

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

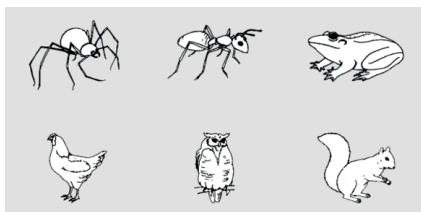


FIGURA 1. Ejemplo de estímulo con imágenes de objetos animados

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 ($DE = 1.52$).

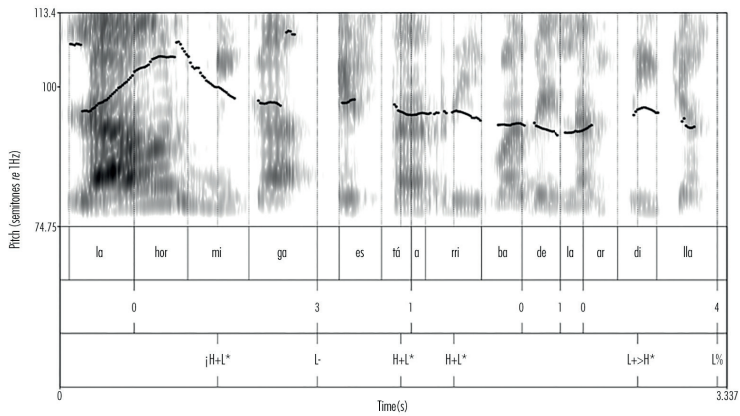


FIGURA 2. Ejemplo de respuesta dada por un niño en la Condición Sujeto para el enunciado "La hormiga está arriba de la ardilla"

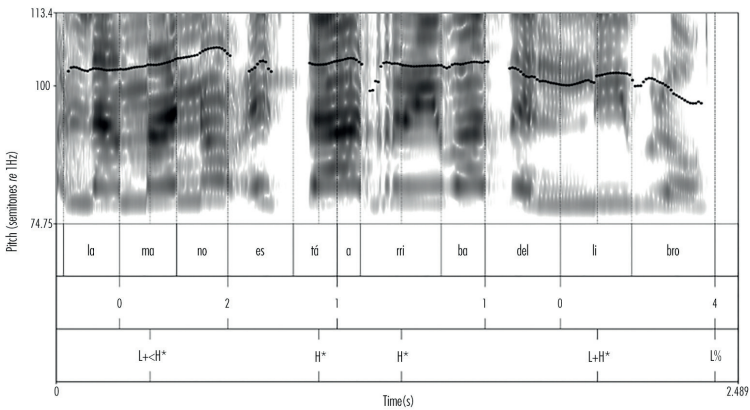


FIGURA 3. Ejemplo de respuesta dada por un niño en la Condición Atributo para el enunciado "La mano está arriba del libro"

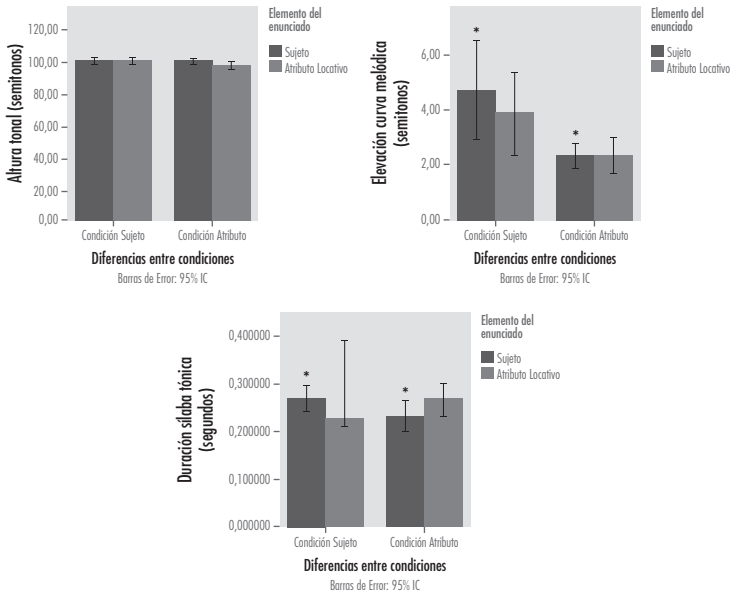


FIGURA 4. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por niños

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).

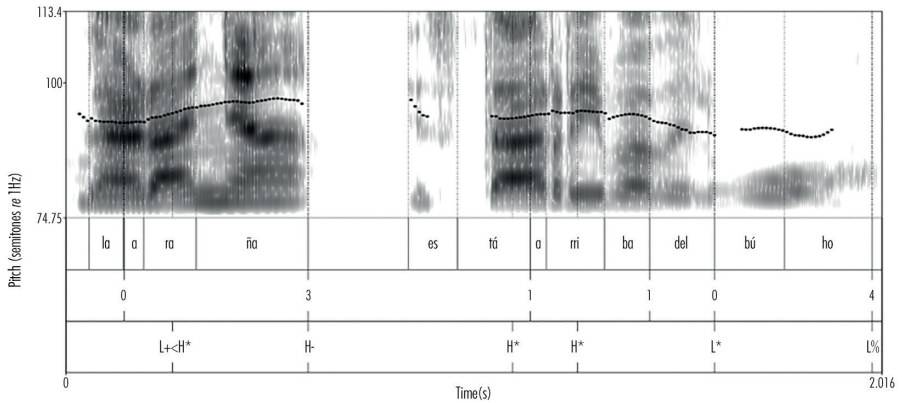


FIGURA 5. Ejemplo de respuesta dada por adulto en la Condición Sujeto para el enunciado “La araña está arriba del búho”, con cesura nivel 3 después del sujeto

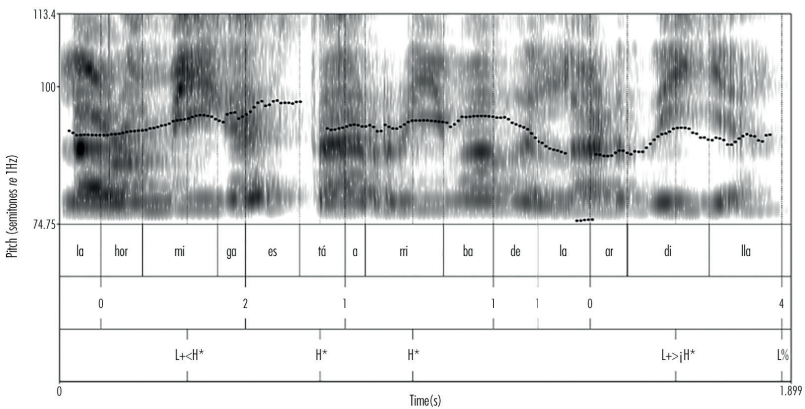


FIGURA 6. Ejemplo de respuesta dada por adulto en la Condición Atributo para el enunciado “La hormiga está arriba de la ardilla”, con cesura nivel 2 después del sujeto

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.

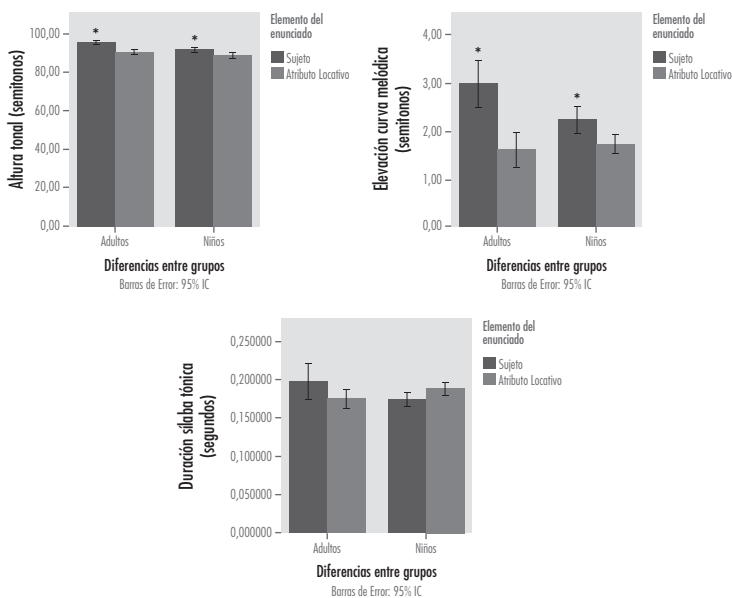


FIGURA 7. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos

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-

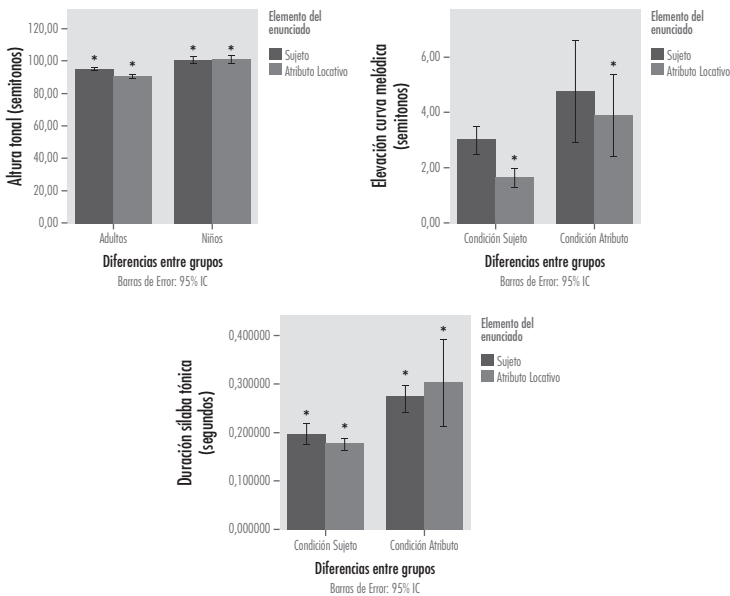


FIGURA 8. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos y niños en la Condición Sujeto

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, $DE = 4.68$;

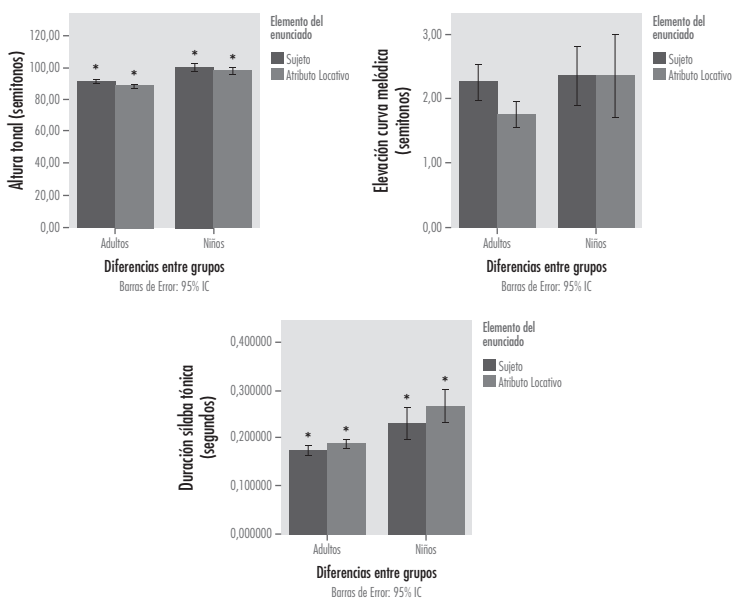


FIGURA 9. Comparación entre altura tonal, elevación de la curva melódica y duración en las respuestas dadas por adultos y niños en la Condición Atributo

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

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

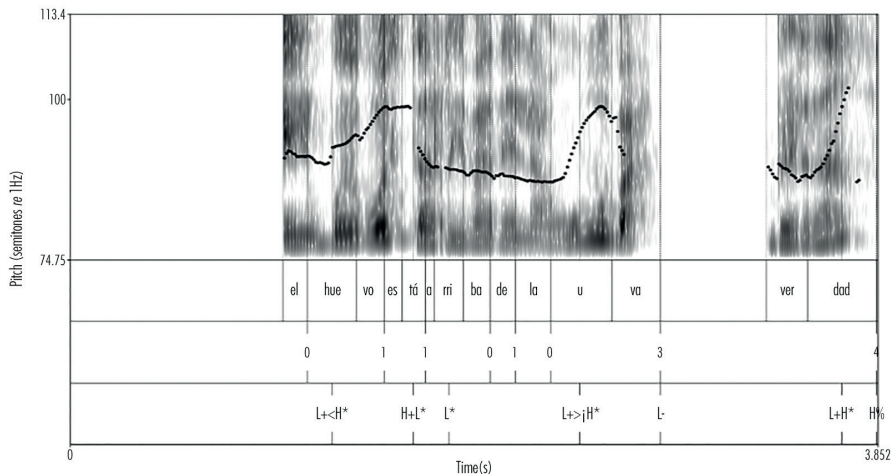
- (Eds.), *Actas del II Congreso Nacional de Fonética Experimental* (pp. 272–278). Sevilla: Universidad de Sevilla.
- Pierrehumbert, Janet Breckenridge (1980). *The phonology and phonetics of English intonation* (Tesis de doctorado). Massachusetts Institute of Technology, Massachusetts. <http://dspace.mit.edu/handle/1721.1/16065>
- 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

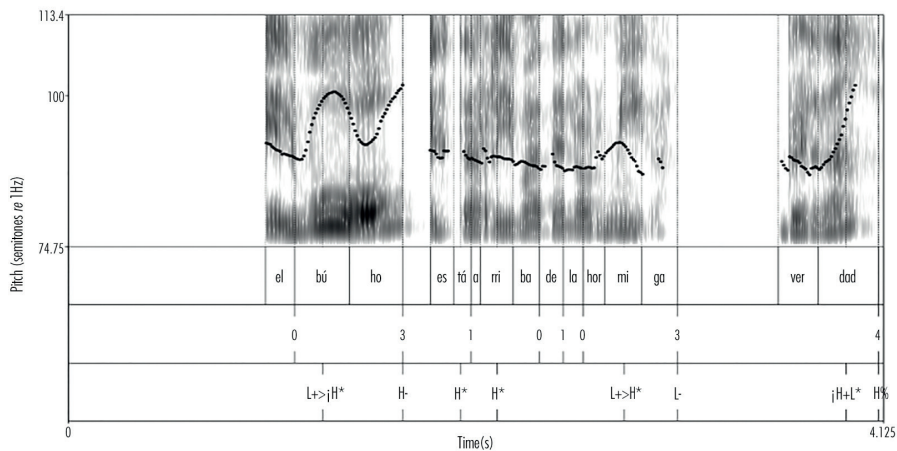
8. Anexos

8.1. Anexo A

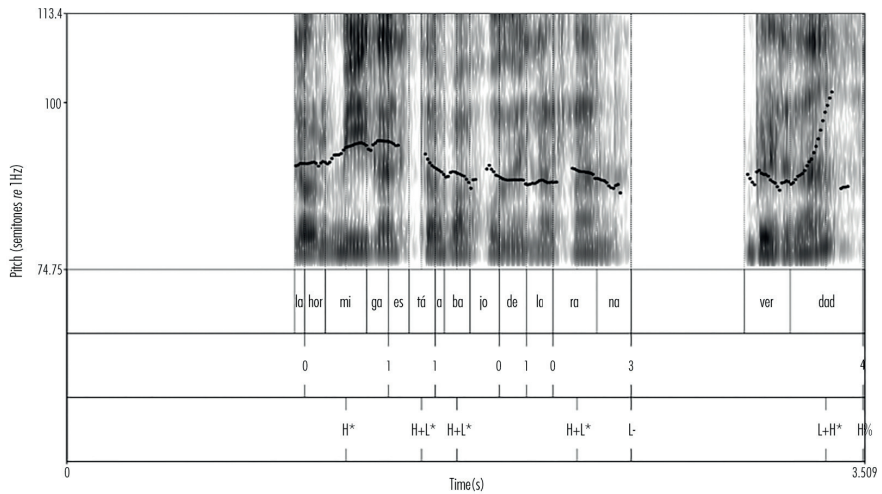
Anexo A1. Ejemplo de estímulo auditivo de la Condición Atributo



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

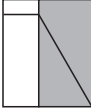
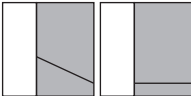
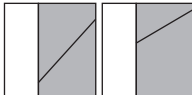



Representación esquemática de los tipos de acentos tonales y tipos de tonos de juntura encontrados en las transcripciones. Adaptación de De la Mota, Martín Butragueño y Prieto (2010), y Villalobos Pedroza (2021).

	Esquema	Transcripción	Descripción
Acentos monotonaes		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.




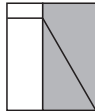
(continuación)

	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.
		$L+;H^*$	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.
		$L+<;H^*$	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.
		$L+>;H^*$	Una elevación $L+>H^*$ mayor a 3 semitonos.

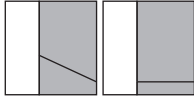
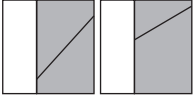
(cont.)

	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.

