1) constituye un tipo de foco informativo completivo, ya que añade información que el oyente desconocía. Sin embargo, el foco puede también usarse para sustituir información que tiene el oyente. Este caso refiere al foco contrastivo correctivo o de reemplazo (*replacement focus* en Dik, 1997: 333). Por ejemplo, el foco en (1b) corregiría a alguien que indicara erróneamente que Pablo compró el periódico, reemplazando “Pablo” por “Pedro” en la conversación.

Como se mencionó anteriormente, en español la palabra focalizada suele colocarse al final de las oraciones, pues se trata de una lengua de orden libre (Gutiérrez Bravo, 2008: 365). Esto se observa en los ejemplos de (1), en los cuales suele coincidir con el acento nuclear del enunciado.¹ Asimismo, los hablantes pueden

¹ Hualde (2014: 263) señala que “el último acento de la frase entonativa se percibe como portador de mayor prominencia que los precedentes”, y refiere al

utilizar elementos prosódicos para marcar el foco sin que este se encuentre necesariamente al final de las oraciones. De esta manera, adultos y niños utilizan elevaciones del tono de la voz y la duración vocálica para marcar foco contrastivo al inicio o final de enunciados (Villalobos Pedroza, 2021).

En cuanto a la adquisición del lenguaje, se ha demostrado que niños de dos años hablantes del inglés interpretan la prominencia prosódica al final de las oraciones como la información importante en una conversación, lo cual les ayuda a aprender nuevas palabras (Grassmann & Tomasello, 2007, 2010). En el español de la Ciudad de México, se ha reportado que los adultos suelen enfatizar la palabra que desean que el niño preste atención colocándola al final de las oraciones, guiando su comprensión del discurso (Villalobos Pedroza, 2019). Al respecto, es posible cuestionar: ¿qué sucede en contextos en los que el foco no se localiza al final de las frases? ¿Los niños siguen siendo capaces de identificarlo y entender funciones comunicativas más allá de información nueva, como cuando se contrasta información?

2. Antecedentes

Estudios previos con niños en lenguas como el inglés, francés, alemán y danés, cuyo foco suele marcarse al final de la frase, han llegado a diferentes conclusiones al medir la comprensión del foco en otras posiciones de la oración. Estos han establecido que los infantes lo producen desde los dos años, pero son incapaces de comprenderlo hasta los cinco años o más (Hornby & Hass, 1970; Hornby, 1971; Wieman, 1976; MacWhinney & Bates, 1978; Furrrow, 1984; Cruttenden, 1985; Cutler & Swinney, 1987), lo cual supondría una asimetría en la adquisición de la lengua: los elementos del lenguaje primero se comprenden y luego se producen

papel que tiene una mayor duración de la sílaba nuclear en su prominencia, ya que no siempre constituye la sílaba con mayores valores de F0.

(Hendriks & Koster, 2010: 1887). Sin embargo, dichos resultados pueden deberse a aspectos metodológicos.

Algunos estudios del inglés, cuyas tareas consistían en que los infantes utilizaran objetos para recrear frases escuchadas con foco prosódico en pronombres, partían de diseños experimentales que añadían dificultad a la tarea (Solan, 1980; McDaniel & Maxfield, 1992), ya que requerían procesos cognitivos extra en la identificación de cada objeto con su referente (Chen, 2010: 1929). Otros estudios utilizaron tareas en las que el niño debía elegir la imagen que mejor representara un enunciado escuchado (Hornby, 1971; Cruttenden, 1985); en estas, los niños siempre interpretaban el foco al final de las oraciones, ya que prestaban mayor atención al significado semántico y no a los cambios prosódicos (Chen, 2010: 1927).

La prosodia tampoco ha resultado un factor relevante para resolver tareas en que se debía indicar si una oración era correcta o no. Un ejemplo es el trabajo de Lahey (1974), que presentó, a niños de entre tres y cinco años, oraciones con y sin entonación (usando un sintetizador de voz para lograrlo), sin obtener diferencias entre condiciones. Esto fue interpretado como una incapacidad de los niños en edad preescolar para identificar cambios prosódicos, aunque estos pudieron no tener relevancia en la tarea. Este también fue el caso del estudio de Cutler y Swinney (1987), en el cual los niños debían presionar un botón al oír las palabras focalizadas en oraciones gramaticales y agramaticales. Los resultados mostraron que los niños menores de cinco años tuvieron mayores tiempos de reacción en las oraciones gramaticales, indicando que el significado es una pista a la que se le prestó mayor atención que a la prosodia en este tipo de tareas y, por lo tanto, identificar el elemento enfatizado llevaba más tiempo.

Por este motivo, otros estudios han intentado desarrollar estrategias en las que la prosodia sea un elemento clave para resolver tareas experimentales. Szendrői, Bernard, Berger, Gervain y Höhle (2018) crearon una tarea que consistía en lo siguiente: el investigador decía una oración sintácticamente idéntica con énfasis

sis en el sujeto u objeto que describía erróneamente una imagen, posteriormente, el niño debía corregir el sujeto u objeto dependiendo de dónde se encontraba el énfasis. Los autores concluyeron que los niños de tres a seis años, hablantes del inglés, francés y alemán, podían comprender e identificar el foco tanto al principio como al final de las oraciones. El éxito de la tarea se debió a que integraba una situación experimental en un contexto natural de conversación en la que el foco era relevante, ya que de eso dependía que los niños corrigieran al investigador.

Cabe destacar que algunos de los estudios anteriores incluyeron en sus experimentos un grupo de control integrado por población adulta (Cruttenden, 1985; McDaniel & Maxfield, 1992; Szendrői *et al.*, 2018). En todos los casos, los adultos tuvieron una mayor cantidad de aciertos al identificar el foco en diferentes posiciones del enunciado, es decir, arriba del azar. Sin embargo, en el caso de Szendrői *et al.* (2018), tanto los adultos como los niños hablantes del francés obtuvieron un menor número de respuestas correctas, en comparación con los demás idiomas (alemán e inglés), cuando el foco era marcado con prominencia tonal al principio de las oraciones, ya que, al igual que el español, esta lengua también puede cambiar el orden de los elementos del enunciado para colocar la palabra enfatizada al final. A pesar de eso, los adultos tuvieron una mayor cantidad de respuestas correctas que los niños y, en ambas poblaciones, se comprobó su comprensión del foco en diferentes posiciones del enunciado.

Según nuestro conocimiento, la comparación de la comprensión del foco al inicio o final de enunciados por parte de adultos y niños hablantes del español mexicano se ha hecho únicamente en estudios de producción. Es el caso de Villalobos Pedroza (2021), en el cual niños de 23 a 28 meses fueron capaces de usar el foco contrastivo al principio y final de oraciones, tal como se ha visto en investigaciones previas de otros idiomas. En este estudio los niños utilizaron estrategias distintas a los adultos para marcar énfasis, favoreciendo la duración silábica sobre una mayor prominencia de la F0 y una mayor excursión tonal, lo que es preferido por los

adultos, sin embargo, ambas poblaciones produjeron de la misma forma el pico de la curva melódica en medio de la sílaba tónica (identificado como prealineamiento). Asimismo, tanto niños como adultos produjeron el foco en posición no final, lo cual sugería que eran capaces de comprenderlo. Siguiendo esta línea, nuestro estudio busca comparar la comprensión del foco al principio o al final de oraciones por parte de niños preescolares y adultos hablantes del español del centro de México, con el fin de encontrar diferencias en su interpretación de los enunciados.

Se tomó como base el estudio de Szendrői *et al.* (2018), que se adaptó al español mexicano. Adicionalmente, se empleó un método enfocado en el habla para analizar la comprensión del foco prosódico: el análisis acústico. Esto deriva de hallazgos en estudios previos, los cuales empleaban métodos como el rastreo visual o el análisis de tiempos de respuesta; estos mostraban que, aunque los niños marcaban el foco al final de las oraciones, eran capaces de identificarlo en otras posiciones (Chen, 2010; Höhle, Fritzsche & Müller, 2016). Con esta metodología, se esperó encontrar un patrón similar al de los estudios antes mencionados, según el cual los niños colocarían el foco al final, pero marcarían diferencias prosódicas dependiendo de qué constituyente de la oración interpretan como focalizado. Asimismo, este método podría develar la comprensión de lo escuchado a partir del análisis de las respuestas de los participantes en el contexto natural de una conversación.

Tomando en cuenta lo anterior, las preguntas que guían esta investigación son: ¿los niños en edad preescolar (4–5 años) hablantes del español del centro de México pueden distinguir la prominencia prosódica según aparezca al principio o al final de oraciones declarativas idénticas? Si es así, ¿cómo se implementa el foco contrastivo (el foco correctivo en términos de Dik, 1997) en una situación de controversia en el constituyente sujeto y en el constituyente de atributo locativo (en adelante, llamado atributo)? Aunado a esto, ¿los adultos presentarán la misma interpretación e implementación del foco contrastivo? Se espera que los niños y

adultos marquen de manera distinta el foco contrastivo en el sujeto o atributo de la oración, dependiendo de su percepción del foco en oraciones declarativas, sintácticamente idénticas, con prominencia prosódica al principio o al final; tal como lo muestran estudios de producción (Villalobos Pedroza, 2021). Además, se espera que ambos muestren una comprensión del foco al inicio o al final de los enunciados (Szendrői *et al.*, 2018).

3. Método

3.1. Muestra

Se evaluó a ocho niños (tres mujeres y cinco hombres) de cuatro a cinco años de edad ($\bar{x} = 5;2$, $DE = 0.36$) y a 19 adultos (15 mujeres y cuatro hombres) de entre 24 y 40 años ($\bar{x} = 31;6$, $SD = 8.8$) del centro y centro-occidente de México. Los niños se repartieron aleatoriamente en dos grupos: cuatro en la Condición Atributo (CA), cuyo énfasis se encontraba en el atributo de la oración, y cuatro en la Condición Sujeto (CS), que focalizaba el sujeto. Los adultos también se repartieron de manera aleatoria en ambas condiciones (CA y CS), sin embargo, debido a un problema con la plataforma, 11 fueron asignados a la CA y ocho a la CS.

Los participantes fueron contactados mediante las redes sociales del Laboratorio de Psicolingüística de la Universidad Nacional Autónoma de México (UNAM), y tenían acceso a una computadora con cámara. La vía remota fue la opción más viable debido a la pandemia de COVID-19. Los participantes, sus padres o tutores, confirmaron que no tenían problemas auditivos, de visión, cognitivos o de lenguaje en una encuesta inicial.

3.2. Diseño de estímulos experimentales

En el experimento, basado en el propuesto por Szendrői *et al.* (2018) y adaptado a hablantes mexicanos, se mostraba una serie de imágenes, cuya posición en la pantalla era descrita errónea-

mente por el experimentador, quien señalaba que el elemento se encontraba arriba o abajo de un elemento equivocado, usando un mayor énfasis prosódico al principio (en el sujeto) o final (en el atributo) de los enunciados. Se esperaba que los participantes niños y adultos, dependiendo de su percepción de dicha prominencia, usaran el foco contrastivo para corregir el elemento focalizado de las oraciones escuchadas (sujeto en la CS y atributo en la CA). De este modo, aunque se medía la producción como respuesta a los estímulos escuchados, la prominencia prosódica en dichas respuestas dependía de la posición en la que los participantes percibían el énfasis (sujeto o atributo), por lo que se podía evaluar también su percepción del foco en un elemento u otro al momento en el que los participantes determinaran qué se debía corregir.

Para este fin, se utilizaron enunciados distintos a los empleados por Szendrői *et al.* (2018), quienes usaron enunciados transitivos (como “El puercoespín tiene una botella”) en un diseño en el que un objeto animado estaba a un lado de uno inanimado para mostrar posesión (como el puercoespín junto a una botella en el ejemplo anterior). En el caso de esta investigación, se prefirió el uso de oraciones que describieran de manera simple las imágenes (como “La araña está arriba del búho”). Se tomó dicha decisión ya que el presente estudio forma parte de uno más extenso que involucra una tarea con rastreador visual, por lo que la distribución de las imágenes debía realizarse de manera equitativa en la pantalla para determinar áreas de interés. Por este motivo, la relación entre las imágenes no sería clara para usar oraciones transitivas, ya que están separadas una de la otra. Del mismo modo, los objetos animados e inanimados se presentaron por separado, dificultando aún más el uso de verbos transitivos como “tener”. Por lo tanto, se consideró simplificar la tarea, de manera que los niños hicieran una descripción de lo que veían en la pantalla (elementos arriba o abajo de otros).

Los estímulos consistieron en imágenes de seis objetos animados o inanimados mostrados simultáneamente en pantalla, acom-

pañadas de enunciados que hacían referencia a ellas. Las palabras que referían a las imágenes son conocidas por niños hablantes del español mexicano de al menos 30 meses de edad, según el inventario MacArthur-Bates (Jackson-Maldonado, Thal, Marchman, Newton, Fenson & Conboy, 2003; Weisleder & Fernald, 2013). Estas palabras representaban el sujeto y el atributo de los enunciados, y contaban mínimamente con dos sílabas, además, se trataba de palabras graves o esdrújulas, lo cual servía para medir la frecuencia fundamental entre una sílaba y otra; asimismo, contaban solo con consonantes sonoras para evitar una interrupción en la curva melódica (Vanrell Bosch, Stella, Gili Favela & Prieto, 2013). Las palabras, compiladas del Wordbank (<http://wordbank.stanford.edu>), fueron las siguientes: araña [a.'ra.ɲa], hormiga [or.'mi.ɣa], rana [ra.na], gallina [ga.'ji.na], búho [bu.o] y ardilla [ar.'ði.ja], en el caso de los objetos animados; y mano [ma.no], huevo [we.βo], uva [u.βa], globo [glo.βo], llave [ja.βe] y libro [li.βro], en el de los objetos inanimados.

Las imágenes en blanco y negro que representan estas 12 palabras fueron seleccionadas de la base de datos de Snodgrass y Vanderwart (1980), con excepción de la imagen del huevo, que se obtuvo de una base gratuita (<https://www.istockphoto.com>). En cuanto a sus características, ninguna imagen resaltaba más que otra en color o tamaño. Las imágenes, según representaran objetos animados o inanimados, se presentaron por separado para eliminar el sesgo en el que el sujeto suele ser un objeto animado (Szendrői *et al.*, 2018).

Las imágenes se repetían tres veces en seis diferentes posiciones de la pantalla (derecha, centro o izquierda; arriba o abajo). Cada una se centró en una de las seis posiciones y ninguna apareció al lado de la misma imagen en la misma posición. Además, dos pares eran comparados mientras que uno era el distractor. Esto resultó en 12 combinaciones en total. Un ejemplo de estímulo se presenta en la Figura 1.

Cada combinación de imágenes se acompañó de tres diferentes enunciados. Para esto, se utilizaron oraciones declarativas con

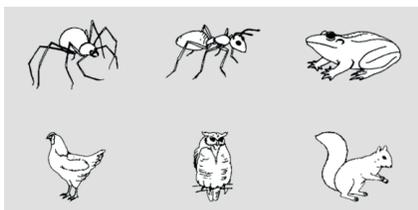


FIGURA 1. Ejemplo de estímulo con imágenes de objetos animados

el orden Sujeto, Verbo y Atributo, en el que la prominencia podía caer en el inicio del enunciado (sujeto) o en el final (atributo), con el fin de contrastar la comprensión del foco en dos elementos opuestos de la oración. Estos enunciados podían ser estímulos experimentales, en los cuales se decía algo erróneo de las imágenes y que el participante debía corregir (en la Figura 1, “La ARAÑA está arriba de la ardilla, ¿verdad?”, según la CS, o, “La araña está arriba de la ARDILLA, ¿verdad?”, según la CA); estímulos control, en los que se decía algo correcto sobre las imágenes (siguiendo el ejemplo de la Figura 1, “La ARAÑA/araña está arriba de la GALLINA/gallina, ¿verdad?”, con énfasis según la condición); o estímulos de relleno, que no presentaban foco contrastivo en ningún elemento, de los cuales una mitad describía correctamente las imágenes y la otra incorrectamente (en este caso, quien grabó los estímulos leyó las oraciones sin contexto, a modo de lista). En total, se presentaron 12 de cada tipo, con un total de 36 ensayos experimentales divididos en cuatro bloques.

3.3. Criterios de análisis fonético

Los criterios previamente estipulados se tomaron en cuenta tanto para el análisis y creación de estímulos auditivos como para el posterior análisis de datos. Se compararon las palabras objetivo, el sujeto y el atributo, en ambas condiciones, CS y CA. De dichas palabras, se consideró la tonía y la duración, al ser pistas relevantes en la marcación del foco en español (Villalobos Pedroza, 2021).

Para el análisis prosódico se usó el modelo métrico autosegmental (Pierrehumbert, 1980; Ladd, 2008). Según Hualde (2003: 157), los componentes básicos de este modelo son los tonos de juntura, asociados con los lindes de los enunciados, así como los acentos tonales, asociados en español a las sílabas tónicas (para una descripción de los tipos de acentos y tonos de juntura analizados, véase Anexo B). Dichas sílabas tónicas se distinguen del resto por una mayor duración vocálica, intensidad (volumen de la voz) o tono (frecuencia en la que vibran las cuerdas vocales) (Face, 2002: 31). En el nivel post-léxico, y en función del contexto comunicativo, en esta sílaba recae el énfasis de la palabra focalizada.

El modelo métrico autosegmental fue la base para el desarrollo del sistema de transcripción ToBI (Tones and Breaks Indices); para el español la primera versión Sp_ToBI fue propuesta por Beckman, Díaz Campos, Tevis McGory y Morgan (2002). Para este trabajo se empleó la versión de Hualde y Prieto (2015), la cual ha sido ampliamente utilizada en el análisis prosódico y también en estudios de producción con niños hablantes del español mexicano (Villalobos Pedroza, 2021). Del mismo modo, se utilizaron las notaciones establecidas por Sun-Ah Jun (2005), en las que el símbolo > indica prealineamiento, mientras que <, el pico del acento tonal en la sílaba postónica, esto en acentos bitonales. Con base en Martín Butragueño (2019), se utiliza el diacrítico de *upstep* (j) para representar un ascenso mayor a tres semitonos (st); además se utiliza el de *downstep* (!) cuando el tono tiene un nivel menor de 3 st del resto del enunciado. En esta primera aproximación se siguió un criterio cualitativo para la marcación de prealineamiento.

Asimismo, se marcaron los niveles de separación prosódica utilizando las convenciones de transcripción Sp_ToBI, las cuales, a su vez, equivalen a los niveles de la jerarquía prosódica de Nespor y Vogel (1994): el nivel 0, el cual indica que dos palabras forman una sola palabra fonológica cuando una de ellas no tiene acento y puede unirse a la siguiente (por ejemplo, “la” y “mano”, formando [la'mano]); el 1, que marca el límite entre palabras fonológicas (por ejemplo, [lamano]₁[es'ta]); el 2, que señala cuando se pro-

duce un descenso abrupto, percibido como una cesura dentro del enunciado, menor a la de una frase intermedia; el 3, que indica una pausa dentro del enunciado (por ejemplo, la que se representa con coma en “Entonces, la araña corrió”), y el 4, que marca el fin del enunciado fonológico.

El análisis acústico se llevó a cabo con el programa Praat (Boersma & Weenink, 2006). Se midió directamente en semitonos los puntos más bajos y altos de la sílaba tónica, pretónica y postónica en el sujeto o atributo de la oración. Posteriormente, se calculó la diferencia entre estos para obtener el valor de la excursión tonal. En la asignación de acentos bitonales se tomó el criterio de 1.5 st para distinguir un cambio en la tonía, al ser el mínimo perceptible (Pamies Bertrán, Fernández Planas, Martínez Celdrán, Ortega Escandell & Amorós Céspedes, 2001). También se midió la duración silábica.

Para el análisis de los enunciados de acuerdo con los criterios antes estipulados, en esta primera aproximación se buscó que las respuestas de los participantes incluyeran los tres constituyentes del enunciado (sujeto, verbo y atributo) con el fin de analizar la comprensión y producción del foco contrastivo al final o principio del enunciado. Así, se seleccionaron aquellas respuestas en las que el participante enunciara la oración completa para el análisis acústico, por ejemplo, “No, la ardilla está arriba de la rana”, y no una respuesta como “No, arriba de la rana”.

3.4. Creación de estímulos auditivos

Para la creación de los estímulos, todos los enunciados fueron grabados por una hablante del español de la Ciudad de México usando habla dirigida a infantes, ya que este tipo de habla atrae la atención de los niños. Específicamente, se hizo hincapié en la prosodia, con una mayor duración vocálica de la sílaba tónica de la palabra enfatizada, un tono más alto y una frecuencia fundamental que se mueve en un rango mayor (Villalobos Pedroza, 2022).

Para grabar los enunciados, se utilizó un micrófono condensador Shure MV51 USB colocado a una distancia aproximada de 30 cm de la hablante. La grabación se realizó mediante el programa Adobe Audition en estéreo a 44100 Hz en 16 bits. Todos los enunciados comenzaban luego de un segmento de 1000 ms que permitía la visualización de las imágenes antes de escucharlos. En primer lugar, se grabaron los estímulos de relleno. Se le entregó a la informante una lista de oraciones sin contexto y sin imágenes para evitar sesgos que generaran énfasis en algún elemento. Enseguida, se le presentaron los estímulos experimentales y de control en ambas condiciones (CS o CA), que fueron grabados uno tras otro sin presentar contexto, excepto tres imágenes para que la hablante entendiera que estaba comparando un elemento con los otros. Se le indicó que leyera en voz alta las oraciones teniendo en mente que comparaba el elemento de la oración marcado en negritas (sujeto o atributo) con los demás de las imágenes.

Se verificó que los enunciados fueran similares en cada condición utilizando Praat (Boersma & Weenink, 2006); dichos estímulos no se manipularon para mantener la naturalidad de la emisión de la hablante, pero se aseguró la marcación de la prominencia prosódica dependiendo de la condición. En el caso de la CA, se observó que la hablante realizó una elevación de la curva melódica de 8.3 st en promedio en la sílaba tónica del atributo del enunciado, con un acento tonal $L+>_jH^*$ en la mayoría de los casos. El sujeto, a su vez, tuvo un acento tonal $L+<H^*$ con una elevación de la curva melódica de 2.5 st en promedio. Además, la sílaba tónica del atributo tuvo una duración mayor ($\bar{x} = 295$ ms) en comparación con la sílaba tónica del sujeto ($\bar{x} = 168$ ms) (véase Anexo A1).

En la CS, se observó una elevación de la curva melódica en la sílaba tónica del sujeto de 6.1 st en promedio, con la realización bitonal $L+>_jH^*$. A su vez, el atributo tuvo una forma monotonal H^* o bitonal $L+>H^*$, con una elevación de la curva melódica de 1.8 st en promedio. Además, la sílaba tónica del sujeto tuvo una mayor duración ($\bar{x} = 246$ ms) en comparación con la del atributo ($\bar{x} = 188$ ms). Asimismo, se registró un tono de juntura alto H^- , ya

que se observó una pausa entre el sujeto y el resto del enunciado (véase Anexo A2); esto marcaría una diferencia entre condiciones, con dos frases entonativas en la cs.

Finalmente, los estímulos de relleno se produjeron con elevaciones de la curva melódica de 1.5 st en el sujeto y 1.6 st en el atributo en promedio, y tuvieron una duración silábica de 1.8 ms en promedio en el sujeto y de 1.7 en el atributo. El pico de la curva melódica del sujeto culminaba en la sílaba postónica, y dicha curva tenía un descenso progresivo, como ocurre con oraciones de foco amplio (Hualde, 2014: 268) (véase Anexo A3).

3.5. Procedimiento del experimento: niños

El investigador enviaba al padre o madre del niño un enlace para una videoconferencia en la plataforma Zoom. El día de la videoconferencia, el experimentador se presentaba y luego iniciaba con una sesión de familiarización. En esta se mostraban dos diapositivas con las 12 imágenes del experimento, seis en cada una, en un orden diferente al de la tarea experimental, combinando objetos animados e inanimados. Esto ayudaba a que el niño se familiarizara con las imágenes, se confirmara que era capaz de nombrarlas y que comprendía los conceptos de *arriba* y *abajo* utilizados en los enunciados experimentales. Cuando el niño corregía correctamente al experimentador tres veces consecutivas, se le proporcionaba el enlace de la plataforma Gorilla, que desplegaba el experimento (<https://gorilla.sc/>). El investigador permanecía en la llamada como apoyo.

Antes de la presentación de los estímulos, aparecía en la pantalla la caricatura de un gato, el cual indicaba al niño que necesitaba ayuda para una amiga que debía recordar las imágenes que el niño había visto antes. Luego, aparecía la imagen de otro personaje, un conejo, tapándose los ojos, que funcionaba como punto de fijación para recordar al niño quién estaba diciendo los enunciados. A continuación, se presentaban los 36 ensayos, divididos en cuatro bloques, entre los cuales los personajes felicitaban al niño

y le presentaban un trofeo. En cada ensayo, el padre o la madre debía hacer clic en un botón de “continuar” después de que el niño respondiera. Al finalizar, los personajes felicitaban al participante.

3.6. *Procedimiento del experimento: adultos*

En el caso de los adultos, el experimento era similar. Sin embargo, se omitían los personajes que se utilizaron en la versión para niños. Además, los participantes accedían al experimento a través de un enlace compartido en las redes sociales. Antes de comenzar, se les presentaba un texto y una grabación que les instruía sobre el proceso. Se les informaba que verían una serie de imágenes seguidas de una afirmación sobre ellas y debían indicar si la afirmación era correcta o incorrecta. En el caso de que fuera incorrecta, se les pedía que indicaran en voz alta la respuesta correcta. Previo al inicio, se les proporcionaba un ejemplo que mostraba una de las imágenes de práctica utilizadas en el experimento con niños.

En este caso, los estímulos se presentaban sin pausas entre ellos, y se les dio a los participantes siete segundos de silencio para que pudieran responder. La duración de este tiempo se determinó a través de pruebas previas con cuatro miembros del Laboratorio de Psicolingüística, con el objetivo de establecer un tiempo suficiente para dar una respuesta, pero no tan prolongado como para que la tarea se volviera tediosa. Al finalizar, se les agradeció su participación.

4. Resultados

4.1. *Niños*

Se recopiló un total de 288 grabaciones de audio (144 por condición). No hubo grabaciones incompletas o inaudibles. El primer análisis consistió en comparar las respuestas correctas a los estímulos experimentales, que sumaron un total de 96 grabaciones (48 por condición), en los que se requería corregir la descripción erró-

nea de las imágenes. Se esperaba que en la CS se corrigiera al sujeto de la oración, por ejemplo, se corrigiera quién está arriba de la ardilla en “La ARAÑA está arriba de la ardilla, ¿verdad?” En el caso de la CA, se esperaba que se corrigiera el atributo de la oración, p. e., corregir quién está abajo de la araña en “La araña está arriba de la ARDILLA, ¿verdad?” Se observó que el atributo se corrigió en la CA en 64.58% de las ocasiones, mientras que, en la CS, el sujeto se corrigió en un 8.3%. Hubo un sesgo hacia la corrección del atributo en ambas condiciones; la mayoría de los participantes corrigieron el atributo de la oración en la CS, lo cual resultó en algunas respuestas idénticas a las de los participantes de la CA. A pesar de estas respuestas sintácticamente idénticas entre condiciones, perceptualmente parecía haber un énfasis en el sujeto en la CS y en el atributo en la CA en las grabaciones. Por este motivo, se realizó un análisis acústico. Se seleccionaron 24 oraciones en la CA y 12 de la CS (del total de 48 por condición) que mencionaban tanto al sujeto como al atributo para su comparación. En primera instancia, se analizaron los acentos tonales. Los resultados se presentan en el Cuadro 1.

CUADRO 1. Acentos tonales observados en las respuestas dadas por niños

Acentos tonales sujeto		Acentos tonales atributo	
Condición Sujeto	Condición Atributo	Condición Sujeto	Condición Atributo
jH+L* (33.33%)	L+<H* (70.83%)	L+>jH* (33.33%)	L+H* (25%)
L+<jH* (8.33%)	L+<jH* (16.66%)	L+jH* (33.33%)	L+>H* (16.66%)
L+H* (8.33%)	H* (12.5%)	L+>H* (8.33%)	H* (16.66%)
L* (8.33%)		L+H* (8.33%)	L* (16.66%)
H* (8.33%)		H* (8.33%)	jH+L* (8.3%)
			L+>H* (8.3%)
			L+jH* (4.16%)
		L* (8.33%)	H+L* (4.16%)

Como se puede observar, hubo una tendencia en la CS a marcar el sujeto con un acento bitonal $\uparrow\text{H}+\text{L}^*$, mientras que en la CA se utilizó una forma bitonal $\text{L}+\downarrow\text{H}^*$. En el caso del atributo, en la mayoría de los casos se marcó con un acento bitonal $\text{L}+\downarrow\text{H}^*$ o $\text{L}+\uparrow\text{H}^*$ en la CS, y con un acento bitonal $\text{L}+\text{H}^*$ en la CA.

Las respuestas en el Cuadro 1 no mostraban pausa entre el sujeto y el resto de los elementos del enunciado. Sin embargo, en 41.66% de las respuestas en la CS se encontró esta pausa con tono de juntura H-. Al contener frases independientes, se analizaron de manera individual y se observaron dos acentos $\uparrow\text{H}+\text{L}^*$, uno $\text{L}+\downarrow\text{H}^*$ y uno $\text{L}+\downarrow\text{H}^*$, respectivamente. Además, se percibió una cesura en ambas condiciones después del sujeto caracterizada por una caída abrupta en la curva melódica, sin diferencias en el acento tonal, en 33.33% de los casos de la CS. En las Figuras 2 y 3 se presenta un ejemplo de ambas condiciones.

Se comparó la altura tonal, la elevación en semitonos de la curva melódica y la duración de la sílaba tónica del sujeto y atributo de la oración entre ambas condiciones (Figura 4). Se utilizó una prueba U de Mann-Whitney debido al tamaño de la muestra. La diferencia entre condiciones únicamente fue significativa con relación a la elevación de semitonos de la curva tonal en el sujeto ($U = 68$, $p = 0.011$), con una media de 4.74 st en la CS ($DE = 2.88$), en comparación con una media de 2.34 st en la CA ($DE = 1.07$), así como la duración entre las sílabas tónicas de los sujetos ($U = 76$, $p = 0.022$), con una media de 269.83 ms en la CS ($DE = 0.04$), en contraste con una media de 230.04 ms en la CA ($DE = 0.07$). No se encontró una diferencia significativa en la altura tonal del sujeto ($U = 132.50$, $p = 0.699$), ni en la altura tonal ($U = 100.50$, $p = 0.144$) o duración del atributo de la oración ($U = 140$, $p = 0.893$). Sin embargo, se encontró una diferencia marginalmente significativa en la elevación de semitonos de la curva melódica del atributo ($U = 87$, $p = 0.058$), con una media de 3.8 st en la CS ($DE = 2.34$) en comparación con una media de 2.35 st en la CA ($DE = 1.52$).

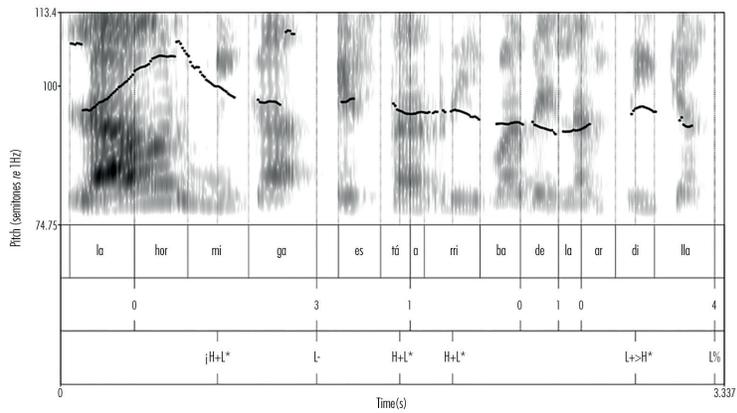


FIGURA 2. Ejemplo de respuesta dada por un niño en la Condición Sujeto para el enunciado "La hormiga está arriba de la ardilla"

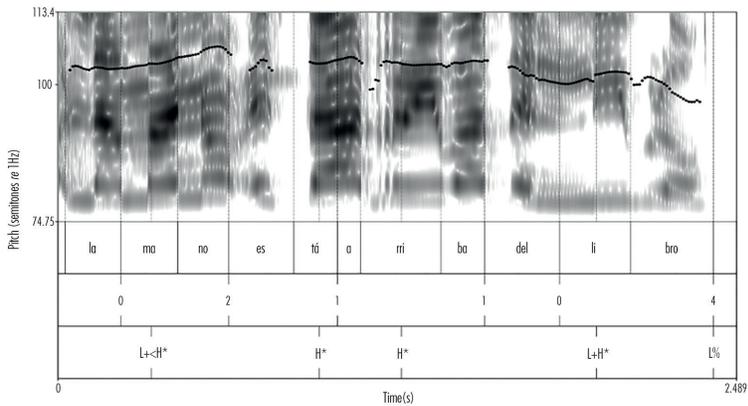


FIGURA 3. Ejemplo de respuesta dada por un niño en la Condición Atributo para el enunciado "La mano está arriba del libro"

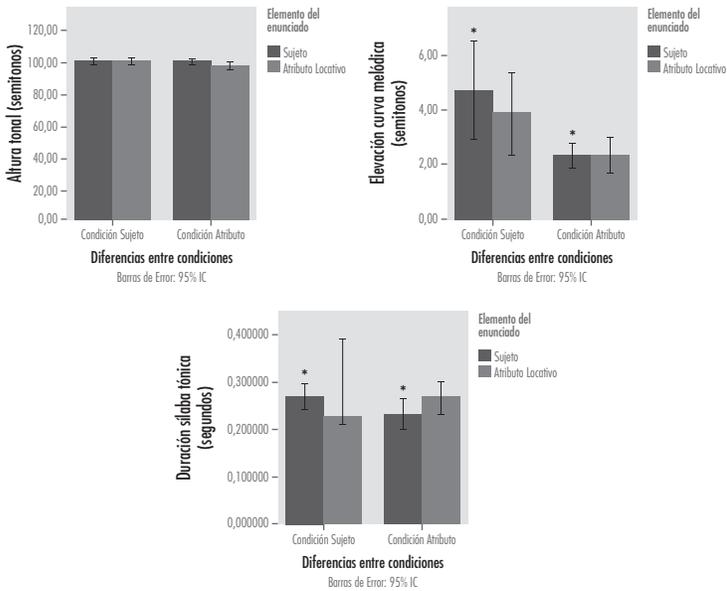


FIGURA 4. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por niños

4.2. Adultos

Se recopilaron 684 grabaciones (CS, $n = 288$; CA, $n = 396$). Se realizó una selección de datos, descartando aquellos que eran inaudibles (22 en la CA) y las respuestas incompletas, que fueron seis (2 en CS y 4 en CA). Esto resultó en un total final de 428 grabaciones (CS, $n = 190$; CA, $n = 238$).

El primer análisis consistió en comparar las respuestas correctas a los estímulos experimentales, en los que debía corregirse la descripción errónea del experimentador. Se examinó un total de 173 audios (CS, $n = 69$; CA, $n = 104$). En la CA, el atributo se corrigió el 100% de las ocasiones, mientras que, en la CS, el sujeto se corrigió solo 14.49% de las ocasiones. Al igual que en el caso de los niños, a pesar de las respuestas sintácticamente idénticas, perceptualmente parecía haber una diferencia entre las condiciones,

por lo que se realizó un análisis acústico. Los resultados del análisis de acentos tonales se muestran en el Cuadro 2.

CUADRO 2. Acentos tonales observados en las respuestas dadas por adultos

Acentos tonales sujeto		Acentos tonales atributo	
Condición Sujeto	Condición Atributo	Condición Sujeto	Condición Atributo
L+<H* (17.64%)	L+<H* (43.90%)	L* (44.11%)	L+>H* (30.48%)
H* (11.8%)	H* (28%)	L+>H* (14.70%)	L* (31.7%)
L+>¡H* (8.82%)	L+<¡H* (23.17%)	L+>¡H* (11.76%)	H+L* (21.9%)
L+<¡H* (8.82%)	L+¡H* (2.4%)	L+H* (8.8%)	L+>¡H* (8.53%)
L+H* (5.88)	H+L* (1.2%)	L+H* (8.8%)	H* (4.9%)
L+¡H* (5.88)	H+>L* (1.2%)	H* (8.8%)	L+H* (2.4%)
¡H+L* (2.9%)			

Como puede observarse, no hubo diferencia en la marcación de acentos tonales en el sujeto de la oración entre condiciones (ambos L+<H*). Sin embargo, en el atributo, la mayoría de los participantes marcó un tono L* en la CS y un bitono L+>H* en la CA.

También se observó una pausa entre sujeto y el resto del enunciado con tono H- en la CS en 38.23% de los casos. Estos últimos se analizaron de manera independiente y se encontraron los siguientes acentos: cuatro L+<H*, dos L+<¡H*, cuatro L+>H*, uno L+>¡H* y dos L+¡H*. Además, en ambas condiciones se registró una cesura después del sujeto, con un descenso abrupto de la curva melódica, en 26% de los casos totales, principalmente después de un bitono L+<H*. Se presentan ejemplos de ambas condiciones en las Figuras 5 y 6, con un tono de juntura en el sujeto y con una cesura 2, respectivamente.

A continuación, se examinó la altura tonal, la elevación en semitonos de la curva melódica y la duración de la sílaba tónica del sujeto y atributo de ambas condiciones (Figura 7).

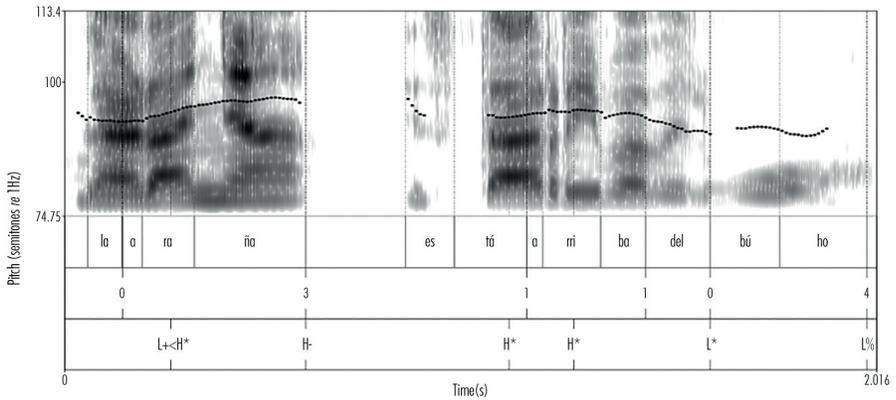


FIGURA 5. Ejemplo de respuesta dada por adulto en la Condición Sujeto para el enunciado “La araña está arriba del búho”, con cesura nivel 3 después del sujeto

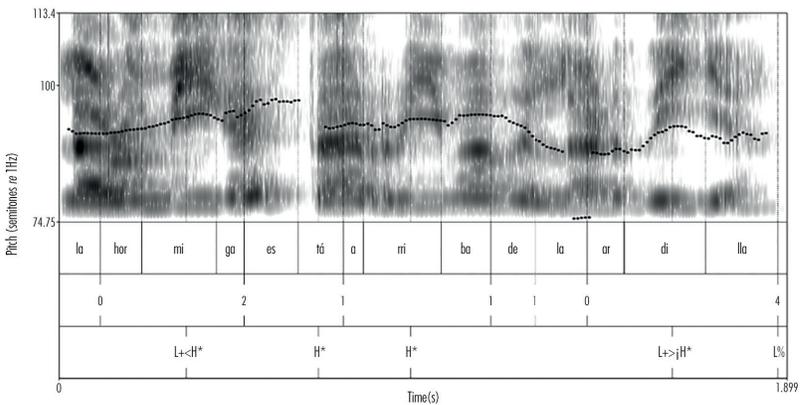


FIGURA 6. Ejemplo de respuesta dada por adulto en la Condición Atributo para el enunciado “La hormiga está arriba de la ardilla”, con cesura nivel 2 después del sujeto

La diferencia entre condiciones solo fue significativa con respecto a la elevación de semitonos de la curva melódica en el sujeto ($U = 919$, $p = 0.004$) con una mayor elevación en la CS ($\bar{x} = 2.98$ st, $DE = 1.39$) en comparación con la CA ($\bar{x} = 2.25$ st, $DE = 1.25$). También hubo una diferencia significativa en la altura tonal en el

sujeto ($U = 706$, $p < 0.001$), con una mayor altura tonal en la cs ($\bar{x} = 95.42$ st, $DE = 2.07$) en comparación con la CA ($\bar{x} = 91.58$ st, $DE = 5.11$). No se encontraron diferencias significativas en la duración de la sílaba tónica del sujeto ($U = 1093$, $p = 0.068$) ni en la altura tonal ($U = 1085$, $p = 0.061$), elevación de la curva melódica ($U = 1181$, $p = 0.196$) o duración de la sílaba tónica ($U = 1147$, $p = 0.134$) del atributo de la oración.

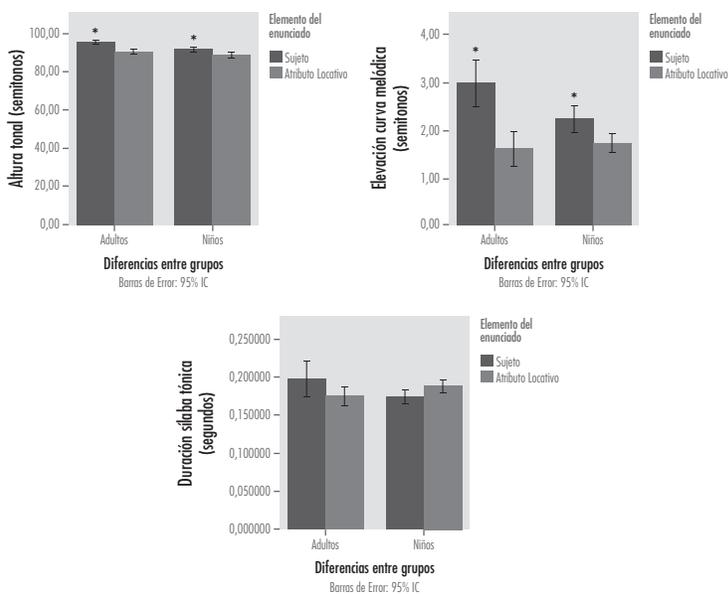


FIGURA 7. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos

4.3. Comparación entre niños y adultos

Se realizó una comparación entre niños y adultos en cuanto a la elevación de la curva melódica y la duración de la sílaba tónica en el sujeto y atributo en ambas condiciones. En la cs (Figura 8), se encontró una diferencia significativa entre adultos y niños en la altura tonal del sujeto y del atributo (sujeto, $U = 29$,

$p < 0.001$; atributo, $U = 0$, $p < 0.001$). Los niños mostraron una mayor altura tonal tanto en el sujeto (niños, $\bar{x} = 100.96$ st, $DE = 3.48$; adultos, $\bar{x} = 95.42$ st, $DE = 2.07$) como en el atributo (niños, $\bar{x} = 101.33$ st, $DE = 3.73$; adultos, $\bar{x} = 90.02$ st, $DE = 3.10$). Asimismo, se observó una diferencia significativa en la elevación de la curva melódica en el atributo ($U = 94$, $p < 0.006$), siendo mayor en el caso de los niños (niños, $\bar{x} = 3.88$ st, $DE = 2.34$; adultos, $\bar{x} = 1.62$ st, $DE = 0.99$). No hubo diferencia significativa en la elevación de la curva melódica del sujeto ($U = 134.50$, $p = 0.082$). En cuanto a la duración de la sílaba tónica, tanto en el sujeto como en el atributo hubo una diferencia significativa entre participantes (sujeto, $U = 76$, $p < 0.001$; atributo, $U = 93$, $p = 0.005$); los niños presentaron una mayor duración tanto en el sujeto (niños, $\bar{x} = 269.83$ ms, $DE = 43.37$; adultos, $\bar{x} = 196.47$ ms, $DE = 65.95$) como en el atri-

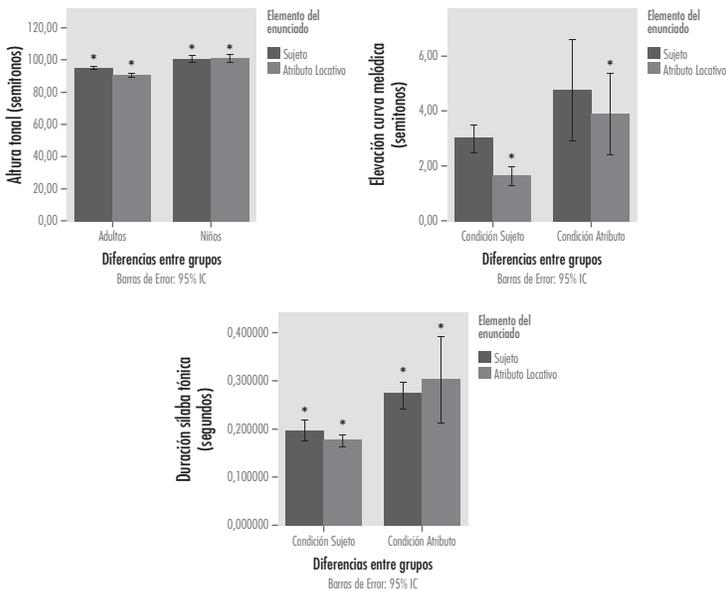


FIGURA 8. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos y niños en la Condición Sujeto

buto (niños, $\bar{x} = 302.20$ ms, $DE = 14.06$; adultos, $\bar{x} = 175.20$ st, $DE = 35.70$).

En la CA (Figura 9), también se encontró una diferencia significativa entre participantes adultos y niños en la altura tonal tanto del sujeto como del atributo (sujeto, $U = 189$, $p < 0.001$; atributo, $U = 146$, $p < 0.001$), siendo los niños quienes presentaron una mayor altura tonal en el sujeto (niños, $\bar{x} = 100.45$ st, $DE = 4.68$; adultos, $\bar{x} = 91.58$ st, $DE = 5.11$) y en el atributo (niños, $\bar{x} = 97.98$ st, $DE = 4.75$; adultos, $\bar{x} = 88.53$ st, $DE = 5.18$). No hubo diferencias significativas en la elevación de la curva melódica del sujeto ($U = 894.50$, $p = 0.449$) ni del atributo ($U = 765.50$, $p = 0.099$). En cuanto a la duración de la sílaba tónica, se observó una diferencia significativa en el sujeto y el atributo (sujeto, $U = 445$, $p < 0.001$; atributo, $U = 369$, $p < 0.001$); los niños presentaron una mayor duración tanto en el sujeto (niños, $\bar{x} = 230.94$ ms, $DE = 4.68$;

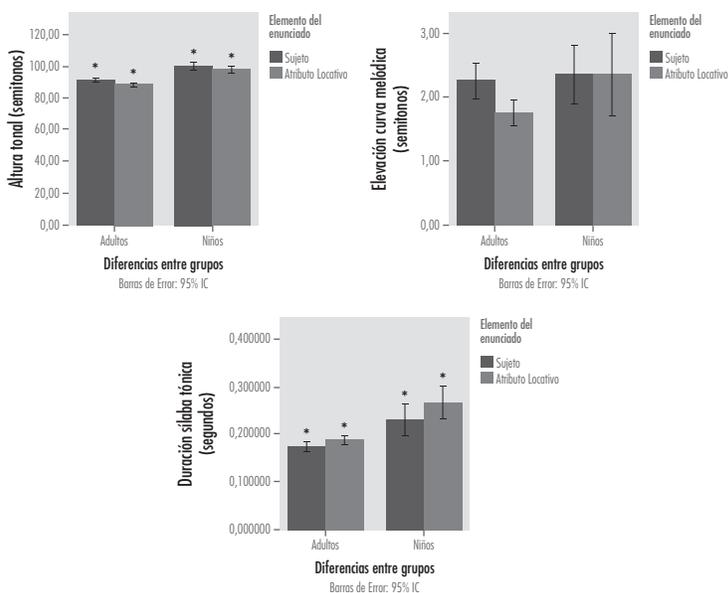


FIGURA 9. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos y niños en la Condición Atributo

adultos, $\bar{x} = 173.86$ ms, $DE = 42.60$) como en el atributo (niños, $\bar{x} = 267.07$ ms, $DE = 80.23$; adultos, $\bar{x} = 180.80$ st, $DE = 41.12$).

5. Discusión

Los resultados indican que tanto los adultos como los niños preescolares hablantes del español mexicano fueron capaces de comprender el foco en diferentes posiciones del enunciado. Si bien ambos grupos dieron respuestas idénticas en ambas condiciones, marcaron énfasis en el sujeto o atributo dependiendo de la condición. Ambas poblaciones utilizaron distintas pistas acústicas para marcar este énfasis.

En el caso de los niños, solo en la CA el sujeto tendió a marcarse con un acento tonal usual del tópico del enunciado en español, $L+<H^*$ (Hualde, 2014), lo cual sugiere que el elemento enfatizado fue interpretado como el tema de conversación y no constituía un foco informativo o contrastivo. Este patrón no se presentó en la CS, en la que predominó el acento bitonal $H+L^*$, característico del final de una frase entonativa y reflejo de la cesura entre el sujeto y el resto del enunciado, lo que también indicaría un tipo de énfasis propio del foco contrastivo (Martín Butragueño, 2005: 134).²

En ambas condiciones, el atributo se marcó con $L+>H^*$ o $L+H^*$, tonos usados para marcar foco informativo en español (Hualde, 2014: 268–271; Hualde & Prieto, 2015), sin una distinción predominante entre condiciones. En cuanto a los adultos, el sujeto se marcó, en ambas condiciones sin distinción, con $L+<H^*$, posiblemente como tópico (Hualde, 2014). En cambio, el atributo fue enfatizado solo en la CA con un patrón mucho más enfático, $L+>H^*$, que se encuentra en focos contrastivos del español usados para corregir información (Hualde, 2013: 271).

Además, los niños mostraron una elevación mayor en semitonos de la curva melódica, así como una mayor duración de la sílaba tónica, en el sujeto en la CS en comparación con la CA, sin dis-

² Se agradece la observación de uno de los dictaminadores acerca de este punto.

tinguir atributos entre condiciones. Los adultos, en cambio, mostraron una mayor distinción en el atributo de la oración y no en el sujeto, al enfatizar el atributo en la CA con un tono L+>H*, aunque también distinguieron al sujeto con una mayor elevación de la curva melódica y la altura tonal en el sujeto en la CS, este resultado también podría deberse a que la energía al final del enunciado es menor que al inicio (como puede apreciarse en la F0 en Hualde, 2014: 265). Asimismo, en contraste con los niños, distinguieron ambas condiciones con diferentes rasgos acústicos.

Además, se observó una pausa que separaba el sujeto del resto del enunciado en algunos casos tanto en niños como en adultos, pero solo en la CS. Esto podría ser otro indicio de focalización por fraseo encontrado en el foco contrastivo (Dorta Luis, 2008). Aunque en los dos grupos se encontró una cesura entre sujeto y atributo en ambas condiciones, caracterizado por un tono alto en la sílaba postónica del sujeto, seguida por un descenso abrupto, la pausa y, por lo tanto, un énfasis mayor, solo se encontró en la CS.

Finalmente, al comparar los rasgos acústicos empleados por ambos grupos, se encontró que los niños producían una mayor altura tonal y duración silábica tanto en los sujetos como en los atributos de la oración en ambas condiciones. Esta falta de distinción entre condiciones podría explicarse por una mayor expresividad por parte de los niños ya que, además, el experimento fue presentado como un juego.

Al contrastar estos resultados con estudios anteriores, también se encontró un sesgo por corregir el último elemento del enunciado (el atributo en nuestro caso) en ambas condiciones por parte de hablantes del español mexicano, al igual que en los hablantes del francés en Szendrői *et al.* (2018). Estos resultados se explican porque, en ambas lenguas, además de las pistas acústicas, el elemento enfatizado se coloca al final de las oraciones, lo que puede dificultar la identificación del foco utilizando únicamente pistas acústicas para algunos hablantes. Sin embargo, al igual que en dicho estudio previo, la presente investigación mostró que tanto adultos

como niños fueron capaces de comprender el foco al inicio de los enunciados a pesar de esta dificultad.

Del mismo modo, en el presente estudio, los niños utilizaron la duración de la sílaba tónica para marcar énfasis al usar un foco contrastivo en el sujeto, mientras que los adultos prefirieron la tónica. Este uso diferente del foco coincide con los resultados de producción de Villalobos Pedroza (2021).

6. Conclusiones

En general, se encontró evidencia de la capacidad de los niños preescolares hablantes del español del centro de México de cuatro a cinco años para distinguir el énfasis prosódico al inicio o al final de oraciones declarativas idénticas. En especial, identificaron el foco contrastivo al principio, utilizando un mayor énfasis en el sujeto para corregirlo cuando este era enfatizado. Por otro lado, los adultos de esta población también fueron capaces de identificar el foco contrastivo al principio o al final de oraciones, aunque su manera de señalarlo implicó el uso de pistas acústicas diferentes a las de los niños, enfatizando de distinta manera tanto el sujeto en la Condición Sujeto como el atributo en la Condición Atributo.

Estos resultados deben tomarse con precaución, ya que son producto de una muestra pequeña y corresponden a un estudio preliminar. En esta primera aproximación, se observaron elementos relevantes, como el prealineamiento, que requieren de un análisis acústico y estadístico más detallado en estudios posteriores. Adicionalmente, las condiciones experimentales difirieron entre niños y adultos, al igual que el número desigual de participantes en ambas muestras. Sería recomendable considerar el uso de oraciones transitivas con estímulos visuales que las acompañen a fin de comparar de manera más precisa los resultados con los encontrados por Szendrői *et al.* (2018). Por último, también podrían modificarse los estímulos auditivos para asegurar su estandarización entre condiciones, eliminando la pausa entre el sujeto y el resto del enunciado en la CS para homogeneizar las condiciones experimen-

tales, considerando que esta pista adicional pudo haber influido en las diferencias encontradas entre condiciones.

A pesar de lo anterior, este trabajo demuestra, de manera preliminar y por primera vez, la comprensión del foco prosódico en diferentes elementos de la oración por parte de una población preescolar hablante del español del centro y centro-occidente de México, utilizando un análisis acústico de sus respuestas ante oraciones con énfasis al inicio o al final de la oración.

7. Referencias

- Beckman, Mary; Díaz Campos, Manuel; Tevis McGory, Julia, & Morgan, Terrell (2002). Intonation across Spanish, in the tones and break indices framework. *Probus*, 14, 9–36. doi: 10.1515/prbs.2002.008
- Boersma, Paul, & Weenink, David (2006). *Praat: Doing phonetics by computer* (Versión 6.1.16) [Software]. <http://www.praat.org/>
- Chafe, Wallace L. (1974). Language and consciousness. *Language*, 50(1), 111–133. doi: 10.2307/412014
- Chen, Aaju (2010). Is there really an asymmetry in the acquisition of the focus-to-accentuation mapping? *Lingua*, 120(8), 1926–1939. doi: 10.1016/j.lingua.2010.02.012
- Cruttenden, Alan (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language*, 12(3), 643–661. doi: 10.1017/S030500090000670X
- Cutler, Anne, & Swinney, David A. (1987). Prosody and the development of comprehension. *Journal of Child Language*, 14, 145–167. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.487.3744&rep=rep1&type=pdf>
- De la Mota, Carmen; Martín Butragueño, Pedro, & Prieto, Pilar. (2010). Mexican Spanish intonation. En Pilar Prieto & Paolo Roseano (Eds.), *Transcription of intonation of the Spanish language* (pp. 319–350). Múnich: Lincom. https://prosodia.upf.edu/home/arxiu/publicacions/delamota/delamota_mexican_spanish_intonation.pdf
- Dik, Simon C. (1997). *The theory of functional grammar. Part 1: The structure of the clause*. Berlín: De Gruyter Mouton.

- Dorta Luis, Josefa (2008). La focalización prosódica: funcionalidad en los niveles lingüístico y pragmático. *Estudios de Fonética Experimental*, 17, 105–138. <https://core.ac.uk/download/pdf/39111163.pdf>
- Face, Timothy L. (2002). El foco y la altura tonal en el español. *Boletín de Lingüística*, 17, 30–52. <https://www.redalyc.org/pdf/347/34701703.pdf>
- Fifer, William P., & Moon, Christine M. (1994). The role of mother's voice in the organization of brain function in the newborn. *Acta Paediatrica*, 83(s397), 89–93. doi: 10.1111/j.1651-2227.1994.tb13270.x
- Furrow, David (1984). Young children's use of prosody. *Journal of Child Language*, 11(1), 203–213. doi: 10.1017/S0305000900005663
- Grassmann, Susanne, & Tomasello, Michael (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language*, 34(3), 677–687. doi: 10.1017/S0305000907008021
- Grassmann, Susanne, & Tomasello, Michael (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics*, 42(11), 3098–3105. doi: 10.1016/j.pragma.2010.04.019
- Goswami, Usha (2008). *Cognitive development: The learning brain*. Hove: Psychology Press.
- Gundel, Jeanette K., & Fretheim, Thorstein (2004). Topic and focus. En Laurence R. Horn & Gregory Ward (Eds.), *The handbook of pragmatics* (pp. 175–196). Massachusetts: Blackwell. doi: 10.1002/9780470756959.ch8
- Gutiérrez Bravo, Rodrigo (2008). La identificación de los tópicos y los focos. *Nueva Revista de Filología Hispánica*, 56(2), 363–401. doi: 10.24201/nrfh.v56i2.969
- Hendriks, Petra, & Koster, Charlotte (2010). Production/comprehension asymmetries in language acquisition. *Lingua*, 120(8), 1887–1897. doi: 10.1016/j.lingua.2010.02.002
- Höhle, Barbara; Fritzsche, Tom, & Müller, Anja (2016). Children's comprehension of sentences with focus particles and the role of cognitive control: An eye tracking study with German-learning 4-year-olds. *PLoS One*, 11(3), 1–27. doi: 10.1371/journal.pone.0149870
- Hornby, Peter A. (1971). Surface structure and the topic-comment distinction: A developmental study. *Child Development*, 42(6), 1975–1988. doi: 10.2307/1127600

- Hornby, Peter A., & Hass, Wilbur A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research*, 13(2), 359–399. doi: 10.1044/jshr.1302.395
- Hualde, José Ignacio (2003). El modelo métrico y autosegmental. En Pilar Prieto (Ed.), *Teorías de la entonación* (pp. 155–184). Barcelona: Ariel.
- Hualde, José Ignacio (2014). *Los sonidos del español: Spanish language edition*. Cambridge: Cambridge University Press.
- Hualde, José Ignacio, & Prieto, Pilar (2015). Intonational variation in Spanish: European and American varieties. En Sonia Frota & Pilar Prieto (Eds.), *Intonation in Romance* (pp. 350–391). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199685332.003.0010
- Jackson-Maldonado, Donna; Thal, Donna; Marchman, Virginia; Newton, Tyler; Fenson, Larry, & Conboy, Barbara (2003). *MacArthur inventarios del desarrollo de habilidades comunicativas: User's guide and technical manual*. Baltimore: P. H. Brookes.
- Johnson, Elizabeth K., & Jusczyk, Peter W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44(4), 548–567. doi: 10.1006/jmla.2000.2755
- Jun, Sun-Ah (2005). Prosodic typology. En Sun-Ah Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430–458). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199249633.001.0001
- Kisilevsky, Barbara S.; Hains, Sylvia M. J.; Lee, Kang; Xie, Xing; Huang, Hefeng; Ye, Hai He; Zhang, Ke, & Wang, Zengping (2003). Effects of experience on fetal voice recognition. *American Psychological Society*, 14(3), 220–224. doi: 10.1111/1467-9280.02435
- Kisilevsky, Barbara S.; Hains, Sylvia M. J.; Brown, C. Ann; Lee, Charlotte T.; Cowperthwaite, Bernardine; Stutzman, Sherri S.; Swansburg, Melissa L.; Lee, Kang; Xie, Xing; Huang, Hefeng; Ye, Hai He; Zhang, Ke, & Wang, Zengping (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71. doi: 10.1016/j.infbeh.2008.10.002
- Krifka, Manfred (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3-4), 243–276. <https://www.jstor.org/stable/10.2307/26190905>

- Ladd, Robert (2008). *Intonational phonology*. Cambridge: Cambridge University Press. doi: 10.1075/fol.17.2.08ogr
- Lahey, Margaret (1974). Use of prosody and syntactic markers in children's comprehension of spoken sentences. *Journal of Speech and Hearing Research*, 17(4), 656–668. doi: 10.1044/jshr.1704.656
- Lambrecht, Knud (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- MacWhinney, Brian, & Bates, Elizabeth (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior*, 17(5), 539–558. doi: 10.1016/S0022-5371(78)90326-2
- Martín Butragueño, Pedro (2005). La construcción prosódica de la estructura focal en español. En Gabriele Knauer & Valeriano Bellosta von Colbe (Eds.), *Variación sintáctica en español: un reto para las teorías de la sintaxis* (pp. 117–144). Berlín: De Gruyter. doi: 10.1515/9783110923353
- Martín Butragueño, Pedro (2019). *Fonología variable del español de México. Volumen II: prosodia enunciativa. Tomo I*. Ciudad de México: El Colegio de México.
- McDaniel, Dana, & Maxfield, Thomas L. (1992). Principle B and contrastive stress. *Language Acquisition*, 2(4), 337–358. doi: 10.1207/s15327817la0204_4
- Mora Bustos, Armando (2010). Construcciones escindidas, foco y entonación. En Sergio Bogard (Ed.), *Semántica, pragmática y prosodia. Reflejos en el orden de palabras en español* (pp. 219–248). México: Centro de Estudios Lingüísticos y Literarios, El Colegio de México.
- Nazzi, Thierry; Bertoncini, Josiane, & Mehler, Jacques (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766. doi: 10.1037//0096-1523.24.3.756
- Nespor, Marina, & Vogel, Irene (1994). *La prosodia*. Madrid: Visor.
- Pamies Bertrán, Antonio; Fernández Planas, Ana María; Martínez Celdrán, Eugenio; Ortega Escandell, Alicia, & Amorós Céspedes, Ma. Cruz (2001). Umbrales tonales en español peninsular. En Marina M. Barrio Parra, María Heliadora Cuenca Villain, Jesús Díaz, Luis Rodríguez & José Agustín Vidal Domínguez

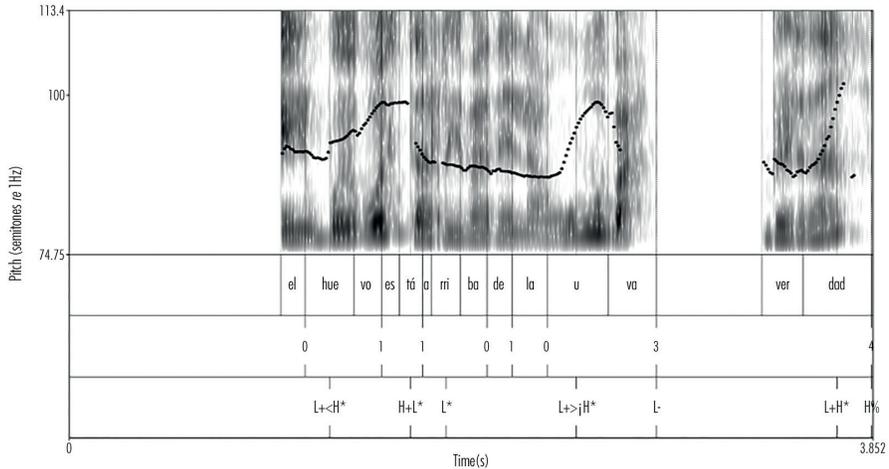
- (Eds.), *Actas del II Congreso Nacional de Fonética Experimental* (pp. 272–278). Sevilla: Universidad de Sevilla.
- Pierrehumbert, Janet Breckenridge (1980). *The phonology and phonetics of English intonation* (Tesis de doctorado). Massachusetts Institute of Technology, Massachusetts. <http://dspace.mit.edu/handle/1721.1/16065>
- Pynte, Joel (1998). The role of prosody in semantic interpretation. *Music Perception: An Interdisciplinary Journal*, 16(1), 79–97. doi: 10.2307/40285779
- Reinhart, Tanya (2006). Focus: The PF interface. En Tanya Reinhart, *Interface Strategies: Optimal and Costly Computations* (pp. 125–163). Cambridge: MIT Press. doi: 10.7551/mitpress/3846.003.0005
- Shahidullah, Sara, & Hepper, Peter G. (1994). Frequency discrimination by the fetus. *Early Human Development*, 36(1), 13–26. doi: 10.1016/0378-3782(94)90029-9
- Solan, Lawrence (1980). Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research*, 23(3), 688–698. doi: 10.1044/jshr.2303.688
- Snodgrass, Joan G., & Vanderwart, Mary (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. doi: 10.1037/0278-7393.6.2.174
- Szendrői, Kriszta; Bernard, Carline; Berger, Frauke; Gervain, Judit, & Höhle, Barbara (2018). Acquisition of informative focus marking by English, French, and German three-, four-, five- and six-year-olds. *Journal of Child Language*, 45(1), 219–241. doi: 10.1017/S0305000917000071
- Vallduví, Enric, & Engdahl, Elisabet (1996). The linguistic realization of information packaging. *Linguistics*, 34(3), 459–519. doi: 10.1515/ling.1996.34.3.459
- Vanrell Bosch, Maria del Mar; Stella, Antonio; Gili Favela, Barbara, & Prieto, Pilar (2013). Prosodic manifestations of the effort code in Catalan, Italian and Spanish contrastive focus. *Journal of the International Phonetic Association*, 43(2), 195–220. <https://www.jstor.org/stable/26351947>
- Villalobos Pedroza, Laura Cristina (2019). La marcación de foco en el habla dirigida a niños: marcos léxicos y estrategias prosódicas. En Valeria A.

- Belloro (Ed.), *La interfaz sintaxis-pragmática: estudios teóricos, descriptivos y experimentales* (pp. 283–310). Boston: De Gruyter.
- Villalobos Pedroza, Laura Cristina (2021). *Prosodia de la expresión temprana del foco informativo y del foco contrastivo en español* (Tesis de doctorado). Universidad Nacional Autónoma de México, Ciudad de México. http://ru.atheneadigital.filos.unam.mx/jspui/handle/FFYL_UNAM/4058
- Villalobos Pedroza, Laura Cristina (2022). La melodía del habla dirigida a niños y la adquisición de la lengua materna. *Punto Cunorte*, 14, 104–132. doi: 10.32870/punto.v1i14.136
- Weisleder, Adriana, & Fernald, Anne (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143–2152. doi: 10.1177/0956797613488145
- Wieman, Leslie (1976). Stress patterns of early child language. *Journal of Child Language*, 3(2), 283–286. doi: 10.1017/S0305000900001501

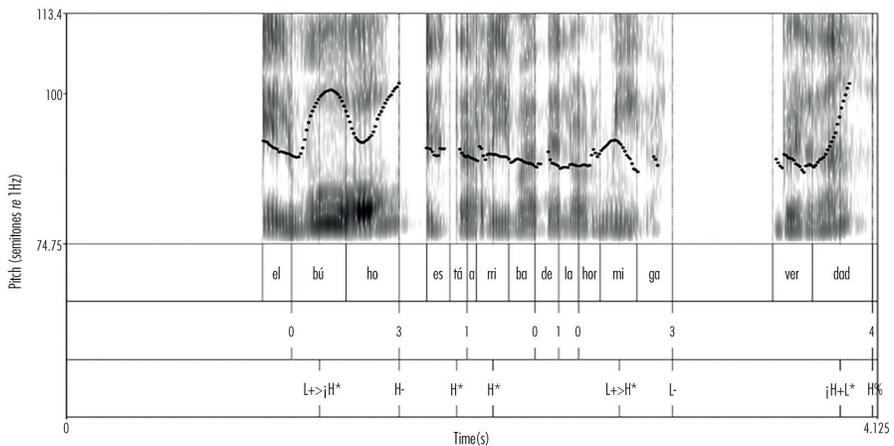
8. Anexos

8.1. Anexo A

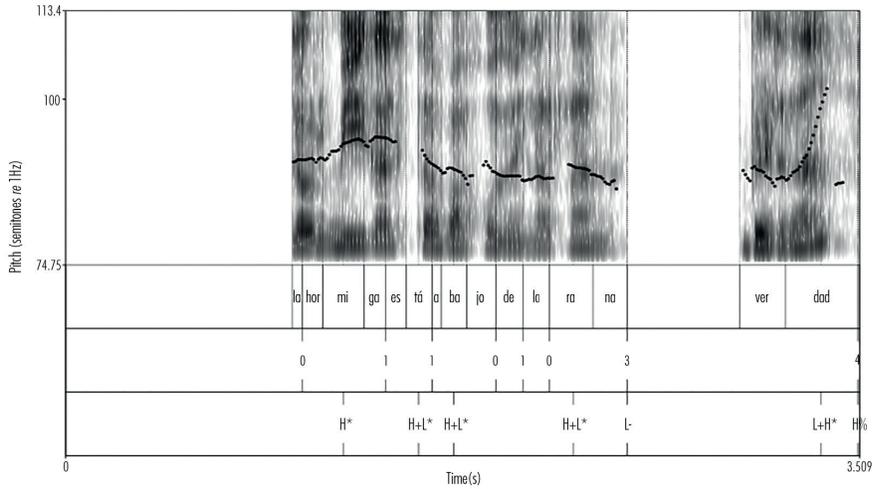
Anexo A1. Ejemplo de estímulo auditivo de la Condición Atributo



Anexo A2. Ejemplo de estímulo auditivo de la Condición Sujeto



Anexo A3. Ejemplo de estímulo auditivo de los estímulos de relleno



8.2. Anexo B

Representación esquemática de los tipos de acentos tonales y tipos de tonos de juntura encontrados en las transcripciones. Adaptación de De la Mota, Martín Butragueño y Prieto (2010), y Villalobos Pedroza (2021).

Esquema	Transcripción	Descripción
	L*	Caída progresiva o mantenimiento de un valor bajo de la FO en la sílaba tónica.
	H*	FO alta y sostenida en la sílaba tónica.

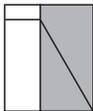
(continuación)

	Esquema	Transcripción	Descripción
Acentos bitonales		L*+H	Tono bajo en la sílaba tónica con un ascenso en la sílaba postónica.
		L+H*	Ascenso al inicio de la sílaba tónica con el pico tonal alineado al final de la misma.
		L+ _j H*	Una elevación L+H* mayor a 3 semitonos.
		L+<H*	Ascenso progresivo de la FO en la sílaba tónica con el pico tonal en la sílaba postónica.
		L+< _j H*	Una elevación L+<H* mayor a 3 semitonos.
		L+>H*	Ascenso de la FO con el pico tonal dentro de la sílaba tónica; en medio o antes de su final.
		L+> _j H*	Una elevación L+>H* mayor a 3 semitonos.

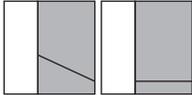
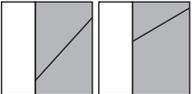
(cont.)

	Esquema	Transcripción	Descripción
		H+L*	Descenso de la FO dentro de la sílaba tónica.
		¡H+L*	Movimiento tonal mayor a 3 semitonos.
Tonos de juntura		L% (final del enunciado) L- (final de frase entonativa dentro del enunciado)	En la sílaba, un descenso de la FO menor a 1.5 semitonos o un tono bajo sostenido.
		H% (final del enunciado) H- (final de frase entonativa dentro del enunciado)	En la sílaba tónica, un ascenso de la FO desde un tono bajo.
Acentos monotonaes		L*	Caída progresiva o mantenimiento de un valor bajo de la FO en la sílaba tónica.
		H*	FO alta y mantenida en la sílaba tónica.
Acentos bitonaes		L*+H	Tono bajo en la sílaba tónica con un ascenso de la FO en la sílaba postónica.

(cont.)

Esquema	Transcripción	Descripción
	L+H*	Ascenso de la FO al inicio de la sílaba tónica con el pico tonal al final de la misma.
	L+ _j H*	Una elevación L+H* mayor a 3 semitonos.
	L+<H*	Ascenso progresivo de la FO en la sílaba tónica con el pico tonal en la sílaba postónica.
	L+< _j H*	Una elevación L+<H* mayor a 3 semitonos.
	L+>H*	Ascenso de la FO con el pico tonal dentro de la sílaba tónica; en medio o antes de su final.
	L+> _j H*	Una elevación L+>H* mayor a 3 semitonos.
	H+L*	Descenso de la FO dentro de la sílaba tónica.
	_j H+L*	Una elevación H+L* mayor a 3 semitonos.

(cont.)

	Esquema	Transcripción	Descripción
Tonos de junctura		L% (final del enunciado) L- (final de frase entonativa dentro del enunciado)	En la sílaba, un descenso de la FO menor a 1.5 semitonos o un tono bajo mantenido.
		H% (final del enunciado) H- (final de frase entonativa dentro del enunciado)	En la sílaba tónica, un ascenso de la FO desde un tono bajo.

Nota. En el esquema, las divisiones en los cuadros representan las sílabas. La sílaba tónica se presenta en color gris, y las líneas representan la curva melódica y sus movimientos dentro de las sílabas.

La activación de la información fonológica en aprendientes de segundas lenguas por medio de imágenes

Activation of the phonological information through images in second language learners

Alma Luz Rodríguez Lázaro
Universidad Nacional Autónoma de México, Escuela Nacional de Lenguas, Lingüística y Traducción, Departamento de Lingüística Aplicada
almaluzrl@enallt.unam.mx

Natalia Arias Trejo
Universidad Nacional Autónoma de México, Facultad de Psicología, Laboratorio de Psicolingüística
nariast@unam.mx



Recepción: 1 de diciembre del 2022

Aceptación: 25 de abril del 2023

doi: 10.22201/enallt.01852647p.2023.77.1049

Resumen

Investigaciones sobre activación del léxico han mostrado que cuando se escucha una palabra existe una activación de la fonología. Asimismo, estudios en bilingües señalan que la fonología se activa inicialmente para decodificar e integrar la información de las palabras en la segunda lengua (L2). El objetivo central de este trabajo fue describir si la información fonológica de aprendientes universitarios de una segunda lengua sin experiencia de inmersión es activada inicialmente en una tarea de preferencia visual. En la tarea se presentaban de forma simultánea cuatro estímulos visuales (un competidor fonológico, un competidor semántico, un competidor de rasgos visuales y un distractor) relacionados con una palabra en L2 percibida auditivamente. Los resultados revelaron que la información fonológica era activada inicialmente ante la presentación de imágenes, excepto cuando se exhibían palabras escritas. En este caso, los participantes mostraron una preferencia hacia la información de rasgos visuales. Se concluye que la activación de la fonología en la L2 se observa en aprendientes que no han tenido experiencias de inmersión, al igual que lo reportado en estudios con bilingües.

Palabras clave: activación del léxico; aprendientes de L2; preferencia visual; procesamiento fonológico

Abstract

Research on lexical activation has shown that phonological information is activated when a word is heard. Likewise, studies in bilingual individuals indicate that phonology is initially activated to decode and integrate word information in the second language (L2). This study aimed to describe whether phonological information in university-level learners of a second language without immersion experience is initially activated in a visual preference task. In this assignment, four visual stimuli were simultaneously presented (a phonological competitor, a semantic competitor, a shape competitor, and a distractor), all related to an L2 word perceived auditorily. The results revealed that phonological information is initially activated upon the presentation of images, except when written words were displayed. In this case, participants showed a preference for the visual shape competitor. It is concluded that phonological activation in L2 is observed in learners without immersion experiences, as reported in previous studies with bilinguals.

Keywords: lexical activation; L2 learners; visual preference; phonological processing

1. Introducción. La activación del léxico*

Diversos estudios han examinado los procesos cognitivos que dan pauta al entendimiento de las palabras (p. ej., Allopenna, Magnuson & Tanenhaus, 1998; Carreiras, Ferrand, Grainger & Perea, 2005; Thomas & van Heuven, 2005; Yee & Sedivy, 2006; Huettig & Altmann, 2007; Li, 2013; de Groot & Hagoort, 2018; Shirai, 2018). Se ha planteado que, al escuchar una palabra como *plato*, se decodifica inicialmente la información fonológica, posteriormente la información semántica e incluso se activa información no lingüística como los rasgos visuales, esto es, las principales características físicas de la forma prototípica del plato (Huettig & McQueen, 2007; Shook & Marian, 2013). Los procesos mencionados anteriormente se refieren a la activación del léxico, es decir, al reconocimiento de las palabras que percibimos.

Esta activación del léxico implica un conocimiento previo de las palabras que hemos aprendido. En este sentido, la memoria funciona como un dispositivo que ayuda a almacenar y recuperar información, así, cuando escuchamos la sílaba inicial de *plato* activamos la información lingüística y no lingüística relacionada con ese objeto (Huettig & Altmann, 2007). Entonces, la activación del léxico es un proceso de reconocimiento de palabras, en el cual la información fonológica y semántica, por ejemplo, se activan por la percepción auditiva de palabras (Harley, 2005).

Uno de los modelos más aceptados para explicar la activación del léxico es el denominado TRACE (McClelland & Elman, 1986). Este modelo describe cómo distintas unidades activan e inhiben diferentes conexiones que permiten el reconocimiento de la información que se percibe de manera visual o auditiva. El nombre

* Este manuscrito tuvo financiamiento de la DGAPA-PASPA-UNAM (2018-2019) brindado a la primera autora. Agradecemos a los participantes y a los miembros del Laboratorio de Psicolingüística de la Facultad de Psicología por su apoyo y comentarios para mejorar este trabajo. También agradecemos a Armando Q. Angulo Chavira y Alejandra Mitzi Castellón Flores por su orientación en los análisis estadísticos.

del modelo hace referencia a la huella de la información percibida que se procesa en distintas unidades y que facilita la distinción de fonemas y posibles palabras con una fonología similar a la percibida. Por ejemplo, cuando se escucha la primera sílaba de *plato* se pueden activar otras palabras en español que suenen de manera parecida, como *playera*, *playa*, *platino*, por mencionar algunas. Sin embargo, conforme se va percibiendo la segunda sílaba de *plato*, se descartan las palabras similares y se mantiene activa *plato*.

Ahora bien, existe evidencia que demuestra que cuando percibimos visualmente imágenes de objetos, entre ellos un plato, y simultáneamente escuchamos la sílaba inicial de *plato*, existe una preferencia visual hacia la imagen del plato que hacia las otras imágenes presentadas (Allopenna *et al.*, 1998; Yee & Sedivy, 2006; Huettig & Altmann, 2007). Esto es, la comprensión de la palabra a través de sus rasgos fonológicos ayuda a discriminar, entre otros objetos presentados simultáneamente, que la palabra que se estaba percibiendo era la de un plato, y por ello la mirada es más dirigida hacia dicho objeto.

Inclusive se ha evidenciado que la activación del léxico puede observarse cuando se presentan imágenes relacionadas semánticamente. Tal es el caso del estudio de Huettig y Altmann (2007), en el cual se mostraron imágenes de objetos que no tenían relación semántica entre ellos y en estas imágenes se incluyó la de un velero. De manera simultánea se expuso auditivamente la palabra *lago*, que no figuraba entre las imágenes. Los resultados demostraron que al escuchar la palabra *lago*, relacionada semánticamente con el velero, los participantes manifestaron una mayor preferencia visual hacia el velero que hacia las otras imágenes.

Otro ejemplo de la activación de léxico es el estudio de Huettig y McQueen (2007), en el cual se utilizaron cuatro experimentos para explorar si la preferencia visual se orientaba hacia la información fonológica, semántica o la de rasgos visuales, cuando se presentaba una palabra de manera auditiva y, simultáneamente, objetos y los nombres de los objetos. Estos experimentos se realizaron con hablantes del holandés y las principales manipulaciones

entre los experimentos fueron dos. La primera manipulación consistió en que la palabra auditiva no estuvo representada entre los estímulos visuales, solo fue percibida auditivamente. La segunda manipulación fue el tiempo de aparición entre la palabra auditiva y los estímulos visuales, como se detalla en las siguientes líneas.

Los participantes escucharon una palabra objetivo dentro de una oración, como *beker* (vaso), mientras se presentaban tres imágenes de objetos relacionados con *vaso*, entre ellos, la imagen de un castor (*bever*), por la relación fonológica con vaso en holandés, una imagen con rasgos visuales semejantes a la de un vaso, como canilla (en holandés *klos*), una imagen de un objeto relacionado semánticamente con vaso, como *tenedor* (en holandés *fork*) y, por último, un objeto distractor que no tenía relación con ninguno de los estímulos, como *paraguas* (en holandés *paraplu*). Es importante indicar que la palabra objetivo *vaso* solamente fue presentada auditivamente y no se mostró entre los estímulos visuales, por lo que se evaluó cómo se relacionaba la palabra objetivo con los objetos o, en este caso, competidores y distractor. Se les llama competidores porque se busca que la preferencia visual del participante recaiga sobre alguno de estos estímulos, y no sobre el distractor, debido a la nula relación entre este y la palabra objetivo (Huettig & McQueen, 2007).

Los principales resultados del estudio de Huettig y McQueen (2007) mostraron que la manipulación del tiempo y el tipo de modalidad de presentación de los estímulos visuales (imágenes o palabras escritas) generaron una activación del léxico distinta en cada uno de los cuatro experimentos. En el Experimento 1, los estímulos (oración con palabra objetivo e imágenes) se presentaron de manera simultánea, lo cual produjo que la preferencia visual de los participantes se dirigiera primero hacia el competidor fonológico, luego al competidor de los rasgos visuales semejantes y, al final, al competidor semántico. En el Experimento 2, la presentación del estímulo auditivo y las imágenes fue desfasada, esto es, primero se mostraron las imágenes y luego se escuchó la oración con la palabra objetivo, por lo cual se observó una marcada preferencia

visual hacia el competidor de rasgos visuales. Por otra parte, en el caso de los experimentos 3 y 4 se presentaron palabras escritas, es decir, los nombres de las imágenes de los experimentos 1 y 2. En el Experimento 3 se expusieron de manera simultánea los estímulos y en el Experimento 4 se mostraron de manera desfasada, o sea, primero las palabras escritas y posteriormente la oración con la palabra objetivo. En ambos casos la preferencia visual fue hacia el competidor fonológico, luego al competidor de rasgos visuales y, finalmente, al competidor semántico, independientemente de la manipulación temporal en la presentación de los estímulos.

Como se puede observar en el estudio de Huettig y McQueen (2007), la presentación de las imágenes (Experimentos 1 y 2) activó la información fonológica y la de los rasgos visuales de manera diferente. Esto es, cuando la manipulación de los estímulos (visuales y auditivos) es presentada de manera simultánea, la fonología es activada inicialmente, sin embargo, cuando los estímulos se presentan de manera asincrónica (primero el estímulo visual y posteriormente el estímulo auditivo), la información de rasgos visuales es activada inicialmente y por mayor tiempo. Por su parte, al exhibirse las palabras escritas (Experimentos 3 y 4) existió una mayor activación de la información fonológica sin importar la manipulación en la presentación de los estímulos. Por lo tanto, el estudio de Huettig y McQueen (2007) permite destacar cómo se realiza la activación del léxico en una lengua por medio de imágenes y palabras escritas.

Al respecto, el modelo TRACE (McClelland & Elman, 1986) postula que los procesos de activación del léxico donde se reconocen las letras que conforman las palabras, implican que se recupere la ortografía y la forma fonológica (p. ej., la pronunciación de la palabra) para que se pueda acceder al significado de la palabra que se encuentra almacenado en la memoria.

Asimismo, la activación del léxico a través de las palabras escritas conlleva que las representaciones de los grafemas (o letras) activen las correspondencias entre estos y los fonemas (los sonidos de los grafemas) (Lindner, Wijekumar & Joshi, 2022). Es decir, al

percibir palabras escritas, la activación de la ortografía produce la activación de la información de la fonología de las palabras que conocemos de una lengua (Smith, Monaghan & Huettig, 2021). Ahora bien, la correspondencia individual entre grafema y fonema depende de los sistemas alfabéticos (Smith *et al.*, 2021). Por ejemplo, el español es una lengua considerada transparente debido a la clara correspondencia entre los grafemas y fonemas de las palabras, lo cual implica que no se modifica la pronunciación (p. ej., *plato* /'pla.to/). Sucede lo contrario en el caso del inglés, considerada una lengua opaca debido a las irregularidades de la correspondencia entre grafemas y fonemas (p. ej., *mouth* /maʊθ/ [boca]; Smith *et al.*, 2021; Lindner *et al.*, 2022). Cabe señalar que el reconocimiento de las palabras es un proceso previo al de la lectura, ya que este último implica el entendimiento de la información de un texto extenso como un libro (Gough & Tunmer, 1986; Gaskins, Ehri, Cress, O'Hara & Donnelly, 1996; Henbest & Apel, 2018).

En esta primera sección se ha descrito cómo se puede explorar la activación del léxico en una lengua materna (L1), por lo que surge la duda de si dicha activación se realiza de forma similar en una L2. Particularmente, es de interés si la activación de la información fonológica en una L2 se puede observar ante la presentación de imágenes y palabras escritas. El presente estudio tiene como objetivo general explorar la activación del léxico en aprendientes universitarios de inglés sin experiencia de inmersión, en particular, la activación de la información fonológica por medio de imágenes y palabras escritas en una L2.

1.1. *La activación del léxico en una L2*

Existen modelos que tratan de ejemplificar los procesos de recuperación de palabras previamente conocidas en la L2 que dependen de la forma en la que se perciben, ya sea de manera auditiva o visual. Un ejemplo de estos modelos es el Revised Hierarchical Model (RHM; Kroll & Stewart, 1994), el cual propone que la activación

de las palabras de una L2 dependerá del nivel de proficiencia¹ del aprendiente. Además, el RHM sugiere que la lengua materna (L1) es un medio para la adquisición de la L2, debido a que ayuda a almacenar y recuperar la información fonológica y semántica de las palabras de la L2 a través de la traducción. No obstante, conforme aumenta la proficiencia, el reconocimiento y activación de palabras en la L2 se realiza de manera automática. Es decir, la L1 deja de ser un mediador para activar la L2, por lo que al momento de percibir palabras en la L2 únicamente esta se activa.

Otro modelo sobre activación de léxico en bilingües es el Bilingual Interactive Model (BIA) y su versión más reciente BIA+ (Dijkstra & van Heuven, 2002), el cual sugiere que la L1 se activa de manera momentánea cuando las palabras de la L2 se decodifican. En este modelo se asume que una lengua está activada de manera sustancial mientras que la otra se encuentra “apagada” durante el reconocimiento de palabras (Kroll & Bogulski, 2012).

Por su parte, Shook y Marian (2013) generaron el modelo Bilingual Language Interaction Network for Comprehension of Speech (BLINCS), que incorpora los postulados de modelos anteriores como el TRACE, el RHM y el BIA+. Este modelo trata de simular el procesamiento del léxico en los bilingües cuando se percibe información lingüística (p. ej., palabras escritas) y no lingüística (p. ej., imágenes) a través de la audición, la visión o ambos. BLINCS se compone de un nivel fonológico que comparten las dos lenguas (L1 y L2), un nivel fonológico y otro ortológico, ambos especializados, que distinguen los fonemas y grafemas, respectivamente, de la L1 y la L2, y, por último, un nivel semántico que comparten la L1 y la L2, ya que se plantea que el significado de las palabras es equivalente entre las dos lenguas. Una de las principales contribuciones de este modelo es la especialización de los módulos fonoló-

¹ En este trabajo, la *proficiencia* se entiende como el nivel de adquisición de la L2 que un aprendiente ha logrado después de la exposición a contextos donde esta se usa, por ejemplo, el estudio del inglés como materia curricular en la escuela (The Douglas Fir Group, 2016).

xico y ortoléxico, lo cual es relevante para explicar las diferencias en la decodificación de las palabras de cada una de las lenguas del bilingüe. Igualmente, la incorporación de la activación del léxico a través de imágenes en la L1 o L2 es una premisa que en los modelos de activación de léxico mencionados anteriormente no se había desarrollado (Shook & Marian, 2013). En suma, este modelo ofrece un acercamiento sobre cómo la activación del léxico en las dos lenguas se puede generar según la modalidad de presentación de los estímulos que se perciben (p. ej., imágenes).

A su vez, los estudios con bilingües han intentado demostrar la activación del léxico en la L1 y la L2. Ejemplo de ello es el estudio de Mishra y Singh (2016) realizado en la India con bilingües de hindi (L1) e inglés (L2), que explora si la información fonológica de la L2 se activa cuando se escucha una palabra en la L1 y viceversa. Sus principales resultados mostraron que según el nivel de proficiencia en la L2 puede observarse la activación de la L2 hacia la L1 o viceversa, esto es, los participantes con una proficiencia alta en la L2 tuvieron una mayor preferencia visual hacia los competidores fonológicos que los participantes con una proficiencia baja. Este estudio resalta la importancia de la proficiencia de los participantes hablantes de una L2 para reportarla como una variable que pudiera influir en los resultados de la activación del léxico en una L2.

Por otro lado, Sunderman y Kroll (2006) realizaron un estudio en el cual se les pidió a los bilingües responder si las palabras escritas que se les presentaban eran equivalentes en traducción de la L1 a la L2. Los principales resultados mostraron que los bilingües, sin importar su proficiencia en la L2, mostraron interferencias (influencia de una lengua sobre otra) en sus respuestas al tratarse de palabras con similitud fonológica y ortográfica (p. ej., *carta-cart* [carro]), porque contestaron incorrectamente a las palabras presentadas que eran equivalentes en traducción. No obstante, los participantes tuvieron una mayor precisión al identificar las relaciones semánticas entre las palabras de la L1 y la L2 (p. ej., *cama-blanket* [cobija]). Este estudio de Sunderman y Kroll (2006)

permite observar que la decodificación fonológica y ortográfica de las palabras de la L2 requiere de una especialización de la activación de léxico para procesar adecuadamente una palabra en la L2. Asimismo, el procesamiento de la información semántica, al integrar un nivel compartido entre las dos lenguas, como lo indica el modelo BLINCS (Shook & Marian, 2013), puede implicar una menor complejidad en la decodificación de una palabra en la L2.

Los estudios descritos anteriormente son una muestra de cómo se puede explorar la activación del léxico en una L2. Sin embargo, debe considerarse que existen algunas diferencias en cuanto al tipo de participantes. Por ejemplo, en el estudio de Mishra y Singh (2016) se consideraron participantes bilingües en un contexto de inmersión, esto es, los hablantes están expuestos a dos lenguas que son utilizadas en contextos escolares o de socialización, como las interacciones familiares. En contraste, en el estudio de Sunderman y Kroll (2006) los participantes eran aprendientes de una L2, en este sentido, una forma en determinar si los participantes tenían una proficiencia baja o alta fueron las experiencias previas de inmersión en la L2; en caso de que los participantes no hubieran tenido dichas experiencias, se les consideraba como aprendientes con proficiencia baja. Esto último representa una variable confusa y, como consecuencia, los resultados de Sunderman y Kroll (2006) podrían ser contradictorios, ya que es necesario evitar la subjetividad en la medición de la proficiencia de una L2 (Bice & Kroll, 2021).

A su vez, se ha planteado que existe una diferencia entre el aprendiente de una L2 y un bilingüe. Esto es, el bilingüe está inmerso en la L2 en contextos escolares formales y de socialización, los cuales ayudan a mejorar la proficiencia en la L2 (The Douglas Fir Group, 2019). En contraste, el aprendiente de una L2 únicamente ha adquirido la L1 a través de contextos escolares y de socialización (como la interacción con la familia y amigos), por consiguiente, la L2 ha sido aprendida en un contexto escolar o, inclusive, de manera informal en algún momento de su vida (The Douglas Fir Group, 2019). Como se puede observar, un bilingüe

tiene mayores oportunidades de aprendizaje e interacción que un aprendiente, el cual tiene contextos más limitados. En esta investigación denominamos aprendientes de una L2 a aquellos que la han aprendido en contextos formales escolares y sin experiencias de inmersión en la L2.

Siguiendo esta línea, surge la pregunta de investigación de este trabajo: los aprendientes de una L2 que no han tenido una experiencia de inmersión, ¿muestran una activación de la información fonológica en la L2?

Como mencionamos anteriormente, el objetivo de este estudio fue explorar si la información fonológica de una palabra de la L2 es activada inicialmente al presentarse imágenes y palabras escritas relacionadas con dicha palabra —percibida auditivamente— a aprendientes de inglés como L2, a pesar de no tener una experiencia de inmersión. Se hipotetizó que existiría una activación inicial de la información fonológica de las palabras de la L2 al momento de presentar de forma simultánea un competidor fonológico, un competidor semántico, un competidor de rasgos visuales y un distractor. Se sugiere que, dadas las irregularidades en la correspondencia entre grafemas y fonemas que existe en inglés (Smith *et al.*, 2021; Lindner *et al.*, 2022), se observe una activación de la información fonológica debido a la diferenciación fonológica y ortográfica que deben desarrollar los aprendientes cuando perciben palabras del inglés. Existe evidencia de que los bilingües activan la información fonológica en la L2 en contextos similares (Mishra & Singh, 2016), por lo que se esperaría que los aprendientes, a pesar de no tener experiencias de inmersión, también activen la información fonológica de la L2.

Además, como objetivo secundario se examinó si la información fonológica de la L2 es activada en mayor proporción de tiempo que la información semántica y la de rasgos visuales cuando se presentan de manera simultánea estímulos visuales (p. ej., imágenes o palabras escritas) relacionados con una palabra percibida auditivamente en la L2. Se hipotetizó que la información fonológica de las palabras en la L2 tendría una mayor activación y, por ende,

una mayor atención de mirada, que la información semántica y la de rasgos visuales, ya que la integración de la información lingüística (p. ej., fonemas y grafemas) requiere de mayor diferenciación que la información no lingüística (p. ej., rasgos visuales de objetos) de las palabras del inglés como L2, como lo sugiere el modelo BLINCS (Shook & Marian, 2013).

Para conocer la activación del léxico se registró la mirada a través de una tarea de preferencia visual, ya que se ha evidenciado una interacción entre la activación del léxico y la mirada, como se explicó en los apartados anteriores (Allopenna *et al.*, 1998; Yee & Sedivy, 2006; Huettig & Altmann, 2007; Huettig & McQueen, 2007; Mishra & Singh, 2016). A su vez, la tarea de preferencia visual siguió los postulados del paradigma visual global (Visual World Paradigm o vwp; Huettig, Olivers & Hartsuiker, 2011), el cual tiene como objetivo registrar únicamente los movimientos oculares sin necesidad de que los participantes emitan otra conducta (p. ej., presionar un botón). Esto es, se les pide a los participantes prestar atención a lo que escuchan y no retirar la mirada de la pantalla que presenta los estímulos visuales (Huettig & McQueen, 2007; Huettig *et al.*, 2011). En las siguientes secciones se brinda una mayor explicación del método del presente estudio.

2. Método

2.1. Participantes

La muestra de este estudio fue por conveniencia como se describe a continuación. Los participantes fueron alumnos de inglés de la Escuela Nacional de Lenguas, Lingüística y Traducción (ENALLT) de la UNAM. Los alumnos en estos cursos deben demostrar un nivel A2 (Common European Framework of Reference for Languages; Council of Europe, 2001) conforme a un examen de admisión. Debido a la pandemia de COVID-19, los alumnos tomaban sus cursos en línea. Se decidió que se invitaría a los alumnos del nivel B1 ya que en este nivel de adquisición de la L2 se espera que la compren-

sión de palabras se realice de manera más precisa que en niveles inferiores (Council of Europe, 2001). Los alumnos fueron invitados por sus profesores mediante un correo electrónico que incluía la información del estudio y los datos de la investigadora principal.

Los criterios de inclusión fueron que los alumnos estuvieran inscritos en los cursos de inglés del nivel B1, que hablaran español como L1 y que tuvieran un rango de edad de 18 a 35 años. Los criterios de exclusión fueron tener una L1 diferente al español y que los participantes tuvieran información faltante en las tareas del estudio. La muestra estuvo conformada por 63 alumnos, de los cuales 46 fueron mujeres. Además, ninguno de los alumnos reportó tener experiencias de inmersión en un país anglófono para estudiar inglés.

Con el fin de corroborar el nivel de adquisición de la L2 de los participantes, se aplicó el test LexTALE (Lemhöfer & Broersma, 2012), que se trata de una prueba en línea de decisión léxica (p. ej., decir si una palabra pertenece a una lengua) que mide el conocimiento de vocabulario y que indica el nivel de adquisición de la L2 de los participantes como: principiante, intermedio y avanzado. Esta prueba ha sido validada contra otras pruebas de medición de proficiencia de inglés como el Quick Placement Test (QPT) y el TOEIC test (Lemhöfer & Broersma, 2012). El Cuadro 1 muestra la información más relevante de los participantes.

CUADRO 1. Información de los participantes

	Media (<i>DE</i>)	Rango
Edad	21.76 (2.4)	19–31
Edad de adquisición de la L2	10.97 (4.9)	2–23
Años de exposición a la L2	10.79 (4.6)	1–21
LexTALE	63.77 (6.9)	50–76*

Nota: $n = 63$. *DE* = Desviación Estándar.

* Estos puntajes representan un nivel de adquisición del inglés de intermedio-bajo a intermedio-alto.

2.2. *Consentimiento informado*

El Comité de Ética del Posgrado en Psicología de la UNAM aprobó el protocolo presentado para realizar este estudio en los alumnos de la ENALLT. Los participantes interesados recibieron el protocolo del estudio en español, en el cual se indicaba el propósito, la utilización de una plataforma para registrar los movimientos de los ojos y las características técnicas de las computadoras personales que eran necesarias para llevar a cabo la tarea de rastreo visual. Este documento era enviado por correo electrónico a los alumnos, solicitando que firmaran digitalmente el consentimiento si estaban de acuerdo en participar, de conformidad con la información proporcionada.

2.3. *Tarea de rastreo visual*

Se utilizó la plataforma digital RealEye (<https://www.realeye.io/>) para registrar los movimientos oculares de los participantes. Esta plataforma utiliza las cámaras de las computadoras para realizar el rastreo visual, siempre y cuando el participante esté de acuerdo en permitir que se active su cámara únicamente con estos fines. A fin de llevar a cabo la tarea de preferencia visual se envía a los participantes una liga, generada automáticamente por la plataforma, la cual dirige al video de la tarea; durante la visualización del video la cámara del participante estará activa para registrar sus movimientos oculares.

2.4. *Diseño*

Este estudio tuvo dos condiciones para la presentación de estímulos visuales (imágenes de objetos y palabras escritas). Para la Condición 1. Imágenes, se presentó de manera auditiva una palabra en inglés y, de modo visual, imágenes que representaban objetos relacionados con la palabra; la palabra escuchada no aparecía entre las imágenes. Para la Condición 2. Palabras escritas, se presentó auditivamente una palabra en inglés (la misma utilizada en

la Condición 1. Imágenes), pero, en vez de exponerse imágenes, se mostraron como estímulos visuales los nombres de los objetos, esto es, palabras escritas. La palabra escuchada tampoco figuraba entre estas.

En la siguiente sección se describe la selección de los estímulos visuales y auditivos.

2.4.1. Selección de estímulos

Se seleccionaron sustantivos de la base SubtlexUS (Brysbaert & New, 2009) con una frecuencia alta de uso, que fungieron como los estímulos auditivos de este estudio. En específico, se eligieron palabras con un valor Zipf de 4 a 7, tomando en cuenta que en esta base de datos un valor alto indica una frecuencia elevada de uso, siendo 7 el valor más alto. Asimismo, se escogieron palabras cortas, con una longitud de nueve letras, o tres sílabas, para facilitar el procesamiento (Carreiras *et al.*, 2005). En total, se seleccionaron 28 palabras que fueron grabadas digitalmente por un hablante nativo del inglés, al cual se le pidió no remarcar la pronunciación. Las palabras se grabaron por medio de Adobe Audition *software* y los archivos se normalizaron para ajustar la amplitud, volumen y reducción de ruido.

Después de elegir las palabras que conformarían los estímulos auditivos, se seleccionaron los *competidores fonológicos*, es decir, palabras que compartieran la sílaba inicial o el núcleo vocálico con los estímulos auditivos. Para dicho propósito se utilizó el *Macmillan English Dictionary for Advanced Learners* (Rundell & Fox, 2002). Se revisó en las Normas de Asociación de la Universidad del Sur de Florida (Nelson, McEvoy & Schreiber, 1998) que los competidores fonológicos no estuvieran relacionados semánticamente con los estímulos auditivos. Para validar los competidores fonológicos se comparó la distancia ortográfica y fonológica a través de la medida de Levenshtein, la cual describe la diferencia de caracteres, así como el solapamiento fonológico y ortográfico cuando se comparan dos palabras (Carrasco-Ortiz, Midgley & Frenck-Mestre, 2012;

Levenshtein, 1975). Según esta medida, una distancia fonológica corta significa que existe menos diferencia entre dos palabras. La media de distancia entre los estímulos auditivos y los competidores fonológicos fue de 2.50 (rango: 1–5, $DE = 1.24$). Esta media indica que existe una distancia entre una y cinco letras de diferencia entre la palabra auditiva y los competidores fonológicos. Igualmente, estos competidores fueron estadísticamente diferentes de los competidores semánticos, de rasgos visuales y distractores. Se revisó también que el solapamiento fonológico entre los estímulos auditivos y los competidores fonológicos siguieran una distribución normal en el inicio de la consonante y el núcleo vocálico, por lo que dicha distribución se cotejó mediante la prueba Shapiro-Wilk, con el fin de seleccionar la prueba estadística adecuada. Se encontró que los estímulos auditivos ($W(18) = .638, p < .001$) y los competidores fonológicos ($W(18) = .767, p < .001$) no seguían una distribución normal. Por ello se tomó la decisión de analizar el solapamiento fonológico mediante la prueba no paramétrica de Friedman, en la que no se encontraron diferencias estadísticamente significativas entre la cantidad de sílabas de los competidores fonológicos, los competidores semánticos, de rasgos visuales y los distractores ($X^2(3) = .310, p = .958$). Tampoco se observaron diferencias estadísticamente significativas entre la cantidad de caracteres de los competidores fonológicos, competidores semánticos, de rasgos visuales y distractores ($X^2(3) = .248, p = .969$).

Adicionalmente, se revisó con la prueba de Friedman la comparación total de la distancia de Levenshtein entre los estímulos auditivos contra los competidores semánticos, los competidores de rasgos visuales y los distractores; se encontró una diferencia significativa entre la distribución de la distancia entre estas palabras ($X^2(3) = 26.65, p < .001$). El análisis *post hoc* con la prueba de signos de Wilcoxon reveló diferencias significativas en la distancia entre estímulos auditivos y competidores semánticos ($Mdn = 4.00$) contra estímulos auditivos y competidores fonológicos ($Mdn = 2.00; Z = 3.687, p < .001$), la distancia entre estímulos auditivos y distractores ($Mdn = 5.00$) contra estímulos

auditivos y competidores fonológicos ($Mdn = 2.00$; $Z = 3.42$, $p = .001$), y la distancia entre estímulos auditivos y competidores de rasgos visuales ($Mdn = 4.00$) contra estímulos auditivos y competidores fonológicos ($Mdn = 2.00$; $Z = 3.43$, $p = .001$). Estos resultados muestran que los estímulos auditivos estuvieron controlados en relación con las otras palabras presentadas como competidores y distractores, tanto en la ortografía como en la fonología.

Los *competidores semánticos* se escogieron siguiendo las Normas de Asociación de la Universidad del Sur de Florida (Nelson *et al.*, 1998), de acuerdo con una asociación semántica moderada (FSG: 0.010 a 0.050), esto para evitar que una alta asociación (FSG > 0.050) del competidor semántico con el estímulo auditivo atrajera más la mirada sobre los otros competidores (Aschenbrenner & Yap, 2019). De esta manera, todos los competidores tenían la misma oportunidad de atraer la atención al momento de su visualización.

Los *competidores de rasgos visuales o forma* se seleccionaron conforme a las características visuales compartidas con los estímulos auditivos. Por ejemplo, el estímulo auditivo *plate* (plato) comparte características visuales con *disk* (disco). Se verificó que estos competidores de forma no estuvieran relacionados fonológica ni semánticamente con los otros competidores. Asimismo, se realizó un estudio de validación para corroborar que las imágenes seleccionadas de los competidores de forma compartieran características visuales con los estímulos auditivos. En este estudio participaron aprendientes de inglés con características similares a la muestra final de esta investigación, pero que no fueron incluidos en ella. En la validación se les preguntaba a los participantes, por ejemplo, qué tanto la imagen de *plate* (plato) comparte características visuales con *disk* (disco). Se escogieron las imágenes que tuvieron una alta aceptación entre los participantes de compartir características visuales para integrarlos como competidores de forma.²

² La información sobre este estudio se puede consultar en: https://osf.io/z3cgx/?view_only=6f710b0b17e94f8bebe6cf7041c6c75c0

Por otra parte, se cuidó que los *distractores* no se relacionaran fonológica, visual, ni semánticamente con los estímulos auditivos.

Finalmente, se corroboró que las palabras seleccionadas (estímulos auditivos competidores y distractores) fueran conocidas por los participantes mediante un estudio de validación. Esto permitió incluir únicamente palabras conocidas por los alumnos que participaron en la muestra final. Cabe aclarar que en este estudio se incluyeron aprendientes de inglés con características similares a la muestra de esta investigación, pero que no formaron parte de ella.³

2.4.2. Estímulos visuales

Se seleccionaron objetos que fungieron como estímulos visuales, los cuales representaron los competidores y los distractores elegidos. Dichos objetos se presentaron en vectores en formato blanco y negro, esto es, las siluetas de los objetos fueron de color negro mientras que el fondo y el relleno fue blanco. Los objetos se eligieron de bases de imágenes (Snodgrass & Vanderwart, 1980; Pérez & Navalón, 2003; Szekely, Jacobsen, D'Amico, Devescovi, Andonova, Herron, Lug, Pechmann, Pléh, Wicha, Federmeier, Gerdjikova, Gutierrez, Hung, Hsu, Iyer, Kohnert, Mehotcheva, Orozco-Figueroa, Tzeng, Tzeng, Arévalo, Vargha, Butler, Buffington & Bates, 2004). Los objetos que no se encontraron en dichas bases se obtuvieron de internet a través de Shutterstock. Las imágenes tuvieron un tamaño de 1080 x 1440 pixeles y se utilizó una resolución de 72 pixeles por pulgada.

2.5. *Diseño experimental*

Se realizaron 28 ensayos para cada condición, de los cuales 18 fueron de interés y 10 de relleno. Los estímulos visuales se presentaron de manera aleatoria y contrabalanceada en cada condición

³ Los detalles de esta validación se pueden revisar en: https://osf.io/z3cgx/?view_only=6f710b0b17e94fbebe6cf7041c6c75c0

para evitar efectos de aprendizaje por parte de los participantes. De igual manera, en las dos condiciones se exhibieron los estímulos visuales antes del estímulo auditivo. A través de un cuadrante se mostraron cuatro estímulos visuales: competidor fonológico, competidor semántico, competidor de forma y distractor, los cuales tuvieron una probabilidad de mirada de .25. Cada ensayo tuvo una duración de 4250 ms. En la Figura 1 se presenta el diseño experimental y la distribución temporal de los ensayos. En el Anexo A se muestran las palabras y estímulos utilizados en esta investigación.

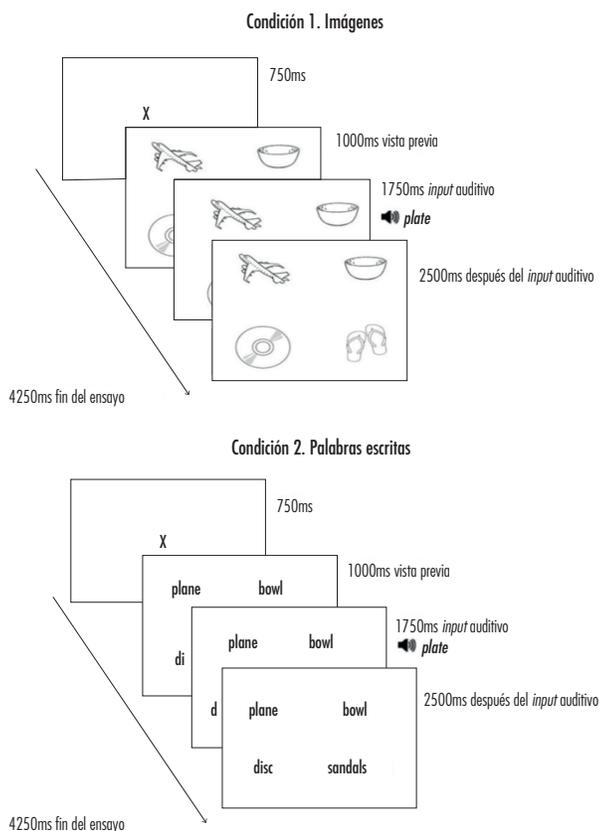


FIGURA 1. Diseño experimental y distribución temporal

2.6. *Procedimiento general*

Este estudio se realizó en línea durante la pandemia de COVID-19, por lo que los participantes ejecutaron las tareas desde casa. Los participantes eran alumnos inscritos en los cursos de inglés de la ENALLT, los cuales fueron invitados a participar por parte de sus profesores mediante un mensaje de correo electrónico. En dicho mensaje se incluía el consentimiento informado y los requisitos técnicos de la tarea en la plataforma RealEye para colaborar en el estudio.

Si los participantes entregaban por correo electrónico el consentimiento informado firmado, se les enviaba un video que explicaba la tarea de preferencia visual y también una liga para ingresar a la plataforma de RealEye y realizar dicha tarea. Una vez que los participantes accedían a la plataforma, se les presentaba por escrito las instrucciones de la tarea de preferencia visual, en las cuales se les pedía que se sentaran cómodamente a una distancia cercana a la pantalla, usaran audífonos, evitaran movimientos fuertes con la cabeza y mantuvieran una buena iluminación ambiental. De igual forma, se les recordaba que solo tenían que mirar la pantalla y no apartar la vista para que se registraran adecuadamente sus movimientos oculares. Al final de las instrucciones se incluía una casilla de “OK” para indicar que las habían entendido. Una vez que oprimían esa casilla, la tarea de preferencia visual comenzaba. Finalizada la tarea visual, la plataforma los dirigía automáticamente a la prueba LexTALE, realizada mediante un formulario de Google. Al terminar la prueba, los participantes debían introducir el puntaje que habían obtenido. La duración de la aplicación de la tarea de preferencia visual y del LexTALE fue de quince minutos aproximadamente.

2.7. *Tarea de rastreo visual*

Se analizó la proporción total de la fijación de la mirada hacia los estímulos visuales. Esta medida permite examinar la proporción de

la atención que un participante dirige mientras se presentan otros estímulos visuales de manera simultánea. La proporción total de la fijación de la mirada hacia los competidores y el distractor se calculó cada 100 ms. Se obtuvieron 3528 ensayos, sin embargo, se decidió eliminar aquellos que tuvieran datos perdidos, quedando un total de 1008 ensayos que fueron analizados.

3. Resultados

El objetivo de este estudio fue explorar la activación de la información fonológica en aprendientes de inglés sin experiencias de inmersión en la L2, a través de dos condiciones: imágenes y palabras escritas.

En esta sección presentamos los análisis de las ventanas de tiempo que reflejan la temporalidad de la activación del léxico de los participantes (los resultados estadísticos de cada condición pueden consultarse en el Anexo B). Cada ventana de tiempo representa 100 ms. La Figura 2 muestra los datos desde la previsualización de los estímulos visuales hasta los 1500 ms posteriores al inicio de la palabra auditiva. Se decidió dividir las ventanas temporales en dos secciones similares para explorar si la activación del léxico sucedía de forma temprana en la Ventana 1 (0 a 800 ms) o de forma tardía en la Ventana 2 (800 a 1500 ms). Esta decisión se sustenta en los resultados de los estudios mencionados anteriormente. Por ejemplo, el estudio de Huettig y McQueen (2007) cuenta con una ventana temporal de registro de 0 a 1000 ms, en la cual los procesos de activación de léxico se pueden observar dentro de los 200 a los 500 ms. Por su parte, en la investigación de Mishra y Singh (2016) la ventana temporal de registro abarca de los 0 a los 1200 ms, en cuyo caso se puede percibir la activación del léxico desde los 300 ms hasta los 1000 ms. Debido a estas diferencias en las ventanas temporales decidimos analizar 1500 ms para revisar nuestros resultados con la tarea de rastreo visual.

Igualmente, se revisó que los datos cumplieran con una distribución normal para decidir si el análisis de datos se realizaría

con pruebas paramétricas o no paramétricas. La prueba Kolmogorov-Smirnoff mostró que los datos no tenían una distribución normal, por lo que se utilizó la prueba no paramétrica de rangos de Wilcoxon para medidas repetidas. Se utilizó un valor hipotético de .25, que corresponde a la probabilidad de la fijación de la mirada (PFM) hacia los estímulos visuales. Los valores $>.25$ representaron una mayor atención de la mirada hacia un estímulo.

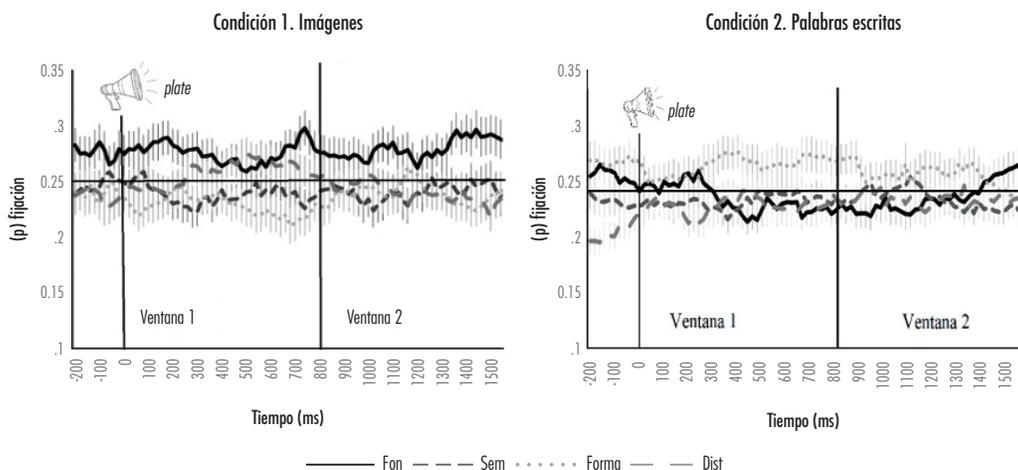


FIGURA 2. Ventanas de tiempo de cada condición

Nota: En el eje de la Y se muestra el PFM hacia los competidores. En el eje de la X se presentan las ventanas de tiempo. El inicio de la presentación de la palabra auditiva se ubica en la ventana de 0 ms. Fon: competidor fonológico, Sem: competidor semántico, Forma: competidor de forma, y Dist: Distractor. La línea horizontal indica la probabilidad de preferencia visual, es decir, los competidores arriba de esta línea tuvieron una mayor preferencia visual. Debajo de esta línea horizontal la preferencia visual fue azarosa. $n = 63$. Error estándar: 0.01.

La Figura 2 muestra que en la Condición 1. Imágenes, después del inicio de la palabra auditiva, como *plate* (plato), los competidores fonológicos, como la imagen de *plane* (avión), atrajeron miradas tempranamente, mientras las otras imágenes se mostraron sin efectos. De igual manera, se observa en la Ventana 2 que la proporción de la fijación de la mirada, en este caso la atención, se sostuvo

hasta la finalización de dicha ventana. A pesar de que se muestra una proporción de la fijación de la mirada $>.25$ en la Ventana 1 en el caso de los distractores, hay una disminución de esta correspondencia en la Ventana 2, lo cual sugiere que existió un efecto de inhibición hacia los distractores en esta última ventana de tiempo.

En contraste, en la Condición 2. Palabras escritas, después de que se escucha la palabra auditiva, como *plate* (plato), se observa una mayor proporción de la fijación de la mirada hacia las palabras escritas que fungieron como competidores de forma, como la palabra escrita *disk* (disco). A su vez, se muestra que en la Ventana 1, de los 0 a los 200 ms, existió una atención temprana hacia las palabras escritas que fueron competidores fonológicos. No obstante, después de los 300 ms se observa una inhibición de esta información fonológica y la proporción de la mirada fue mayor hacia el competidor de forma.

4. Discusión

Este estudio utilizó una tarea de preferencia visual para explorar la activación de la información fonológica en aprendientes de inglés como L2 sin experiencias de inmersión. En las condiciones utilizadas, se observó una mayor atención hacia los competidores fonológicos de la L2 en la Condición 1. Imágenes, mientras que en la Condición 2. Palabras escritas, se mostró una mayor atención hacia los competidores de forma.

Se esperaba que en ambas condiciones la atención hacia el competidor fonológico fuera mayor que hacia los otros competidores. La hipótesis se cumplió únicamente en la Condición 1. Estas diferencias de activación de la fonología en las condiciones experimentales podrían estar relacionadas con las inconsistencias ortográficas que afectan la fonología en el inglés. Como habíamos mencionado, el inglés es una lengua que se considera opaca a diferencia del español, que se concibe como una lengua transparente debido a que la relación entre la ortografía y la fonología es consistente, esto es, cada letra corresponde a un sonido (Suárez-Coalla,

Martínez-García & Carnota, 2020; Linder *et al.*, 2022). Entonces, para los aprendientes de inglés de nuestro estudio, esta relación ortográfica y fonológica pudiera ser desafiante debido a que su L1 tiene un sistema lingüístico distinto.

Ante esto último, resulta relevante el modelo BLINCS (Shook & Marian, 2013), que señala cómo la percepción, auditiva o visual, de una palabra en la L2 puede activar distintos niveles que decodifican e integran información ortográfica y fonológica almacenada para distinguir correctamente una palabra, incluso entre otras que pudieran ser similares en cuanto a letras y sonidos. En nuestro estudio, es interesante observar que los aprendientes de inglés mantuvieron su atención hacia la información fonológica en la Condición 1 sin cambios atencionales a los otros competidores, únicamente hacia el distractor. Esta atención a la imagen del distractor por parte de los participantes pudiera indicar que estaban confirmando que dicha imagen no estaba relacionada fonológicamente con la palabra percibida de modo auditivo, ya que posteriormente la atención se redirigió hacia el competidor fonológico.

Asimismo, nuestros resultados sugieren que los aprendientes de inglés como L2 distinguen las inconsistencias entre la ortografía y la fonología en inglés debido a su proficiencia en dicha lengua, lo cual confirma que la proficiencia es importante para activar la información fonológica de la L2. Además, se encuentra evidencia de que las imágenes también activan información a nivel lingüístico en una L2, como lo postula el modelo BLINCS (Shook & Marian, 2013).

En contraste, en la Condición 2. Palabras escritas, encontramos una atención sostenida, a través de la preferencia visual, hacia el competidor de forma, esto es, hacia la palabra escrita del competidor, por ejemplo, *disk* (disco), que está relacionada con el objeto percibido auditivamente, en este caso, *plate* (plato). Esta atención al competidor de forma estuvo precedida por el competidor fonológico, sin embargo, nuestra hipótesis no se cumplió en esta condición, lo cual es distinto a lo demostrado por Huettig y McQueen (2007), quienes observaron que el competidor fonoló-

gico tuvo mayor atención, por encima de los otros competidores, desde el inicio de la emisión de la palabra percibida auditivamente.

Los resultados de ambas condiciones sugieren que los enlaces léxicos de las palabras presentadas tienen una activación distinta según la modalidad de los estímulos: imágenes o palabras escritas. Es importante recordar que los participantes mostraron una proficiencia intermedia, lo que les permite comprender las palabras de uso frecuente en el inglés, como las que se ocuparon en este estudio. Sin embargo, la calidad en los enlaces léxicos (Perfetti, 2007) pudiera estar en construcción continua; se esperaría que a mayor proficiencia en la L2, se activarán los enlaces léxicos en mayor medida, como se ha reportado en el caso de la fonología, que se ha observado con mayor activación en bilingües con alta proficiencia en contextos de inmersión en la L2 (Mishra y Singh, 2016).

A su vez, y también en concordancia con el modelo BLINCS, se observa que los participantes activaron las propiedades físicas que asemejan a las de la palabra percibida auditivamente en la Condición 2. Palabras escritas. Los resultados muestran que las diferencias de presentación de los estímulos en las dos condiciones pueden afectar la activación del léxico, al menos en la L2 de nuestros participantes. Esto es, la activación de la fonología en la L2 se realiza en mayor medida cuando se presentan imágenes que cuando se hace con palabras escritas. En contraste, cuando se muestran palabras escritas se activa en mayor medida la información no lingüística, como la de los rasgos visuales de los competidores de forma. No obstante, en este estudio se pudo observar una activación de la fonología de la L2 por medio de imágenes, que solo se había demostrado en poblaciones bilingües (Mishra & Singh, 2016; Botezatu Guo, Kroll, Peterson & Garcia, 2022).

Las diferencias mostradas por los participantes en cuanto al acceso al léxico también pudieran estar relacionadas con la experiencia en la L2. Es decir, la calidad con la que se almacena la información, tanto lingüística como no lingüística, relacionada con una palabra en la L2 (Perfetti, 2007) dependerá de factores como la exposición y práctica formal e informal de la lengua de interés.

Por ejemplo, la experiencia en una lengua puede beneficiarse por actividades como la lectura por placer o la escritura de diarios en una L2 (Choi & Nunan, 2018). Por último, este estudio pudo demostrar una activación del léxico en dos condiciones diferentes en aprendientes de una L2 sin experiencias de inmersión.

5. Limitaciones

Se necesita realizar más estudios con distintas metodologías para enriquecer el campo de la activación del léxico. Una de las limitaciones de nuestro estudio es el tipo de aprendizaje de la L2 que se utiliza en el salón de clases, como el uso del diccionario (Lopera, 2019) o de material auténtico para aprender una palabra (Macedonia, 2015). Igualmente, pudiera ser de interés para futuras investigaciones contemplar una muestra equilibrada con respecto a la cantidad de participantes por sexo y edad para controlar un posible efecto de estas variables en los resultados. Otra variable metodológica por considerar es el uso limitado de cognados en los estímulos que pudieran afectar los resultados cuando se explora la activación del léxico. Este tipo de variables pueden explorarse en futuros estudios para conocer si existen similitudes con los resultados de la presente investigación.

6. Conclusión

Por medio de dos condiciones experimentales este estudio demostró la activación del léxico en una L2 a través de la preferencia visual en participantes sin experiencias de inmersión. Se hipotetizó que la activación de la información fonológica de las palabras en la L2 sería visible mediante la presentación simultánea de imágenes y una palabra auditiva relacionada con dichas imágenes. Los resultados demostraron que la información fonológica en una L2 es activada en mayor medida cuando se presentan imágenes relacionadas con una palabra en la L2. Por otro lado, como segunda hipótesis se consideró que la información fonológica de las palabras de la

L2 tendría una mayor proporción de atención que la información semántica y la de rasgos visuales. Nuestros resultados encontraron evidencia a favor de la segunda predicción en la Condición 1. Imágenes, pero no en la Condición 2. Palabras escritas. Esto es, cuando se presentan palabras escritas en una L2 y se escucha una palabra relacionada con dichas palabras, existe una activación hacia la información no lingüística como la de los rasgos visuales de los objetos. En conclusión, este estudio aporta evidencia sobre la activación de la fonología en una L2 en una población diferente a los bilingües en contexto de inmersión. Adicionalmente, se sustentan algunos de los modelos teóricos (p. ej., modelo BLINCS) que postulan a la fonología como un proceso inicial de la activación del léxico en una L2.

7. Referencias

- Alloppenna, Paul D.; Magnuson, James S., & Tanenhaus, Michael K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38(4), 419–439. doi: 10.1006/jmla.1997.2558
- Aschenbrenner, Andrew, & Yap, Melvin (2019). The influence of relatedness proportion on the joint relationship among word frequency, stimulus quality, and semantic priming in the lexical decision task. *Quarterly Journal of Experimental Psychology*, 72(10), 2452–2461. doi:10.1177/1747021819845317
- Bartz, Albert E. (1999). *Basic statistical concepts* (4a. ed.). Merrill: Merrill.
- Bice, Kinsey, & Kroll, Judith F. (2021). Grammatical processing in two languages: How individual differences in language experience and cognitive abilities shape comprehension in heritage bilinguals. *Journal of Neurolinguistics*, 58. doi: 10.1016/j.jneuroling.2020.100963
- Botezatu, Mona Roxana; Guo, Taomei; Kroll, Judith F.; Peterson, Sarah, & Garcia, Dalia (2022). Sources of variation in second and native language speaking proficiency among college-aged second language learners. *Studies in Second Language Acquisition*, 44(2), 305–330. doi: 10.1017/S0272263121000188

- Brybaert, Marc, & New, Boris (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. doi:10.3758/BRM.41.4.977
- Carascas-Ortiz, Haydée; Midgley, Katherine, & Frenck-Mestre, Cheryl (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones. *Psychophysiology*, 49(4), 531–543. doi: 10.1111/j.1469-8986.2011.01333.x
- Carreiras, Manuel; Ferrand, Ludovic; Grainger, Jonathan, & Perea, Manuel (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science*, 16(8), 585–589. doi: 10.1111/j.1467-9280.2005.01579.x
- Choi, Julie, & Nunan, David (2018). Language learning and activation in and beyond the classroom. *Australian Journal of Applied Linguistics*, 1(2), 49–63. doi: 10.29140/ajal.v1n2.34
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Estrasburgo: Language Policy Unit, CEFR. <https://rm.coe.int/16802fc1bf>
- de Groot, Annette M. B., & Hagoort, Peter (Eds.) (2018). *Research methods in Psycholinguistics and the neurobiology of language: A practical guide*. Oxford: Wiley Blackwell.
- Dijkstra, Ton, & van Heuven, Walter J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175–197. doi: 10.1017/S1366728902003012
- Gaskins, Irene W.; Ehri, Linnea C.; Cress, Cheryl; O'Hara, Collen, & Donnelly, Katharine (1996). Procedures for word learning: Making discoveries about words. *The Reading Teacher*, 50(4), 312–327. <http://www.jstor.org/stable/20201768>
- Gough, Philip B., & Tunmer, William E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10. doi: 10.1177/074193258600700104

- Harley, Trevor A. (2005). *The Psychology of language: From data to theory*. Hove: Psychology Press. <https://mohammedaljohani.files.wordpress.com/2014/10/psychology-of-language-from-data-to-theory.pdf>
- Henbest, Victoria S., & Apel, Kenn (2018). Orthographic fast-mapping across time in 5-and 6-year-old children. *Journal of Speech, Language, and Hearing Research*, 61(8), 2015–2027. doi: 10.1044/2018_JSLHR-L-17-0379
- Huettig, Falk, & Altmann, Gerry (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, 15(8), 985–1018. doi: 10.1080/13506280601130875
- Huettig, Falk, & McQueen, James M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460–482. doi: 10.1016/j.jml.2007.02.001
- Huettig, Falk; Olivers, Christian N. L., & Hartsuiker, Robert J. (2011). Looking, language, and memory: Bridging research from the visual word and visual search paradigms. *Acta Psychologica*, 137(2), 138–150. doi: 10.1016/j.actpsy.2010.07.013
- Kroll, Judith F., & Bogulski, Cari A. (2012). Organization of the second language lexicon. En Carol A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 4322–4330). Oxford: John Wiley. doi: 10.1002/9781405198431.wbeal0886
- Kroll, Judith F., & Stewart, Erika (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. doi: 10.1006/jmla.1994.1008
- Lemhöfer, Kristin, & Broersma, Mirjam (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavioral Research Methods*, 44, 325–343. doi: 10.3758/s13428-011-0146-0
- Levenshtein, Vladimir (1975). On the minimal redundancy of binary error-correcting codes. *Information and Control*, 28(4), 268–291. doi: 10.1016/S0019-9958(75)90300-9

- Li, Ping (2013). Computational modeling of bilingualism: How can models tell us more about the bilingual mind? *Bilingualism: Language and Cognition*, 16(2), 241–245. doi: 10.1017/S1366728913000059
- Lindner, Amanda L.; Wijekumar, Kausalai, & Joshi, R. Malatesha (2022). English spelling performance in writing samples among Spanish-speaking ELLs. *Journal of Learning Disabilities*, 55(2), 114–122. doi: 10.1177/0022219420982995
- Lopera, Sergio Alonso (2019). Effects of reading strategy and dictionary instruction in an undergraduate foreign language reading comprehension group. *Folios*, 50, 127–138. doi: 10.17227/Folios.50-10226
- Macedonia, Manuela (2015). Learning styles and vocabulary acquisition in second language: How the brain learns. *Frontiers in Psychology*, 6. doi: 10.3389/fpsyg.2015.01800
- McClelland, James L., & Elman, Jeffrey L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1–86. doi: 10.1016/0010-0285(86)90015-0
- Mishra, Ramesh Kumar, & Singh, Niharika (2016). The influence of second language proficiency on bilingual parallel language activation in Hindi-English bilinguals. *Journal of Cognitive Psychology*, 28(4), 396–411. doi: 10.1080/20445911.2016.1146725
- Nelson, Douglas L.; McEvoy, Cathy L., & Schreiber, Thomas A. (1998). The University of South Florida word association, rhyme, and word fragment norms. <http://w3.usf.edu/FreeAssociation/>
- Pérez, Miguel Ángel, & Navalón, Conrado (2003). Normas españolas de 290 nuevos dibujos: acuerdo en la denominación, concordancia de la imagen, familiaridad, complejidad visual y variabilidad de la imagen. *Psicológica*, 24(2), 215–241.
- Perfetti, Charles (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. doi: 10.1080/10888430701530730
- Rundell, Michael, & Fox, Gwyneth (2002). *Macmillan English dictionary for advanced learners of American English*. Oxford: Macmillan Education.
- Shirai, Yasuhiro (2018). *Connectionism and second language acquisition*. Nueva York: Routledge. doi: 10.4324/9780203118085

- Shook, Anthony, & Marian, Viorica (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition*, 16(2), 304–324. doi: 10.1017/S1366728912000466
- Smith, Alastair C.; Monaghan, Padraic, & Huettig, Falk (2021). The effect of orthographic systems on the developing reading system: Typological and computational analyses. *Psychological Review*, 128(1), 125–159. doi: 10.1037/rev0000257
- Snodgrass, Joan G., & Vanderwart, Mary (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. doi: 10.1037/0278-7393.6.2.174
- Suárez-Coalla, Paz; Martínez-García, Cristina, & Carnota, Andrés (2020). Reading in English as a foreign language by Spanish children with dyslexia. *Frontiers in Psychology*, 11, 1–13. doi: 10.3389/fpsyg.2020.00019
- Sunderman, Gretchen, & Kroll, Judith F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, 28(3), 387–422. doi: 10.1017/S0272263106060177
- Szekely, Anna; Jacobsen, Thomas; D'Amico, Simona; Devescovi, Antonella; Andonova, Elena; Herron, Daniel; Lug, Ching Ching; Pechmann, Thomas; Pléh, Csaba; Wicha, Nicole; Federmeier, Kara; Gerdjikova, Irina; Gutierrez, Gabriel; Hung, Daisy; Hsu, Jeanne; Iyer, Gowri; Kohnert, Kathryn; Mehotcheva, Teodora; Orazco-Figueroa, Araceli; Tzeng, Angela; Tzeng, Ovid; Arévalo, Analía; Vargha, Andras; Butler, Andrew C.; Buffington, Robert, & Bates, Elizabeth (2004). A new on-line resource for Psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247–250. doi: 10.1016/j.jml.2004.03.002
- The Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, 100(S1), 19–47. doi: 10.1111/modl.12301
- Thomas, Michael S. C., & van Heuven, Walter J. B. (2005). Computational models of bilingual comprehension. En Judith F. Kroll & Annette M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 202–225). Nueva York: Oxford University Press.
- Yee, Eiling, & Sedivy, Julie C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Ex-*

perimental Psychology, Learning, Memory, and Cognition, 32(1),
1–14. doi: 10.1037/0278-7393.32.1.1

8. Anexos

Anexo A. Palabras y estímulos utilizados en Condiciones 1 y 2

Palabra auditiva	Competidor fonológico	Competidor semántico	Competidor de forma	Distractor
1. beer /bɪr/ [cerveza]	bee /bi/ [abeja]	nut [nuez]	cup [taza]	dress [vestido]
2. bell /bɛl/ [campana]	bed /bed/ [cama]	door [puerta]	skirt [falda]	frog [rana]
3. candy /kændi/ [dulce]	candle /kændəl/ [vela]	almond [almendra]	fish [pescado]	shoe [zapato]
4. car /kɑːr/ [carro]	carrot /kærət/ [zanahoria]	bridge [puente]	rabbit [conejo]	pen [pluma]
5. castle /kæsl/ [castillo]	cassette /kæ'sɛt/ [cassette]	throne [trono]	hat [sombrero]	ruler [regla]
6. cat /kæt/ [gato]	catapult / kæ'tæ.pʌlt/ [catapulta]	bird [pájaro]	pear [pera]	ship [barco]
7. cheese /tʃiːz/ [queso]	chicken /tʃɪkən/ [pollo]	milk [leche]	sponge [esponja]	hammer [martillo]
8. cherry /tʃɛri/ [cereza]	chair /tʃɛr/ [silla]	strawberry [fresa]	bomb [bomba]	pencil [lápiz]
9. coke /koʊk/ [coca-cola]	cone /koʊn/ [cono]	pizza [pizza]	violin [violín]	shovel [pala]
10. letter /lɛ'tɛr/ [carta]	lettuce /lɛ'tɪs/ [lechuga]	box [caja]	mirror [espejo]	umbrella [paraguas]
11. peanut /piːnʌt/ [cacahuete]	piano /piːənoʊ/ [piano]	elephant [elefante]	snowman [hombre de nieve]	clip [sujetapapeles]
12. photo /fəʊto/ [fotografía]	phone /fəʊn/ [teléfono]	camera [cámara]	screen [pantalla]	soup [sopa]

(continuación)

Palabra auditiva	Competidor fonológico	Competidor semántico	Competidor de forma	Distractor
13. planet /'plænit/ [planeta]	plant /plænt/ [planta]	star [estrella]	button [botón]	cab [taxi]
14. plate /pleit/ [plato]	plane /plein/ [avión]	bowl [tazón]	disc [disco]	sandals [sandalias]
15. road /roud/ [camino]	rose /rouz/ [rosa]	truck [camión]	snake [serpiente]	coin [moneda]
16. soap /soʊp/ [jabón]	sofa /'soʊfə/ [sofá]	towel [toalla]	eraser [goma]	giraffe [jirafa]
17. medal /'medəl/ [medalla]	medicine / 'medisɪn/ [medicina]	trophy [trofeo]	ball [pelota]	sock [calcetín]
18. mouth /maʊθ/ [boca]	mouse /maʊs/ [ratón]	lipstick [lápiz labial]	banana [plátano]	scissors [tijeras]

Anexo B. Resultados estadísticos de los análisis de las ventanas de tiempo para cada condición

Ventana de tiempo	Condición 1. Imágenes															
	Fonológico						Competidor									
	Semántico			De forma visual			Semántico			De forma visual						
	<i>Mdn</i>	<i>Z</i>	<i>p</i>	<i>r</i>	<i>Mdn</i>	<i>Z</i>	<i>p</i>	<i>r</i>	<i>Mdn</i>	<i>Z</i>	<i>p</i>	<i>r</i>	<i>Mdn</i>	<i>Z</i>	<i>p</i>	<i>r</i>
Inicio de la palabra auditiva	0.23	-.884	.377	0.11	0.25	-.117	.907	0.02	0.27	.936	.349	0.12	0.21	-.420	.675	0.05
100 ms	0.25	-1.00	.316	0.13	0.27	-.192	.848	0.02	0.25	1.08	.279	0.14	0.23	-.520	.603	0.07
200 ms	0.23	-.888	.375	0.11	0.25	-.040	.968	0.01	0.23	.354	.724	0.04	0.27	-.219	.826	0.03
300 ms	0.25	-.634	.526	0.08	0.23	-.504	.614	0.06	0.27	2.06	.039*	0.26	0.21	-1.50	.134	0.19
400 ms	0.23	-.046	.964	0.01	0.27	.302	.762	0.04	0.27	1.53	.126	0.19	0.21	-2.56	.010*	0.32
500 ms	0.22	-.856	.392	0.11	0.27	.906	.365	0.11	0.27	1.29	.196	0.16	0.21	-2.47	.013*	0.31
600 ms	0.27	-.129	.897	0.02	0.23	-.576	.565	0.07	0.27	1.16	.246	0.15	0.21	-2.47	.013*	0.31
700 ms	0.22	-.703	.482	0.09	0.25	.435	.663	0.06	0.27	.619	.536	0.08	0.21	-2.17	.030*	0.27
800 ms	0.23	-.676	.499	0.09	0.29	1.63	.103	0.21	0.25	.786	.432	0.10	0.21	-2.31	.021*	0.29
900 ms	0.23	-.510	.610	0.06	0.25	.676	.499	0.09	0.27	1.73	.082	0.22	0.20	-2.61	.009*	0.33
1000 ms	0.23	-.716	.474	0.09	0.25	.158	.874	0.02	0.27	1.05	.294	0.13	0.27	-1.11	.264	0.14
1100 ms	0.23	-1.62	.105	0.20	0.29	1.40	.160	0.18	0.25	.378	.706	0.05	0.23	-1.40	.160	0.18
1200 ms	0.25	-.873	.383	0.11	0.25	.410	.682	0.05	0.27	.845	.398	0.11	0.21	-1.92	.054	0.24
1300 ms	0.21	-1.01	.311	0.13	0.27	.770	.442	0.10	0.27	1.39	.163	0.18	0.21	-1.66	.095	0.21
1400 ms	0.23	.004	.997	0.00	0.23	.388	.698	0.05	0.27	1.08	.277	0.14	0.21	-2.52	.012*	0.32
1500 ms	0.23	-1.11	.264	0.14	0.27	.339	.734	0.04	0.29	1.29	.197	0.16	0.23	-1.52	.128	0.19

Nota: ms = milisegundos; *Mdn* = Mediana; **p* ≤ .05; *r* = tamaño del efecto obtenido a través del coeficiente de correlación de Rosenthal (Bartz, 1999: 19).

Condición 2. Palabras escritas

Ventana de tiempo	Competidor															
	Fonológico				Semántico				De forma visual				Distractor			
	Mdh	Z	p	r	Mdh	Z	p	r	Mdh	Z	p	r	Mdh	Z	p	r
Inicio de la palabra auditiva	0.25	.013	.989	0.00	0.20	-1.46	.143	0.18	0.27	2.13	.032*	0.27	0.21	-1.46	.142	0.18
100 ms	0.27	.667	.505	0.08	0.21	-1.55	.120	0.20	0.27	.632	.528	0.08	0.23	-.883	.377	0.11
200 ms	0.27	1.20	.228	0.15	0.21	-1.22	.221	0.15	0.27	1.25	.210	0.16	0.21	-1.69	.090	0.21
300 ms	0.25	-.060	.952	0.01	0.21	-1.60	.110	0.20	0.27	1.70	.088	0.21	0.21	-1.09	.275	0.14
400 ms	0.20	-1.43	.152	0.18	0.21	-1.62	.105	0.20	0.29	2.41	.016*	0.30	0.25	-.282	.778	0.04
500 ms	0.20	-2.31	.020*	0.29	0.21	-.476	.634	0.06	0.27	1.49	.135	0.19	0.20	-.225	.822	0.03
600 ms	0.21	-1.81	.069	0.23	0.25	-.646	.518	0.08	0.27	1.71	.086	0.22	0.25	-.416	.677	0.05
700 ms	0.21	-1.27	.203	0.16	0.25	-.768	.443	0.10	0.29	2.00	.045*	0.25	0.20	-.877	.380	0.11
800 ms	0.25	-1.27	.203	0.16	0.21	-.525	.599	0.07	0.27	2.15	.031*	0.27	0.25	-.217	.828	0.03
900 ms	0.21	-1.98	.047*	0.25	0.25	-.199	.842	0.03	0.27	1.01	.311	0.13	0.27	.604	.546	0.08
1000 ms	0.23	-1.09	.274	0.14	0.25	-.454	.650	0.06	0.27	1.50	.133	0.19	0.25	-.267	.789	0.03
1100 ms	0.23	-1.82	.068	0.23	0.25	.585	.558	0.07	0.25	.534	.593	0.07	0.25	-.269	.788	0.03
1200 ms	0.23	-.834	.404	0.10	0.25	-.804	.421	0.10	0.27	1.70	.088	0.21	0.23	-1.30	.192	0.16
1300 ms	0.25	-.810	.418	0.10	0.23	-1.21	.224	0.15	0.27	2.17	.030*	0.27	0.25	-.899	.368	0.11
1400 ms	0.25	.136	.892	0.02	0.25	-1.62	.105	0.20	0.27	1.26	.207	0.16	0.25	-.497	.619	0.06
1500 ms	0.25	.768	.443	0.10	0.21	-1.61	.105	0.20	0.25	.306	.759	0.04	0.25	-.597	.551	0.08

Nota: ms = milisegundos; Mdh = Mediana; *p ≤ .05; r = tamaño del efecto obtenido a través del coeficiente de correlación de Rosenthal (Bartz, 1999: 19).

Can foreign language learners discover the rules of non-salient forms?

¿Pueden los aprendices de una lengua extranjera descubrir las reglas de formas lingüísticas no sobresalientes?

José Luis Moreno-Vega

Universidad de Sonora, Departamento
de Lenguas Extranjeras,
Licenciatura en Enseñanza del Inglés
joseluis.moreno@unison.mx

Ana Mónica Preciado-Sánchez

Universidad de Sonora, Departamento
de Lenguas Extranjeras,
Licenciatura en Enseñanza del Inglés
ana.preciado@unison.mx



Recepción: 13 de abril del 2023

Aceptación: 22 de junio del 2023

doi: 10.22201/enallt.01852647p.2023.77.1059

Abstract

Learners more easily notice some linguistic forms than others. This quasi-experiment investigates whether the effects of explicit-deductive instruction and explicit-inductive instruction are equally mediated by the perceptual salience level of target forms. It also examines whether each of these types of instruction promotes explicit knowledge, implicit knowledge, or both. The study was conducted with 65 university students aged 21 years old on average. Three conditions were tested: explicit-deductive instruction, explicit-inductive instruction, and a control. The results revealed minor learning effects due to explicit-inductive instruction of a target form with higher perceptual salience but no learning for a less salient target form. Explicit-deductive instruction resulted in learning both forms regardless of their salience level. The findings suggest that explicit-inductive instruction is suitable for teaching more salient forms, while explicit-deductive instruction seems necessary for teaching less salient forms. The pedagogical and theoretical implications of these findings are discussed.

Keywords: explicit instruction; deductive instruction; inductive instruction; implicit knowledge; explicit knowledge

Resumen

Para los aprendices, ciertas formas lingüísticas sobresalen más que otras. Este cuasiexperimento investiga si los efectos de la instrucción explícita-deductiva y explícita-inductiva son igualmente afectados por el nivel de sobresalencia perceptual de las formas lingüísticas meta, y si cada tipo de instrucción promueve el conocimiento explícito, implícito o ambos. El estudio se llevó a cabo con 65 estudiantes universitarios con una edad promedio de 21 años. Se examinaron tres condiciones: la instrucción explícita-deductiva, la explícita-inductiva y una condición control. Los resultados revelan que la instrucción explícita-inductiva generó un aprendizaje menor en el caso de una forma lingüística meta perceptualmente más sobresaliente, pero no hubo aprendizaje de la menos sobresaliente. La instrucción explícita-deductiva resultó en el aprendizaje de ambas formas independientemente de qué tanto sobresalen. Los resultados sugieren que la instrucción explícita-inductiva es más adecuada para enseñar formas más sobresalientes y la explícita-deductiva para las menos sobresalientes. Las implicaciones pedagógicas y teóricas de estos hallazgos son explicadas a detalle.

Palabras clave: instrucción explícita; instrucción deductiva; instrucción inductiva; conocimiento implícito; conocimiento explícito

1. Introduction

The role of instruction has attracted the attention of researchers studying second-language acquisition (SLA) and teachers over the past four decades. Specifically, there has been wide interest in knowing whether teaching formal rules of a second or foreign language (L2) allows students to use the target language more fluently and accurately during actual conversations or whether learners must discover rules by themselves to gain spontaneous knowledge. According to Roehr (2008), explaining a target rule tends to lead learners to better understand the form-meaning mapping of target items than instructing them to discover a rule on their own. However, this type of instruction does not encourage learners to create and test their own hypotheses about L2 forms; therefore, learners could potentially be less likely to play an active role in establishing form-meaning connections (Herron & Tomasello, 1992; Vogel, Herron, Cole & York, 2011). Learners are more likely to become active and engaged in processing target forms when they are guided to pay attention to a target linguistic form within a context and try to discover the underlying rule by themselves (Cerezo, Caras & Leow, 2016). Nevertheless, this also has its drawbacks as it requires learners to make inferences (Erlam, 2003), and not all learners are equally capable of discovering rules on their own (Vogel *et al.*, 2011). In addition, learners need to encounter a form many times over a long period to successfully discover its underlying rule by themselves (refer to Ellis, 1993).

2. Background literature

Basically, there are two types of explicit instruction in L2 learning: deductive and inductive. According to Decoo (1996), explicit-deductive instruction refers to the pedagogical intervention through which the language teacher explains a target form to students at the beginning of a lesson, and then the learners practice the form. On the other hand, explicit-inductive instruction entails providing

learners with examples of a target form and directing their attention to those forms while guiding them so that they can discover the underlying rules of the target form on their own.

There are two theoretical perspectives on the role of each type of instruction. One proposes that inductive instruction is more effective when teaching target forms that are difficult to explain (Krashen, 1981, 1982; Larsen-Freeman, 2009), have a higher complexity level, do not stand out much, and have many exceptions to their underlying rule (Cerezo *et al.*, 2016; Lai, Qi, Lü & Lyu, 2020: 498). If the underlying rule is too challenging for learners to understand, there is no point in teaching it; instead, learners may benefit more by engaging in input and output activities that lead them to discover the target rule. The opposing view proposes that explicit-inductive instruction is more suitable for teaching easy rules, and explicit-deductive instruction should only be adopted for difficult forms because learners may feel frustrated if they are not able to discover the target rule (Russel, 2014; Shirav & Nagai, 2022). If the rules of the target form are easy to understand and perceptually salient to the learner, there is a higher probability that learners will consciously and intentionally implement a strategy leading them to search for the target rule (Reber, 1993).

Empirical evidence has consistently shown that both types of instruction are effective at increasing learners' accuracy of target forms, but explicit-inductive instruction has been predominantly more effective than explicit-deductive instruction in cases that comprise an explanation of the target form after learners have attempted to discover its underlying rule (refer to Benitez-Correa, Gonzalez-Torres, Ochoa-Cueva & Vargas-Saritama, 2019; Cerezo *et al.*, 2016; Haight, Herron & Cole, 2007; Lai *et al.*, 2020; Shirav & Nagai, 2022; Takimoto, 2008; Vogel *et al.*, 2011). To our knowledge, only two studies (refer to Erlam, 2003; Robinson, 1995) have shown that explicit-deductive instruction was superior to guided-inductive instruction, but the beneficial effects were short-lasting. Additionally, two studies have revealed that explic-

it-deductive and explicit-inductive instructions are equally effective (refer to Hejvani & Farahani, 2018; Rosa & O'Neill, 1999).

These experiments have only assessed explicit knowledge, except for the one by Erlam (2003), which additionally assessed implicit knowledge. Explicit knowledge refers to factual knowledge about the language and requires awareness and controlled attention. It can be declarative if the learner is able to explain the use of a target form; it can be used in a conversation if the speaker deliberately monitors his/her own speech. By contrast, implicit knowledge is incidental, procedural, and unconscious and requires minimal effort from the learner; thus, it is primarily used during spontaneous talk. The distinction between these two types of knowledge is important because L2 learners require both to cope with the demands of using the target language.

3. The current study

This quasi-experiment is a follow-up of a previous study that tested the effectiveness of textual enhancement and explicit-deductive instruction (refer to Moreno-Vega & Preciado-Sánchez, 2023) assessed by implicit and explicit knowledge measures. Their findings revealed that explicit-deductive instruction was superior to textual enhancement. Textual enhancement was ineffective at helping learners gain implicit or explicit knowledge of the prepositions *in*, *on*, and *by* in the context of means of transportation. The study by (Moreno-Vega & Preciado-Sánchez, 2023), therefore, sheds some light on the effectiveness of providing learners with an explanation of a non-salient form. However, it does not focus on the effectiveness of explicit-inductive instruction.

Whether inductive instruction should be implemented for more or less salient rules is still an empirical question that deserves further research, as only a few SLA studies have attempted to answer it, with inconclusive findings. Besides, there is no consensus on whether the type of instruction (i.e., deductive or inductive) mediates the knowledge (explicit or implicit) acquired by learners.

The present study addressed these issues by attempting to answer the following research questions:

- 1) Are explicit-inductive and explicit-deductive instruction effective at promoting either implicit or explicit knowledge, respectively, of the prepositions *in*, *on*, and *by* in the context of means of transportation? If so, does the salience level of the target forms mediate the effectiveness of either type of instruction?
- 2) Can learners figure out any of the two target rules because of explicit-inductive instruction?
- 3) Can students receiving explicit-inductive instruction improve from pre- to post-test without being able to explain any of the target rules?

4. Method

4.1. Design

This exploratory study had a quasi-experimental design to test the effectiveness of two pedagogical treatments over four treatment sessions: a) an explicit-deductive condition in which the target forms were explained to the students at the beginning of each training session, followed by a reading containing the target forms; b) an explicit-inductive condition in which students were instructed at the beginning of each training session to pay attention to the target forms that were colorized in a text and try to discover two underlying rules; and c) a control condition in which students received no instruction to pay attention to the target forms; they just read the same enhanced texts as in the explicit-inductive condition. Implicit and explicit knowledge was tested before and after the treatment. None of the groups was informed about the specific purpose of the study, nor were they given any information about the treatment to which they had been assigned.

4.2. Participants

Data collection was conducted at a public university in Sonora. A total of 65 university students participated in this study. There were 41 female and 24 male participants aged 21 years on average. All participants were Mexican native speakers of Spanish. They were all enrolled in an intermediate English course taught five times a week in 50-minute sessions, and they all had a B1 level of English proficiency according to the Common European Framework of References for Languages. The exposure of participants to English outside of class was limited, as shown by a linguistic background questionnaire. Eight intact groups were considered in this quasi-experiment. Three of the intact groups were assigned to the explicit-deductive condition ($n = 25$), two to the explicit-inductive FFI condition ($n = 15$), and three to the control condition ($n = 25$). Initially, the explicit-inductive condition included 26 students, but 11 dropped out of the course. The explicit-deductive condition included 28 students at the beginning of the study, and three stopped attending the class; the control condition initially included 27 students, two of whom did not finish the course. Thus, the explicit-inductive condition ended with an unequal number of participants, in contrast to the other two conditions. Methodologically, having a more even distribution of participants across each condition would have yielded more robust and generalizable findings. Therefore, the nature of this study is exploratory, and it attempts to shed some light on this topic, but the findings should be interpreted with caution.

4.3. Target structures

Following the same criteria as in the quasi-experiment by (Moreno-Vega & Preciado-Sánchez, 2023), the linguistic target of this study was the prepositions *in*, *on*, and *by* used in the context of means of transportation. These means of transportation were selected because, despite their high frequency (refer to Lindstromberg,

2010), it takes most learners many years to learn how to use them properly, and many students do not achieve this goal partially because of their low perceptual salience. Perceptual salience refers to how difficult it is to notice a target form (Goldschneider & DeKeyser, 2001). A linguistic form can be perceptually salient depending on its typographical features (typographical salience) and can also be salient depending on whether its underlying rule is present or absent in learners' L1.

Typographical salience is driven by bottom-up processing because it is related to stimulus features (Shiffrin & Schneider, 1977). For example, the target forms in this study are considered to have different levels of typographical salience depending on how much they differ from one another. In the field of Psychology, Tudge, Brandt, and Schubert (2018: 677) have noted that “if a scene is composed mostly of identical objects but contains one object that is different from the others, then we are more likely to direct our gaze to the odd one out” (also refer to Donk & van Zoest, 2008; van Zoest, Donk & van der Stigchel, 2012). In addition, Tudge *et al.* (2018: 677) also explained that “salience is greater when the differences between an object and its surroundings are greater”. As such, this study was based on the premise that it is more difficult for learners to notice and understand the differences in form-meaning mapping of two target forms if they have typographical resemblance. For example, the prepositions *in* and *on* are very similar, so it is reasonable to assume that learners may not even detect that these are two different forms. However, *by* is typographically very different from *in* and *on*, making it more salient, and this may increase learners' chances of noticing its form-meaning function.

Some prepositions in L2 can be more salient than others if their meaning is relatively transparent in terms of learners' L1 and if learners can rely on their previous knowledge to understand the way the preposition is used. This type of salience is triggered by top-down processes because it is caused by “emotional, cognitive, and motivational” factors (Ellis, 2016: 343). That is, a form can also be more salient to some learners than others for reasons re-

lated to the learner and not only to the form. In the case of prepositions, L2 learners have to rely heavily on abstract concepts that are not necessarily evident contextually or available in their L1. For example, when using the target prepositions, there are basically two distinctions that should be made; the learner of English must know a) whether the means of transportation is specific or generic and b) whether the means of transportation is small or large. According to Lindstromberg (2010: 148), we can use *by* if we are referring to “a generic means of transportation... when we are not thinking of any particular machine” such as in the example: *I go to school every day by bus*. However, if we are referring to a specific means of transportation and both the speaker and the listener know it, then it is appropriate to use the prepositions *in* or *on*, as in the example: *We commuted to work in my car today*. If the learner knows that he/she is referring to a particular means of transportation, he/she must decide whether to use *in* or *on*. The preposition *in* is used when the speaker refers to a means of transportation small and privately owned. For example, it is correct to say: *in the car*. On the other hand, if the speaker refers to a means of transportation large and public, like a bus or a train, it is correct to use the preposition *on*.

There is an equivalent translation in Spanish for the use of each of the three target prepositions; for example, in Spanish, the preposition *en* can carry the same meaning as the three target forms in English, as in these examples: *en el automóvil = in the car*, *en el tren = on the train*, *by ship = en barco*. However, some rules in English that govern the use of these prepositions in the context of means of transportation are not applicable in Spanish. The rule governing the use of *by* vs. *in* or *on* exists in Spanish. It is plausible to say: *viajé en mi automóvil (I traveled in my car)* to express that you are referring to a specific car, and it is also plausible to say: *viajé en automóvil (I traveled by car)* to refer to any car. However, the rule governing the use of prepositions *in* and *on* in this context is missing in Spanish; instead, *en* is used to refer to any means of transportation regardless of its size or whether it is privately owned

or public, as in the examples: a) *voy a la playa en mi auto*, b) *voy al trabajo en el tren directo*. (For additional examples, refer to Díaz & Yagüe, 2019; Sampedro Mella & Estévez-Rionegro, 2021). By contrast, the preposition *by* has an equivalent meaning in Spanish, as shown in the following example: *I go to school by bus*, *Voy a la escuela en autobus*. The target form *by* translates to Spanish as *en* without an article before the noun that follows.

A pilot study was conducted before the present study to test the tasks and ensure that the students would have enough time to complete them during each training session. Fifteen university students (six females, nine males) aged between 19 and 23 years (mean = 20.93 years, *SD* = 1.87) participated in this pilot study. Five participants were assigned to each training condition. The training conditions and the pre- and post-tests in the pilot study were the same as those in the present study. The pilot study revealed that learners were indeed more likely to notice the preposition *by* than the prepositions *in* and *on*.

4.4. Measures

Reliable ways to assess explicit and implicit knowledge are currently available (refer to Elder, 2009; Erlam, 2009; Godfroid, Loewen, Jung, Park, Gass & Ellis, 2015; Loewen, 2009). One limitation to keep in mind, as Ellis (2009) points out, is that there are no pure measures of explicit and implicit knowledge. This study measured explicit knowledge by implementing the same fill-in-the-blank (FIB) test and the metalinguistic knowledge test (MKT) used in the study by (Moreno-Vega & Preciado-Sánchez, 2023) and the same timed grammaticality judgment test (TGJT) to measure implicit knowledge.

4.4.1. Fill-in-the-blank test

According to Peters (2016), the fill-in-the-blank (FIB) test allows learners to think about the form-meaning function of target forms.

It is an untimed test, potentially enabling students to think about language rules. Thus, it is considered a measure of explicit knowledge. According to Tremblay (2011), learners' lower grammar and lexical proficiency levels can be assessed through fill-in-the-blank tests. Thus, it is plausible for learners to gain some explicit knowledge of a target form even if they are not fully capable of explaining it. Learners may have a learning effect on a FIB test without having any learning effect on a metalinguistic knowledge test. This was why the FIB test was implemented as part of the explicit knowledge assessment. We wanted to ensure that learners' explicit knowledge was tested at both lower and higher metalinguistic levels. The FIB test comprised 25 items. Fifteen were target items and ten were distractor items. Five target items assessed the use of *by* and ten items the use of *in* vs. *on*; additionally, five items evaluated the use of *in* and five the use of *on*.

4.4.2. Metalinguistic knowledge test

The participants in each condition completed the metalinguistic knowledge test (MKT) at the end of each training session. They were given a handout requiring them to write the rules of each target item found in the text they had read. More specifically, participants in the control and the explicit-inductive instruction conditions were instructed to write down the cases in which it was appropriate to use each colorized word (*in*, *on*, and *by*). They could use simple language and had no time limit to complete this test. To check whether participants in the explicit-deductive condition were able to understand the explanation of the target rules and remember them, they were told to write down the rules they had been explained at the beginning of each training session.

4.4.3. Timed grammaticality judgment test

The timed grammaticality judgment test (TGJT) mainly assesses implicit knowledge according to several psychometric studies that

have tested its construct validity through factor analysis (refer to Bowles, 2011; Ellis, 2005, 2009; Gutiérrez, 2013; Kim & Nam, 2017). However, Suzuki (2017) has claimed that TGJTs assess explicit procedural knowledge because they attract attention to form. Nevertheless, even if learners relied partially on explicit procedural knowledge to make their judgments, this does not mean they are not using their implicit knowledge to some extent. In addition, implicit and explicit procedural knowledge are functionally similar, and either approach can enable learners to cope with the demands of using L2 in real time.

The rationale for using the TGJT instead of other measures of implicit knowledge was that, according to Kim and Nam (2017), it is cognitively less demanding on students than other tests, such as the oral elicited imitation test (OEIT), and students at Beginning or Lower Intermediate stages are usually developmentally ready to use the automatic cognitive processes needed to complete a TGJT. Our participants had an intermediate proficiency level but had limited opportunities to practice the L2 outside the classroom. Therefore, we considered that the OEIT would be too challenging for them as it would have likely required higher L2 processing levels compared to those normally used by students when they practice English in the classroom.

The TGJT contained 26 items, 16 of which were target items and ten were distractors. Six items (three grammatical and three ungrammatical) evaluated the use of *by*. Similarly, ten items (five grammatical and five ungrammatical) assessed the use of *in* vs. *on*; five items assessed the use of *in* and the other five the use of *on*. Distractor items (50% grammatical and 50% ungrammatical) assessed the use of tag questions, yes/no questions, simple present conjugation, and the past progressive tense. We recorded the voice of a native English speaker while he read each TGJT item. We decided to use an aural TGJT because, according to Kim and Nam (2017), this modality triggers stronger implicit knowledge than written TGJTs. Also, a factor analysis conducted by Spada, Shiu, and Tomita (2015) revealed that aural TGJTs correlate strong-

ly with implicit knowledge, whereas written TGJTs tend to correlate closer with explicit knowledge.

Participants were given six seconds to listen to each target item and judge whether it was grammatical or ungrammatical. We decided that six seconds were sufficient based on a previous pilot study with five native English speakers. In addition, we followed the criteria implemented by Ellis (2005) and Zhang (2015), who established the time limit for the TGJT based on the mean number of seconds it took a group of native speakers to make a judgment, plus an extra 20% of time. In our pilot study, participants took an average of five seconds to listen to each sentence and select one of two options: grammatical or ungrammatical. Following the same procedure as in the studies by Ellis and Zhang, we agreed that we would give the participants one extra second for each item, considering that they were not native English speakers like the participants in the pilot study. Besides, we pilot-tested the TGJTs a second time with L2 learners to ensure that six seconds would be a reasonable time. We found that, on average, six seconds enabled learners to listen to a sentence and make a judgment while limiting their access to linguistic rules.

4.5. Procedures

The data collection process was completed in six sessions. Each session lasted twenty minutes and was separated from the others by a one-week interval. The participants were given an information sheet and a consent form at the beginning of their first training session, and every participant read and signed it. Next, the participants were given a pre-test, which included a fill-in-the-blank (FIB) test and a timed grammaticality judgment test (TGJT). Participants were unaware that they were completing a test. Instead, they were told that they were going to complete an exercise. In the four sessions that followed, participants in the control and explicit inductive groups were given a text with the target forms colorized. Each text included seven instances of *in* or *on* and seven instances of *by*.

The prepositions *in* or *on* in the context of means of transportation were marked in red, and the preposition *by* was marked in blue. To know whether participants had focused primarily on the meaning while reading the texts, they had to write down a summary after reading each text. The only difference between the control and the explicit-inductive condition was that in the latter, participants were told at the beginning of each training session to discover two rules for using the colorized forms. The participants in the explicit-deductive group were given a brief explanation of the target forms at the beginning of each training session before they read each text. The target forms in the texts given to the explicit-deductive group were not colorized. After the participants in the three groups had read the text, they were asked to write down the target rules. In session six, all participants completed a post-test FIB, a TGJT, and a background questionnaire about their English acquisition process.

5. Results

The pre- and post-test score datasets for each group were tested for data distribution with the Shapiro-Wilk test. Participants' FIB scores of the target items *in* vs. *on* were tested separately from the target item *by*. The pre- and post-test FIB datasets corresponding to the target items *in* vs. *on* of the three groups were normally distributed, except for the pre-test dataset for the target items *in* vs. *on* of the control group. The pre- and post-test FIB datasets for the target item *by* were not normally distributed for any of the three groups. The same procedure was followed with the TGJT scores. The Shapiro-Wilk tests revealed that for the target items *in* vs. *on*, the pre- and post-test datasets for the control group and the post-test TGJT datasets for the inductive group were not normally distributed. By contrast, pre- and post-test TGJT datasets for the target item *by* were normally distributed for the three groups.

5.1. Within-group comparisons

To discover whether there were significant gains from pre- to post-test scores, the data of items corresponding to the *in vs. on* rule were analyzed separately from the items that tested the *by* rule. The FIB test comprised 15 items; of these, ten tested the *in vs. on* rule and five the *by* rule. A Wilcoxon signed-ranks test was conducted with the pre- and post-test FIB scores corresponding to items *in vs. on* for the control group, and paired samples *t*-tests were conducted with the scores for the inductive and the explicit-deductive groups. Within-group contrasts were analyzed with a Wilcoxon signed-rank test with the pre- and post-test FIB scores corresponding to the items *by* for the three groups. Tables 1 and 2 and Figures 1 and 2 illustrate the results of the FIB test for each target form.

Within-group contrasts of pre- to post-test TGJT scores for the *in vs. on* items were also analyzed using Wilcoxon signed-rank tests with the datasets for the control and explicit-inductive groups, and a paired samples *t*-test was used with pre- to post-test datasets for the explicit-deductive group. Paired samples *t*-tests were also used to test pre- to post-test scores corresponding to the target items *by* for the three groups. Tables 3 and 4 and Figures 3 and 4 illustrate the results of the TGJT for each target form.

5.2. Between-group comparisons

Since the datasets of pre- and post-test FIB mean scores for target items *in vs. on* failed to meet the assumption of a normal distribution, a Kruskal-Wallis test was conducted. No significant difference was found on pre-test datasets between any of the three groups ($X^2_{(2, 65)} = 0.464, p = 0.793$), nor on the post-test ($X^2_{(2, 65)} = 4.491, p = 0.106$). The datasets of pre- and post-test FIB mean scores for items *by* were also not normally distributed; again, a Kruskal-Wallis test was conducted. Pre-test scores were not significantly different between groups ($X^2_{(2, 65)} = 1.292, p = 0.524$), but the opposite was found for post-test scores, i.e., a significant difference be-

tween groups ($X^2_{(2, 65)} = 8.215, p = 0.016$). Mann-Whitney U tests were conducted to identify the specific groups that differed. After performing a Bonferroni adjustment for pairwise comparisons, alpha was set at < 0.016 . No significant difference was found between the control and the inductive groups ($U = 181.50, z = -0.176, p = 0.860, R = -0.021, r^2 = < 0.001$) nor between the inductive and the deductive groups ($U = 120.00, z = -1.93, p = 0.053, R = -0.239, r^2 = 0.057$), although in the latter case, the difference approached significance. A significant difference was also found between the control and the explicit-deductive instruction groups ($U = 174.00, z = 2.75, p = 0.006, R = -0.341, r^2 = 0.116$).

The dataset of TGJT mean scores for items *in vs. on* did not meet the normality assumption either; therefore, a Kruskal-Wallis test was conducted. No significant difference was observed between groups regarding pre-test scores ($X^2_{(2, 65)} = 1.84, p = 0.399$) or post-test scores ($X^2_{(2, 65)} = 4.808, p = 0.090$).

Since the dataset of TGJT mean scores for the target item *by* met the normality assumption, a repeated-measures MANOVA was conducted to test for differences between groups. A significant difference was found over time between the groups (Wilk's Lambda = 0.870, $F_{(2, 65)} = 9.230, p = 0.003$). A post-hoc Tukey's HSD test showed that the pre- and post-test mean scores were significantly higher ($p = 0.045$) in the inductive group versus the control group. However, mean scores did not increase significantly ($p = 0.111$) in the deductive group versus the control group.

5.3. Answers to the research questions

R1. Are explicit-inductive and explicit-deductive instruction effective at promoting either implicit or explicit knowledge, respectively, of the prepositions *in*, *on*, and *by* in the context of means of transportation? If so, does the salience level of the target forms mediate the effectiveness of either instruction?

Explicit-inductive instruction did not promote either implicit or explicit knowledge of the target form *in vs. on*, but explicit-de-

ductive instruction was effective at significantly increasing learners' implicit and explicit knowledge of this target form. Table 1 shows that neither the control group nor the explicit-inductive group significantly increased the learners' explicit knowledge of the underlying rule for *in* vs. *on* from pre- to post-test on the FIB test, and their effect sizes were small in contrast to the significant increase in the explicit-deductive group, with a moderate effect size. Moreover, the explicit-deductive condition attained a larger standard deviation of post-test scores than the other two conditions, indicating that not all students benefited equally from the explanation given by the teacher.

TABLE 1. Pre- to post-test FIB scores for the *in* vs. *on* rule

Condition	Pre-test		Post-test		Pre- to post-test			Min. score	Max. score	Test
	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>P</i>	<i>z</i> -score	<i>d</i>			
Control	25	4.44 (1.71)	25	4.92 (1.63)	0.248	-1.16	0.26	1	9	B
Inductive	15	4.80 (1.52)	15	5.40 (1.72)	0.346	-0.91	0.25	2	8	A
Deductive	25	4.36 (1.63)	25	6.20 (2.4)	0.001	-3.01	0.76	1	10	A

* Confidence interval = 95% * A = Paired-samples *t*-test B = Wilcoxon signed-rank test; max. score = 10 points

In the case of the preposition *by*, Table 2 illustrates how both explicit-inductive and explicit-deductive instruction led to gains of explicit knowledge of this form on the FIB test, but the explicit-deductive instruction had a larger effect size, and only the gains by the explicit-deductive group were significantly greater than those by the control group. In addition, since the control group also promoted a significant increase of explicit knowledge of the target form *by* with a moderate effect size, these results suggest that the preposition *by* can be learned through mere exposure over four sessions without requiring any type of form-focused instruction. An important observation shown in Figure 2 is that most participant scores of *by* in the three conditions had a floor effect on the pre-

test FIB, confirming that participants had no explicit knowledge of this form at the beginning of the study.

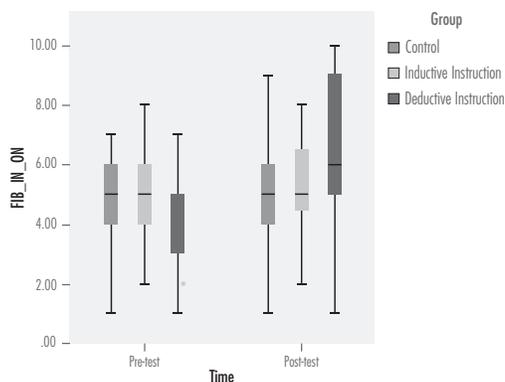


FIGURE 1. Comparison of pre- and post-test FIB scores for items *in vs. on*

TABLE 2. Pre- to post-test FIB scores for the *by* rule

Condition	Pre-test		Post-test		Pre- to post-test			Min. score	Max. score	Test
	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>P</i>	<i>z</i> -score	<i>d</i>			
Control	25	0.68 (1.22)	25	1.52 (1.85)	0.013	-2.47	0.55	1	5	B
Inductive	15	0.60 (1.40)	15	1.80 (2.08)	0.035	-2.11	0.6	0	5	B
Deductive	25	0.52 (1.23)	25	3.12 (1.81)	<0.001	-3.89	1.33	0	5	B

*Confidence interval = 95% *A = Paired-samples *t*-test B = Wilcoxon signed-rank test; max. score = 10 points

The salience level of target forms appears to mediate the effectiveness of inductive and deductive instruction, as illustrated in Tables 1 and 2. Evidence of this is that explicit-inductive instruction did not promote explicit knowledge of the *in vs. on* rule, but explicit-deductive instruction did. Further evidence is that both types of form-focused instruction (inductive and deductive) were effective at promoting explicit knowledge of the preposition *by*. Howev-

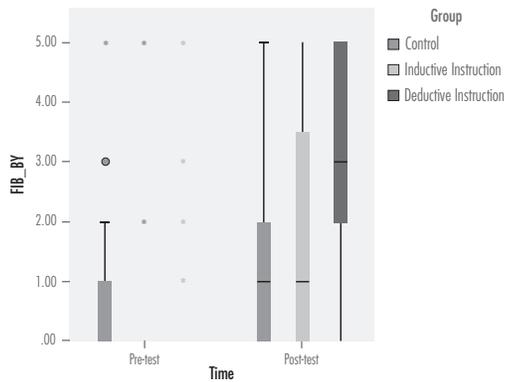


FIGURE 2. Comparison of pre- and post-test FIB scores for items *by*

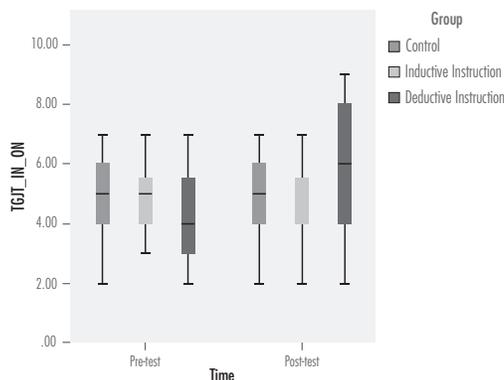
er, only explicit-deductive instruction facilitated the acquisition of implicit knowledge of the preposition *by*, as shown in Table 4. In this study, the less salient target rule (*in* vs. *on*) required explicit-deductive instruction for learners to pick it up. Furthermore, for learners to gain explicit knowledge of the more salient target form (*by*), they required neither grammar explanation nor any guidance from the teacher to discover the underlying rule; instead, they just needed to receive meaning-based input containing the target form over four training sessions.

A similar pattern occurred with the acquisition of implicit knowledge. Table 3 and Figure 3 show no learning effect on the TGJT from pre- to post-test in the control group or the explicit-inductive condition. By contrast, the explicit-inductive condition significantly improved the learners' scores with a moderate effect size. As in the post-test FIB, the post-test TGJT scores by the explicit-inductive group yielded a higher standard deviation than the control and the explicit-inductive groups, indicating that the score range was broader than for the other two conditions.

TABLE 3. Pre- and post-test TGJT scores for the *in vs. on* rule

Condition	Pre-test		Post-test		Pre- to post-test			Min. score	Max. score	Test
	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>P</i>	<i>z</i> -score	<i>d</i>			
Control	25	4.84 (1.28)	25	4.83 (1.06)	0.96	-0.05	0.01	2	7	B
Inductive	15	4.93 (1.22)	15	5.00 (1.46)	0.68	-0.41	0.03	3	7	B
Deductive	25	4.44 (1.30)	25	5.92 (2.14)	0.008	-2.48	0.58	2	9	A

* Confidence interval = 95% * A = Paired-samples *t*-test B = Wilcoxon signed-rank test; max. score = 10 points

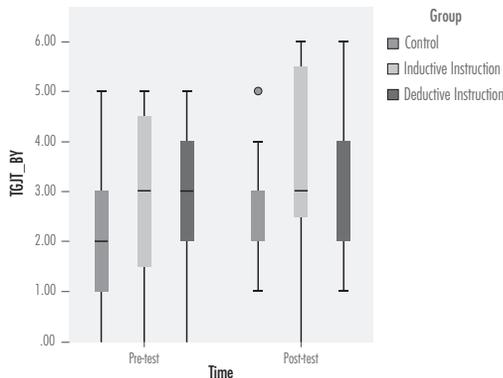
FIGURE 3. Comparison of pre- and post-test TGJT scores for items *in vs. on*

In the case of the preposition *by*, Table 4 and Figure 4 show that only the explicit-deductive group significantly improved its TGJT scores from pre- to post-test. However, the gains by the explicit-deductive group had a small effect size, indicating that the explanation by the teacher may not always result in increased implicit knowledge of the target form.

TABLE 4. Pre- and post-test TGIT scores for the *by* rule

Condition	Pre-test		Post-test		Pre- to post-test			Min. score	Max. score	Test
	N	M (SD)	N	M (SD)	P	z-score	d			
Control	25	2.12 (1.27)	25	2.85 (.97)	0.89	-2.03	0.46	0	5	A
Inductive	15	2.93 (1.71)	15	3.73 (1.87)	0.27	-0.89	0.3	0	6	A
Deductive	25	2.72 (1.54)	25	3.48 (1.48)	0.046	-1.82	0.42	0	6	A

* Confidence interval = 95% * A = Paired-samples t-test B = Wilcoxon signed-rank test; max. score = 10 points

FIGURE 4. Comparison of pre- and post-test TGIT scores for items *by*

R2. Can learners figure out any of the two target rules through explicit-inductive instruction?

As illustrated in Tables 5 and 6, none of the learners who received explicit-inductive instruction was able to discover the target rule for the prepositions *in* vs. *on* as measured by a metalinguistic knowledge test, and only one-third of the learners in this condition were able to discover the target rule for the preposition *by*. This indicates that both rules are challenging for learners. The underlying rule for *in* vs. *on* is particularly difficult to discover by

students without previous assistance. The rule for preposition *by* is more likely to be discovered by learners; however, even then, only a few students receiving explicit-inductive instruction on this target form were able to figure it out, as shown in Tables 5 and 6.

TABLE 5. Learning patterns of explicit knowledge and metalinguistic knowledge

Condition	N	Scores	FIB test <i>in vs. on</i>			FIB test <i>by</i>		
			Number of participants	Able to explain the target rule	Not able to explain the target rule	Number of Participants	Able to explain the target rule	Not able to explain the target rule
Control	25	Increased	9	0	9	12	0	12
		Decreased	7	0	7	2	0	2
		Didn't change	9	0	9	11	0	11
Inductive	15	Increased	6	0	6	6	4	2
		Decreased	3	0	3	2	1	1
		Didn't change	6	0	6	7	0	7
Deductive	25	Increased	18	16	2	20	18	2
		Decreased	5	3	2	1	0	1
		Didn't change	2	2	0	4	2	2

TABLE 6. Learning patterns of implicit and metalinguistic knowledge

Condition	N	Scores	TGIT <i>in vs. on</i>			TGIT <i>by</i>		
			Number of participants	Able to explain the target rule	Unable to explain the target rule	Number of Participants	Able to explain the target rule	Unable to explain the target rule
Control	25	Increased	8	0	0	12	0	12
		Decreased	9	0	0	6	0	6
		No change	8	0	0	7	0	7
Inductive	15	Increased	8	0	0	9	3	6
		Decreased	6	0	0	5	1	4
		No change	1	0	0	1	1	0

(continued)

TABLE 6. Learning patterns of implicit and metalinguistic knowledge

Condition	N	Scores	TGJT <i>in vs. on</i>			TGJT <i>by</i>		
			Number of participants	Able to explain the target rule	Unable to explain the target rule	Number of Participants	Able to explain the target rule	Unable to explain the target rule
Deductive	25	Increased	17	15	2	15	12	3
		Decreased	6	5	1	6	5	1
		No change	2	1	1	4	3	1

R3. Can students receiving explicit-inductive instruction improve from pre- to post-test without being able to explain any target rules?

Tables 5 and 6 show 15 learners that received explicit-inductive instruction. Six of them had gains in the test items *by* from pre- to post-test on the FIB test and nine on the TGJT. Two of the former and six of the latter were not able to explain the target rule during any of the sessions. This result shows that it is plausible for explicit-inductive instruction to enable learners to gain some explicit and implicit knowledge of target prepositions without being able to verbalize them. However, as previously mentioned, salience probably plays a role. Evidence of this is that most participants who increased their FIB and TGJT test scores for items *by* were able to explain the target rule, but none of those who increased their FIB and TGJT test scores for *in vs. on* was able to express the underlying rule. This suggests that they were aware of the target forms at the noticing level but probably not at the understanding level.

6. Discussion

The lack of a significant increase in the score of prepositions *in* and *on* obtained by the control group in both tests indicates that not all forms can be learned without explicit instruction. By contrast, the significant increase in the score of the preposition *by* on the FIB

test by this same group indicates that it is plausible for learners to pick up some forms through repeated exposure. Our results showed that it was easier for L2 learners to understand when *by* is needed in the context of means of transportation instead of *in* or *on*. One plausible explanation is that the forms *in* and *on* are typographically very similar, making the distinction between the two forms not salient for the learner. If learners cannot notice the distinction between the two forms, this may prevent them from engaging in syntactic processing. Instead, they may rely on semantic processing only. Since the difference in the meaning of the prepositions *in* vs. *on* does not affect the overall meaning of the sentence, it does not seem to be plausible for students to discover their referential meaning of the rule governing the use of *in* vs. *on* without receiving assistance or instruction from the teacher.

Another plausible reason could be that the rule for *by* is present in learners' L1 and L2 whereas the rule for *in* vs. *on* is not. Students could have had potentially greater concept availability of the meaning conveyed by the preposition *by* than by the prepositions *in* and *on*. As Russel (2014) stated, learners can pick up the rules of linguistic forms with a very transparent form-meaning connection simply by receiving abundant input containing the forms.

The results indicate that no participant in the inductive group was able to explain the rule governing the use of *in* vs. *on* at the end of any of the four training sessions. Four of the six participants who improved their scores of *by* in the FIB from pre- to post-test were able to explain the rule for *by* at the end of at least one of the training sessions. The remaining two participants were not able to explain the rule. This indicates that almost one-half of the students gained explicit knowledge of the form *by* due to explicit-inductive instruction, and more than half of these students acquired metalinguistic knowledge. One potential explanation as to why none of the students in the inductive condition was able to gain metalinguistic knowledge of the rule *in* vs. *on* is that there are no semantic cues that lead students to interpret the difference in use between *in* and *on* in this context. By contrast, when to use *by* instead of *in* or *on*

appears to be more obvious for students. The participants may have faced greater subjective difficulty when trying to identify the rule for *in* vs. *on* than for *by*. According to DeKeyser (2016), subjective difficulty refers to “the degree of difficulty experienced, for a given structure in a given context, by different learners”. Perhaps the participants lacked the semantic cues or background knowledge to engage in the top-down processing required to understand the underlying rule for *in* vs. *on*.

These results suggest that although it is plausible for inductive instruction to help learners figure out some target rules, there is no guarantee that all learners will identify them even if these rules have bottom-up and top-down salience. Additionally, inductive instruction does not seem to help learners gain metalinguistic knowledge of a rule governing a low-salience form (the use of *in* vs. *on*).

In the present study, explicit-deductive instruction was more effective than explicit-inductive instruction. Tables 1–4 illustrate that explicit-deductive instruction had larger effect sizes in all the tests. In addition, explicit-deductive instruction led to significant gains of explicit and implicit knowledge of both target forms from pre- to post-test, whereas explicit-inductive instruction only caused significant explicit knowledge gains of the more salient form *by*. Further evidence is that a larger percentage of the participants in the explicit-deductive group improved their mean scores from pre- to post-test. Eighteen participants (of 25) in the explicit-deductive group increased their FIB mean score of *in* vs. *on* items from pre- to post-test, and 20 increased their FIB scores of *by*. These results support Morgan-Short, Sanz, Steinhauer, & Ullman (2010), who highlighted that a particular type of instruction can be successful partly because of the characteristics of the target form. However, these findings should be interpreted with caution. Since the inductive group was much smaller than the deductive and control groups, these findings cannot be extrapolated to the teaching of other forms in different settings. Instead, this study is exploratory, and as such, it underscores the sometimes-necessary role of deductive instruction. Our findings also seem to indicate

that the type of instruction can be mediated by the salience level of the target forms.

These findings are partially consistent with those by Robinson (1995) in that explicit-deductive instruction was superior to explicit-inductive instruction. However, in our study explicit-inductive instruction was partially effective for one of the forms, whereas in Robinson's study it was not.

In this study, the finding that explicit-deductive instruction was superior to explicit-inductive instruction contrasts with those by Shirav and Nagai (2022), Lai *et al.* (2020), Cerezo *et al.* (2016), Vogel *et al.* (2011), Haight *et al.* (2007), Benitez-Correa *et al.* (2019). One likely reason is that in our study, explicit-inductive instruction consisted only of asking students to find the underlying rule of the colorized forms, and the instructor never gave learners an explanation of the target rules. By contrast, in the aforementioned studies, except for those by Haight *et al.* (2007) and Benitez-Correa *et al.* (2019), the participants received information on the target rules after trying to figure them out. In the study by Haight *et al.* (2007), participants were given feedback by the instructor, which also helped them to discover the target rules. Similarly, in the study by Benitez-Correa *et al.* (2019), students received indirect feedback in the form of recasts. By contrast, in our study, participants had to rely entirely on their ability to discover the underlying rules.

Therefore, the results of this exploratory quasi-experiment suggest that explicit-inductive instruction may be more effective than explicit-deductive instruction only if learners are given an explanation of the target rules at the end of the task or lesson or if students get feedback to help them know whether their discovery of the target rule is accurate. This type of inductive instruction is deemed guided-inductive instruction. These findings strongly emphasize the important role of explaining target rules of forms with low perceptual salience because they do not stand out typographically and their rule is nonexistent in learners' L1. Future research

could focus on other target forms that vary in their salience level and should ensure that all groups be of similar size.

7. Limitations

This study has some limitations. Although a timed grammaticality judgment test was used to assess learners' implicit knowledge, it would have been helpful to include confidence rating tests such as those used by Rebuschat (2013). However, it was not possible to implement them due to time restrictions. Also, having a larger sample size and giving learners a longer treatment would have been beneficial, likely yielding more robust results. Nevertheless, it was not possible because not all the groups in the Foreign Languages department were able to participate in the study, and those who did, had limited available time. Also, in the present study, it was not possible to determine which of the two factors — typographical salience and the existence of a rule in learner's L1 and L2 — has a greater impact in each type of instruction. Future studies may address this aspect by isolating these constructs.

8. Conclusion

The findings of the present study suggest that explicit-inductive instruction was less effective at enabling learners to determine the underlying rule of a form with low perceptual salience driven by bottom-up and top-down processing. Students were more successful at discovering a form with higher salience for being typographically different from other forms and governed by a rule available in learners' L1, such as the preposition *by*; even then, only very few of them did. With forms that lack these characteristics, explicit-deductive instruction is probably necessary. Also, there are instances when explicit-deductive or explicit-inductive instruction may be unnecessary, as some forms do not require the teacher's intervention beyond simply providing learners with comprehensible input.

However, because of the lower number of students in the explicit-inductive instruction condition, these findings cannot be extrapolated to the teaching of additional target forms in other settings. Further research is needed to confirm whether these findings transfer when teaching additional target forms in other languages to learners with diverse L1 backgrounds. This exploratory quasi-experimental study suggests that the salience level of target forms appears to be an important factor for L2 teachers to consider when deciding on the type of instruction to implement.

9. Supplementary material

The readings and tests used in this study can be requested at the following email address: joseluis.moreno@unison.mx

10. Acknowledgements

We appreciate the valuable feedback by two anonymous reviewers on an earlier version of this paper. We thank the students at Universidad de Sonora who participated in this study for their cooperation and time. We are also very grateful to the teachers who allowed us to collect the data.

11. References

- Benitez-Correa, Carmen; Gonzalez-Torres, Paul; Ochoa-Cueva, Cesar, & Vargas-Saritama, Alba (2019). A comparison between deductive and inductive approaches for teaching EFL grammar to high school students. *International Journal of Instruction*, 12(1), 225–236.
- Bowles, Melissa A. (2011). Measuring implicit and explicit linguistic knowledge. What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33, 247–271. doi: 10.1017/S0272263110000756
- Cerezo, Luis; Caras, Allison, & Leow, Ronald P. (2016). The effectiveness of guided induction versus deductive instruction on the development of complex

- Spanish *gustar* structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, 38(2), 265–291.
- Decoo, Wilfried (1996). The induction-deduction opposition: Ambiguities and complexities of the didactic reality. *International Review of Applied Linguistics in Language Teaching*, 34(2), 95–118.
- DeKeyser, Robert (2016). Of moving targets and chameleons: Why the concept of difficulty is so hard to pin down. *Studies in Second Language Acquisition*, 38(2), 353–363. doi: 10.1017/S0272263116000024
- Díaz, Lourdes, & Yagüe, Agustín (2019). Preposiciones. *Revista de Didáctica ELE*, 19, 1–28.
- Donk, Mieke, & van Zoest, Wieske (2008). Effects of salience are short-lived. *Psychological Science*, 19(7), 733–739.
- Elder, Catherine (2009). Validating a test of metalinguistic knowledge. In Rod Ellis, Shawn Loewen, Catherine Elder, Rosemary Erlam, Jenefer Philp & Hayo Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 113–138). Bristol: Multilingual Matters.
- Ellis, Nick C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5(3), 289–318.
- Ellis, Nick C. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition*, 38(2), 341–351.
- Ellis, Rod (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141–172.
- Ellis, Rod (2009). Implicit and explicit learning, knowledge and instruction. In Rod Ellis, Shawn Loewen, Catherine Elder, Rosemary Erlam, Jenefer Philp & Hayo Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3–26). Bristol: Multilingual Matters.
- Erlam, Rosemary (2003). The effects of deductive and inductive instruction on the acquisition of direct object pronouns in French as a second language. *The Modern Language Journal*, 87(2), 242–260.

- Erlam, Rosemary (2009). The elicited oral imitation test as a measure of implicit knowledge. In Rod Ellis, Shawn Loewen, Catherine Elder, Rosemary Erlam, Jenefer Philp & Hayo Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 65–93). Bristol: Multilingual Matters.
- Godfroid, Aline; Loewen, Shawn; Jung, Sehoon; Park, Ji-Hyun; Gass, Susan, & Ellis, Rod (2015). Timed and untimed grammaticality judgements measure distinct types of knowledge. Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, 37(2), 269–297. doi: 10.1017/S0272263114000850
- Goldschneider, Jennifer M., & DeKeyser, Robert (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 51(1), 1–50.
- Gutiérrez, Xavier (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, 35(3), 423–449. doi: 10.1017/S0272263113000041
- Haight, Carrie E.; Herron, Carol, & Cole, Steven P. (2007). The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom. *Foreign Language Annals*, 40(2), 288–310.
- Hejvani, Fatemeh, & Farahani, Mehrdad V. (2018). An investigation into inductive and deductive methods in teaching grammar to German EFL learners: A comparative study. *Global Journal of Foreign Language Teaching*, 8(2), 76–91.
- Herron, Carol, & Tomasello, Michael (1992). Acquiring grammatical structures by guided induction. *The French Review*, 65(5), 708–718.
- Kim, Jeong-eun, & Nam, Hosung (2017). Measures of implicit knowledge revisited. Processing modes, time pressure, and modality. *Studies in Second Language Acquisition*, 39(3), 431–457.
- Krashen, Stephen (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krashen, Stephen (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

- Lai, Chun; Qi, Xuedan; Lü, Chan, & Lyu, Boning (2020). The effectiveness of guided inductive instruction and deductive instruction on semantic radical development in Chinese character processing. *Language Teaching Research*, 24(4), 496–518. doi: 10.1177/1362168818805265
- Larsen-Freeman, Diane (2009). Teaching and testing grammar. In Michael Long & Catherine Doughty (Eds.), *The handbook of language teaching* (pp. 518–542). Hoboken: Wiley-Blackwell.
- Lindstromberg, Seth (2010). *English prepositions explained: Revised edition*. Amsterdam: John Benjamins.
- Loewen, Shawn (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In Rod Ellis, Shawn Loewen, Catherine Elder, Rosemary Erlam, Jenefer Philp & Hayo Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94–112). Bristol: Multilingual Matters.
- Moreno-Vega, José Luis, & Preciado-Sánchez, Ana Mónica (2023) The Role of Salience in Form-Focused Instruction in L2 Learning. [Manuscript in preparation].
- Morgan-Short, Kara; Sanz, Cristina; Steinhauer, Karsten, & Ullman, Michael T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, 60(1), 154–193.
- Peters, Elke (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113–138.
- Reber, Arthur S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York. Oxford University Press.
- Rebuschat, Patrick (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595–626.
- Robinson, Peter (1995). Aptitude, awareness and the fundamental similarity of implicit and explicit second language learning. In Richard Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 303–357). Honolulu: University of Hawaii Press.
- Roehr, Karen (2008). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, 29(2), 173–199. doi: 10.1093/applin/amm037

- Rosa, Elena, & O'Neill, Michael (1999). Explicitness, intake, and the issue of awareness: another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511-556.
- Russel, Victoria (2014). A closer look at the output hypothesis: The effect of pushed output on noticing and inductive learning of the Spanish future tense. *Foreign Language Annals*, 47(1), 25-47.
- Sampedro Mella, María, & Estévez-Rionegro, Noelia (2021). Teoría y didáctica de las preposiciones en español. *Revista Electrónica de Didáctica del Español*, 33, 162-184.
- Shiffrin, Richard M., & Schneider, Walter (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127-190.
- Shirav, Anastasia, & Nagai, Emi (2022). The effects of deductive and inductive grammar instructions in communicative teaching. *English Language Teaching*, 15(6), 102-123.
- Spada, Nina; Shiu, Julie L., & Tomita, Yasuyo (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning*, 65(3), 723-751.
- Suzuki, Yuichi (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, 38(5), 1229-1261.
- Takimoto, Masahiro (2008). The effects of deductive and inductive instruction on the development of language learners' pragmatic competence. *The Modern Language Journal*, 92(3), 369-386.
- Tremblay, Annie (2011). Proficiency assessment standards in second language acquisition research. "Clozing" the gap. *Studies in Second Language Acquisition*, 33(3), 339-372.
- Tudge, Luke; Brandt, Stephan A., & Schubert, Torsten (2018). Salience from multiple feature contrast: Evidence from saccade trajectories. *Attention, Perception, & Psychophysics*, 80(3), 677-690. doi: 10.3758/s13414-017-1480-9
- van Zoest, Wieske; Donk, Mieke, & van der Stigchel, Stephan (2012). Stimulus salience and the time-course of saccade trajectory deviations. *Journal of Vision*, 12(8), 16.

- Vogel, Séverine; Herron, Carol; Cole, Steven P., & York, Holly (2011). Effectiveness of a guided inductive versus a deductive approach on the learning of grammar in the intermediate-level college French classroom. *Foreign Language Annals*, 44(2), 353–380.
- Zhang, Runhan (2015). *Investigating linguistic knowledge of a second language*. Bern: Peter Lang.



Metáfora enactiva en la música didáctica virtual: una mirada crítica decolonial

Enactive metaphor in a virtual didactic musical performance: A decolonial critical look

Luis Escobar L.-Dellamary
Escuela de Ciencias Antropológicas
Universidad Autónoma de Sinaloa
luisescobar@uas.edu.mx

Luis Quintero Castro
Posgrado en Educación
Facultad de Ciencias de la Educación
Universidad Autónoma de Sinaloa
contactoluisquintero@gmail.com



Recepción: 1 de diciembre del 2022
Aceptación: 10 de julio del 2023
doi: 10.22201/enallt.01852647p.2023.77.1048

Resumen

A diferencia de la metáfora conceptual, la metáfora enactiva se define como un acto deliberado, la voluntad de enriquecer el mensaje para otorgar un papel activo al interlocutor en la integración del significado. El concierto didáctico virtual reúne la intención de ilustrar la música y las ventajas del medio digital. Presentamos el marco conceptual que nos ayuda a comprender el papel de la metáfora enactiva en la educación, los medios digitales y, en particular, en la educación musical. En la demostración de la pertinencia de una perspectiva postcognitiva para la descripción lingüística, integramos el lenguaje figurativo didáctico, la modalidad digital y la conceptualización corporizada de la música. En el análisis de los componentes multimodales del concierto, fincamos una visión crítica que argumenta una asociación ineludible entre el racionalismo como sustrato de la agenda colonialista en el pensamiento científico y la idea de lenguaje construida en el cognitismo estándar. Al evidenciar el sesgo ideológico en la construcción del objeto de estudio, buscamos advertir sobre la necesidad de transformar nuestras prácticas discursivas en las ciencias educativas y del lenguaje.

Palabras clave: enactivismo; metáfora; educación musical; YouTube; colonialismo científico

Abstract

Unlike conceptual metaphor, enactive metaphor is defined as a deliberate act. That is, the will to enrich the message to give the speaker an active role in the integration of meaning. In a virtual educational concert, the aim of illustrating music and the advantages of digital media converge. We outline the conceptual framework that helps understand the role of enactive metaphor in education, digital media, and music education. In demonstrating the relevance of a post-cognitive view for linguistic description, we articulate the didactic figurative language, digital modality, and embodied conceptualization of music. In the analysis of the musical performance multimodal components, we establish a critical vision that asserts the close association between rationalism as a substrate of the colonialist agenda in scientific thought and the idea of language constructed in standard cognitivism. By clearly establishing the ideological bias in this construction of the object of study, we seek to warn about the need to transform our discursive practices in language and education sciences.

Keywords: enactivism; metaphor; musical education; YouTube; scientific colonialism

1. Introducción

Nuestro propósito es presentar el concierto didáctico en su modalidad digital como territorio nativo para la metáfora enactiva que, junto con otros recursos audiovisuales, tiene un papel crítico en la ruptura de las prácticas discursivas racionalistas en la educación y el lenguaje. Nos enfocamos en la virtualidad por la forma deliberada en la que están dispuestos sus recursos comunicativos cuando hay una intención didáctica clara. Es decir, existe una selección consciente por parte del creador del video sobre la información y su presentación audiovisual.

Cuando decimos *racionalismo*, no nos referimos al sustrato del pensamiento colonialista (Flores & Rosa, 2022; Kallio, 2020; Posholi, 2020; Tupas, 2019) en el que se basa la conceptualización de la música como una manifestación lógica que puede representarse en una partitura, o que representa al lenguaje humano como gramática. La crítica no proviene de una controversia teórica, preferencia metodológica o de una rivalidad histórica entre escuelas de pensamiento, sino de un cuestionamiento a las consecuencias normalizadas de dicha perspectiva en temas de discriminación y supremacía cultural.

Durante mucho tiempo, hemos entendido a las posturas teóricas y a los marcos conceptuales como valorativa y políticamente neutros, meras preferencias temáticas o metodológicas, en el mejor de los casos. No obstante, cada vez es más claro que el juego de la normalidad y el privilegio en las ciencias sociales y humanas está sentado sobre las bases de una versión eurocéntrica de los objetos de estudio, entre ellos, la música, la cognición y el lenguaje. Estas prácticas discursivas son a veces producto de la costumbre y no necesariamente provienen de una postura teórica deliberada por parte de las personas que las reproducen. Sin embargo, el hecho de que se repitan frecuentemente sostiene la cosificación intelectual de la experiencia musical y del lenguaje en general.

Proponemos el análisis de un concierto didáctico virtual realizado en el marco de esta investigación para mostrar el papel de

la comunicación enactiva y las particularidades del medio audiovisual que tiene fines educativos. Aunque distintos trabajos, como revisaremos más adelante, han elaborado argumentos a favor de estas aproximaciones didácticas, así como de los medios digitales y los recursos multimodales (gestuales y verbales), consideramos que la articulación de los distintos conceptos que presentamos aquí según la perspectiva lingüística es una contribución necesaria para defender la adecuación descriptiva de este enfoque teórico crítico (que trata de ser socialmente responsable¹) y su papel en la transformación de nuestras prácticas discursivas.

Definimos *metáfora enactiva* (Gallagher & Lindgren, 2015) como un acto deliberado de reclutamiento de recursos expresivos para la integración de escenarios sensoriomotores que son más diversos que aquellos comúnmente proyectados en la comunicación cotidiana. Este acto articula intenciones lúdicas, ilustrativas y didácticas con el fin de entretener y promover la comprensión de un mensaje o concepto. A diferencia de otras nociones de metáfora y lenguaje figurativo (Dancygier & Sweetser, 2014; Lakoff, 1993), la perspectiva enactiva se basa en un principio fenomenológico (Gallagher, 2018) de adecuación empírica, es decir, describe al lenguaje como un evento de comunicación interactiva en el cual ningún recurso tiene, en principio, mayor relevancia que otro.

Como mencionamos al inicio, el caso de análisis seleccionado nos interesa debido al *control* que el creador del video puede ejercer sobre el guion, los ángulos de las cámaras de grabación y los efectos de posproducción. Sin embargo, estos no son los únicos elementos relevantes, puesto que el análisis multimodal revela los otros aspectos fuera del margen de control, lo cual cuestiona

¹ Asumimos que partir de la agenda postcognitivist (corpórea y enactiva) para realizar análisis del lenguaje se trata de una postura crítica cuando se hace conscientemente desde un enfoque decolonial, como lo muestran trabajos que utilizan las nociones de *epistemologías nativas*, *ontologías críticas* o *decolonización del conocimiento* (Kincheloe, 2011; Macrine & Fugate, 2021; Oyetade & Adegoke, 2017; Welch, 2019; Werner, 2020).

un enfoque racionalista de la comunicación. Podríamos decir que, al mostrar un análisis de la gestualidad en el marco de este planteamiento teórico-conceptual, también estamos reforzando la evidencia lingüística sobre por qué los medios digitales son más atractivos en temas educativos, y, por qué el enfoque racionalista no debería informar nuestro análisis ni nuestra planeación didáctica.

Estas tres líneas de exploración analítica y conceptual (la metáfora enactiva, la multimodalidad y, en particular, el papel de la gestualidad en el lenguaje) demuestran el valor didáctico del concierto virtual deliberadamente ilustrado y su papel en la transformación del discurso de la educación musical mediante plataformas virtuales o repositorios audiovisuales.

El resto del artículo está estructurado de la siguiente manera. En las siguientes subsecciones de la primera parte, describimos las nociones centrales de la investigación y justificamos la relación entre la elección de la perspectiva teórica y el caso de análisis. En §2, introducimos brevemente la metodología y mostramos las líneas de observación empírica que integran la interpretación del análisis, el cual articula la teoría, descripción y crítica decolonial (§2.2). En §3, concluimos con un repaso del papel de la metáfora enactiva en la educación musical, el lenguaje y la deconstrucción de la perspectiva eurocéntrica de estos fenómenos.

1.1. *El concierto didáctico*

Un concierto didáctico (Pérez Rodríguez, 2003; Richter, der Künste, Kraemer, Yob & Herzig, 1996) es la interpretación de piezas musicales enriquecidas con un conjunto de recursos ilustrativos diversos de un escenario narrativo. Típicamente, se trata de música instrumental de corte académico o la llamada “música clásica” en su sentido eurocéntrico. En el concierto el intérprete utiliza narración oral, efectos especiales, escenografía y otros medios, para proveer elementos a la integración creativa del espacio mental del espectador, en general, un público no experto en la materia. La composición sinfónica de Sergei Prokofiev *Pedro y el lobo* es

un ejemplo clásico, no obstante, se podría pensar que el concierto didáctico tiene sus raíces en todas las alusiones ostensivas y humorísticas que se han representado a lo largo de la historia de la música; por ejemplo, durante el periodo del Barroco (1600-1750) (Kliuchynska, 2022).

Una vez que reconocemos el desarrollo histórico de la música, desde su interpretación exclusivamente en vivo hasta los conciertos disponibles en línea, pasando por las grabaciones magnéticas y el disco digital (Thibeault, 2012), la noción de música didáctica se amplía. Podríamos incluir en esta categoría a películas animadas como *Fantasia* y *Fantasia 2000* de Walt Disney Pictures, o la famosa animación, basada en la obra de Richard Wagner, “¿Qué es la ópera, Doc?” (Warner Bros., *Merrie Melodies*).

Hoy en día, YouTube ofrece un amplio repertorio de música instrumental con elaboradas producciones audiovisuales (como el canal @thepianoguys, o el de la violinista @lindseystirling) que podrían incluirse en la categoría de didáctica musical. No obstante, no es el propósito de este trabajo realizar un ejercicio de clasificación en dicha plataforma. Por esta razón, se ha elaborado un concierto didáctico virtual con producción básica y sin efectos especiales para simplificar el análisis.

A continuación, justificamos nuestro interés en el medio digital y su papel en la educación desde la perspectiva de la multimodalidad y el análisis del lenguaje.

1.2. Los medios digitales en la educación musical

La utilidad de los medios digitales en la educación y, en particular, del repositorio y red social YouTube está ampliamente documentada (Jones & Cuthrell, 2011; Moghavvemi, Sulaiman, Jaafar & Kasem, 2018; Rahmatika, Yusuf & Agung, 2021). Existen trabajos que afirman su papel positivo en correlación con el desarrollo cognitivo y la activación integral de las modalidades de la inteligencia humana (Asmiarti & Winangun, 2018; Berk, 2009). Esto aplica

también a la educación artística, en específico, la dancística (Hong, Chen & Ye, 2020) y, por supuesto, la musical (Cayari, 2018; Rahmaturrizki & Sukmayadi, 2021; Waldron, 2012).

Entre las nociones que explican su atractivo y eficiencia didáctica destacan la accesibilidad, la aceptación (Rosenthal, 2018), la posibilidad de repetición del video (Jebe, Konietzko, Lichtschlag & Liebau, 2019, citado en Beautemps & Bresges, 2021) y la riqueza multimodal en la exposición de los contenidos. Es decir, gran cantidad de personas en entornos educativos aceptan, y a veces prefieren, las lecciones en video porque lo consideran un medio que les permite estudiar a su ritmo, que en general está estructurado como una lección corta (7~15 min.) y claramente expuesta, y, sobre todo, que utiliza deliberadamente los recursos multimodales.

Beautemps y Bresges (2021), que realizaron un estudio con más de 5000 participantes, llegaron a estas mismas conclusiones. Los autores analizaron el efecto que el presentador y sus características semióticas, sociales y enactivas tienen sobre las preferencias de las personas. Encontraron que cerca de la mitad de los participantes del estudio (49%) elegían y permanecían atentos a un video o canal de YouTube solo por las características del presentador. Al integrar una lista de catorce elementos que sintetizaban sus resultados acerca de qué caracteriza a un video educativo exitoso en ciencia, cuatro de ellos se relacionaban directamente con el aspecto multimodal del lenguaje: la riqueza semiótica del video, la calidad del video, la personalidad del presentador, esto es, gestualidad y claves de presentación de sí mismo (ing. *self-presentation*, ver p. ej. Huang & Jaszczolt, 2018; Kapidzic & Herring, 2015) y su uso del humor (lenguaje figurativo).

Estos hallazgos no son nuevos. Existe también una amplia literatura sobre el papel del lenguaje figurativo multimodal y el atractivo de la publicidad y los anuncios (Pérez-Sobrino, 2016; Pérez-Sobrino, Littlemore & Houghton, 2019; Pérez-Sobrino & Littlemore, 2020; Sweetser, 2017). Incluso, se ha estudiado cómo las particularidades culturales impactan en la interpretación de los

mensajes implícitos en los recursos expresivos de la publicidad impresa y en video (Forceville, 2017).

Aunque diversos autores parten de la idea que existe una estructura básica metafórica en el fondo de toda esa complejidad semiótica (Kövecses, 2013; Lakoff, 1993), también afirman que la comunicación enactiva va mucho más allá de la lógica de los recursos verbales, su composicionalidad o, incluso, de la explicitud de los mensajes.

En este sentido, el lenguaje multimodal y enactivo se entiende mejor en un entorno multidimensional y de espacios mentales integrados (Fauconnier & Turner, 1998). Su comprensión deriva, por decirlo así, menos de la imagen computacional del lenguaje y más de la idea que tenemos de la expresión poética (Sweetser, 2017),² según la cual los símbolos se disponen con el fin de motivar un papel activo en el interlocutor (espectador-participante), quien integra sus propios escenarios complejos de significación.

De tal suerte, podemos partir de una presuposición general: un video con fines educativos en YouTube tiene como uno de sus atractivos principales el recurso de la metáfora enactiva, según el amplio sentido en que la hemos definido. Asimismo, se posiciona como un recurso con ventajas didácticas con respecto a la exposición en los ámbitos presenciales (físicos) educativos, por motivos cognitivos, psicosociales y culturales. Las personas se sienten

² “Poetry may not be unique in such uses of partially explicit expressions and open meaning possibilities, creating multiple interactions between meaning spaces. Advertisers who evoke potential source and target frames certainly do not have to lay out the mappings between them to prompt construction of a metaphoric blend. Often parallel inferential structures guide the blending process...” [“La poesía puede no ser la única en los usos de expresiones parcialmente explícitas y con posibilidades interpretativas abiertas, creando interacciones múltiples entre distintos espacios de significado. Los publicistas que evocan esos potenciales marcos de fuente y meta ciertamente no tienen que desplegar el mapeo entre ellos para incentivar la integración de una metáfora. Frecuentemente las estructuras inferenciales paralelas guían el proceso de integración...”] (Sweetser, 2017, p. 5; las traducciones de las citas son nuestras).

menos estresadas frente a la pantalla de su computadora, además, la posibilidad de repetición del video ayuda a descargar la tensión en el proceso de aprendizaje. Sin duda, tanto la evidencia sobre el tema como el panorama presentado por la pandemia del COVID-19 lo posicionan como un medio de alta relevancia en la educación (Vladova, Ullrich, Bender & Gronau, 2021).

Con respecto al tema que desarrolla el presente trabajo, la literatura señala que existe una cuestión pedagógica en el modo en que los videos se estructuran (Beautemps & Bresges, 2021; Iskrú & Schulz, 2020), por lo que el análisis de sus componentes enactivos y multimodales, así como su relación con la transformación del discurso educativo en un nivel ideológico y teórico son fundamentales. Esto, como se ha visto, es un tema que atraviesa a las ciencias cognitivas y del lenguaje.

1.3. *Lenguaje, música y la agenda postcognitiva*

Parafraseando a Lawrence Zbikowski (2008), la comunicación musical no está contenida en las partituras de Bach, Beethoven o Glass, sino en los patrones sonoros y sensoriomotores que adquieren significado para las personas en la interacción colectiva. Ese acto de significación, por tanto, no se realiza en la bidimensionalidad y pentagramaticalidad de una partitura, sino en la emisión y orquestación de sonidos, gestos, posturas, pausas e intenciones. Sucede en un contexto específico, en una reunión de escenarios subjetivos que se articulan.

Desde diferentes perspectivas, algunos investigadores del lenguaje y la cognición han coincidido en una premisa de fondo: la experiencia del lenguaje es irreductible a su representación (Enfield & Sidnell, 2014; Gibbs, 2019; Kendon, 2014; Mondada, 2016). Ya sea que se hable de gramática, transcripción o comentarios sobre las generalidades de las prácticas comunicativas de las personas, esta máxima se sostiene incluso a un nivel básico. Es decir, constituye una versión lingüística de la postura ontológica según la cual las cualidades de la experiencia son irreductibles a

la materia, los procesos neurológicos o cualquier otra versión de aquello-que-existe-con-independencia-del-sujeto (Gallagher, 2020: 43; Kastrup, 2019: 22).

Las aplicaciones conceptuales y metodológicas en las que esta premisa se diversifica se pueden agrupar bajo la etiqueta de “post-cognitismo” o “cognición 4E” (Newen, Gallagher & De Bruin 2018). Se trata de una agenda común que cuestiona la presuposición del cognitivismo tradicional que señala que la interacción comunicativa es, esencialmente, un programa de computadora en el cerebro. En su versión más común, las 4 “E” provienen de las palabras en inglés *embodied*, *enactive*, *embedded* y *extended*, que se pueden traducir como corpórea, enactiva, mutualista y extendida. En otras versiones (Stilwell & Harman, 2021) se sustituye una de las “E” por *emotional* (emocional) e, incluso, se habla de una cognición 5E o 6E al añadirse la palabra *ecological* (ecológica).

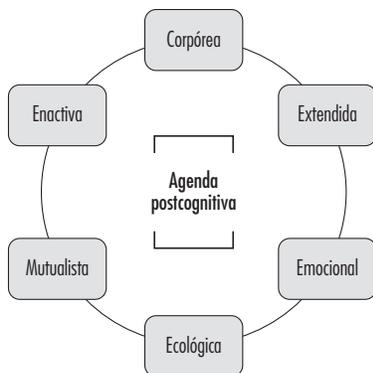


FIGURA 1. Representación 6E de la agenda postcognitiva

Todos estos términos redundan en la concepción, con distintos grados de compromiso ontológico, de la cognición (y el lenguaje) como algo que sucede y se define en un contexto particular. Es *corpórea* porque el cuerpo es el medio de interacción con el mundo, y ha evolucionado con él. Es *enactiva* porque, en tanto no es reductible a una estructura o programa de computadora, existe cuando se

realiza y no como una representación fija y abstracta. Es *mutualista* (Szokolszky, 2019: 18; Withagen, 2022: 116–120) porque no solo se trata de lenguaje en un contexto, sino de una interacción comunicativa situada cuyas dinámicas se explican únicamente en virtud de las características integrales de su contexto.

Es *extendida* porque la cognición y el lenguaje suceden a todo lo largo y ancho de la experiencia subjetiva concreta (no adentro ni afuera). Es *ecológica* porque todos los recursos de la interacción comunicativa situada la constituyen sin que podamos considerar, en principio, a unos primarios y a otros secundarios. Y, finalmente, es *emocional* porque la suposición de que la base explicativa es racionalista o lógica conlleva una expectativa eurocentrista (Flores & Rosa, 2022) y sin base empírica (Kastrup, 2018). Resulta más adecuado modelar la cognición según las dimensiones del fenómeno que representa y no como un sustrato arbitrario e ideológicamente sesgado hacia el racionalismo europeo del siglo XVIII. En otras palabras, postular una cognición lógica, emotiva, espacial y sensoriomotora, advirtiendo que el lenguaje, entonces, no puede ser ni sola ni primordialmente lógico-verbal.

Siguiendo esta línea, en nuestro trabajo nos enfocamos en la noción de *enacción* (Varela, Thompson & Rosch, 1991; Ward, Silverman & Villalobos, 2017). Nos proponemos argumentar a favor de una mirada situada y dinámica del estudio de la figuración o la representación metafórica en la enseñanza musical. Aunque puede sonar a cliché, cuestionarse sobre el estatus lingüístico de la música es un ejercicio conceptual inducido por estas nuevas formas de concebir la experiencia subjetiva que incluye a la comunicación. Es decir, si al lenguaje hablado o en señas no lo delimita una representación estructural y se trata de una inmersión integral en un contexto comunicativo intersubjetivo ecológico y global (que incluye la comunicación emocional), no presenta ningún problema conceptual o metodológico caracterizar a la música como un lenguaje.

Los estudios sobre gestualidad, por ejemplo, ya han introducido en la discusión sobre qué es y cómo funciona la comunicación el concepto de las *modalidades semióticas paralelas* (Kendon,

2014; McNeill, 2017; Okrent, 2002), esto es, las distintas maneras de expresarse en la interacción lingüística y —aún más relevante— los diversos modos de interpretar o construir el mensaje. No todo (o casi nada) se puede reducir a signos lingüísticos con significado que integran cadenas composicionales en las que cada palabra contribuye con un significado identificable, de tipo léxico o, por decirlo de otra manera, un significado básico (más su red de significación asociativa).

Los gestos, como ha mostrado la amplia literatura sobre el tema (McNeill, 2016; Pagán Cánovas, Valenzuela, Alcaraz Carrión, Olza & Ramscar, 2020; Wagner, Malisz & Kopp, 2014), son esenciales para el lenguaje y funcionan con una semiótica global; es decir, su interpretación es sensoriomotora, basada en actos de significación concretos y multidimensionales, más como imágenes corpóreas que como léxico. Según nuestro punto de vista, la observación detallada de diversos ejemplos muestra que el lenguaje es básicamente global y que en dicha globalidad existe un subconjunto de manifestaciones codificadas que se comportan como léxico.

Por su parte, la música conlleva una semiótica que, si bien no puede ser reportada verbalmente en todos los casos, no tiene por qué ser verbal para constituir lenguaje, sino todo lo contrario. Cuando reflexionamos sobre el acto comunicativo de un concierto didáctico, habría que preguntarnos, siguiendo con la lógica planteada desde los estudios gestuales, si no todo el lenguaje es, en un sentido bastante concreto, un concierto.

1.4. *Metáfora enactiva*

La metáfora se definía, en su sentido literario, como una estructura figurativa que enriquece la expresión de una idea en una forma fija del tipo “A es B” (Kövecses, 2010). El primer elemento se denominaba *meta* y, el segundo, *fuentes*. Es decir, A era aquello que se pretendía ilustrar, enriquecer, redimensionar con intenciones retóricas o figurativas, mientras que B era el terreno o el dominio metafórico que sumaba una dimensión de interpretación al elemen-

to meta. En un ejemplo como ‘eres la luz de mi vida’, A, en este caso una persona, es identificada con las propiedades de la luz, B, lo que genera un efecto figurado interesante, pues uno puede imaginarse visualmente el estado mental y emocional de regocijo que hace a una persona expresarse así de otra. Cabe señalar que este proceso no se trata de una simple operación de apropiación de las características físicas de la fuente, sino de la expresión lingüística de la experiencia subjetiva.

La metáfora conceptual, por su parte, es una idea que partió de una transformación histórica en la teoría lingüística, a su vez, considerada una revolución cognitiva (Lakoff, 1993). En pocas palabras, se trató de la reubicación (o localización) del programa central con el que funciona el lenguaje, esto es, su gramática, que pasó de ser una estructura abstracta y deslocalizada a una estructura en el cerebro con propiedades conceptuales, representaciones esquemáticas y funciones de integración de conceptos. El cerebro, a partir de entonces, se consideró una especie de computadora orientada a representaciones tipo “objetos conceptuales”; por ejemplo, el concepto [MANZANA], y no la palabra ‘manzana’, como unidad del léxico mental.

Es en dicho terreno del léxico conceptual que Lakoff y Johnson (1980) propusieron abandonar la definición tradicional de metáfora, basada en su identidad escrita o retórica formulada bajo una lógica simple (A es B), y sustituirla según la nueva representación cognitiva de la gramática: si a las palabras subyacen conceptos como su cara representativa y operativa en el cerebro, entonces la metáfora es una operación conceptual. Dicha operación sigue teniendo la estructura “A es B”, pero a nivel conceptual, por lo que en la expresión lingüística es necesario que presente esa forma. A su vez, la característica figurativa de la expresión constituye el resultado de la integración del dominio conceptual meta en el fuente: el dominio [PERSONA] queda integrado en el dominio [LUZ].

En este sentido, afirmar que la frase ‘Lucecita, ven conmigo’ también es metafórica porque el nombramiento de una persona con el diminutivo de *luz* implica el patrón metafórico conceptual

[UNA PERSONA ES LUZ (A es B)], representa una explicación que ya no tiene anclaje empírico “directo” en la expresión lingüística. Si bien esto no es ilegítimo como argumentación teórica, su pertinencia explicativa debe estar cuidadosamente justificada. Al proponer, por otra parte, que se trata de una operación conceptual, queda claro que dicha pertinencia depende de la noción de la computadora conceptual, ya que la identidad de las fichas de ajedrez, es decir, los conceptos, y sus reglas de interacción, existen o tienen pertinencia solamente en el entorno del tablero, el cerebro.

Por esta razón, la definición de metáfora de Lakoff y Johnson (1980) es dependiente del solipsismo cognitivo y pierde su pertinencia explicativa “fuera” de él. Es decir, depende del supuesto de que todo aquello necesario para que el lenguaje funcione debe ocurrir en el cerebro. Cualquier versión alternativa o más laxa en cuanto a los límites de la cognición echaría por la borda su consistencia como operación básica. Diversos autores han reconocido la ineludible obsolescencia de la metáfora conceptual como operación lingüística cognitiva (Cserép, 2014; Escobar, 2021; McGlone, 2007) una vez que se abandona el estricto confinamiento del lenguaje como resultado del cómputo neurológico.

De hecho, la noción misma de metáfora pierde sentido analítico cuando abarca prácticamente por entero el potencial figurativo del lenguaje y deja de distinguirse de otras figuras como la metonimia (Littlemore, 2015). Por ejemplo, es común encontrar en la literatura sobre el tema la interpretación de verbos como ‘entrar’ en ‘me entran unas ganas de llorar’ como metáforas conceptuales, esto según el argumento de que se está utilizando el patrón metafórico [LAS EMOCIONES SON SUSTANCIAS] que entran en el cuerpo, puesto que el verbo ‘entrar’ tiene un significado de base que es físico, como en ‘entrar en la tienda’ o ‘le entró agua al vaso’.

En contraste con esta postura, en nuestro trabajo asumimos que el potencial figurativo del lenguaje se manifiesta en la situación comunicativa. En este sentido, la metáfora enactiva (Gallagher & Lindgren, 2015), como acto deliberado de reclutamiento

de recursos expresivos, permite la integración de escenarios sensoriomotores más ricos en comparación con aquellos comúnmente proyectados en la comunicación cotidiana. Además, tiene intenciones lúdicas, ilustrativas y didácticas. Nos expresamos así para entretener, promover la comprensión de un mensaje o concepto abstracto, o procurar que nuestros interlocutores vivan una experiencia más completa y comunicativamente más eficiente.

Desde el enactivismo, el acto de *significación* es el acto comunicativo en sí y no subyace a él una estructura mental donde el significado se asigne antes de expresarse con palabras, señas o gestos. El término *metáfora*, entonces, no hace referencia a una operación mental, sino al acto de metaforización en sí (Müller, 2019). Paraphrasing Haan (2020), el término *metáfora enactiva*, como la denominación del acto deliberado de significación multimodal y figurativa, no se opone a la noción de metáfora como una operación interna o mental, se opone a una situación comunicativa cotidiana en la que no exista tal voluntad ilustrativa de parte del hablante. En pocas palabras, las metáforas son enactivas o no son metáforas.

Por tanto, el análisis de la metáfora enactiva en un concierto didáctico virtual que proponemos en la siguiente sección resulta pertinente para examinar las dinámicas de la figuración del lenguaje en el favorecimiento de la comprensión de contenidos musicales. Como hemos señalado, este formato promueve el interés público y la ruptura de barreras socioculturales propias de la colonización del arte y la creatividad, que descansan en prejuicios que fomentan la división social, como el prestigio o el esnobismo (Kallio, 2020; Leech-Wilkinson, 2020). Como pocos formatos, el concierto didáctico virtual es un espacio que favorece las enacciones metafóricas al nivel del teatro didáctico (Grajales-Acevedo & Posada-Silva, 2020; Stagg & Verde, 2019) y que además cuenta con todas las ventajas del medio digital.

2. Análisis y discusión

Para el análisis del componente gestual, retomamos una representación simplificada de sus propiedades semióticas basada en lo que McNeill (2005) denomina el continuo de Kendon. Este continuo es la propuesta de Adam Kendon (2004) de comparar las propiedades de la gestualidad con las atribuidas en la teoría lingüística a la lengua, en cuanto a su autonomía como signo lingüístico, su estabilidad forma-significado y sus otras características como símbolos expresivos del lenguaje, con el fin de clasificarlas. Con el paso del tiempo, la observación empírica y el cambio teórico nos han llevado a comprender que esta clasificación no funciona descriptivamente como grupos de gestos, sino como *focalizaciones semióticas* (ing. *semiotic foregrounding*). Es decir, intenciones o efectos ostensivos que enfatizan una posibilidad semiótica entre otras. Convencionalmente, esta clasificación resulta en, al menos, cinco focos semióticos en la gestualidad:

- Rítmicos. Aquellos cuyo foco principal es el ritmo sensorio-motor.
- Deícticos. Aquellos cuyo foco es la referencia a una entidad en el contexto de la interacción comunicativa vía un vector de señalamiento de algún tipo.
- Emblemas. Aquellos que tratan de ser un mensaje en sí mismos, como el gesto de ‘hola’ o ‘adiós’.
- Representativos icónicos. Aquellos cuyo foco está en la simulación de propiedades analógicas del escenario comunicativo, como cuando las manos forman ‘una esfera’ para representar una pelota.
- Representativos metafóricos. Aquellos que en la misma medida simulan propiedades, pero lo hacen en términos de una representación más abstracta, como cuando las manos forman ‘una esfera’ para representar no una pelota sino un tema o concepto.

Cada uno de estos focos tiene propiedades particulares en el discurso. Ahora bien, aunque puede haber gestos claramente rítmicos-icónicos-deícticos y solo unos pocos que no parezcan tener más de un foco semiótico, la integración de los recursos en la interacción comunicativa, incluyendo los intersubjetivos o de co-construcción del significado, define ciertos usos convencionales. Por ejemplo, los gestos de palma abierta hacia arriba (PAA) son, como su nombre indica, gestos hechos con la (o las) palmas abiertas hacia arriba.

Otros focos semióticos han sido divididos en *epistémicos* y *presentativos* (Cooperrider, Abner & Goldin-Meadow, 2018; Müller, 2004). En el primer caso, su función es expresar distancia del hablante con respecto a la veracidad/seguridad/evidencialidad de la información, y se acompañan comúnmente con un alzamiento de hombros (ing. *shoulder shrug*). En el segundo caso, proyectan un escenario ostensivo, ya que presentan al interlocutor información relevante o, incluso, distinguen niveles de presentación ostensiva del mensaje. Se trata, en suma, de un deíctico ostensivo con una forma emblemática de invitación, permiso o cesión de la palabra.

Antes de proceder con el análisis, es importante aclarar otro detalle técnico. Los gestos tienen una estructura articulatoria. La manera de nombrar esta estructura sigue convencionalmente estos pasos. La *unidad gestual* es el nombre de la actividad articulatoria entre dos estados de reposo, desde que las manos se empiezan a mover hasta que se detienen, por ejemplo. La *frase gestual*, por su parte, son los segmentos de actividad gestual con cierta autonomía semiótica, aunque no siempre con distinta forma; por ejemplo, cuando una figura pública saluda (ing. *hand-waving*) a las personas frente a ella y luego a las personas a su derecha. Cada frase gestual tiene un *núcleo* compuesto por un estado de *preparación*, uno de *movimiento* y uno de *detención*. Adicionalmente, se habla de un estado de *retracción* cuando la articulación regresa a un estado de reposo o de inicio de un nuevo movimiento. En este trabajo, combinamos este análisis con el de la lengua de señas (Escobar, 2008; Johnson & Liddell, 2010; Liddell & Johnson, 1989), y etiquetamos

con X a los estados transitorios (preparación y retracción), con M a los movimientos y con D a las detenciones, que deben ser estados articulatoriamente estables durante, al menos, tres cuadros de video.

2.1. *Multimodalidad y enacción*

El concierto didáctico que analizaremos, en el cual se interpreta el vals opus 8 n. 4 de Agustín Barrios, está disponible en YouTube (<https://youtu.be/ca4HtmOS0fk>). Segmentos ilustrativos de los argumentos presentados en este trabajo fueron transcritos por los autores en ELAN 6.4 (Lausberg & Sloetjes, 2009; The Language Archive, 2022), como se muestra en las Figuras 2, 3, 4 y 5.

Hemos dividido la exposición de los ejemplos en tres temáticas relacionadas con la integración y relevancia de la retórica en-activa deliberada del concierto:

- (I) la diferencia entre la representación lineal como racionalización del ritmo musical y su presentación sensoriomotora;
- (II) la importancia de la enacción en la construcción de los niveles ostensivos de la comunicación, y
- (III) el aspecto intersubjetivo, en este caso, a través de los recursos de autopresentación e involucramiento del presentador.

Como se muestra en la Figura 2, la transcripción toma en cuenta la gestualidad, definida previamente en términos metodológicos, y la expresión oral como fuentes primarias para los datos presentados, aunque en futuras exploraciones no se descarta la posibilidad de incluir más rasgos multimodales, como la dirección de la mirada o la postura del hablante, entre otros. La intención de este primer ejemplo es argumentar cómo ninguna representación escrita o explicación lineal, en este caso de la base rítmica de un vals, podría integrar la información que se muestra a partir de la metáfora en-activa o la enacción sensoriomotora.

En la Figura 3 se muestra la misma transcripción de la Figura 2, pero señalando: los bordes de las palabras orales basados en el



FIGURA 2. Ejemplo de transcripción en ELAN 6.4

espectrograma (A1-A9),³ los bordes de las frases gestuales [G1-G7] y los picos de intensidad basados en la curva generada en Praat 6.3.01 (1, 2, 3) (Boersma & Weenink, 1992-2022). La lógica de nuestra argumentación es la siguiente: decir que la base rítmica del vals es “a tres” no incluye información sobre coordinación compleja, la cual sí se transmite de forma multimodal; por ejemplo, no se trata de únicamente mover los pies con cada pulso, si no de señalar que también existen tiempos débiles y fuertes. Adicionalmente, podríamos preguntarnos si realmente creemos que sea posible reducir la información rítmico-musical a una representación verbal o a una gramática, como la expresada en las partituras (Harnoncourt, 2006).

Podemos observar, entre otras cosas, cómo en casi todas las ocasiones hay una distancia entre el borde de la frase gestual y el

³ Hemos propuesto la distinción con paréntesis y corchetes en correspondencia icónica con la forma en la que aparecen estas etiquetas en la Figura 3.

borde de palabra, una asincronía rítmica que queda clara cuando uno observa el segmento del video transcrito. Son excepciones los bordes de palabra (A6) y (A8), que coinciden con [G4] y [G6] respectivamente. Este orden asincrónico de la expresión enactiva se muestra también en la coincidencia entre bordes de palabra, bordes de frase gestual y picos de intensidad, los más claros son [G2], [G4] y [G5]. Mientras que, en otros casos, como [G1] y el borde de palabra en (A3), parecen más bien coincidir con un valle o contrapicado en la intensidad de la expresión oral.⁴

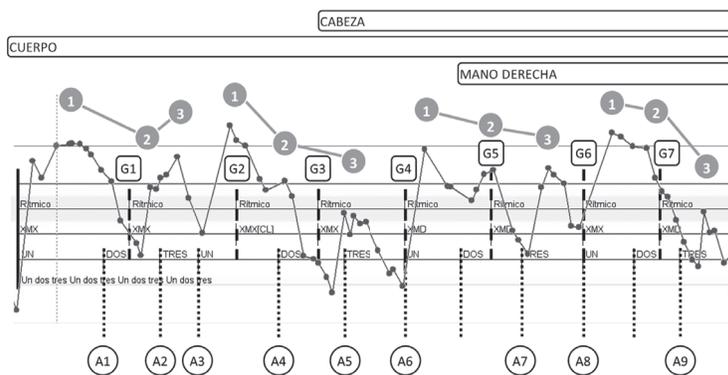


FIGURA 3. Análisis de la relación entre los picos de intensidad acústica, los bordes de palabra y los límites de las frases gestuales en una instrucción rítmica musical

Asimismo, en la Figura 3 se muestra cómo la expresión gestual del ritmo del vals transita por distintos focos semióticos. Al inicio, en la primera secuencia 1–2–3, el hablante mueve solo el torso y los

⁴ Está ampliamente documentado (Wagner *et al.*, 2014) cómo la coincidencia más frecuente se da entre el pico entonativo, y no el de intensidad, y el ápice gestual, es decir, el momento de mayor actividad articulatoria. Dado que nuestro argumento busca mostrar la complejidad que es irreductible a la lógica racionalista de las instrucciones escritas o puramente verbales no hemos enfatizado esta regularidad, que parecería permitir la simplificación algorítmica de la coordinación multimodal.

hombros, después, en la segunda secuencia, involucra más activamente la cabeza, con un movimiento lado a lado, y, finalmente, en la tercera ocasión que esta secuencia se repite, mueve la mano derecha. Este fenómeno constituye un patrón general en la gestualidad según el principio de no redundancia (Escobar, 2022: 23–24; McNeill, 2005), que afirma que ningún recurso en la multimodalidad enactiva es equivalente a otro ni, por tanto, a la repetición de una secuencia. Los hablantes, incluso en una aparente reiteración, usamos distintos énfasis o, como hemos dicho previamente, focalizamos diversa información semiótica.

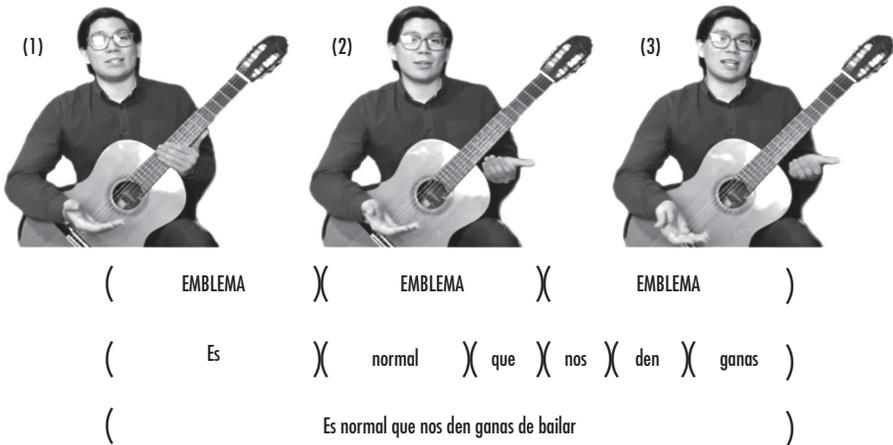


FIGURA 4. Tres niveles de ostensión gestual

En la Figura 4, se muestra el papel de la ostensión en la forma del gesto palma abierta hacia arriba (PAA) junto con los niveles de organización de este mensaje multimodal. Las tres apariciones de PAA son consideradas emblemas, como se observa en las líneas de transcripción representadas en la figura como corchetes redondeados. Se puede observar la coincidencia entre las frases gestuales y las palabras orales. El gesto PAA (1) que presentamos ocurre inmediatamente después de la enacción del ritmo de vals que se exhibe en las Figuras 2 y 3. El hablante abre un espacio semiótico-

discursivo distinto con este gesto presentativo para decir ‘es’. Posteriormente, con la mano izquierda PAA (2), acompaña ostensivamente a la expresión ‘normal’ como una forma de “mostrar un estado de cosas” (Cooperrider *et al.*, 2018: 10–11). En tercer lugar, integra ambas manos en un PAA bimanual (3), para expresar ‘nos den’, una ostensión presentativa de una concepción colectiva integradora de la dimensión pronominal de la expresión (a nosotros).



FIGURA 5. Relaciones multimodales *no lineales*

Por último, en la Figura 5, observamos otro tipo de asincronicidades que serían impredecibles desde la lógica racionalista de análisis lingüístico y que, adicionalmente, expresan matices sensoriomotores que resultan irreductibles a lo escrito o lo puramente verbal. El hablante dice ‘les voy a platicar la historia que yo personalmente me imagino’ para introducir un escenario metafórico enactivo en su presentación de la pieza musical en cuestión, no obstante, al decirlo sus gestos tienen un orden contrario al que se supondría desde una interpretación ‘secundaria’ o ‘dependiente’ de la gestualidad con respecto a la lengua oral (Wagner *et al.*, 2014). Este es un fenómeno ya observado en otras ocasiones (Escobar, 2022: 251; Goldin-Meadow, 2014: 8; McNeill, 2016: 115).

En este caso, el gesto metafórico (1), que parece asociarse conceptualmente con la expresión oral ‘imagino’, coincide tem-

poralmente con la expresión oral ‘la historia que’; mientras que el gesto deíctico (2) coincide, como es esperable, con el pronombre ‘yo’. Posteriormente, el hablante realiza el gesto deíctico en (3), con la palma abierta, para representar el ámbito de lo ‘personal’ (Cooperrider, 2014). Y, por último, en (4), acompaña a la expresión con un emblema tipo palma abierta que se dirige ‘hacia abajo’ como un atenuador (Caffi, 1999; Fraser, 1980), que expresa epistémicamente la postura del hablante con respecto al estatus de la narración que presentará, una especie de “es solamente mi versión de esa historia”.

2.2. Enacción y didáctica musical

Partiendo de las tres líneas de evidencia descriptiva presentadas, podemos ver cómo un ejercicio relativamente sencillo de análisis devela un contundente soporte empírico, según el enfoque lingüístico, para las transformaciones que busca la agenda postcognitiva en su asociación con la crítica decolonial. Asumimos, como se ha hecho desde los inicios de la lingüística cognitiva (Evans, Bergen & Zinken, 2007; Lakoff & Johnson, 1980; Langacker, 1991), que la base del lenguaje es la comunicación o, por decirlo en términos analíticos, que todos los recursos tienen un impacto en el significado de la expresión.

CUADRO 1. Contribuciones del componente gestual según las líneas de análisis presentadas

Rítmico	Evidencia de la complejidad no lineal de la interacción rítmica de los recursos multimodales y su impacto en la semántica global (sensoriomotora) del mensaje (Figura 3).
Ostensivo	Evidencia de la distribución de la atención en la interacción comunicativa, así como del escenario sensoriomotor de la expresión lingüística (“ser”, “ser normal”, “nosotros”, Figura 4).
Autopresentativo	Evidencia de la construcción del “sí mismo” narrativo (involucramiento subjetivo) e interlocutivo* (Figura 5).

* Esta diferencia se puede formalizar como, por un lado, lo *interlocutivo*, es decir, la llamada “pragmática indexical” (Jaszczolt, 2018), que alude concretamente a los recursos de referencia (en términos amplios) relacionados con la primera persona del discurso. Y, por el otro lado, el *involucramiento subjetivo* como la construcción de la narrativa del “sí mismo” en el escenario experiencial. Más cercano, en este segundo sentido, a lo que en la literatura se le conoce como “narrativa del yo” en su sentido filosófico y, particularmente, fenomenológico (Gallagher, 2017).

Como se muestra en el Cuadro 1, incluso la danza rítmica de los recursos tiene un impacto en la semántica cuando la concebimos como una integración global (*imaginística* o de espacios mentales) y sensoriomotora (corpórea) con un sustrato abstracto lógico-racional (semántico, en su sentido componencial), y no al contrario. Lo mismo sucede con la construcción ostensiva de los escenarios en frases como ‘es normal que nos den ganas de bailar’ en los distintos niveles de presentación del involucramiento subjetivo, así como en la integración de la narrativa del “sí mismo”.

Parte de las consecuencias interpretativas de la evidencia presentada por otras investigaciones, es que la comprensión de la música no puede seguir las mismas prácticas discursivas, según las cuales es básicamente matemáticas, partitura, reproducción de un guion. No se trata tampoco, como hemos observado en la Figura 3, de una sucesión azarosa de recursos multimodales. Existen finos patrones de coordinación gestual. La conciencia, por ejemplo, de que lingüísticamente hablando no existe tal cosa como la repetición de una expresión si se consideran los distintos focos semióticos que adquiere el canal sensoriomotor, debería transformar nuestra idea sobre la comunicación.

Esto, claro, sucede automáticamente en un medio audiovisual. No obstante, es importante insistir en que estas propiedades de la didáctica dejen de invisibilizarse en el discurso o se piensen como enriquecimientos accesorios o secundarios del mensaje, como ha sucedido en la historia de los estudios gestuales (de Beer Hogrefe, Hielscher-Fastabend & de Ruiter, 2020; McNeill, 2005; So, Kita & Goldin-Meadow, 2009).⁵ Juzgar ciertas propiedades del fenómeno

⁵ Resulta interesante cómo siguen coexistiendo perspectivas sobre el papel de la gestualidad en el lenguaje que no advierten las consecuencias ideológicas de sus posturas. Cuando se habla de redundancia entre la lengua oral y la gestualidad, se está perdiendo de vista que es empíricamente inadecuado juzgar que dos expresiones con condiciones materiales y semióticas distintas estén significando lo mismo. Por ejemplo: “there are many times when gesture and speech convey information that appears to be redundant: the gesture contributes information that is already fully specified in the speech” [“hay muchas

como centrales no en términos de la acotación del análisis, donde sería adecuado, sino en términos de su misma concepción (o naturaleza), se trata de una clara manifestación de un patrón generalizado de discriminación cultural.

La observación detallada de lo que sucede formalmente en la expresión de una intención didáctica musical nos hace cuestionar la historia y la tradición de la representación de la música, así como sus prácticas institucionales y costumbres educativas, que muchas veces han normalizado todo tipo de violencia (Fernández-Morante, 2018).

Por otra parte, en el análisis de las Figuras 4 y 5, podemos observar la importancia comunicativa de las expresiones ostensivas y de los significados globales que se menosprecian en la educación formal. En el concierto didáctico virtual, el presentador, partiendo de una noción enactiva de figuración lingüística vía su intención didáctica, provee al espectador de recursos que completan los niveles de comunicación del mensaje que se está construyendo: el escenario propuesto para la comprensión de la narrativa musical, las propiedades sensoriomotoras de un vals y no solamente su reducción numérica a tres y a un compás de $\frac{3}{4}$, así como las claves emocionales que legitiman la experiencia subjetiva de este acercamiento a la música académica.

La descripción de estos recursos de ostensión gestual de la complejidad rítmica, sobre todo al tratarse de una pieza asociada al baile,⁶ los niveles de la semántica global tanto de la descripción

ocasiones cuando la gestualidad y el habla expresan información que parece redundante: el gesto contribuye con la información que ya está completa en el habla”] (Abner, Cooperrider & Goldin-Meadow, 2015: 6). Mientras que, por otra parte, se afirma que “Speech and gesture are co-expressive but nonredundant in that each has its own means for packaging meanings” [“El habla y el gesto son co-expresivos pero no redundantes dado que cada uno tiene su propia estrategia de significación”] (McNeill, 2005: 91).

⁶ También la separación histórico-cultural de la música instrumental del entorno corpóreo (el baile, la danza y el festejo) durante los periodos romántico y moderno es motivo de análisis en cuanto a su sesgo colonial y racionalista (Haynes,

por parte del presentador del escenario narrativo de cada pieza como del involucramiento de sí mismo como sujeto que también experimenta la música que interpreta, son la regla en toda interacción lingüística. En el concierto didáctico virtual, estos recursos se focalizan como parte de una estrategia metafórica enactiva. Su carácter esencial en la comunicación debería modificar el papel que tienen comúnmente en los temas lingüísticos, cognitivos y educativos: pasar de ser enriquecimientos “obvios” y “secundarios” a ser elementos transformadores del discurso de la teoría y la práctica.

3. Conclusiones

En términos del lenguaje figurativo en general, Pérez-Sobrino y Littlemore (2020) ya habían identificado las características de un predictor consistente de la viralidad de los videos en redes sociales. Adicionalmente, estos autores, junto con la amplia literatura citada en su trabajo, encuentran efectos claros con respecto a la multimodalidad del lenguaje empleado (imágenes y palabras). Desde nuestro punto de vista, se trata de dos caras de la misma moneda enactivista. Cuando los videos hacen un uso intencionalmente bien organizado de los recursos multimedia, resultan más atractivos.

Gallagher y Lindgren (2015) focalizan, siguiendo la definición de metáfora enactiva, el papel de los contenidos multimodales en los contextos educativos formales y, sobre todo, su orientación semiótica, que favorece una comprensión integral de conceptos relevantes para distintos cursos y programas. Citan también un buen número de autores que muestran datos experimentales y observan criterios de significancia estadística de la correlación entre el acto deliberado de metáfora, la comprensión y el aprendizaje. Entre ellos, Glenberg (2008), quien señala los efectos positivos que

2007). No podemos negar que parte del discurso de la supremacía cultural que justificó las invasiones imperialistas se relacionaba con la separación ritual de las conductas que eran percibidas como “salvajes” o “incivilizadas”; entre ellas, la íntima asociación de la música y el cuerpo, sobre todo “el cuerpo” libre de la normatividad social.

se generan cuando los alumnos de materias como, por ejemplo, ciencia y matemáticas actúan y representan de manera sensorio-motora (corpórea) los experimentos y resultados que comparten en clase.

La consistente evidencia que clama una transformación radical de nuestra idea de aprendizaje, sobre todo en el contexto de la educación formal, que vaya de los contenidos estáticos (según la noción de *sitting metaphor* o metáfora estática de Gallagher y Lindgren, 2015) a los contenidos dinámicos y enactivos, no ha sido ignorada por la educación musical. Al contrario, desde aproximaciones como el método Dalcroze (Juntunen, 2016; Sutela, Juntunen & Ojala, 2020) y el reconocimiento específico de la perspectiva enactiva (Bremmer & Nijs, 2020; Gutierrez, 2019; Hayes, 2019), se ha discutido y aplicado el principio de situación, contextualidad, corporeidad y actuación como mecanismos de integración del significado en una interacción que tiene como objetivo específico la formación de los alumnos en temas de teoría e historia musical.

Schiavio y van der Schyff (2018), y, más recientemente, van der Schyff, Schiavio y Elliott (2022), han reflexionado, desde la pedagogía musical y las ciencias cognitivas, sobre el continuo sensoriomotor del aprendizaje cotidiano de las personas y la relación corpórea y sensorial del músico y su instrumento. Argumentan que al caracterizar el aprendizaje como una dinámica autopoietica, esto es, de constante reescritura de significados o emergencia situada, es posible comprender cómo surgen naturalmente principios de autoorganización en la relación del individuo y su aprendizaje de acuerdo con intenciones expresivas musicales. Su aproximación, como la del resto de la literatura citada, aunque centrada en formulaciones pedagógicas, implica la promoción de ambientes de aprendizaje que favorezcan la integración generosa y creativa de recursos expresivos diádicos, tanto de quien guía el aprendizaje como de quien está en papel de aprendiz.

La metáfora en el concierto didáctico virtual es enactiva en dos sentidos, uno ontológico y otro epistemológico o hasta metodo-

lógico. En el primer sentido, es enactiva porque no puede ser otra cosa si concedemos que el lenguaje es acción situada, experiencia subjetiva o, mejor dicho, intersubjetiva. En el segundo, y posible tercer sentido, es enactiva porque la sutileza de la integración de los recursos mediante los cuales se expresa existe en un contexto situado y corpóreo. Queda claro que la adecuación descriptiva en el ejercicio de transcripción del lenguaje es pobre si no se consideran, al menos, cuatro o hasta seis dominios de modalidad expresiva: (1) oral, (2) oral-gestual (Dingemanse, 2020; Okrent, 2002), (3) viso-gestual, (4) interactiva (Enfield & Sidnell, 2014; Fusaroli, Gangopadhyay & Tylén, 2014; Linell, 2017), (5) proxémica y (6) háptica (Pezzulo, Roche & Saint-Bauzel, 2021).

El baile de la interactividad es aún más rico que el concierto de la multimodalidad individual que hemos mostrado en este trabajo, en particular, en la Figura 3. No obstante, la evidencia de esta enacción metafórica que es esencial en la ostensión del concierto didáctico virtual basta para comprender cómo, cuando decimos que la gestualidad, la metáfora multimodal o el lenguaje figurativo tienen efectos positivos en el atractivo de un video de YouTube, en el aprendizaje o en la comprensión de conceptos abstractos, en realidad, estamos diciendo que el lenguaje completo es mejor que el lenguaje empobrecido.

4. Referencias

- Abner, Natasha; Cooperrider, Kensy, & Goldin-Meadow, Susan (2015). Gesture for linguists: A handy primer. *Language and Linguistics Compass*, 9(11), 437–451. doi: 10.1111/lnc3.12168
- Asmiarti, Dwi, & Winangun, Guntur (2018). The role of YouTube media as a means to optimize early childhood cognitive development. *MATEC Web of Conferences*, 205, 00002. doi: 10.1051/mateconf/201820500002
- Beautemps, Jacob, & Bresges, André (2021). What comprises a successful educational science YouTube video? A five-thousand user survey on viewing behaviors and self-perceived importance of various variables con-

- trolled by content creators. *Frontiers in Communication*, 5. <https://www.frontiersin.org/articles/10.3389/fcomm.2020.600595>
- Berk, Ronald (2009). Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. *International Journal of Technology in Teaching and Learning*, 5(1), 1–21.
- Boersma, Paul, & Weenink, David (1992–2022). *Praat: Doing phonetics by computer* [Computer program]. Version 6.3.01. <https://www.fon.hum.uva.nl/praat/>
- Bremmer, Melissa, & Nijs, Luc (2020). The role of the body in instrumental and vocal music pedagogy: A dynamical systems theory perspective on the music teacher's bodily engagement in teaching and learning. *Frontiers in Education*, 5. <https://www.frontiersin.org/articles/10.3389/educ.2020.00079>
- Caffi, Claudia (1999). On mitigation. *Journal of Pragmatics*, 31(7), 881–909. doi: 10.1016/S0378-2166(98)00098-8
- Pagán Cánovas, Cristóbal; Valenzuela, Javier; Alcaraz Carrión, Daniel; Olza, Inés, & Ramscar, Michael (2020). Quantifying the speech-gesture relation with massive multimodal datasets: Informativity in time expressions. *PLOS ONE*, 15(6), e0233892. doi: 10.1371/journal.pone.0233892
- Cayari, Christopher (2018). Connecting music education and virtual performance practices from YouTube. *Music Education Research*, 20(3), 360–376. doi: 10.1080/14613808.2017.1383374
- Cooperrider, Kensy (2014). Body-directed gestures: Pointing to the self and beyond. *Journal of Pragmatics*, 71, 1–16.
- Cooperrider, Kensy; Abner, Natasha, & Goldin-Meadow, Susan (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Frontiers in Communication*, 3(23). doi: 10.3389/fcomm.2018.00023
- Cserép, Attila (2014). Conceptual Metaphor Theory: In defence or on the fence? *Argumentum*, 10, 261–288.
- Dancygier, Barbara, & Sweetser, Eve (2014). *Figurative language*. Nueva York: Cambridge University Press.
- de Beer, Carola; Hogrefe, Katharina; Hielscher-Fastabend, Martina, & de Ruiter, Jan P. (2020). Evaluating models of gesture and speech production for people with aphasia. *Cognitive Science*, 44(9), e12890. doi: 10.1111/cogs.12890

- de Haan, Sanneke (2020). An enactive approach to Psychiatry. *Philosophy, Psychiatry, & Psychology*, 27(1), 3–25. doi: 10.1353/ppp.2020.0001
- Dingemans, Mark (2020). Between sound and speech: Liminal signs in interaction. *Research on Language and Social Interaction*, 53(1), 188–196. doi: 10.1080/08351813.2020.1712967
- Enfield, Nick, & Sidnell, Jack (2014). Language presupposes an enchronic infrastructure for social interaction. En Daniel Dor, Chris Knight & Jerome Lewis (Eds.), *The social origins of language* (pp. 92–104). Oxford: Oxford University Press.
- Escobar, Luis (2008). *Estructura silábica de la lengua de señas mexicana* (Tesis de licenciatura inédita). Escuela Nacional de Antropología e Historia, México.
- Escobar, Luis (2021). Y ahora, ¿cómo encuentro una metáfora? La teoría de la metáfora conceptual 40 años después. *Semas*, 2(4), 99–132.
- Escobar, Luis (2022). *Los gestos del tiempo: expresión gestual y representación espacial*. Ciudad de México: Instituto Nacional de Antropología e Historia, Escuela de Antropología e Historia del Norte de México.
- Evans, Vyvyan; Bergen, Benjamin, & Zinken, Jörg (2007). The cognitive linguistics enterprise: An overview. En Vyvyan Evans, Benjamin Bergen & Jörg Zinken (Eds.), *The cognitive linguistics reader* (pp. 1–36). Londres: Equinox.
- Fauconnier, Gilles, & Turner, Mark (1998). Conceptual integration networks. *Cognitive Science*, 22(2), 133–187.
- Fernández-Morante, Basilio (2018). Psychological violence in current musical education at conservatoires. *Revista Internacional de Educación Musical*, 6(1), 13–24. doi: 10.12967/RIEM-2018-6-p013-024
- Flores, Nelson, & Rosa, Jonathan (2022). Undoing competence: Coloniality, homogeneity, and the overrepresentation of whiteness in applied linguistics. *Language Learning*. doi: 10.1111/lang.12528
- Forceville, Charles (2017). Visual and multimodal metaphor in advertising: Cultural perspectives. *Styles of Communication*, 9(2), 26–41.
- Fraser, Bruce (1980). Conversational mitigation. *Journal of Pragmatics*, 4(4), 341–350. doi: 10.1016/0378-2166(80)90029-6

- Fusaroli, Riccardo; Gangopadhyay, Nivedita, & Tylén, Kristian (2014). The dialogically extended mind: Language as skillful intersubjective engagement. *Cognitive Systems Research*, 29-30, 31–39.
- Gallagher, Shaun (2017). Self and narrative. En Jeff Malpas & Gander Hans-Helmuth (Eds.), *The Routledge Companion to Philosophical Hermeneutics* (pp. 403–414). Londres: Routledge.
- Gallagher, Shaun (2018). Rethinking nature: Phenomenology and a non-reductionist cognitive science. *Australasian Philosophical Review*, 2(2), 125–137. doi: 10.1080/24740500.2018.1552074
- Gallagher, Shaun (2020). *Action and interaction*. Oxford: Oxford University Press.
- Gallagher, Shaun, & Lindgren, Robb (2015). Enactive metaphors: Learning through full-body engagement. *Educational Psychology Review*, 27, 391–404.
- Gibbs, Raymond W. (2019). Metaphor as dynamical–ecological performance. *Metaphor and Symbol*, 34(1), 33–44.
- Glenberg, Arthur (2008). Embodiment for education. En Paco Calvo & Antoni Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 355–372). San Diego: Elsevier. doi: 10.1016/B978-0-08-046616-3.00018-9
- Goldin-Meadow, Susan (2014). Widening the lens: What the manual modality reveals about language, learning and cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130295. doi: 10.1098/rstb.2013.0295
- Grajales-Acevedo, Carolina, & Posada-Silva, Wadis Yovany (2020). El trasfondo didáctico del teatro. *Revista Latinoamericana de Estudios Educativos (Colombia)*, 16(1), 187–210.
- Gutierrez, James (2019). An enactive approach to learning music theory? Obstacles and openings. *Frontiers in Education*, 4. <https://www.frontiersin.org/articles/10.3389/feduc.2019.00133>
- Harnoncourt, Nikolaus (2006). *La música como discurso sonoro*. Barcelona: Acantilado.
- Hayes, Lauren (2019). Beyond skill acquisition: Improvisation, interdisciplinarity, and enactive music cognition. *Contemporary Music Review*, 38(5), 446–462. doi: 10.1080/07494467.2019.1684059

- Haynes, Bruce (2007). *The end of early music: A period performer's history of music for the twenty-first century*. Oxford: Oxford University Press.
- Hong, Jon-Chao; Chen, Mei-Lien, & Ye, Jian-Hong (2020). Acceptance of YouTube applied to dance learning. *International Journal of Information and Education Technology*, 10(1), 7–13. doi: 10.18178/ijiet.2020.10.1.1331
- Huang, Minyao, & Jaszczolt, Kasia M. (Eds.) (2018). *Expressing the self: Cultural diversity and cognitive universals*. Oxford: Oxford University Press.
- Iskru, Victoria V., & Schulz, John (2020). How postgraduate students use video to help them learn. *Contemporary Educational Technology*, 12(2), ep276. doi: 10.30935/cedtech/8400
- Jaszczolt, Kasia M. (2018). Pragmatic Indexicals. En Minyao Huang & Kasia M. Jaszczolt (Eds.), *Expressing the Self: Cultural Diversity and Cognitive Universals* (pp. 260–311). Oxford: Oxford University Press.
- Jebe, Frank; Konietzko, Sebastian; Lichtschlag, Margrit, & Liebau, Eckart (2019). Jugend / YouTube / Kulturelle Bildung. Horizont 2019. [Juventud / YouTube / Educación Cultural. Horizonte 2019]. Alemania: Rat für Kulturelle Bildung e.V. <https://www.bosch-stiftung.de/de/publikation/jugend-youtube-kulturelle-bildung-horizont-2019>
- Johnson, Robert, & Liddell, Scott (2010). Toward a phonetic representation of signs: Sequentiality and contrast. *Sign Language Studies*, 11(2), 241–274. doi: 10.1353/sls.2010.0008
- Jones, Troy, & Cuthrell, Kristen (2011). YouTube: Educational potentials and pitfalls. *Computers in the Schools*, 28(1), 75–85. doi: 10.1080/07380569.2011.553149
- Juntunen, Marja-Leena (2016). The Dalcroze approach: Experiencing and knowing music through embodied exploration. En Carlos R. Abril & Brent M. Gault (Eds.), *Teaching general music: Approaches, issues, and viewpoints* (pp. 141–167). Oxford: Oxford University Press.
- Kallio, Alexis Anja (2020). Decolonizing music education research and the (im) possibility of methodological responsibility. *Research Studies in Music Education*, 42(2), 177–191. doi: 10.1177/1321103X19845690
- Kapidzic, Sanja, & Herring, Susan C. (2015). Race, gender, and self-presentation in teen profile photographs. *New Media & Society*, 17(6), 958–976. doi: 10.1177/1461444813520301

- Kastrup, Bernardo (2018). Conflating abstraction with empirical observation: The false mind-matter dichotomy. *Constructivist Foundations*, 13(3), 341–361.
- Kastrup, Bernardo (2019). *The idea of the world: A multi-disciplinary argument for the mental nature of reality*. Winchester: Iff Books.
- Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam (2014). Semiotic diversity in utterance production and the concept of ‘language’. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, 369, 20130293. doi: 10.1098/rstb.2013.0293
- Kincheloe, Joe L. (2011). Critical ontology and indigenous ways of being: Forging a postcolonial curriculum. En Kecia Hayes, Shirley R. Steinberg & Kenneth Tobin (Eds.), *Key works in critical pedagogy* (pp. 333–349). Róterdam: SensePublishers. <https://brill.com/display/book/edcoll/9789460913976/BP000026.xml>
- Kliuchynska, Nataliaia (2022). Humor in the musical-didactic works of the baroque period. *Scientific Journal of Polonia University*, 52(3), 60–65. doi: 10.23856/5207
- Kövecses, Zoltán (2010). *Metaphor: A practical introduction*. Nueva York: Oxford University Press.
- Kövecses, Zoltán (2013). The metaphor–metonymy relationship: Correlation metaphors are based on metonymy. *Metaphor and Symbol*, 28(2), 75–88. doi: 10.1080/10926488.2013.768498
- Lakoff, George (1993). The contemporary theory of metaphor. En Andrew Ortony, *Metaphor and thought* (2a. ed., pp. 202–250). Cambridge: Cambridge University Press.
- Lakoff, George, & Johnson, Mark (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Langacker, Ronald W. (1991). *Concept, image, and symbol: The cognitive basis of grammar*. Berlín: De Gruyter Mouton.
- Lausberg, Hedda, & Sloetjes, Han (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41(3), 841–849. doi: 10.3758/BRM.41.3.841

- Leech-Wilkinson, David (2020). *Challenging performance: Classical music performance norms and how to escape them*. <https://challengingperformance.com/the-book/>
- Liddell, Scott, & Johnson, Robert (1989). American Sign Language: The phonological base. *Sign Language Studies*, 64, 195–278.
- Linell, Per (2017). Intersubjectivity in dialogue. En Edda Weigand (Ed.), *The Routledge handbook of language and dialogue* (pp. 109–126). Nueva York: Routledge.
- Littlemore, Jeannette (2015). *Metonymy: Hidden shortcuts in language, thought and communication*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781107338814
- Macrine, Sheila L., & Fugate, Jennifer M. B. (2021). Translating embodied cognition for embodied learning in the classroom. *Frontiers in Education*, 6. <https://www.frontiersin.org/articles/10.3389/educ.2021.712626>
- McGlone, Matthew S. (2007). What is the explanatory value of a conceptual metaphor? *Language & Communication*, 27(2), 109–126.
- McNeill, David (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- McNeill, David (2016). *Why we gesture?* Cambridge: Cambridge University Press.
- McNeill, David (2017). Gesture-speech unity: What it is, where it came from. En Ruth Breckinridge Church, Martha W. Alibali & Spencer D. Kelly (Eds.), *Why gesture?* (pp. 77–101). Ámsterdam: John Benjamins.
- Moghavverni, Sedigheh; Sulaiman, Ainin; Jaafar, Noor Ismawati, & Kasem, Nafisa (2018). Social media as a complementary learning tool for teaching and learning: The case of YouTube. *The International Journal of Management Education*, 16(1), 37–42. doi: 10.1016/j.ijme.2017.12.001
- Mondada, Lorenza (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20(3), 336–366. doi: 10.1111/josl.1_12177
- Müller, Cornelia (2004). Forms and uses of the Palm Up Open Hand: A case of a gesture family? En Cornelia Müller & Roland Posner (Eds.), *The semantics and pragmatics of everyday gestures* (pp. 233–256). Berlín: Weidler.

- Müller, Cornelia (2019). Metaphorizing as embodied interactivity: What gesturing and film viewing can tell us about an ecological view on metaphor. *Metaphor and Symbol*, 34(1), 61–79.
- Newen, Albert; Gallagher, Shaun, & De Bruin, Leon (2018). 4E cognition: Historical roots, key concepts, and central issues. En Albert Newen, Leon De Bruin & Shaun Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 3–16). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780198735410.013.1
- Okrent, Arika (2002). A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (or at least give us some criteria to work with). En Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 175–198). Cambridge: Cambridge University Press.
- Oyetade, Eunice M., & Adegoke, Ojeniyi (2017). Educational embodied cognition: Nigerian education change agent. *Nigerian Journal of Social Studies*, 20(2), 1–22.
- Pérez Rodríguez, Julián (2003). Hacia una educación musical para el futuro: el concierto didáctico. *Revista Electrónica de LEEME*, 11.
- Pérez-Sobrino, Paula (2016). Multimodal metaphor and metonymy in advertising: A corpus-based account. *Metaphor and Symbol*, 31(2), 73–90. doi: 10.1080/10926488.2016.1150759
- Pérez-Sobrino, Paula, & Littlemore, Jeannette (2020). What makes an advert go viral? The role of figurative operations in the success of Internet videos. En Laura Hidalgo-Downing & Blanca Kraljevic Mujic (Eds.), *Performing metaphorical creativity across modes and contexts* (pp. 119–152). Amsterdam: John Benjamins.
- Pérez-Sobrino, Paula; Littlemore, Jeannette, & Houghton, David (2019). The role of figurative complexity in the comprehension and appreciation of advertisements. *Applied Linguistics*, 40(6), 957–991. doi: 10.1093/applin/amy039
- Pezzulo, Giovanni; Roche, Lucas, & Saint-Bauzel, Ludovic (2021). Haptic communication optimises joint decisions and affords implicit confidence sharing. *Scientific Reports*, 11, 1051. doi: 10.1038/s41598-020-80041-6

- Posholi, Lerato (2020). Epistemic decolonization as overcoming the hermeneutical injustice of Eurocentrism. *Philosophical Papers*, 49(2), 279–304. doi: 10.1080/05568641.2020.1779604
- Rahmatika, Rahmatika; Yusuf, Munawir, & Agung, Leo (2021). The effectiveness of YouTube as an online learning media. *Journal of Education Technology*, 5(1), 152–158. doi: 10.23887/jet.v5i1.33628
- Rahmaturrizki, Muhammad Ilham, & Sukmayadi, Yudi (2021). YouTube as audio visual media learning in music education. En Yudi Sukmayadi, Zakarias S. Soeteja, Tati Narawati, Juju Masunah, Rita Milyartini & Trianti Nugraheni (Eds.), *Proceedings of the 3rd International Conference on Arts and Design Education (ICADE 2020)* (pp. 297–303). Dordrecht: Atlantis Press. doi: 10.2991/assehr.k.210203.064
- Richter, Christoph; der Künste, Hochschule; Kraemer, Oliver; Yob, Iris, & Herzig, Monika (1996). The didactic interpretation of music. *Philosophy of Music Education Review*, 4(1), 3349.
- Rosenthal, Sonny (2018). Motivations to seek science videos on YouTube: Free-choice learning in a connected society. *International Journal of Science Education, Part B: Communication and Public Engagement*, 8(1), 2239. doi: 10.1080/21548455.2017.1371357
- Schiavio, Andrea, & van der Schyff, Dylan (2018). 4E music pedagogy and the principles of self-organization. *Behavioral Sciences*, 8(8), 72. doi: 10.3390/bs8080072
- So, Wing Chee; Kita, Sotaro, & Goldin-Meadow, Susan (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, 33(1), 115–125. doi: 10.1111/j.1551-6709.2008.01006.x
- Stagg, Bethan C., & Verde, Michael F. (2019). Story of a seed: Educational theatre improves students' comprehension of plant reproduction and attitudes to plants in primary science education. *Research in Science & Technological Education*, 37(1), 15–35.
- Stilwell, Peter, & Harman, Katherine (2021). Phenomenological research needs to be renewed: Time to integrate enactivism as a flexible resource. *International Journal of Qualitative Methods*, 20. doi: 10.1177/1609406921995299

- Sutela, Katja; Juntunen, Marja-Leena, & Ojala, Juha (2020). Applying music-and-movement to promote agency development in music education: A case study in a special school. *British Journal of Music Education*, 37(1), 71–85.
- Sweetser, Eve (2017). Metaphor and metonymy in advertising: Building viewpoint in multimodal multi-space blends. *Journal of Pragmatics*, 122, 65–76. doi: 10.1016/j.pragma.2017.10.012
- Szokolszky, Agnes (2019). Perceiving metaphors: An approach from developmental ecological Psychology. *Metaphor and Symbol*, 34(1), 17–32.
- The Language Archive (2022). ELAN 6.4. *The Language Archive. Max Planck Institute for Psycholinguistics*. <https://tla.mpi.nl/tools/tla-tools/elan/>
- Thibeault, Matthew D. (2012). Ubiquitous music learning in a postperformance world. *Teachers College Record: The Voice of Scholarship in Education*, 114(13), 196–215. doi: 10.1177/016146811211401312
- Tupas, Ruanni (2019). Entanglements of colonialism, social class, and ‘Unequal Englishes’. *Journal of Sociolinguistics*, 23(5), 529–542.
- van der Schyff, Dylan; Schiavio, Andrea, & Elliott, David J. (2022). *Musical bodies, musical minds: Enactive cognitive science and the meaning of human musicality*. Cambridge: The MIT Press.
- Varela, Francisco; Thompson, Evan, & Rosch, Eleanor (1991). *The embodied mind: Cognitive science and human experience*. Cambridge: The MIT Press.
- Vladova, Gergana; Ullrich, André; Bender, Benedict, & Gronau, Norbert (2021). Students’ acceptance of technology-mediated teaching – How it was influenced during the COVID-19 pandemic in 2020: A study from Germany. *Frontiers in Psychology*, 12. <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.636086>
- Wagner, Petra; Malisz, Zofia, & Kopp, Stefan (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209–232.
- Waldron, Janice (2012). Conceptual frameworks, theoretical models and the role of YouTube: Investigating informal music learning and teaching in online music community. *Journal of Music, Technology & Education*, 4(2–3), 189–200. doi: 10.1386/jmte.4.2-3.189_1
- Ward, Dave; Silverman, David, & Villalobos, Mario (2017). Introduction: The varieties of enactivism. *Topoi*, 36(3), 365–375. doi: 10.1007/s11245-017-9484-6
- Welch, Shay (2019). Native epistemology and embodied cognitive theory. En Shay Welch (Ed.), *The phenomenology of a performative knowledge*

- system: Dancing with Native American epistemology* (pp. 53–89). Cham: Palgrave Macmillan. doi: 10.1007/978-3-030-04936-2_3
- Werner, Konrad (2020). Cognitive confinement, embodied sense-making, and the (de)colonization of knowledge. *Philosophical Papers*, 49(2), 339–364. doi: 10.1080/05568641.2020.1779603
- Withagen, Rob (2022). *Affective Gibsonian Psychology*. Nueva York: Routledge.
- Zbikowski, Lawrence M. (2008). Cognitive science, music theory, and music analysis. En Jan Philipp Sprick, Reinhard Bahr & Michael von Trotschke (Eds.), *Musiktheorie im context: Kongress der Gesellschaft für Musiktheorie* (pp. 447–463). Berlín: Weidler.



Formas apelativas de segunda persona en entrevistas político-presidenciales mexicanas

Address forms in Mexican
political-presidential interviews

Melanie del Carmen
Salgado López

Universidad Nacional Autónoma de
México, Facultad de Filosofía y Letras,
Colegio de Letras Hispánicas
melaniesalgadol@filos.unam.mx



Recepción: 2 de noviembre del 2022

Aceptación: 28 de junio del 2023

doi: 10.22201/enallt.01852647p.2023.77.1041

Resumen

A partir de lo sugerido en otros estudios (Cuenca, 2013; Cuenca & Marín, 2015), analizaremos el uso de formas apelativas de segunda persona (vocativos, formas de tratamiento y formas pronominales en función de sujeto) que impactan en la construcción de las imágenes discursivas de los interlocutores, con el objetivo de establecer si los resultados aportan datos que permitan comprender con mayor precisión el género de la entrevista política presidencial en México. Los resultados muestran que dichas formas funcionan como una marca de transgresiones que reflejan la pugna por el poder en este dispositivo discursivo, impactan en el grado de transgresión del género y constituyen una manifestación lingüística que permite analizar la construcción del adversario político en el plano interaccional. Asimismo, es posible que las estrategias en el uso de las formas apelativas de segunda persona presenten patrones distintos con construcción del adversario correlacionados con los modelos presidenciales que encarnan los entrevistados, por lo que su análisis aportaría también al conocimiento del comportamiento de este género en el discurso político mexicano.

Palabras clave: vocativos; fórmulas de tratamiento; construcción discursiva del enemigo; entrevista presidencial; cortesía

Abstract

Based on previous studies (Cuenca, 2013; Cuenca & Marín, 2015), we will analyze the use of some second-person appeal forms (vocatives, forms of address, and pronominal forms as subjects) as an index of transgressive and courteous/discourteous strategies impacting the construction of discursive images of speakers. The objective is to establish whether the results provide data to better understand the presidential political interview genre in Mexico. The results reveal that the studied forms can indeed function as markers of such transgressions, reflecting the power struggle within this discursive device, impacting the degree of genre transgression and constituting one of the linguistic manifestations through which the construction of the political adversary can be analyzed on the interactional level. Our observations suggest that strategies in the use of second-person appeal forms show distinct patterns in the construction of the adversary, correlated with the presidential models represented by the interviewees. Therefore, their analysis may also contribute to understanding the behavior of this genre in Mexican political discourse.

Keywords: vocatives; forms of address; discursive construction of the political other; presidential interview; politeness

1. Introducción

Son numerosos los trabajos en los que se han analizado las entrevistas políticas hechas a presidentes. Varios de ellos se centran en las estrategias argumentativas de contenido que impactan en la construcción de la imagen tanto del entrevistado como de los aliados y enemigos políticos. Estas investigaciones han aportado conocimiento sobre los patrones y el comportamiento de tales manifestaciones discursivas. En esta línea se ubica un trabajo previo (Salgado López, 2019; 2020) en el cual, al analizar el uso de las formas adversativas en la construcción discursiva de aliados y enemigos, se detectó que las entrevistas a presidentes priistas¹ mexicanos mostraban variaciones en la aparición y preferencia de recursos lingüísticos, así como de estrategias extralingüísticas correlacionadas con el modelo de presidencialismo ejercido y con el grado de desgaste de la figura presidencial:

la entrevista a los presidentes pasa de ser un dispositivo que activan y manejan como iniciativa de ellos mismos en la construcción de su imagen a convertirse en un dispositivo que es activado por los entrevistadores para cuestionar a la figura presidencial. [...] La figura presidencial ha ido perdiendo el control de la organización y se ha ido convirtiendo en un dispositivo del que se tiene que hacer uso frente al desgaste de la misma. Frente a eso, el dispositivo ha estado fuertemente controlado por la presidencia siempre. [...] Así, el modelo presidencialista se caracteriza por una figura presidencial más legítima y con mayor fuerza (incluso a meses de dejar la presidencia), con mayor presencia de heteroglosia en su discurso, con presencia

¹ El Partido Revolucionario Institucional (PRI), en un inicio Partido Nacional Revolucionario, es un partido político mexicano fundado en 1929. En él se aglutinaron los sectores que adquirieron poder político y territorial durante y después de la Revolución mexicana. El PRI mantuvo el control y el dominio hegemónico del poder político presidencial en México hasta el año 2000.

de marcas de afecto en el discurso y con una figura presidencial más autónoma. En el caso del modelo tecnocrático, es una figura con menos legitimidad, que hace gala de menos fuerza, con presencia de heteroglosia y con una figura presidencial autónoma. Finalmente, el nuevo PRI presenta una figura presidencial con menos legitimidad, con menos gala de fuerza, con mayor presencia de monoglosia y con una figura presidencial sumamente dependiente de su equipo de trabajo. [...] El desgaste de la figura presidencial también se puede observar por medio del aumento en las interrupciones que los entrevistadores pueden hacer y hacen a los presidentes. En el caso de Peña Nieto el bajo número de unidades pregunta-respuesta tiene que ver con la monoglosia de su discurso y, también con la estrategia de desviación de tema, lo que hace que tenga respuestas excesivamente largas en las que se va cambiando el tema planteado en la pregunta. Ningún entrevistador ha tratado de tú a un presidente, sin embargo, Peña Nieto sí tuteó a su entrevistadora e incluso en dos ocasiones la llamó por su nombre (Salgado López, 2020: 289–290).

▶ Sin embargo, debido a que estaba fuera de los objetivos del trabajo, no se analizó con detalle el uso de vocativos, formas de tratamiento y marcas de segunda persona, aunque se señaló que parecían estar vinculadas también a la estrategia de construcción discursiva del oponente y del locutor mismo (Salgado López, 2020). En investigaciones previas se ha mostrado el uso de las formas apelativas de segunda persona como un recurso argumentativo del nivel implícito o interactivo de ciertos géneros discursivos. Por ejemplo, el trabajo de Cuenca (2013) apunta el uso de vocativos como índice de la transgresión y cortesía/descortesía en una entrevista política presidencial, mientras que el de Cuenca y Marín (2015) afirma que el uso de las formas apelativas de segunda persona en un debate político está conectado con la construcción de las imágenes discursivas del adversario.

El objetivo de este trabajo es analizar dos recursos de marcas del receptor (vocativos y formas pronominales explícitas en función de sujeto) en tres entrevistas realizadas a presidentes mexicanos priistas en distintos momentos de la historia reciente mexicana. La discusión de los datos encontrados busca establecer, de manera exploratoria, si su estudio aporta nuevos conocimientos y posibilidades de aproximación a este género y a algunos efectos discursivos asociados a él. La hipótesis que guio este acercamiento es que el uso de estos recursos lingüísticos puede ser un indicador de las estrategias de cortesía y descortesía en las entrevistas político-presidenciales, las cuales presentarían variaciones derivadas de los modelos presidenciales que represente cada uno de los presidentes entrevistados. Asimismo, se prevé que en las marcas analizadas se podría encontrar un índice del grado de confrontación que existe en la interacción entre entrevistador(es) y entrevistado.

La importancia de los objetivos de esta investigación se debe a que no se han realizado análisis pormenorizados del comportamiento del uso de formas apelativas explícitas y los cambios que estas han experimentado dependiendo de las transformaciones en los modelos presidenciales. Además, dicho análisis podría aportar al conocimiento mismo de las características del género discursivo (entrevista política y política presidencial), debido a que las marcas examinadas podrían indicar distintos grados de confrontación en el género. Por último, cabe señalar que el abordaje de este fenómeno es complejo debido a la subjetividad que supone.

2. La entrevista política: breve caracterización discursiva y marco teórico

2.1. Sobre las entrevistas políticas

En tanto que el corpus que se analizará en esta investigación se compone de tres entrevistas a presidentes mexicanos, estas pertenecen al género del discurso político, por lo que presentarán las siguientes características (Verón, 1987):

- a) Una dimensión polémica, que implica la construcción de un adversario y de un destinatario complejo (integrado por paradestinatarios, contradestinatarios y prodestinatarios).
- b) Una multifunción del discurso político, que abarca la polémica, el refuerzo y la persuasión.
- c) Una multipresencia de entidades discursivas.

En el caso de esta investigación, el discurso político se presenta bajo el género discursivo textual conocido como *entrevista política*. A continuación, resumimos algunas de las características que muestra la entrevista política como escenario genérico.²

- a) Tiene un carácter dialógico con formato preestablecido.
- b) La confrontación es uno de los objetivos de la entrevista.
- c) Existe una contradicción con la supuesta neutralidad del entrevistador.

La entrevista política constituye entonces un género discursivo textual que puede adoptar el discurso político (en su sentido extensivo) cuando en él se presentan materialidades semióticas y discursivas del ejercicio del poder bajo el formato o dispositivo de entrevista. En ella, se presentan características específicas que derivan tanto del discurso político como del formato entrevista, entre las que destaca que en la dinámica entrevistador-entrevistado existe una relación asimétrica de poder, ya que el entrevistador cuestionará, representando inquietudes de paradestinatarios, prodestinatarios y contradestinatarios, a su entrevistado; según esta dinámica se construyen distintas escenografías.³

Si comparamos el uso del dispositivo *entrevista* en el discurso presidencial frente a otras manifestaciones político-presidenciales

² Estas características han sido desarrolladas de manera más amplia en un trabajo anterior (Salgado López, 2020).

³ La entrevista política es un escenario genérico donde el entrevistador tiene, en teoría, poco control de la escenografía que quiere construir, igual que en el debate político.

(véase Salgado López, 2020), como la toma de protesta o el mensaje a la nación, notaremos que la entrevista aparece relativamente menos institucionalizada, y su tema o contenido resultan más cotidianos (menos especializados). Asimismo, ciertos roles, al menos el del entrevistado, son institucionales, lo que le confiere un carácter institucional pero menos formal que otras manifestaciones discursivas presidenciales. Los roles en la interacción suponen un poder desigual, y el formato intenta construir una interacción que parezca más íntima. Además, la entrevista es un formato muy anclado al contexto y menos multimodal que otros.⁴

Varias investigaciones han definido algunas de las características de la entrevista, por ejemplo, Halperín (1995) señala cierto tono de “proximidad, intercambio, exposición discursiva con interrupciones, un tono marcado por la espontaneidad, presencia de lo personal y atmósfera de intimidad” (37). La entrevista, añade, “es la más pública de las conversaciones privadas” (Halperín, 1995: 12). A este rasgo se suma su formalidad y alto grado de convención y rigidez (Stenström, 1994). Sin embargo, la buena recepción que tiene se debe, en cierta medida, a que genera una apariencia menos acartonada que otros géneros (Depetris & García, 2008), pues se muestra como una conversación informal. Esta imagen no debe evitar que se tenga presente que “dicho diálogo se rige por convenciones o normas que los medios de comunicación imponen, como son, por ejemplo, los roles claramente delimitados de entrevistador y entrevistado, los temas que se abordarán y sus límites y perspectivas” (Araya Seguel, 2011: 47).

Por otra parte, si bien la entrevista es una subclasificación que pertenece al género periodístico, cuando es realizada a un presidente, por ejemplo, se habla de entrevista política. Diversos autores

⁴ Las características mencionadas hacen referencia a las que se observan en las entrevistas político-presidenciales si se les compara con dispositivos más institucionalizados y ritualizados del discurso presidencial, como la toma de protesta. No contamos con espacio suficiente ni es objetivo de este trabajo por menorizar estos detalles.

coinciden en que bastan criterios restrictivos para usar el término, por lo tanto, se considera como entrevista política aquella en la que participa un actor político en el ejercicio del poder (Cortés Rodríguez & Bañón Hernández, 1997; Yanes Mesa, 2006). Otros han intentado describir especificidades que diferenciarían a una entrevista política de cualquier otro tipo de entrevista. Por ejemplo, algunos autores señalan que en una entrevista política debe haber una relación asimétrica entre los interlocutores (Mullany, 1999; Pérez Minchola, 2007); mientras que otros postulan que solo en la entrevista política el entrevistado tiene una fuerte conciencia del desdoblamiento del nivel de recepción, dado que sabe y tiene presente que no está hablando únicamente con su entrevistador, sino que alguien más será su auditorio (Fetzer, 2000, 2006; Halperín, 1995). A su vez, Arfuch (1995) ha especificado que solo en la entrevista política el entrevistador puede representar a uno o a los tres tipos de interlocutores que construye el discurso político.

En el caso específico de la entrevista política, se ha afirmado que el entrevistado está a merced del entrevistador para ser guiado en una conversación en la cual se le podrá cuestionar o criticar (Halperín, 1995), ya que la distribución de roles otorga a estos últimos “la libertad para penetrar en su vida. [Están] autorizados para cuestionarlo públicamente y poner en duda sus declaraciones” (Halperín, 1995: 68).

Ciertas investigaciones han focalizado otras dimensiones discursivas en entrevistas políticas; no obstante, es importante señalar que pocas fuentes han abordado específicamente los mecanismos lingüísticos que nos interesan en este trabajo, al menos para el caso de México (Bermúdez Chaves, 1986; Cuenca, 2013; Gómez Sánchez, 2008; Gutiérrez, 2006; Laborda, 2000). Por otra parte, en el intento por caracterizar a la entrevista política se ha discutido y debatido sobre la confrontación o el grado de confrontación que hay en ella. Consideramos que examinar algunas formas apelativas de segunda persona podría contribuir a estas discusiones, así como a la definición y caracterización del género.

2.2. Marco teórico

Más allá de estudiar el discurso político-presidencial en la manifestación genérica de la entrevista, nos interesa en particular observar el uso de vocativos y de las formas de tratamiento como índices de a) la construcción discursiva de la autoimagen del locutor y la imagen del interlocutor, y b) el uso de estos mecanismos como herramientas para confrontar.

En términos discursivos, el poder refiere a “la capacidad de controlar —más o menos— en su propio interés, los actos y las mentes de los (miembros de) otros grupos” (van Dijk, 1999: 26), así, el poder y la lucha por él en las entrevistas político-presidenciales está relacionado con la posibilidad de ejercer o no esta facultad durante la interacción.

Con respecto al fenómeno de cortesía, se trata del “conjunto de estrategias conversacionales destinadas a evitar o mitigar las tensiones que aparecen cuando el hablante se enfrenta a un conflicto creado entre sus objetivos y los del destinatario” (Arjonilla, Atienza, Castro, Cortés, González, Inglés, Iruela, Lahuerta, López, Montmany, Pueyo, Puig, Sánchez, Torner, Vañó, Wesenaar y Martín, s. f., cortesía). La teoría de Leech (1983) explica este fenómeno en términos de máximas:

Según G. Leech, la cortesía lingüística es un principio regulador de la conducta verbal que se sitúa a medio camino entre la distancia social y la intención del emisor; persigue el equilibrio social entre los interlocutores, a pesar de que la intención comunicativa del emisor suponga una molestia para el destinatario. G. Leech evalúa la cortesía en términos de coste y beneficio, de manera que una acción verbal es más “descortés” cuanto mayor es el coste del destinatario y menor su beneficio, y más cortés en el caso contrario. De acuerdo con esta relación, establece una clasificación de intenciones en cuatro categorías generales:

- 1) Acciones verbales que apoyan la cortesía, como un cumplido, un agradecimiento o una felicitación.
- 2) Acciones prácticamente indiferentes a la cortesía, por ejemplo, una declaración.
- 3) Acciones que entran en conflicto con la cortesía, como una petición o una queja.
- 4) Acciones dirigidas frontalmente contra el mantenimiento de la relación entre los interlocutores, por ejemplo, un insulto, un reproche o una burla. Y una serie de máximas: máxima de tacto, de generosidad, de aprobación, de modestia, de acuerdo y de simpatía. (Arjonilla *et al.*, s. f.: párr. 5)

En ese sentido, consideramos que las entrevistas político-presidenciales pueden o no desplegar un efecto confrontativo. Las que no lo hacen se denominan “entrevistas a modo”, en ellas el comportamiento no marcado consiste en proteger la imagen de los interlocutores mutuamente. En contraste, las entrevistas en las que se despliega un tono confrontativo se caracterizan por actos que pueden ir desde la descortesía cooperativa (Martín Rojo, 2000) hasta actos abiertamente descorteses. Algunos de esos actos pueden analizarse en las marcas de segunda persona y constituyen un elemento clave de este tipo de manifestación genérica, ya que forman parte de la estrategia de construcción de la imagen discursiva tanto del entrevistado como del entrevistador.

Ahora bien, hemos afirmado en otras investigaciones que “lo que hace uno o varios sujetos por medio del discurso es construir una esquematización, que se logra a partir ‘de un cierto número de operaciones que uno bien puede llamar lógico-discursivas. Son lógicas porque son operaciones del pensamiento y discursivas porque el pensamiento se manifiesta a través del discurso’ (Grize 1993, p. 4.)” (Salgado López, 2020: 114). El uso de las formas apelativas de segunda persona funcionaría en las intervenciones de los interlocutores de las entrevistas político-presidenciales como operaciones de construcción de los mismos interlocutores —lo que en otras tradiciones se conoce como construcción de los referentes y, para

el caso de la escuela de análisis del discurso francesa, supone la construcción del *ethos* del interlocutor. En la entrevista, este tipo de mecanismos lingüísticos contribuyen a la construcción discursiva de, al menos, la imagen del interlocutor, la cual se encuentra en pugna o discusión en los casos de tensión. Mediante esta construcción, tanto el entrevistado como el entrevistador despliegan mecanismos que generan efectos de cortesía y descortesía en constante pugna. Nuestro objetivo es examinar detalladamente cómo se comporta el uso de estos mecanismos lingüísticos en el corpus para establecer las estrategias de confrontación de las cuales son índice.

Finalmente, entendemos por *formas apelativas de segunda persona* al uso de recursos lingüísticos por medio de los cuales tanto el entrevistador como el entrevistado identifican como referente a su interlocutor. Las formas estudiadas incluyen los vocativos, formas pronominales explícitas en función de sujeto y las formas de tratamiento. Ahora bien, los pronombres que cumplen con estos criterios pueden desempeñar distintas funciones sintácticas, sin embargo, nos concentraremos solo en aquellos que tienen la función de sujeto (Cuenca & Marín, 2015: 2).

De acuerdo con la Real Academia de la Lengua Española y la Asociación de Academias de la Lengua Española (2010), las expresiones vocativas:

son pronombres personales o grupos nominales que se usan para dirigirse a alguien, generalmente solicitando una respuesta o una reacción. Se emplean también para saludar o iniciar una conversación (¡Buenas tardes, doña Encarna!), para llamar la atención (¡Eh, tú!), pedir u ordenar algo (¡Acércate, muchacho!), para disculparse (¡Perdone, señor!) o para dirigirse a alguien con muy diversos propósitos [...] Se emplean como vocativos los pronombres, los nombres propios de persona, los de parentesco, oficios y profesiones, títulos honoríficos, y otros sustantivos análogos. (813).

A su vez, se señala que las formas de tratamiento son

las variantes pronominales que se eligen para dirigirse a alguien en función de la relación social que existe entre el emisor y el receptor (tú, usted, vos, os, le, te, etc.). Se incluyen también entre ellas los grupos nominales usados para dirigirse a algún destinatario, tanto cuando constituyen fórmulas genéricas del trato cortés (don Francisco, señor Martínez) como cuando se eligen en función de la posición del destinatario en alguna jerarquía (Su Majestad, Vuestra Ilustrísima, Su Señoría, Su Emi-nencia, Su Santidad).

16.7.1b En el uso de los tratamientos intervienen circunstan-cias sociales y situacionales de tipo diverso que pueden variar dependiendo de áreas geográficas y comunidades particulares. (Real Academia de la Lengua Española & Asale, 2010: 321).

Finalmente, se entiende por *formas pronominales en función de sujeto* a los pronombres *usted, tú y ustedes* cuando aparecen explícita o tácitamente (expresados en el verbo) cumpliendo la función sintáctica de sujeto, la única que conserva concordancia con el verbo de la oración.

Cada forma de las mencionadas fue considerada como unidad de análisis, criterio aplicado también en el llenado de las tablas de análisis.

3. Metodología de análisis y constitución del corpus

3.1. Metodología

Con base en las definiciones ya planteadas, se localizaron y eti-quetaron todos los pronombres que pasaran la prueba de sujeto, así como los sintagmas nominales que constituyeran un vocativo o forma de tratamiento, por medio de los cuales entrevistador y entrevistado se dirigen el uno al otro, solicitan respuestas, o bien el cambio de turno, durante la entrevista.

Es importante señalar que las formas apelativas que nos inte-resan pertenecen a los mecanismos de cohesión textual, pues for-

man parte de las múltiples manifestaciones lingüísticas de la construcción del receptor como un referente discursivo. Por esto, en algunas ocasiones no es sencillo establecer cuáles de estas formas son solo vocativos, ya que constituyen también manifestaciones de la forma de tratamiento y de la forma empleada para construir al receptor como referente. Por ejemplo, en (1) se muestran varias formas pronominales o sustantivas, sin embargo, el pronombre que aparece en el segmento “razones que le atañen a usted”, aunque permite localizar la forma de tratamiento de la entrevistadora al entrevistado (formal), no cumple con la función de sujeto, a diferencia del caso del *usted* que aparece en cursivas y que, por tanto, sí fue considerado en los datos.

- (1) Hace unas semanas exactamente leímos en los periódicos, después lo confirmó Relaciones exteriores, que el embajador de México en Francia había renunciado, que había dado razones que le atañen a usted, yo quería preguntarle, *señor presidente*, si *¿cree usted*, como político, como mexicano, como ideólogo, que esta persona tomó decisiones propias o fue maquinada su renuncia desde la misma capital de la República Mexicana?

Ahora bien, en varios casos, como el de “señor presidente”, se observa una forma de tratamiento que funciona como vocativo para dirigir una pregunta al entrevistado. Por esta razón, las formas de tratamiento también se incluyeron en las tablas de análisis de los vocativos. No obstante, es evidente que ambos pronombres permiten, debido a su cualidad déctica, actualizar como referente al presidente, quien es el receptor de la pregunta de la entrevistadora. En esta investigación solo se consideraron las formas que cumplen con las características ya descritas, sin embargo, las demás formas de construcción del referente podrían estudiarse en un trabajo posterior.

Dado que este análisis es cualitativo y discursivo, la metodología de trabajo implicó los siguientes pasos:

- 1) Localización y marcaje de vocativos y formas de tratamiento en el corpus. Los criterios operativos empleados se definieron en el apartado anterior. A cada unidad que cumplía con los criterios operativos de vocativo se le asignó un número y una clave para incluirla en una tabla-matriz de análisis.
- 2) Vaciado de los datos encontrados en tablas, se destacaron los siguientes elementos: quién los usó (entrevistador o entrevistado), forma concreta, parte de la entrevista y del turno de habla en la que se presentó.
- 3) Recuperación de datos y tablas de análisis de una investigación previa (Salgado López, 2020) que permitiera el cruce de los datos de vocativos y formas de tratamiento con otros elementos discursivos, como el número de intervenciones o unidades pregunta-respuesta de las entrevistas, interrupciones, interacción establecida, así como la estructura informativa (tema y rema). Además, se tomaron en cuenta otros elementos, como el tipo de modelo presidencial que representaba el entrevistado.
- 4) Revisión de los vocativos a la luz del fenómeno de la confrontación, la lucha por el poder y la cortesía, tal y como se definieron en §2, con el objetivo de analizar su función en la construcción de estrategias de cortesía y descortesía, la construcción discursiva del receptor, los mecanismos de ejercicio del poder en la dinámica conversacional y la construcción de la imagen de los interlocutores.

3.2. *Sobre la constitución del corpus*

El corpus de trabajo estuvo integrado por tres entrevistas realizadas a presidentes mexicanos, como lo indica el Cuadro 1.

CUADRO 1. Entrevistas que componen el corpus de trabajo

Manifestación discursiva	Fechas	Características
Entrevista con Gustavo Díaz Ordaz realizada por Ernesto Sodi y otros	1970 y 1976	Video del extinto Canal 40 que compila los dos entrevistas más importantes que se hicieron a Díaz Ordaz con respecto al tema del movimiento estudiantil de 1968 (50 minutos). https://www.youtube.com/watch?v=5tih7IP-1VM
Entrevista con Ernesto Zedillo realizada por Melisa Vertiz y Jorge Ramos	23 de octubre de 1996	Entrevista realizada en Los Pinos, que trató, fundamentalmente, sobre el Ejército Zapatista de Liberación Nacional (EZLN) y el Ejército Popular Revolucionario (EPR). La transcripción íntegra de la entrevista fue difundida por Univisión y por presidencia (8617 palabras).
Entrevista con Enrique Peña Nieto realizada por Rosa Elvira Vargas	Mayo de 2016	La entrevista completa es inédita. Un resumen fue publicado en <i>La Jornada</i> , “No hay motivos para el mal humor social”, por Rosa Elvira Vargas. Se tuvo acceso a una copia en audio de la entrevista completa. El tema principal fueron las reformas estructurales (117 minutos). http://www.jornada.unam.mx/2016/05/23/politica/004e1pol

Adaptado de Salgado López (2020: 70).

Para la constitución de dicho corpus (Carbó, 2001; 2007), se tomaron en cuenta ciertos criterios, que coincidían con los empleados en Salgado López (2020). Uno de los más importantes fue el criterio de *homogeneidad*, ya que

[l]os progresos realizados en los últimos años por las diversas disciplinas que se ocupan del lenguaje y del discurso permiten formular una primera respuesta: las variaciones en la relación de los discursos con sus condiciones de producción afectan sobre todo los mecanismos de la enunciación (Verón, 1987: 8).

Siguiendo esta línea, se decidió que todas las manifestaciones seleccionadas fueran entrevistas políticas realizadas a presidentes mexicanos priistas, de modo que compartieran elementos del dispositivo de su formato.

Asimismo, se agregó a este criterio que todas las entrevistas seleccionadas incluyeran preguntas en las que abiertamente —explícitamente— se cuestionara a los presidentes sobre decisiones o políticas de su gobierno. Esto con el fin de evitar que se compararan “entrevistas a modo” con aquellas de “mayor grado de con-

frontación”.⁵ Finalmente, a pesar de que el cargo presidencial ha sido ocupado por mandatarios de distintos partidos políticos, se decidió que en el corpus solo aparecieran presidentes provenientes de uno solo, en este caso, el Partido Revolucionario Institucional (PRI) con el objetivo de mantener cierta semejanza que permitiera hacer comparaciones.

Como criterio de *contraste*, se establecieron tres dimensiones o parámetros: temporalidad, modelo presidencial y situación contextual en la que se activa el dispositivo discursivo de entrevista política.

La delimitación de los cortes temporales fue uno de los aspectos más complicados en el diseño de esta investigación. Se decidió abordar el periodo de la historia contemporánea de México que comienza, de acuerdo con algunos historiadores nacionales, con la masacre del 2 de octubre de 1968. Desde ese punto hasta el triunfo electoral de Andrés Manuel López Obrador —la llamada Cuarta Transformación— hubo un periodo de 50 años, en los que el PRI mantuvo el poder político presidencial la mayor parte del tiempo. Se optó por incluir cortes temporales que abarcaran ambos extremos del periodo estudiado, así como una muestra que se ubicara aproximadamente en el medio. Esto permitiría explorar la evolución, los cambios y las similitudes a lo largo de ese lapso en relación con algunos de los presidentes priistas.

Con respecto al segundo parámetro, el modelo presidencial, en el periodo que hemos mencionado, los candidatos del PRI electos como presidentes han ejercido distintos modelos. Se selecciona-

⁵ No resulta sencillo precisar la diferencia entre una entrevista hecha a modo y una más confrontativa. A esto se suma lo evidente: es muy probable que en los casos revisados Presidencia haya pedido —o dictado— con anterioridad las preguntas que se le iban a realizar al mandatario. No obstante, se parte de que existe un continuo entre más confrontativo y menos confrontativo. En la construcción del acervo de esta investigación se hizo evidente que había entrevistas que claramente estaban “hechas a modo”. Mediante este criterio, se buscó evitar que en la selección se incluyeran manifestaciones más cercanas a lo menos confrontativo.

ron las manifestaciones de Gustavo Díaz Ordaz, Ernesto Zedillo y Enrique Peña Nieto, porque cada uno de ellos es representativo de un modelo diferente. Es decir, mientras Díaz Ordaz encarnó un modelo presidencialista heredado de la institucionalización de la “Familia Revolucionaria” que operaba bajo un esquema de Estado benefactor, Ernesto Zedillo representó el modelo de los tecnócratas en la presidencia mexicana. Finalmente, Peña Nieto personificó un modelo regido bajo un esquema neoliberal, pero que se trató de presentar como la cara del nuevo PRI luego de haber perdido la presidencia por 12 años.

Es importante destacar, además, que la selección incluye momentos contextuales distintos que tuvieron un fuerte impacto en las características de cada entrevista.⁶ Por ejemplo, Díaz Ordaz fue entrevistado poco antes de abandonar su cargo, una vez que ha sido designado el nuevo presidente electo, con quien queda enemistado. En este caso, el dispositivo de la entrevista fue activado en el ocaso de su periodo presidencial. Por su parte, la entrevista con Ernesto Zedillo se realizó casi al inicio de su llegada al poder, debido a los muchos cuestionamientos con los que asumió la presidencia y a la existencia de dos grupos armados que se oponían a su gobierno. A su vez, en el caso de Peña Nieto el dispositivo de la entrevista fue activado a la mitad de su periodo, justo cuando el problema del caso Ayotzinapa y de la reforma educativa generaba manifestaciones de rechazo social.⁷ En todos estos ejemplos se muestra que la entrevista es un dispositivo al que ha recurrido el sistema político ante ciertas crisis.

Como puede apreciarse, la representatividad del corpus y, por lo tanto, de la investigación tiene restricciones claras que corresponden a un estudio exploratorio acerca del uso de vocativos y for-

⁶ Para conocer con mayor detalle el contexto de cada una de ellas, consúltese Salgado López (2020).

⁷ El caso Ayotzinapa hace referencia a la desaparición forzada de 43 estudiantes ocurrida el 26 de septiembre de 2014. Para conocer más, véase <https://centroprodh.org.mx/casos-3/ayotzinapa/>

mas de tratamiento en algunos representantes del presidencialismo priista mexicano en los últimos 50 años de la historia.⁸

4. Resultados generales

El comportamiento de las formas apelativas de segunda persona que nos interesan se resume en los Cuadros 2 (vocativos y formas de tratamiento) y 3 (formas pronominales en función de sujeto).

Cuadro 2. Uso de vocativos y formas de tratamiento

Unidades que conforman la entrevista	Entrevistado	Vocativos empleados para apelar al entrevistador y número de usos	Entrevistadores	Vocativos empleados para apelar al entrevistado y número de usos
52 unidades pregunta-respuesta sin interrupciones en los turnos	Gustavo Díaz Ordaz	<i>doctor</i> (28) <i>muchachita, señorita</i> (2)	Ernesto Sodi Otros	<i>señor presidente</i> (5) <i>señor presidente</i> (1)
55 unidades pregunta-respuesta con interrupciones en los turnos	Ernesto Zedillo	0	Jorge Ramos Melisa Vertiz	<i>señor presidente</i> (6: 5 por Ramos y 1 por Vertiz)
23 unidades pregunta-respuesta con interrupciones en los turnos	Enrique Peña Nieto	<i>Rosa Elvira</i> (1)	Rosa Elvira Vargas	<i>presidente</i> (4)

Cuadro 3. Uso de formas pronominales de segunda persona en función de sujeto explícito

Unidades que conforman la entrevista	Entrevistado	Formas pronominales empleadas para apelar al entrevistador y número de usos	Entrevistador	Formas pronominales empleadas para apelar al entrevistado y número de usos
52 unidades pregunta-respuesta sin interrupciones de turnos	Gustavo Díaz Ordaz	<i>usted</i> (18) <i>ustedes</i> (1)	Ernesto Sodi Otros	<i>usted</i> (21) <i>usted</i> (4)
55 unidades pregunta-respuesta con interrupciones en los turnos	Ernesto Zedillo	<i>usted</i> (10) <i>ustedes</i> (2)	Jorge Ramos Melisa Vertiz	<i>usted</i> (39: 28 por Ramos y 11 por Vertiz)
23 unidades pregunta-respuesta con interrupciones en los turnos	Enrique Peña Nieto	<i>tú</i> (6) <i>ustedes</i> (1)	Rosa Elvira Vargas	<i>usted</i> (13) <i>ustedes</i> (2)

⁸ El corpus completo de la investigación puede consultarse en Salgado López (2020: 339–381).

Los datos incluidos en los cuadros muestran ya ciertos elementos interesantes que se desglosarán en los apartados del análisis. De modo general, llama la atención que en la entrevista con Díaz Ordaz haya más marcas de vocativos por parte del entrevistado que de los entrevistadores. En el caso de Ernesto Zedillo ocurre lo contrario, siendo los entrevistadores los que usaron un mayor número de vocativos. Finalmente, en el caso de Peña Nieto, hay casi el mismo número de uso de vocativos y formas de tratamiento por parte de la entrevistadora como del entrevistado.

En cuanto a las formas de tratamiento, los datos del Cuadro 2 revelan que, en la entrevista con Díaz Ordaz, tanto el presidente como el entrevistador emplean la forma de tratamiento *usted*. Lo mismo se observa en la entrevista con Zedillo, mientras que en la de Peña Nieto es la entrevistadora, Rosa Elvira Vargas, quien siempre utiliza la forma de tratamiento *usted* para dirigirse al presidente, mientras que este último empleó la forma de tratamiento *tú*.

Considerando que la aparición explícita de las formas pronominales de segunda persona con función de sujeto constituye una forma marcada, la fuerte presencia de ellos podría indicar estrategias de ataque o confrontación entre los interlocutores (Cuenca & Marín, 2015: 55). Más adelante se pormenorizará este comportamiento.

Debido a que los datos del análisis de las formas de tratamiento son más homogéneos, consideramos pertinente abordarlos primero.

4.1. *Sobre los vocativos y las fórmulas genéricas de trato cortés y deixis social*

Si bien no se consideró en los objetivos, en el análisis de los vocativos y formas de tratamiento se hizo evidente el fenómeno de la deixis social. Las formas de deixis social empleadas por los entrevistadores para dirigirse al presidente no presentan variación alguna, ya que siempre se trata de la forma *usted*, así como fórmulas genéricas de trato cortés en las que se marca cierta jerarquía, como *señor presidente* o *presidente*. Cabe señalar que solo la entrevista-

tadora de Peña Nieto no usó la forma de tratamiento cortés *señor* antepuesta a *presidente*.

Las formas de tratamiento que utilizan los entrevistados presentan un cambio hacia el pronombre *tú* en el caso de Peña Nieto, quien decide tutear a su entrevistadora. Por su parte, Díaz Ordaz acompañó la forma de tratamiento *usted* con fórmulas genéricas de trato cortés para su entrevistador, como *doctor*, mientras que tanto Zedillo como Peña Nieto no emplearon fórmulas genéricas de este tipo, como se muestra en el Cuadro 4.

CUADRO 4. Uso de formas de tratamiento de los entrevistados a sus entrevistadores

Entrevistados	Díaz Ordaz (45)	Ernesto Zedillo (11)	Peña Nieto (17)
Formas de tratamiento	<i>Usted</i> Fórmulas genéricas de trato cortés (<i>doctor, señorita</i>) Diminutivos peyorativos (<i>muchachito</i>)	<i>Usted / ustedes</i>	<i>Tú</i> Nombre del entrevistador

Debido al protocolo empleado para dirigirse al presidente, lo esperable es que los entrevistadores usen la forma de tratamiento que se presenta. En este sentido, resulta interesante que la forma de tratamiento cortés no aparece en la entrevista hecha a Peña Nieto, ya que únicamente se usa *presidente*. Asimismo, destaca que el uso de la forma pronominal *usted*, de respeto y lejanía, en el caso de Díaz Ordaz se combina con las fórmulas genéricas de trato cortés o descortés. Esto revela un juego de estrategias de confrontación en el uso de las formas de tratamiento del cual solo hace gala el mandatario que encarna el modelo del presidencialismo, como muestran los siguientes ejemplos:

- (2) Gustavo Díaz Ordaz: Disiento totalmente del criterio muy personal de usted de que hay un hecho que ensombreció [sic] la historia de México, hay un hecho que ensombreció la historia de unos cuantos hogares mexicanos. Yo le puedo decir a usted que estoy muy, muy contento de haber podido servir a mi país en tantos cargos como lo

he hecho, estoy muy orgulloso de haber podido ser presidente de la república y haber podido, así, servir a México, pero de lo que estoy más orgulloso de esos seis años es del año de 1968 porque me permitió servir y salvar al país, les guste o no les guste, con algo más que horas de trabajo burocrático, poniéndolo todo: vida, integridad física, horas, peligros, la vida de mi familia, mi honor y el paso de mi nombre a la historia, todo se puso en la balanza, afortunadamente salimos adelante y si no hubiera sido por eso usted no tendría la oportunidad, *muchachito*, de estar aquí preguntando. Lo que pasa es que usted no se acuerda, porque estaba usted muy chavito.

- (3) Gustavo Díaz Ordaz: No, *señorita*, es falso y también nació del mismo lugar, mire usted que si una condesa, uno quién sabe qué una española, que si una noruega, hasta, hasta totonaca, también me cargan [risas]. Le inventan a uno romances, matrimonios, amantes de personas que ni siquiera uno conoce. Hay una gente que dicen que ha sido mi amante, yo nunca la he visto.

Como puede observarse, la combinación de la forma de tratamiento esperada (*usted*) con formas genéricas permite construir cortésmente al interlocutor o descortésmente sin perder la forma de tratamiento esperada. Cabe señalar que, aparentemente, existe una pérdida de la combinación de formas de tratamiento y fórmulas genéricas de trato cortés como estrategia discursiva en los otros presidentes, que no responden al modelo del presidencialismo. Asimismo, es interesante que la pérdida de todas las medidas de control sobre las entrevistas que desplegaba el modelo presidencialista ocurra justo con la pérdida de esta gama de estrategias de combinación.

Por otra parte, en el caso de Peña Nieto podría suponerse que su uso del tuteo es parte de una estrategia actual en los medios informativos, esto debido a las fechas de la entrevista. Sin embargo, la revisión del acervo de entrevistas del cual se obtuvo el corpus desmiente esta idea. En *Conversaciones a Fondo*, programa en el que Peña Nieto participó junto con otros periodistas, el mandatario

usa la forma pronominal *usted* e incluso ciertas formas genéricas de respeto, por ejemplo, al dirigirse a José Carreño Carlón: “Al contrario, don Pepe, muchas gracias a usted”.⁹ En ese mismo espacio, también tutea a Lily Téllez. Otro caso que desmiente esa falsa generalización es el del presidente en turno, Andrés Manuel López Obrador. En una entrevista con Jacobo Zabludovsky —mucho más reciente que la tercera que se analiza en este corpus— se presenta el uso de *usted* y las formas genéricas de cortesía.¹⁰ Esto contrasta con otra entrevista, realizada por Carmen Aristegui cuando López Obrador ya es presidente electo, en la cual, aunque la periodista le pregunta si puede tutearlo y le aclara que cuando tenga la banda presidencial usará el *usted*, López Obrador la tutea.¹¹ Ambos casos evidencian que es necesario realizar un análisis detallado sobre las formas de tratamiento, ya que, si bien es evidente que los entrevistadores deben respetar el protocolo y referirse al presidente de modo formal, estos últimos parecen tener una gama más amplia de opciones para dirigirse a su entrevistador(a), ya sean mediante formas de tratamiento o formas genéricas corteses. Estos usos aportan a la construcción de la imagen de los entrevistadores y, al mismo tiempo, del propio presidente, debido a que posibilitan que estos se muestren de modo más o menos cortés y, por lo tanto, más o menos confrontativo o agresivo.

En el estudio de Cuenca se afirma que los vocativos “permiten observar usos marcados y transgresiones de las normas de género que remiten a la construcción/destrucción de la imagen del interlocutor y el refuerzo implícito de la propia argumentación e imagen de cara al público” (2013: 545). Frente a eso, lo que se observó en una primera instancia —y que no estaba contemplado dentro de los objetivos de este trabajo— es que los vocativos que suponen formas genéricas corteses por parte de los entrevistadores (*señor presidente* y *presidente*) aparecen con frecuencia al inicio

⁹ Véase https://www.youtube.com/watch?v=mAwF0q_dG44

¹⁰ Véase <https://www.youtube.com/watch?v=nLB0hAII8J8>

¹¹ Véase <https://www.youtube.com/watch?v=9gLNPhCKIWg>

de una pregunta cuando esta supone un cambio de tema con respecto a la anterior. En la bibliografía consultada no se reporta este aspecto del vocativo, por lo que resulta un elemento importante a analizar, pues los entrevistadores lo emplean no solo para seguir el protocolo y como marca de uso cortés, sino para delimitar subunidades informativas dentro de la entrevista.

Con respecto a la confrontación, en el corpus examinado no se transgreden las normas de la dinámica de la entrevista (turnos), sin importar el grado de confrontación que exista en la pregunta o en la respuesta. En particular, los turnos son fuertemente respetados en la selección que representa el modelo presidencialista, tanto por parte del entrevistado como del entrevistador. Esto también se observa en el hecho de que las preguntas realizadas a Díaz Ordaz obtuvieron un mayor porcentaje de respuestas en la interacción. Por el contrario, el uso de los vocativos disminuye radicalmente en las entrevistas en donde hay arrebatos de turnos, en los cuales las formas genéricas corteses tampoco aparecen.

Como se ha señalado previamente, el uso de *señor presidente* muestra una amplia cortesía por parte del entrevistador debido a la forma de tratamiento respetuosa y apegada al protocolo, por ejemplo, cuando Díaz Ordaz se refiere a su entrevistador como *doctor*. En contraste, el expresidente emplea *muchachito* cuando el periodista se atreve a preguntarle por su responsabilidad en la masacre del 2 de octubre, lo cual constituye un uso claramente descortés, ya que implica una dimensión peyorativa. En el modelo del presidencialismo, es el entrevistado quien puede cuestionar o transgredir la cortesía por medio de las formas en que se refiere a sus entrevistadores, haciendo uso de los vocativos, formas genéricas corteses, o mediante el empleo de otras que, al no incluirse en este tipo de formas de cortesía, adquieren un efecto despectivo o peyorativo. Por su parte, los entrevistadores no tienen ni han tenido la posibilidad de desplegar tal recurso.

Mientras que el representante del presidencialismo muestra una preferencia por la confrontación mediante las formas de tratamiento, Zedillo y Peña Nieto lo hacen a través de la toma de con-

trol de los turnos en las entrevistas; como ya se dijo, los vocativos no aparecen en este tipo de estrategias.

Desviarse de la progresión temática también supone un punto de tensión. Díaz Ordaz jamás utiliza esta estrategia, sin embargo, tanto Zedillo como Peña Nieto utilizan la evasión, ya que cambian el foco temático de la pregunta con el fin de hablar de otro tema. En esos casos los entrevistadores se ven obligados a tensar la situación al confrontar abiertamente e insistir para anular este recurso del entrevistado. Cabe mencionar que los vocativos no aparecen en la evasión de la temática de una pregunta ni en el intento de los periodistas para que el entrevistado conteste la pregunta realizada; no sucede lo mismo con las formas pronominales de segunda persona, pero esto se discutirá en §4.2. Dado que los vocativos son índices de que existe cooperación en la dinámica, esto nos hace suponer que hay menos confrontación y lucha por el control en las entrevistas, o bien un cambio de temática, en la pregunta que se abre con vocativos. Este comportamiento es diametralmente opuesto al reportado por Cuenca (2013), lo cual vuelve necesario el profundizar en el conocimiento del uso de los vocativos en los contextos mexicanos, como los que se han analizado exploratoriamente en este trabajo. Otro comportamiento que dista de los reportados se relaciona con el uso de vocativos por parte de los entrevistadores. Mientras que en el corpus analizado por Cuenca (2013) “hay más uso de vocativos por parte de la entrevistadora” (531), en las entrevistas examinadas esto ocurre solamente en las dos más recientes (Zedillo y Peña Nieto), pero no en la entrevista a Díaz Ordaz.

Cuenca menciona que en el corpus de su investigación “el tono inicial es distendido y amable, pero la tensión va en aumento y se incrementa cuando se habla de los temas polémicos (relaciones con su hermano, con la prensa y la marcha de los indígenas)” (2013: 531). En las entrevistas analizadas en este trabajo no se observa necesariamente lo mismo, ya que —de acuerdo con una investigación previa (Salgado López, 2020)— en la mayoría de ellas tanto la apertura como el cierre han sido eliminados, lo cual se explica por el control que la presidencia siempre ha mantenido

sobre el material que se hace público. Sin embargo, en el cuerpo central de la entrevista sí se evidencia este aumento en el grado de confrontación; en las unidades pregunta-respuesta hay un aumento de tensión que baja conforme se aproxima el fin de la entrevista.

Como bien se ha señalado, “el vocativo realiza otras funciones que permiten reflejar la posición relativa de los participantes entre sí y respecto a la sociedad: situarlos en un espacio social y definir o construir la relación entre ellos”¹² (Jaworski & Galasiński, 2000: 35). En esta misma línea, Clayman (2010) diferencia dos funciones generales: la desalineación y la expresividad. El autor define la desalineación, en sentido amplio, como un concepto que incluye “respuestas que son desafiliativas, así como aquellas que se apartan del tema o agenda de acción planteada por la pregunta anterior”¹³ (Clayman, 2010: 181). De acuerdo con esta propuesta, también retomada en el análisis de Cuenca (2013), los vocativos se presentan con mayor frecuencia para manifestar desalineación y desacuerdo. No obstante, en las entrevistas mexicanas no parecen comportarse así, pues los vocativos aparecen con mayor fuerza cuando existe más cooperación y menos tensión en la dinámica conversacional.

Finalmente, en el análisis se revisó si existía una tendencia de los vocativos a aparecer en ciertas posiciones, como lo afirman algunos autores que cita Cuenca (2013):

En posición intermedia, el vocativo refuerza desacuerdos formulados aseverativamente o como pregunta orientada. También puede indicar lucha por el turno o intento de marcar la pauta del intercambio, con frecuencia en un contexto de desacuerdo. [...]

¹² “Terms of address do not only reflect the relative positions of interactants vis-a-vis one another and in society as a whole. The speaker’s choice of a particular form of address locates the addressee in social space and defines or constructs, the social actors’ mutual relationship.” [Las traducciones de las citas son de la autora.]

¹³ “[R]esponses that are disaffiliative as well as those that depart from the topic or action agenda set by the previous question.”

En síntesis, se observa un uso diferenciado de los vocativos por parte de los dos interlocutores que no parece responder al uso prototípico en el que los periodistas lo utilizan para organizar el intercambio y el político, para tomar el turno, resolver solapamientos o retrasar una respuesta desalineada (Rendle-Short, 2007; Ilie, 2010: 889).

Si se evalúan los vocativos según la macroestructura y microestructura (turnos en unidades pregunta-respuesta) de las entrevistas políticas analizadas, estos no se comportan del modo reportado por los autores previamente citados. De hecho, en aquellos empleados por Díaz Ordaz —quien más los ha usado— uno de los vocativos más confrontativos (*muchachito*) aparece al final, al igual que la forma cortés *doctor*, que en muchas otras preguntas ocupa esa misma posición.

4.2. Sobre el uso de formas pronominales en función de sujeto

Como se muestra en el Cuadro 3, en la entrevista con Díaz Ordaz el uso de formas pronominales de segunda persona en función de sujeto es igual en términos cuantitativos, tanto de parte del entrevistado como del entrevistador. La mayoría de los usos aparecen acompañados de fórmulas genéricas de trato cortés o descortés (las menos), por lo que la forma marcada de explicitar el pronombre en posición sujeto parece estar amortiguada por la aparición de las fórmulas corteses (*doctor*, *señorita*), mientras que ocurre lo contrario cuando se acompaña de una fórmula descortés (*muchachito*). Como se había mencionado en el apartado de los vocativos, la estrategia de cortesía/descortesía de Díaz Ordaz para la construcción de sus interlocutores mediante formas apelativas de segunda persona se centra en el uso de los vocativos.

A diferencia de este comportamiento, en la entrevista hecha a Ernesto Zedillo, se observó cómo el uso de las formas pronominales de segunda persona en función de sujeto explícito aumenta considerablemente, lo cual, junto con la disminución del uso de

los vocativos, constituye una estrategia distinta que se acompañará de recursos de interacción, como el arrebató de turno, para marcar confrontación.

En la entrevista con Peña Nieto también se presenta un bajo uso de vocativos y un aumento en las formas pronominales en función de sujeto, pero tal tendencia es más marcada en la entrevistadora y no en el entrevistado. Peña Nieto muestra una preferencia por el arrebató de turno y el cambio temático en las preguntas como estrategia ante la confrontación, por lo que la entrevistadora debe apelar al uso constante del pronombre *usted* para insistir en que el expresidente siga la temática de las interacciones pregunta-respuesta. Esto evidencia que en el caso del expresidente hay un debilitamiento en el uso de estos recursos.

5. Conclusiones

El análisis del corpus de esta investigación corrobora lo que Cuenca y Marín (2015) afirmaron, esto es, que las formas apelativas de segunda persona permiten observar cómo “se construyen los interlocutores a ellos mismos como enunciadores que amenazan o ponen en peligro —o no— la imagen cortés de su interlocutor” (55).

Si bien en Cuenca (2013) se advierte que los hallazgos son pertinentes únicamente para la entrevista analizada, el corpus de esta investigación muestra que, para el caso mexicano, solo en la entrevista de Díaz Ordaz, representante del modelo presidencialista, se usan ambas formas apelativas de segunda persona: los vocativos que funcionan como fórmulas genéricas de trato cortés y los vocativos peyorativos para mostrar confrontación. Contrario a lo reportado, en los casos examinados, los vocativos disminuyen su presencia cuando se da mayor confrontación. Además, la confrontación se construye por medio de otros mecanismos, relacionados con el arrebató de turno, el desvío del tema de la pregunta y el uso de marcas de construcción del interlocutor como referente discursivo.

Según el comportamiento detectado en los resultados, una de las vetas de investigación posibles es aplicar procedimientos similares de análisis de las formas apelativas de segunda persona en entrevistas políticas a una cantidad mayor de entrevistas a presidentes mexicanos, con el objetivo de localizar cuándo se dejaron de usar vocativos, fórmulas de trato cortés y vocativos peyorativos, para mostrar confrontación de parte del entrevistado hacia sus entrevistadores, y cómo se comporta el uso de las formas pronominales en función de sujeto cuando disminuye el uso de los vocativos. Además, sería pertinente identificar el momento en que se dejó de usar la fórmula hecha *señor presidente*, pues es un indicador del debilitamiento del protocolo presidencial. Se recomienda también trabajar con corpus más amplios para corroborar o rechazar la función del uso de la fórmula hecha de trato cortés como indicador de cambio de tema en la macroestructura de las entrevistas.

Asimismo, es necesario extender el análisis a las formas pronominales en función de sujeto implícito empleadas para construir discursivamente al interlocutor o receptor. Esto permitiría contrastar los usos explícitos y tácitos, así como comprender con mayor profundidad el grado de confrontación y las estrategias de construcción de los interlocutores.

Por otra parte, sería recomendable examinar entrevistas realizadas a presidentes provenientes de otras asociaciones, como el Partido Acción Nacional (PAN) y el Movimiento de Regeneración Nacional (Morena), con el objetivo de contrastar lo que se ha observado en el caso del Partido Revolucionario Institucional. Otro aspecto interesante que podría compararse con los resultados obtenidos es la posibilidad de que el comportamiento poco confrontativo de los vocativos esté relacionado culturalmente con la fuerte tradición presidencialista que se sostuvo por largo tiempo y el desgaste de la figura presidencial con la llegada de los modelos neoliberales.

6. Referencias

- Araya Seguel, Claudio Alejandro (2011). *Análisis crítico de las representaciones discursivas del gobierno de Chile, a partir de entrevistas al presidente de la república Sebastián Piñera, desde de las perspectivas de las implicaturas convencionales y la teoría de la valoración* (Tesis de maestría inédita). Universidad de Santiago de Chile, Santiago de Chile. <https://glotopolitica.com/2019/07Bosque/01/analisis-critico-de-las-representaciones-discursivas-del-gobierno-de-chile-a-partir-de-entrevistas-al-presidente-de-la-republica-sebastian-pinera-des-de-las-perspectivas-de-las-implicaturas-convencio>
- Arfuch, Leonor (1995). *La entrevista, una invención dialógica*. Barcelona: Paidós.
- Arjonilla Sampedro, Alicia; Atienza Cerezo, Encarnación; Castro Carrillo, María Delia; Cortés Moreno, Maximiano; González Argüello, María Vicenta; Inglés Figueroa, Marta; Iruela Guerrero, Agustín; Lahuerta Galán, Javier; López Ferrero, Carmen; Montmany Molina, Begoña; Pueyo Villa, Silvia; Puig Soler, Fuensanta; Sánchez Quintana, Nuria; Torner Castells, Sergi; Vañó Aymat, Antonio; Wesenaar, Dafne, y Martín Peris, Ernesto (s. f.), *Diccionario de términos clave de ELE (en línea), Centro Virtual Cervantes*. https://cvc.cervantes.es/Ensenanza/biblioteca_ele/diccio_ele/default.htm
- Bermúdez Chaves, Marlen (1986). Aplicación del análisis de contenido a la entrevista. *Ciencias Sociales*, 33, 135–143.
- Carbó, Teresa (2001). Tocar el lenguaje con la mano: experiencias de método. *Revista Latinoamericana de Estudios del Discurso*, 1(1), 43–67.
- Carbó, Teresa (2007). Introducción. La elocuencia de los cuerpos. *Estudios de Lingüística Aplicada*, 46, 13–30.
- Clayman, Steven (2010). Address terms in the service of other actions: The case of news interview talk. *Discourse & Communication*, 4(2), 161–183.
- Cortés Rodríguez, Luis, & Bañón Hernández, Antonio Miguel (1997). *Comentario lingüístico de textos orales II: el debate y la entrevista*. Madrid: Arco Libros.
- Cuenca, María Josep (2013). Uso del vocativo en la entrevista política: género discursivo y (des)cortesía. *Discurso & Sociedad*, 7(3), 522–552.
- Cuenca, María Josep, & Marín, María Josep (2015). La representación discursiva del adversario en el debate electoral. *Oralia: Análisis del discurso oral*, 18, 45–79.

- Depetris, Cielito, & García, María Eugenia (2008). Las emisoras públicas de la ciudad de La Plata: rasgos estilísticos de su programación. El caso de LS 11 Radio Provincia (de la Provincia de Bs. As.) y LR 11 Radio Universidad (de la Universidad Nacional de La Plata). *Tram(p)as de la Comunicación y la Cultura*, 63, 16–56.
- Fetzer, Anita (2000). Negotiating validity claims in political interviews. *Text & Talk*, 20(4), 415–460.
- Fetzer, Anita (2006). “Minister, we will see how the public judges you.”: Media references in political interviews. *Journal of Pragmatics*, 38(2), 180–195.
- Gómez Sánchez, María Elena (2008). Actividades de imagen en la entrevista política: algunos ejemplos de estrategias discursivas para la configuración de la imagen pública. *Anuario de Estudios Filológicos*, 31, 21–35.
- Grize, Jean-Blaise (1993). Logique naturelle et représentations sociales. *Papers on Social Representations*, 2, 151–159.
- Gutiérrez, Silvia (2006). Discurso político y argumentación. https://www.researchgate.net/publication/251886295_DISCURSO_POLITICO_Y_ARGUMENTACION
- Halperín, Jorge (1995). *La entrevista periodística: intimidades de una conversación pública*. Buenos Aires: Paidós.
- Jaworski, Adam, & Galasiński, Dariusz (2000). Vocative address forms and ideological legitimization in political debates. *Discourse Studies*, 2(1), 35–53.
- Laborda, Xavier (2000). La política como conversación: Análisis de un discurso de legitimación. *Círculo de Lingüística Aplicada a la Comunicación*, 2.
- Leech, Geoffrey (1983). *Principles of pragmatics*. Londres: Longman.
- Martín Rojo, Luisa (2000). Enfrentamiento y consenso en los debates parlamentarios sobre la política de inmigración en España. *Oralia*, 3, 113–148.
- Mullany, Louise (1999). Linguistic politeness and sex differences in BBC Radio 4 broadcast interviews. *Leeds Working Papers in Linguistics and Phonetics*, 7, 119–142.
- Pérez Minchola, Ninosca (2007). *El rol de la cortesía en entrevistas políticas conflictivas en la televisión peruana* (Tesis de maestría). Univer-

- sidad de Georgia, Athens. <https://esploro.libs.uga.edu/esploro/outputs/9949333934802959>
- Real Academia de la Lengua Española & Asociación de Academias de la Lengua Española (2010). *Nueva gramática de la lengua española. Manual*. Madrid: Espasa.
- Salgado López, Melanie del Carmen (2019). Concessive constructions, natural logic, and social representations. *Papers on Social Representations*, 28(1), 11.1–11.19.
- Salgado López, Melanie del Carmen (2020). *Presidencialismo, tecnocracia y modernidad: construcción ecoica de aliados y enemigos en entrevistas a Díaz Ordaz, Ernesto Zedillo y Peña Nieto* (Tesis doctoral inédita). Universidad Nacional Autónoma de México, Ciudad de México.
- Stenström, Anna-Brita (1994). *An introduction to spoken interaction*. Londres: Routledge.
- van Dijk, Teun (1999). El análisis crítico del discurso. *Anthropos*, 186, 23–36.
- Verón, Eliseo (1987). La palabra adversativa: observaciones sobre la enunciación política. En Elvira Arnoux (Dir.), *El discurso político: lenguajes y acontecimientos* (pp. 11–27). Buenos Aires: Hachette.
- Yanes Mesa, Rafael (2006). La entrevista como género de la comunicación política. *Chasqui. Revista Latinoamericana de Comunicación*, 96, 52–55.

Lineamientos para publicar en *Estudios de Lingüística Aplicada*

Todos los manuscritos propuestos para publicación en *Estudios de Lingüística Aplicada* (ELA) deben ser enviados a través de su sitio en internet (ela.enallt.unam.mx). Estos lineamientos, así como las instrucciones para hacer los envíos en línea, pueden consultarse en dicho sitio.

De acuerdo con los principios del doble dictamen ciego, en los manuscritos que se envíen por primera vez deberá omitirse el nombre, la adscripción y la dirección de contacto del autor. Además, deberá eliminarse la identificación del autor/a de las propiedades del archivo (véase “Garantizar una evaluación por pares anónima” en el sitio web de la revista). Esta información se solicitará a los autores en el proceso de envío en línea. Posteriormente, y solo si el manuscrito ha sido aceptado, estos datos se incorporarán en el archivo en Word de la versión final del manuscrito.

Los artículos de investigación deberán tener una extensión máxima de 30 cuartillas, las notas de 20 cuartillas y las reseñas de 10 cuartillas, sin incluir referencias bibliográficas y anexos, con interlineado doble, en letra Times New Roman de 12 puntos, en formato carta (21.59 cm x 27.94 cm) y respetando los siguientes márgenes: izquierdo 3 cm, derecho 3 cm, superior 2.5 cm, inferior 2.5 cm. Los cuadros y figuras deberán incluirse dentro del texto en el sitio que les corresponde y presentarse por separado en el formato del programa original en el que fueron creados, así como

* *Estudios de Lingüística Aplicada* sigue la hoja de estilo de *The generic style rules for Linguistics* (Haspelmath, 2014) y, para el formato de las referencias, el manual de publicaciones de la American Psychological Association (APA, 2010), con algunas modificaciones. La elaboración de los lineamientos para publicar en *Estudios de Lingüística Aplicada* estuvo a cargo de Sergio Ibáñez Cerda, Zazil Sobrevilla Moreno y Mariela Elizabeth Pérez Castañeda, con la colaboración de Blanca Estela Solís Valdespino, Armando Rodríguez Briseño y Sandra Contreras Macías.

en archivo de imagen (véase *infra* §11). La versión del artículo en formato PDF y los cuadros y figuras en archivo de imagen son indispensables para controlar los cambios de procesadores y los caracteres fonéticos u otros detalles en el proceso editorial.

Se sugiere a los autores cuya lengua materna no sea el español o el inglés (según la lengua en que esté escrita la contribución) que un hablante nativo revise su manuscrito antes de que sea enviado. Las contribuciones en español deberán seguir las innovaciones ortográficas aprobadas por la Real Academia Española y la Asociación de Academias de la Lengua Española en 2010 (RAE & Asale, 2010).¹

Se recomienda minimizar el uso de marcas tipográficas tales como negritas, versalitas, cursivas, comillas, etc., y conservarlas solo cuando sea indispensable. Cualquier necesidad editorial que esté fuera de las especificaciones de estos lineamientos deberá ser comunicada y puesta a consideración de los editores.

1. Partes del texto

El texto del artículo va precedido por el título y por un resumen, de 1000 a 1200 caracteres con espacios, seguido por cinco palabras clave que no coincidan con las que aparecen en el título del artículo. El título, el resumen o *abstract* y las palabras clave o *keywords* deberán proporcionarse en la lengua principal del texto y en su traducción al español o al inglés, según sea el caso. Tanto las palabras clave como las *keywords* deben ir separadas por punto y coma y no llevar punto final.

¹ Se seguirá, por ejemplo, la eliminación de la tilde diacrítica en el adverbio *solo* y en los pronombres demostrativos (*este, ese, aquel*, etcétera), incluso en casos de posible ambigüedad. Un resumen sobre las principales novedades de la última edición de la *Ortografía* se ofrece en Real Academia Española (RAE, 2010) y puede recuperarse en http://www.rae.es/sites/default/files/Principales_novedades_de_la_Ortografia_de_la_lengua_espanola.pdf

Los artículos se subdividen en secciones numeradas y, de ser necesario, en subsecciones, cada una de las cuales lleva un encabezado. La numeración empieza siempre en 1 (Sección 1: 1.1, 1.2, Sección 2: 2.1, 2.2, etc.). Solo en casos excepcionales se empleará un tercer nivel de división, que llevará también un encabezado. La última sección numerada puede ir seguida de varias secciones opcionales (Corpus o Fuentes, Abreviaturas, etc.), y uno o más Anexos (A, B, etc.). Al final del documento aparece la sección de Referencias. Las normas de estilo de los distintos tipos de referencias bibliográficas se detallan en §15.

Los encabezados de las secciones no llevan punto final ni más mayúsculas que la inicial (véase §2). Si una subsección tiene (sub-)subsecciones, deberá haber al menos dos de ellas. Todo el texto deberá pertenecer a alguna sección, aunque, excepcionalmente, podrá haber un párrafo introductorio.

2. Uso de mayúsculas

Las oraciones, nombres propios, títulos y encabezados inician con mayúscula. Obsérvese que en los manuscritos escritos en inglés no se emplean mayúsculas en el resto de las palabras que componen títulos y encabezados de secciones, figuras, cuadros, etc. (es decir, no se emplea *title case*). Con respecto a las referencias bibliográficas, en los títulos de artículos y libros en inglés también se seguirá este principio, si bien los títulos de las revistas y series respetarán el uso de las mayúsculas de cada publicación, pues se tratan como nombres propios. Si el título de una referencia es en inglés y contiene un subtítulo, el subtítulo iniciará con letra mayúscula solo si se trata de una oración completa; si es en español, siempre iniciará con minúscula. Se presentan a continuación algunos ejemplos:

1.1. Overview of the issues (NO: Overview of the Issues)

Figure 3. A schematic representation of the workflow (NO: A Schematic Representation of the Workflow)

Anderson, Gregory (2006). *Auxiliary verb constructions*. Oxford: Oxford University Press. (NO: *Auxiliary Verb Constructions*)

Clyne, Michael (Ed.) (1991). *Pluricentric languages: different norms in different nations*. Berlín: De Gruyter Mouton.

Cuando se haga referencia a una parte del artículo numerada, esta iniciará con mayúscula (Figura, Cuadro, Apéndice), *e. g.*

como se muestra en el Cuadro 5

En ningún caso se usarán mayúsculas para hacer énfasis o introducir conceptos o como sistema de marcación tipográfica. Tampoco se usarán para iniciar los nombres de leyes, teorías, modelos, procedimientos estadísticos o hipótesis.

3. Cursivas

Las cursivas se emplean para los siguientes propósitos:

- Para todas las formas de la lengua objeto (letras, palabras, frases, oraciones) citadas dentro del texto, a menos que se trate de transcripciones fonéticas o representaciones fonológicas en Alfabeto Fonético Internacional (AFI).
- Cuando los ejemplos en párrafo aparte sean de lenguas que no sean el español o el inglés, y por lo tanto vayan glosados. Los ejemplos en español y en inglés no glosados irán en redondas.
- Para llamar la atención del lector en algún aspecto particular de un ejemplo lingüístico en un ejemplo numerado que aparezca en redondas.
- Para palabras en una lengua distinta a la del texto principal.
- Cuando se hace referencia a un término técnico como forma de la lengua objeto, por ejemplo:

el término *citativo* no es apropiado para esto.

- Cuando se hace referencia a un término técnico en primera mención. En tales contextos, los términos técnicos se manejan como formas de la lengua objeto, por ejemplo:

la noción de *implicatura convencional*, que durante algún tiempo había mantenido un bajo perfil

- Para enfatizar alguna palabra que no es un término técnico, *e. g.*

Esto es posible en este contexto, pero *solo* en este contexto.

- Para enfatizar algo en una cita, siempre con la indicación al final de la misma:

[*énfasis añadido*]

- Para títulos de libros, revistas, tesis y otras fuentes como periódicos, videos y películas.
- Para los símbolos estadísticos (*N, n, ds, etc.*).²

4. Versalitas

Las versalitas se emplean:

- Cuando se considere necesario llamar la atención sobre un término en su primera mención o cuando se ofrece su definición, *e. g.*

Se distinguen dos tipos de alineamiento principales, a saber,
NOMINATIVO-ACUSATIVO Y ERGATIVO-ABSOLUTIVO.

En la medida de lo posible, evite el uso de las versalitas (o de cualquier marcación tipográfica) para introducir términos que sean de

² Excepto las letras griegas, subíndices y exponentes que sirven como identificadores (es decir que no son variables) y las abreviaturas que no son variables.

uso general en la teoría lingüística, por ejemplo, sujeto, objeto directo, pronombre, pretérito, cuando se usen en su sentido convencional.

- También se emplean en las abreviaturas de categorías en las glosas interlineales (véase §10.1) y se pueden usar para indicar foco o prominencia acentual en las oraciones de ejemplo:

(1) PINTURA salía de las tuberías.

- Finalmente, en las contribuciones en español, los siglos se marcarán en versalitas: siglo XVI, XIII, XX, etcétera.

5. Negritas y otras formas de énfasis

Las negritas se utilizan:

- Para llamar la atención del lector en algún aspecto particular de un ejemplo lingüístico, sea en un ejemplo numerado o en un ejemplo que aparezca en el texto cuando estos estén en cursivas:

(2) *das Kind*, **dem** *du* *geholfen hast*
El niño.NOM quien.DAT tú.NOM ayudado has
'el niño a quien ayudaste'

- En general, no se emplean mayúsculas ni subrayado con el fin de destacar algún elemento. De manera excepcional, se podrá emplear el subrayado para resaltar una letra en una palabra de ejemplo y en otros casos donde no se puedan emplear otras formas de énfasis.

6. Uso de comillas

Las comillas dobles se usan en las siguientes situaciones:

- Cuando en el texto se cita un pasaje de otro texto, por ejemplo:³

³ Las citas en párrafo aparte no llevan comillas.

De acuerdo con Takahashi (2009: 33), “los citativos nunca se empleaban en oraciones subordinadas en japonés antiguo”.

- Cuando se menciona un término técnico o alguna expresión que el autor no quiere adoptar, como en el ejemplo de abajo. Alternativamente, en estos casos se puede también usar cursivas:

A esto se le llama comúnmente “pseudo-conservadurismo”, pero no emplearé ese término, pues se presta a confusión.

Las comillas simples se emplean:

- Para significados lingüísticos, por ejemplo:

El latín *habere* ‘tener’ no es cognado con *hafian* ‘tener’ del inglés antiguo.

- Para marcar citas dentro de citas.

7. Empleo de otros signos ortográficos

- En las acotaciones parentéticas —como en este ejemplo— se emplea el guion largo o raya, sin que sea precedido ni seguido de espacios.
- En caso de que el artículo esté escrito en inglés, se emplea mismo signo, antecedido y seguido de espacios, por ejemplo:

The starting hypothesis is that the prosodic expressiveness is divided into two major types — encouragement and discouragement — described through several restrictions within the framework of the Optimality Theory.

- Para indicar rangos numéricos se emplea el guion medio, por ejemplo, 1995–1996.
- La elipsis dentro de una cita se indica con [...], *e. g.*

“Barcelona aceptaba la noche sobre el mar [...] Las ciudades se aceptan porque abrigan.”

- Los corchetes angulares < > se emplean para hacer referencia a grafemas y demás símbolos escritos específicos, por ejemplo, <q>.
- Los millares, millones, etc., no se separan con coma ni con punto (por ejemplo: 35 542 y no 35,542, ni 34.542). La separación se hará con espacios en blanco en los números de más de cuatro cifras, mientras que los números de cuatro cifras se escribirán sin espacios de separación: 1000, 10 000, 100 000, 1 000 000, 100 000 000, etcétera. Los decimales van precedidos por un punto, no por coma (por ejemplo 2.35 y no 2,35).
- En los textos escritos en español no se usará la tilde diacrítica en el adverbio *solo* ni en los pronombres demostrativos (*este, ese, aquel*) incluso en casos de posible ambigüedad (por ejemplo: *solo llevaba un par de monedas en el bolsillo, este es tonto, quiero aquella* y no *sólo llevaba un par de monedas en el bolsillo, éste es tonto, quiero aquélla*). Tampoco se utilizará tilde diacrítica en la conjunción disyuntiva *o* escrita entre cifras (*1 o 2* y no *1 ó 2*).

8. Abreviaturas

- Se deben evitar las abreviaturas de expresiones no comunes.
- Los nombres de lenguas no deberán ser abreviados.
- Las abreviaturas se deben usar para etiquetas de categorías gramaticales en las traducciones morfema por morfema. Las reglas de glosado de Leipzig incluyen una lista de abreviaturas estándar para las categorías gramaticales más usadas (véase §10.1).
- Cuando se hace referencia repetida a un término complejo que no es ampliamente conocido, se podrá abreviar (por ejemplo, CDO para “construcción de doble objeto”). En tal caso, la abreviatura se deberá explicar la primera vez que aparece en el texto y también al final del artículo en la sección Abreviaturas. Las abreviaturas de expresiones no comunes no se usarán en los encabezados ni en los pies de figuras y cuadros y se deben evitar al inicio de las secciones.

9. Citas dentro del texto

Las citas textuales, cuando son menores de 5 líneas se incorporan en el texto entre comillas; cuando son de 5 líneas o más, se despliegan en un bloque independiente (a bando y separado con espacios antes y después), sin comillas y con tipo menor. En cualquier caso debe señalarse la fuente. Las citas en lenguas distintas a las de la lengua principal del texto deberán traducirse, dentro del texto entre corchetes si son cortas y en nota al pie si son largas.

Las fuentes se citan señalando el autor, el año del trabajo y la página específica de la cita. El autor puede aparecer como un elemento dentro del texto principal (como en el primer ejemplo que se muestra abajo) o entre paréntesis (como se muestra en el segundo ejemplo). Los números de página van después del año, separados por dos puntos y un espacio, sin omitir dígitos. Cuando aparece el nombre del autor dentro del paréntesis, este se separa del año con coma.

Thomason y Kaufman (1988: 276–280) señalan que los dialectos del inglés del norte tienen más innovaciones morfológicas que los dialectos de inglés del sur.

La notación que empleamos para representar esto se toma de las teorías que afirman que los rasgos- ϕ aparecen en una geometría de rasgos (Gazdar & Pullum, 1982: 34).

Las referencias bibliográficas completas correspondientes a cada trabajo citado se enlistan alfabéticamente al final del texto. A continuación se hacen algunas especificaciones sobre el formato de las citas dentro del texto:

- En caso de que se cite una obra de dos autores, ambos nombres se coordinarán con *y* cuando aparecen como un elemento del texto principal —como en el primer ejemplo arriba—, y con *&* cuando apa-

recen entre paréntesis —como en el segundo ejemplo. En el primer caso la referencia es a los autores y en el segundo a la obra.

- Cuando hay más de dos autores, la referencia, en la primera mención, incluye a todos y, a partir de la segunda, solo al primero, seguido por la abreviatura *et al.* (véase el ejemplo abajo). En la lista de referencias al final del texto se dará la lista completa de los autores.

López *et al.* (1992) = López, García & Silva (1992)

- Cuando se citan múltiples trabajos de un mismo autor, el nombre no deberá repetirse y los años se separarán por coma.

Mientras que Hawkins (2004, 2014) ha argumentado en favor de un principio de minimización de dominio, otros autores han tratado de explicar los efectos observados en términos puramente gramaticales.

Estudios empíricos previos reportan que la anteposición de objeto en estas lenguas ocurre bajo las mismas condiciones contextuales para los verbos transitivos canónicos y para los verbos de objeto-experimentante (véase Verhoeven, 2008b, 2010a para el turco y el chino).

- Cuando se enlistan varias referencias en paréntesis, estas se separan por punto y coma y normalmente se ordenan cronológicamente (aunque si el autor lo considera adecuado para la exposición, pueden seguirse otros criterios como el alfabético o por orden de importancia).

Los hablantes se basan mucho en trozos formulaicos o “pre-fabricados” durante la producción y comprensión del habla (Pawley & Syder, 1983; Sinclair, 1991; Erman & Warren, 2000; Bybee, 2006; véase Wray, 2002 para un panorama histórico más amplio).

- En algunos casos, cuando se cita una reimpresión de un texto antiguo o una referencia clásica, es conveniente dar el año de la publicación citada, seguida por el año del texto original entre corchetes, por ejemplo:

Correas (1984 [1626])

Lapesa (2000 [1973])

- Las comunicaciones personales (cartas privadas, memorandos, mensajes electrónicos, etc.) no se incluyen en la lista de referencias, pero sí deben aparecer citadas en el texto. Para ello, se debe proporcionar los datos del autor, así como una fecha tan exacta como sea posible:

Thomas Lutes (comunicación personal, 18 de abril, 2001)

- Los números de página deben escribirse completos, sin omitir dígitos; tampoco se usará la abreviatura “y ss.” para ‘y siguientes’.
- Los números de página se podrán omitir solo si la referencia concierne al trabajo entero. Se pueden dar números de capítulo o de sección en lugar de números de página, si se considera pertinente (por ejemplo: *Auer, 2007: capítulo 7; Matras, 2009: §6.2.2*). Asimismo, para la citación de material sin paginación, como en algunas fuentes en línea, puede hacerse referencia a otras secciones como el párrafo, por ejemplo: *Reynolds, 2016: párr. 5*.

Es importante asegurarse de que las referencias bibliográficas correspondientes a cada trabajo citado se encuentren enlistadas al final del texto en la sección de Referencias (véase §15) y viceversa, y que la ortografía, nombres de los autores, fechas y paginado correspondan en ambos lugares.

10. Ejemplos en párrafo aparte

- Todos los ejemplos en párrafo aparte deben ir numerados. Esta numeración debe aparecer entre paréntesis.
- Cuando hay múltiples ejemplos dentro de un mismo número (“sub-ejemplos”), se distinguen por las letras *a*, *b*, etc., seguidas de un punto.

- (3) a. Él la vio a ella.
b. Ella lo vio a él.

- En las referencias cruzadas a ejemplos también se emplearán paréntesis, incluso cuando la referencia está dentro de un comentario en paréntesis:

Como se muestra en (6) y (8–11), esta generalización se extiende a las construcciones transitivas, pero (29b) es una excepción.

En todos los demás contextos, el acento recae en la segunda sílaba (véase (15a)).

- A menos que los ejemplos sean del español o el inglés, deberán ser glosados y traducidos (véase abajo §10.1.). El texto de los ejemplos numerados irá en redondas cuando se trate de ejemplos que requieran glosas. Si el texto principal está escrito en español y el ejemplo está en inglés, el autor debe ofrecer una traducción general, en otra línea entre comillas simples (véase el siguiente ejemplo), y puede omitir la glosa interlineal si no lo considera relevante para la discusión. Este mismo criterio es pertinente para los textos escritos en inglés en los que se ofrezcan ejemplos en español.

- (4) John: Condi is smart but conservative.
‘John: Condi es listo pero conservador.’

- En los ejemplos lingüísticos que no requieren glosas, las fuentes se presentan después del ejemplo; en los ejemplos glosados, directamente después de la traducción idiomática, como se muestra en (5):

(5) Luganda

Maama a-wa-dde taata ssente.
 madre ella.PRS-dar-PRF padre dinero
 ‘La madre le ha dado dinero al padre.’ (Ssekiryango,
 2006: 67)

- Si la fuente no es una referencia bibliográfica, sino el título de un texto (probablemente inédito), como en (6), o el nombre o clave de identificación de un informante o del material elicitado, como en (7), se cita la fuente en corchetes y el artículo deberá incluir una sección especial antes de las Referencias donde se especifique más información sobre dichas fuentes. Esta sección, por lo general, se titula *Corpus* o *Fuentes*.

(6) Jalonke

I sig-aa xon-ee ma.
 2SG ir-IPFV extraño-DEF en
 ‘Vas con el extraño.’ [Mburee 097]

- (7) No entendía la sensación de que era un chiste. Desde el principio pensé que no era y al final no le entendí la gracia. [Alejandra, 5to, chiste 6]

10.1. *Ejemplos glosados*

Cuando un ejemplo proviene de una lengua distinta a la del texto principal, se proporciona una glosa interlineal, alineada palabra por palabra, en la segunda línea. Enseguida se proporciona la traducción idiomática. Véase (8).

(8) Islandés

Storm-ur-inn rak bát-inn.
tormenta-NOM-DEF conducir barca.ACC-DEF
á land.
en tierra
'La tormenta condujo la barca a tierra.'

Las convenciones para el glosado interlineal se pueden encontrar en las Reglas de Glosado de Leipzig,⁴ que son el estándar mundial. El principio básico es que cada elemento del texto primario corresponda a un elemento de la línea de glosa, y los símbolos que marcan los límites entre elementos (especialmente el interno a la palabra (-) y el de límite de clítico (=)) deben estar presentes en el texto primario y la glosa. Las etiquetas de categoría gramatical abreviadas se escriben en versalitas y la traducción idiomática con comillas simples. Se debe proporcionar una lista de abreviaturas al final del trabajo. La correspondencia y alineación entre las palabras del texto primario y las glosas deben ser claras; las palabras y sus glosas deben separarse utilizando la tecla Tab (␣).

En los ejemplos glosados:

- El texto primario del ejemplo va en cursivas, con uso regular de mayúsculas y signos de puntuación: se inician con mayúscula y se cierran con punto final, como en (8) arriba, excepto cuando el ejemplo no es una oración completa, como se muestra en (9).

(9) *das Kind, dem du geholfen hast*
El niño.NOM quien.DAT tú.NOM ayudado has
'el niño a quien ayudaste'

⁴ Las Reglas de Glosado de Leipzig (Comrie, Haspelmath & Bickel, 2008–2015) están disponibles en www.eva.mpg.de/lingua/resources/glossing-rules.php

- Cuando la lengua del ejemplo no suele escribirse, el texto primario puede carecer de mayúsculas iniciales y puntuación normal, por ejemplo:

(10) Hatam

a-yai bi-dani mem di-ngat i
 2SG-tener a-mí para 1SG-ver INT
 ‘¿Me lo darías para que pueda verlo?’ (Reesink, 1999: 69)

- La línea de glosa va en redondas y no lleva mayúsculas ni puntuación. La traducción idiomática tiene nuevamente el uso normal de las mayúsculas y de puntuación, como se ve arriba en (8).
- Cuando se mencionan varias lenguas en un mismo trabajo, se dará el nombre de la lengua en la línea del número del ejemplo, como en (10) y (11). Solo en estos casos el nombre de la lengua comenzará con mayúscula por ser el único elemento en esa línea.

(11) Sakha

a. *En bytaan buol-uoq-uy*
 tú lento ser-FUT-2SG
 ‘Serás lento’ (Baker, 2012: 7)

b. **En bytaan-yaq-yy*
 tú lento-FUT-2SG
 (‘Serás lento’) (Baker, 2012: 7)

- Para los ejemplos agramaticales se proporcionará la traducción idiomática buscada entre paréntesis, como en (11b). También se podrá proporcionar una traducción literal entre paréntesis después de una traducción idiomática, por ejemplo:

(12) Japonés

Tsukue no ue ni hon ga aru.
 mesa GEN alto en libro SUJ estar
 ‘Hay un libro sobre la mesa.’ (Lit. ‘En lo alto de la mesa está un libro.’)

- El texto de la lengua objeto puede aparecer en dos líneas, una no analizada (“superficial”) y una analizada, en caracteres redondos, que puede contener una representación más abstracta, por ejemplo:

(13) Karbi

amatlo la kroikrelo

amāt=lo là krōi-Cē-lò

y.luego=FOC este aceptar-NEG-RL

‘Y luego, ella disintió’ (Konnerth, 2014: 286)

- Los corchetes angulares o diples no van nunca en cursivas, incluso si el texto que aparece dentro de ellos va en cursivas.

11. Cuadros y figuras

- En caso de que se presenten cuadros y/o figuras, estos deberán integrarse en el cuerpo del texto y, además, presentarse por separado en el formato del programa original en el que fueron creadas (por ejemplo, Word, Excel, Praat, ELAN), así como en archivo de imagen (en formato tif, png o jpg). Procure que la resolución de las figuras sea lo más alta posible (no menor a 600 dpi) para que las imágenes sean nítidas en la versión impresa. Asegúrese de que todo elemento gráfico sea pertinente y que contenga los datos necesarios para su lectura.
- El texto debe indicar al lector cuando convenga la consulta de estos objetos gráficos, de aquí que sea necesario darles una numeración. La numeración deberá ser consecutiva y la palabra *figura* o *cuadro* deberá llevar inicial mayúscula (Cuadro 1, Cuadro 2; Figura 1, Figura 2). Los objetos deben insertarse tan cerca como se pueda del punto en el que se los menciona (véase ejemplo *infra*).
- Cada cuadro y figura debe tener un título, idealmente no mayor a una línea, con la palabra *figura* o *cuadro* en versal-versalitas (CUADRO 1, CUADRO 2; FIGURA 1, FIGURA 2) seguida de punto. El título no se cierra con un punto y en él no se hace un uso especial de las

mayúsculas. Este va como encabezado en el caso de los cuadros y al pie en el caso de las figuras.

Ejemplo de una figura y de su mención en el texto principal:

La mayor amplitud la muestra el vocativo de duda, lo que se refleja en un final marcadamente ascendente (Figura 18).

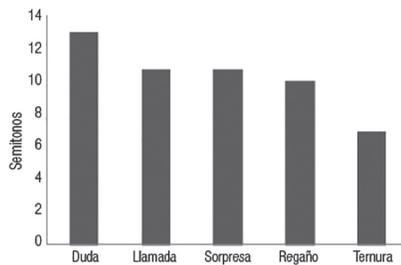


FIGURA 18. Campo tonal de los vocativos en todos los dialectos

FIGURA 18. Campo tonal de los vocativos en todos los dialectos

- Los cuadros generalmente están delimitados por una línea superior y una línea inferior, además de una debajo de los encabezados de columna. Se recomienda editar el cuadro de tal manera que solo sean visibles las líneas necesarias.
- Los cuadros y figuras pueden tener notas y llamadas internas. Estas irán al pie del cuadro y se considerarán independientes a las notas al pie de página en el resto del texto. Bajo el rótulo de *Nota* irán las explicaciones o informaciones generales del cuadro; las llamadas servirán para brindar información específica sobre alguna columna, fila o celda en particular. Estas últimas se indicarán con asterisco(s) (*) o con letras minúsculas (*a*, *b*, *c*) a manera de superíndices (véase ejemplo *infra*). Ejemplo de un cuadro con notas:

CUADRO 1. Porcentaje de la media de respuestas correctas para cada unidad según grado

UNIDAD	POSICIÓN	1ER GRADO ^a		3ER GRADO ^b	
		Media	DE	Media	DE
Sílabas	Inicial	71.49	29.16	94.51	8.6
	Final	57.59	33.56	87.35	12.63
	Rima	62.76	36.14	84.71	22.59

Nota: Los niños de 1^{er} grado son lectores iniciales y los de 3^{er} grado, lectores entrenados.

^a Los niños de 1^{er} grado son lectores iniciales.

^b Los niños de 3^{er} grado son lectores entrenados.

- Como figuras se considerarán las gráficas en eje coordenado o de pastel, diagramas (incluyendo estructuras arbóreas), mapas, dibujos, fotografías e imágenes. Señale, cuando sea necesario, la fuente de las figuras dentro del texto de la nota. Considere que estas irán en blanco y negro en la versión impresa y tome esto en cuenta para su elaboración, sobre todo cuando tenga que mostrar contrastes.

12. Referencias cruzadas en el texto

Para las referencias cruzadas a cuadros, figuras o notas se usan los nombres de estos elementos en mayúsculas (por ejemplo: Figura 3, Cuadro 2, Nota 17). No se emplearán abreviaturas, como “Fig. 3”, “Cap. 4”, o “n. 17”. Las referencias cruzadas a secciones emplearán el carácter § (por ejemplo, §2.3).

13. Notas al pie

El número de referencia para nota al pie debe ir inmediatamente después de un punto o coma, aunque en casos excepcionales irá inmediatamente después de una palabra. Su numeración comenzará con 1. Los agradecimientos se darán en una nota al pie cuya llamada será un asterisco (*) al final del título del trabajo. Las abreviaturas y otras convenciones notacionales, como los sistemas de transcripción, se incluirán en una sección aparte.

Los ejemplos numerados en las notas al pie llevan los números (i), (ii), etc. Si hubiera sub-ejemplos, estos seguirán la numeración (i.a), (i.b), etcétera.

14. Alfabetos no latinos

Todas las formas en lenguas que normalmente no se escriben en alfabeto latino (como japonés o armenio) deben acompañarse de una transcripción o de una transliteración.

Cuando el artículo versa en su totalidad sobre una de estas lenguas, no se omitirá la escritura original, al menos en los ejemplos numerados. Las formas no-latinas no necesitan ir en cursiva.

15. Lista de referencias

La lista de referencias al final del artículo lleva el encabezado Referencias. Las entradas se enlistan alfabéticamente. Se incluirán solo las obras referenciadas explícitamente en el texto.

15.1. Puntos generales

Para el formato de las referencias, *ELA* adopta en casi todos sus aspectos el modelo propuesto por la American Psychological Association (APA, 2010) con algunas excepciones.

Nótese que:

- Se deben incluir los nombres completos de todos los autores y editores, a menos que el autor use habitualmente nombres abreviados (por ejemplo, R. M. W. Dixon). Los nombres se considerarán completos según las convenciones de cada lengua, por ejemplo, los de autores anglosajones irán con un solo apellido y los de autores hispanos con los dos apellidos, salvo cuando sea convención de un autor utilizar uno solo.

- Los números de página son obligatorios tanto en las referencias de artículos de revista como en las referencias de artículos o capítulos en volumen editado. Los números de volumen de las revistas y los títulos de series son opcionales (aunque se recomienda incluirlos).
- Los títulos de revistas, editoriales y asociaciones no se abrevian.
- En el caso de las referencias que hayan sido consultadas en la red, es indispensable incluir el doi o la URL que permita su localización (véase §15.7).

Es responsabilidad del autor corroborar que toda la información de su lista de referencias sea precisa y completa. Se recomienda verificar cada entrada contra la publicación original. Las referencias precisas contribuyen a la credibilidad del autor como investigador cuidadoso y son signo de una ética profesional que valora el trabajo de los demás investigadores.

15.2. Partes estándar y tipos de referencia estándar

Hay cuatro tipos de referencia estándar: artículo de revista, libro, artículo o capítulo en libro editado, tesis. Los distintos tipos de referencia se componen de diferentes partes, como se muestra en el Cuadro 1.

CUADRO 1. Partes estándar de las referencias bibliográficas

	Lista de autores	Año	Título del artículo o capítulo	Lista de editores	Título de la publicación	Número de volumen	Número de páginas	Ciudad	Editorial
Artículo de revista	*	*	*		*	*	*		
Libro	*	*			*			*	*
Capítulo o artículo en libro editado	*	*	*	*	*		*	*	*
Tesis	*	*			*			*	*

Una referencia bibliográfica se conforma por las partes explicadas en el Cuadro 1 (aunque algunas de ellas son opcionales): lista de autores, año, título del artículo o capítulo, lista de editores, título de la publicación, número del volumen, número de páginas, ciudad y editorial. Las partes no estándar pueden añadirse entre paréntesis.

15.3. *Reglas generales de formato*

- Los títulos de artículo o capítulo irán en redondas, sin comillas.
- Los títulos de publicación (tanto de libros como de revistas) aparecerán en cursivas.
- Los títulos inician con mayúscula. En los títulos de artículos y libros en inglés no se emplean mayúsculas en el resto de las palabras (es decir, no se emplea *title case*), salvo en los nombres de revistas y series, y cuando la ortografía de la lengua lo exija. Se sugiere seguir el mismo criterio en títulos en otras lenguas.
- El título principal y el subtítulo se separan con dos puntos. Si es en español, el subtítulo iniciará con letra minúscula; si es en inglés, también iniciará con minúscula a menos que se trate de una oración completa.
- La lista de editores, o compiladores, coordinadores, etc., va seguida de (Ed.) o (Eds.), dependiendo del número de editores, o (Comp.) o (Comps.), etc.
- La ciudad de publicación va seguida de dos puntos.
- Las partes no estándar se pueden añadir al final de una referencia entre paréntesis.
- En las referencias, las abreviaturas, las preposiciones para introducir el volumen editado en el que aparece un capítulo, los nombres de la ciudad de publicación y cualquier información adicional deben ir en la lengua principal del texto. Por ejemplo, en una contribución en español, se usará la preposición “En”, la abreviatura “Ed.” y el nombre “París” en la siguiente referencia:

Erdal, Marcel (2007). Group inflection, morphological ellipsis, affix suspension, clitic sharing. En M. M. Jocelyne Fernandez-Vest (Ed.), *Combat pour les langues du monde: Hommage à Claude Hagège* (pp. 177–189). París: L'Harmattan.

En cambio, cuando la lengua principal del artículo es el inglés, las convenciones de abreviación y especificaciones irán en esta lengua:

Yu, Alan C. L. (2003). *The morphology and phonology of infixation* (Unpublished doctoral dissertation). University of California, Berkeley.

15.4. Tipos estándar de referencia

Estos son algunos ejemplos de los cuatro tipos estándar de referencia: artículos de revistas, libros, artículos en volúmenes editados y tesis:

- Artículo de revista:

 Milewski, Tadeusz (1951). The conception of the word in languages of North American natives. *Lingua Posnaniensis*, 3, 248–268.

En los artículos consultados en línea debe agregarse el doi o el URL:

Ender, Andrea (2016). Implicit and explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics*, 37(4), 536–560. 10.1093/applin/amu051

Cid Uribe, Miriam Elizabeth, & Ortiz-Lira, Héctor (1998). La conducta prosódica del vocativo en el español culto de Santiago de Chile. *Onomazein*, 3, 143–162. http://www.onomazein.net/Articulos/3/8_Cid.pdf

- Libro. El título va seguido de un punto, ya sea que el libro sea de autor o de editor:

Matthews, Peter (1974). *Morphology*. Cambridge: Cambridge University Press.

Lightfoot, David W. (Ed.) (2002). *Syntactic effects of morphological change*. Oxford: Oxford University Press.

- Artículo o capítulo en volumen editado. La lista de editores va precedida de “En” y seguida de “(Ed.)” o “(Eds.)” y coma. Los nombres de los mismos irán en el orden regular, es decir, nombre-apellido. Siempre se incluirá la información completa sobre el volumen, incluso si se enlistan en las referencias otros artículos del mismo libro. No se requiere enlistar el volumen por separado, a menos que así se le cite en el texto.

Erdal, Marcel (2007). Group inflection, morphological ellipsis, affix suspension, clitic sharing. En Marie-Madeleine Jocelyne Fernandez-Vest (Ed.), *Combat pour les langues du monde: Hommage à Claude Hagège* (pp. 177–189). París: L’Harmattan.

Las presentaciones de congresos publicadas en memorias se pueden tratar como artículos en volúmenes editados:

Fridman, Boris (1996). Gestos y oraciones de la lengua de señas de México. En *Memorias de la xxiv Mesa Redonda de la Sociedad Mexicana de Antropología*. México: Sociedad Mexicana de Antropología/Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México.

- Tesis (el tipo de tesis se menciona en paréntesis y se señala el nombre de la institución, seguido del lugar) de acuerdo al siguiente modelo:

Autor, Nombre (año). *Título de la tesis* (Tesis de licenciatura, maestría o doctoral inédita). Nombre de la institución, Lugar.

Yu, Alan C. L. (2003). *The morphology and phonology of infixation* (Tesis doctoral inédita). University of California, Berkeley.

En el caso de las tesis en línea debe agregarse el URL como en el siguiente ejemplo:

Sanz Martín, Blanca Elena (2008). *Construcciones complejas con el verbo tener: caracterización sintáctica y semántica* (Tesis doctoral). Universidad Nacional Autónoma de México, México. <http://132.248.9.195/ptd2009/enero/0638276/Index.html>

15.5. Partes opcionales

- El volumen de la revista puede ir seguido del número, el cual debe aparecer entre paréntesis en redondas:

Coseriu, Eugenio (1964). Pour une sémantique diachronique structurale. *Travaux de linguistique et de littérature*, 2(1), 139–186.

Puesto que los números de revista son muy útiles para rastrear artículos, se recomienda incluirlos.

- El título de un libro puede ir seguido de información sobre la serie (título de la serie y número del volumen en la serie), las ediciones o el número de volumen, proporcionada entre paréntesis:

Lahiri, Aditi (Ed.) (2000). *Analogy, leveling, markedness: Principles of change in phonology and morphology* (Trends in Linguistics, 127). Berlín: De Gruyter Mouton.

En los títulos de las series se usan las mayúsculas de manera especial, como sucede en los títulos de las revistas.

- En el caso de una reimpresión de un texto antiguo o una referencia clásica se usan los datos de la reimpresión o la edición citada y se incluye el año de la publicación original entre corchetes después del año de la reimpresión o edición citada:

Alonso, Amado (1951 [1933]). *Estilística y gramática del artículo en español*. En *Estudios lingüísticos: temas españoles* (pp. 126–160). Madrid: Gredos.

15.6. *Nombres y apellidos de los autores*

- Los nombres de los autores siempre deberán aparecer en el orden “apellido, nombre”, de manera que quede claro cuáles elementos pertenecen al apellido y cuáles al primer nombre. Si el segundo nombre en el ejemplo siguiente estuviera dado en el orden “nombre apellido” (Francisco José Ruiz de Mendoza), la segmentación de los apellidos no sería clara para un lector no hispanohablante.

Pérez Hernández, Lorena, & Ruiz de Mendoza, Francisco José (2002). Grounding, semantic motivation, and conceptual interaction in indirect directive speech acts. *Journal of Pragmatics*, 34(3), 259–284.

En el caso de los traductores, editores, coordinadores, etc., cuyo nombre aparece no al inicio de la referencia, si no en el cuerpo de la misma, el orden corresponde al normal, es decir, nombre-apellido.

- Cuando hay más de dos autores o editores, cada uno se separa por punto y coma, con excepción del último nombre, que es precedido por coma seguida de *et* (&). No se omite el nombre de ningún autor, es decir, en la sección de referencias no se usa *et al.*

Chelliah, Shobhana, & de Reuse, Willem (2010). *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.

Johnson, Kyle; Baker, Mark, & Roberts, Ian (1989). Passive arguments raised. *Linguistic Inquiry*, 20, 219–251.

- Los apellidos compuestos o complejos no se tratan de manera especial y se alfabetizan bajo la primera parte, aunque comiencen con letra minúscula. Así, los apellidos alemanes o neerlandeses que empiezan con *von* o *van* (por ejemplo *van Riemsdijk*) o los apellidos neerlandeses, franceses o españoles que empiezan con *de* (por ejemplo, *de Groot* o *de la Garza*) se tratan como los apellidos belgas (por ejemplo, *De Schutter*) y los apellidos italianos (por ejemplo, *Da Milano*). Los siguientes nombres están ordenados alfabéticamente como se indica (es decir, de manera mecánica):

Da Milano, Federica > de Groot, Casper > de Saussure, Ferdinand > De Schutter, Georges > van der Auwera, Johan > Van Langendonck, Willy > van Riemsdijk, Henk > von Humboldt, Wilhelm

- Los nombres chinos y coreanos se podrán tratar de manera especial. Dado que los apellidos no son muy distintivos, se puede dar el nombre entero dentro del texto, por ejemplo:

la negación neutral *bù* es compatible con verbos estativos y de actividad (cf. Hsieh Miao-Ling, 2001; Lin Jo-wang, 2003; Teng Shou-hsin, 1973)

15.7. Referencias electrónicas (publicaciones en internet)

- Las publicaciones regulares que son accesibles en línea no se tratan de manera especial, excepto que deberán llevar al final de la referencia el número de Digital Object Identifier (doi) o, en caso de no tenerlo, el URL (es decir, la dirección electrónica del recurso referido).

Ender, Andrea (2016). Implicit and explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics*, 37(4), 536–560. 10.1093/applin/amu051

Cid Uribe, Miriam Elizabeth, & Ortiz-Lira, Héctor (1998). La conducta prosódica del vocativo en el español culto de Santiago de Chile. *Onomazein*, 3, 143–162. http://www.onomazein.net/Articulos/3/8_Cid.pdf

- Cuando se cita un recurso de la red que no es una publicación científica regular, por ejemplo, una página de internet, esta se deberá tratar como una referencia impresa, en la medida de lo posible:

Redish, Laura, & Lewis, Orrin (1998–2015). Vocabulary in Native American languages: Salish words. http://www.native-languages.org/salish_words.htm

- Previo al envío y/o publicación del artículo es necesario corroborar que los URL de las referencias electrónicas indiquen la ubicación correcta. Cuando esto no ocurra, es necesario actualizar el URL o sustituirlo por otra fuente.

15.8. *Miscelánea*

- Los libros pueden incluir entre paréntesis información sobre ediciones, números de volúmenes y números de páginas inmediatamente después del título y con el punto después del paréntesis:

Croft, William (2003). *Typology and universals* (2a. ed.). Cambridge: Cambridge University Press.

Rissanen, Matti (1999). Syntax. En Lass, Roger (Ed.), *Cambridge history of the English language* (Vol. 3, pp. 187–331). Cambridge: Cambridge University Press.

- Si una editorial está asociada a varias ciudades, solo es necesario dar la primera, por ejemplo: “Berlín: De Gruyter Mouton” o “Ámsterdam: Benjamins”.
- Los títulos de los trabajos escritos en lenguas distintas al español, inglés o francés pueden ir acompañados de una traducción, que debe aparecer entre corchetes cuadrados después del título en la lengua original:

Li, Rulong (1999). *Minnan fangyan de daici* [Demostrativos y pronombres personales en min del sur]. En Rulong Li & Song-Hing Chang (Eds.), *Daici* [Demostrativos y pronombres personales] (pp. 263–287). Guangzhou: Ji’nan University Press.

- Si el título no solo está en una lengua distinta, sino también en una escritura diferente, se puede proporcionar la escritura original entre paréntesis, además de la transliteración. Del mismo modo, el nombre del autor se podrá dar entre paréntesis después del nombre transliterado, como en el siguiente ejemplo:

Plungian, Vladimir A. (Плунгян, Владимир А.) (2000). *Obščaja morfologija: Vvedenie v problematiku* (Общая морфология: Введение в проблематику) [Morfología general: introducción a los problemas]. Moscú: Editorial URSS.

- En el caso de las ponencias no publicadas, se incluirá el mes y año de la reunión (aunque se recomienda que solo se citen ponencias de congresos recientes, si se puede suponer que el material será publicado eventualmente):

Filppula, Markku (septiembre, 2013). Areal and typological distributions of features as evidence for language contacts in Western Europe. Ponencia presentada en el congreso de la Societas Linguistica Europaea, Split, Croacia.

Para resolver otros casos específicos no contemplados en estos lineamientos, se recomienda a los autores consultar el manual de publicaciones de la American Psychological Association (APA) en su última versión o su sitio web (www.apastyle.org/index.aspx), donde encontrarán un extenso inventario de modelos de fichas para casos especiales (*software*, medios audiovisuales, conjuntos de datos, referencias virtuales, etcétera).

16. Referencias

- American Psychological Association (apa) (2010). *Manual de publicaciones de la American Psychological Association* (3a. ed.). México: El Manual Moderno.
- Comrie, Bernard; Haspelmath, Martin, & Bickel, Balthasar (2008–2015). Leipzig glossing rules. <http://www.eva.mpg.de/lingua/resources/glossing-rules.php>
- Haspelmath, Martin (2014). *The generic style rules for Linguistics*. <http://www.eva.mpg.de/lingua/pdf/GenericStyleRules.pdf>
- Real Academia Española (RAE) (2010). Principales novedades de la última edición de la *Ortografía de la lengua española (2010)*. http://www.rae.es/sites/default/files/Principales_novedades_de_la_Ortografia_de_la_lengua_espanola.pdf
- Real Academia Española (RAE) & Asociación de Academias de la Lengua Española (Asale) (2010). *Ortografía de la lengua española*. Madrid: Espasa.

Distribución, venta y suscripciones

Katsumi Yoshida Hernández
Escuela Nacional de Lenguas, Lingüística
y Traducción (ENALLT), Circuito Interior s/n,
Ciudad Universitaria, Alcaldía Coyoacán,
c. p. 04510, Ciudad de México, México
ela@enallt.unam.mx

Números sueltos y suscripciones	Precio unitario	Precio con envío incluido			
		República Mexicana	Zona 1	Zona 2	Zona 3
Ejemplar regular Cualquier número, excepto <i>ELA</i> número 50	\$100 <small>MXN</small>	\$130 <small>MXN</small>	USD \$22	USD \$25	USD \$29
Ejemplar especial <i>ELA</i> número 50	\$200 <small>MXN</small>	\$230 <small>MXN</small>	USD \$35	USD \$38	USD \$41
Suscripción regular 2 ejemplares regulares	\$160 <small>MXN</small>	\$220 <small>MXN</small>	USD \$40	USD \$45	USD \$53
Suscripción especial 1 ejemplar regular más <i>ELA</i> número 50	\$240 <small>MXN</small>	\$300 <small>MXN</small>	USD \$50	USD \$55	USD \$63
Zona 1: Norteamérica y Centroamérica Zona 2: Sudamérica y Europa Zona 3: Resto del mundo					

Índice de números anteriores
<http://ela.enallt.unam.mx>





Coordinación editorial
Enio Ramírez Campos

Edición y cuidado editorial
Mariela Elizabeth Pérez Castañeda
Claudia Isabel Carrillo Sánchez

Corrección de estilo
Armando Rodríguez Briseño

Corrección de textos en inglés
María Elena Sánchez Salazar

Diseño de portada
Soledad Martínez Piña

Diseño de interiores y formación
Héctor Sandoval Sandoval

ARTÍCULOS ►

- Aproximación al análisis prosódico del foco contrastivo en hablantes preescolares y adultos del español mexicano: datos de producción y percepción 7
An approach to the prosodic analysis of contrastive focus of preschool and adult Mexican Spanish speakers: data from production and perception
Brenda Muñoz Martínez, Natalia Arias Trejo & Erika Mendoza Vázquez
- La activación de la información fonológica en aprendientes de segundas lenguas por medio de imágenes 49
Activation of the phonological information through images in second language learners
Alma Luz Rodríguez Lázaro & Natalia Arias Trejo
- Can foreign language learners discover the rules of non-salient forms? 85
¿Pueden los aprendices de una lengua extranjera descubrir las reglas de formas lingüísticas no sobresalientes?
José Luis Moreno-Vega & Ana Mónica Preciado-Sánchez
- Metáfora enactiva en la música didáctica virtual: una mirada crítica decolonial 119
Enactive metaphor in a virtual didactic musical performance: A decolonial critical look
Luis Escobar L.-Dellamary & Luis Quintero Castro
- Formas apelativas de segunda persona en entrevistas político-presidenciales mexicanas 157
Address forms in Mexican political-presidential interviews
Melanie del Carmen Salgado López

