

# El vocabulario productivo y su relación con la gramática en niños hablantes de castellano peruano entre los 16 y los 30 meses

## Productive vocabulary and its relationship with grammar in Spanish-speaking children from 16 to 30 months of age

Andrea Junyent<sup>1</sup>, María Blume<sup>2</sup>, María Fernández Flecha<sup>3</sup> y Talía Tijero Neyra<sup>4</sup>

<sup>1</sup>PhD. en Psicología del Desarrollo y de los Procesos de Socialización. Investigadora en el Grupo de Investigación en Adquisición del Lenguaje, Departamento de Humanidades, Pontificia Universidad Católica del Perú (PUCP). E-mail: junyent.aa@pucp.pe

<sup>2</sup>PhD. en Lingüística. Profesora Asociada e Investigadora en el Grupo de Investigación en Adquisición del Lenguaje, Departamento de Humanidades, PUCP.

<sup>3</sup>PhD. en Intervención en el Lenguaje. Profesora Asociada e Investigadora en el Grupo de Investigación en Adquisición del Lenguaje, Departamento de Humanidades, PUCP.

<sup>4</sup>PhD. en Lingüística. Coordinadora de Proyectos Académicos, Universidad de Tecnología e Ingeniería (UTEC). Investigadora en el Grupo de Investigación de Adquisición del Lenguaje, Departamento de Humanidades, PUCP.

Este trabajo fue financiado por la Dirección de Gestión de la Investigación de la Pontificia Universidad Católica del Perú (PUCP), a través de las subvenciones DGI-2015-1-0020 y DGI-2016-1-0065.

Las autoras agradecen a los niños participantes en esta investigación y a sus padres, y a las instituciones educativas preescolares que colaboraron para contactarlos. También agradecen al equipo de asistentes de investigación que colaboraron con la realización de esta investigación: Aixa Chaparro, Karina Aguirre, Victoria Maraví, Jorge Narváez, Marcela Damonte, Freyda Schuler, Minerva Cerna, Stefanny Ibarra, Silvia Sinchitullo, Thalya Sarmiento y Geraldine Arias. Finalmente, las autoras expresan su agradecimiento a Fiorella Otiniano por su invaluable ayuda en el análisis estadístico de los datos.

Grupo de Investigación en Adquisición del Lenguaje (GRIAL)  
Pontificia Universidad Católica del Perú.  
Lima, Perú.

### Resumen

Se examinaron datos provenientes de la evaluación del lenguaje de 104 niños con edades entre 16 y 30 meses residentes en Lima con una versión preliminar de la adaptación de la Forma II de los reportes parentales

*MacArthur-Bates Communicative Development Inventories* (CDI) al español del Perú. Se analizó el léxico productivo a partir de la lista de palabras del reporte y se realizaron análisis de regresión para evaluar si el nivel de vocabulario contribuía a predecir el desarrollo gramatical de los niños, empleando

como variables de control la edad del niño y el nivel de escolarización materna (como indicador del nivel socioeconómico) y como variables dependientes la longitud media de los tres enunciados más largos (LME3) y la complejidad morfosintáctica (CM), obtenidas con el CDI. Los sustantivos fueron la clase de palabras más frecuente y que más aumentó con la edad, como en el italiano (Caselli et al., 1995; D'Odorico y Fasolo, 2007) y el español mexicano (Jackson-Maldonado et al., 1993). Además, el nivel de vocabulario predijo el nivel de desarrollo gramatical, como en el español peninsular (Mariscal y Gallego, 2012) y en el mexicano (Thal, Jackson-Maldonado y Acosta, 2000): una vez controlados los efectos de la edad y la escolarización materna, el vocabulario explicó el 20 % de la variabilidad de la LME3 y el 37 % de la variabilidad de la CM. Este estudio ofrece información sobre el desarrollo temprano del vocabulario y la gramática en el español del Perú, una etapa escasamente estudiada en esta variedad.

*Palabras clave:* vocabulario; gramática, desarrollo temprano del lenguaje; inventarios comunicativos MacArthur-Bates; español del Perú.

## Abstract

We present results from the analysis of data from 104 children with ages ranging from 16 to 30 months, living in Lima, Peru. The data was collected using a preliminary version of the MacArthur-Bates Communicative Development Inventories (CDI), Form II, adapted to Peruvian Spanish. The CDIs are parental reports that assess language through the presentation of a series of items to the parents, mainly in the form of a checklist. We aimed to (1) describe the composition of Peruvian children's productive lexicon at this age, (2) identify whether nouns were the most frequent class in their vocabulary, and also the class which increased the most with age, as is the case in Italian (Caselli et al., 1995; D'Odorico y Fasolo, 2007) and Mexican Spanish (Jack-

son-Maldonado et al., 1993); and (3) assess whether vocabulary size explained the level of grammatical development as in Peninsular and Mexican Spanish (Mariscal & Gallego, 2012; Thal, Jackson-Maldonado & Acosta, 2000).

In order to do so, we analyzed the following data: age; maternal education level (as an indicator of socioeconomic status); vocabulary (represented by a list of more than 500 words that parents recognized as produced or not by their children); the mean length of the three longest utterances produced by the child as reported by parents; and sentence complexity, a measure of the level of structures produced by the child based on the parent's report.

To achieve the first and second goals, words in the vocabulary list were classified according to grammatical class. Then, the percentage of production of each class of words was calculated, as well as the correlation between this percentage and age. Children's vocabulary was composed mainly of nouns, which represented 57 % of the production, followed by predicates (verbs and adjectives) and function words (pronouns, determiners, prepositions, conjunctions and connectives). Noun was the class of words that increased the most with age, showing the highest correlation with such variable ( $r = .72$ ). Correlations between other types of words and age varied: it was also high for adjectives; but moderate for some classes and low for others. This showed that each type of word increased at different rates.

Regarding the third goal, to assess the relationship between vocabulary size and level of grammatical development, two regression analyses were conducted to predict the variables mean-length-of-the-three-longest-utterances and sentence-complexity. For both analyses, age and maternal education level were entered in the first step, in order to control for their effect; and vocabulary was entered in the second step. Both models accounted for substantial amounts of variance and showed an important effect on vocabulary. The model with the mean-length-of-

the-three-longest-utterances as the dependent variable accounted for 56 % of the variance. The first step explained 37 % of the variance (due to the contribution of both age and maternal education level), and the second step explained an additional 20 %. The model with sentence-complexity accounted for a higher percentage of the variance, 75 %. The first step explained 38 % of the variance, exclusively due to age; while the second step explained an additional 37 %. These results suggest a strong relationship between vocabulary and grammar.

Further research is needed to confirm these results with direct measures of child production and more representative samples of children acquiring Peruvian Spanish, including speakers of dialects of other parts of the country. However, these results, preliminary as they are, shed light on the composition of vocabulary and its relationship with grammar in a variety of Spanish scarcely studied during early stages of development.

*Keywords: lexicon, grammar, early language development, MacArthur-Bates communicative development inventories, Peruvian Spanish.*

## Introducción

### **Los MacArthur-Bates Communicative Development Inventories (CDI) como medida de evaluación del desarrollo comunicativo y lingüístico temprano**

Los *MacArthur-Bates Communicative Development Inventories* (CDI) (Fenson et al., 1993) son reportes que miden –a través de una serie de ítems presentados a los padres– la comprensión y producción de diferentes conductas comunicativas tales como los gestos, el vocabulario comprensivo y expre-

sivo, la gramática y los usos del lenguaje, entre otros. Los inventarios comprenden dos formas: la Forma I, Palabras y gestos (*Words and Gestures*), mide el desarrollo entre los 8 y 15 meses; y la Forma II, Palabras y oraciones (*Words and Sentences*), lo hace entre los 16 y 30 meses. Estos inventarios han sido adaptados a la fecha a más de 100 lenguas, incluido el español en seis de sus variedades, además de la adaptación peruana que se encuentra actualmente en proceso. La adaptación que se presenta en este artículo, ya autorizada por la junta consultiva de los CDI, comprende las Formas I y II, y parte de la necesidad de contar con una herramienta de evaluación adecuada para el español del Perú<sup>1</sup>. Hasta donde se sabe, no existe para dicha variedad ningún instrumento de evaluación del lenguaje y la comunicación para edades tan tempranas.

Para recolectar los datos de este estudio se han empleado algunas secciones de la Forma II. Por motivos de espacio, se describen aquí solo dichas secciones. En Blume, Fernández-Flecha, Junyent y Tijero (2019), se ofrece una descripción del proceso de adaptación de ambas formas y de la estructura de la versión resultante. La Forma II cuenta con cuatro secciones: Palabras, Gramática, Vocalizaciones (anexo opcional del CDI peruano) y una sección de preguntas que recoge información general sobre el niño y su familia.

La sección Palabras está compuesta por cuatro subsecciones: Producción temprana, Desarrollo del vocabulario, Vocabulario y Usos del lenguaje. En Vocabulario, la única subsección de Palabras que se empleó en este estudio, se mide la producción tanto de palabras como de expresiones que podrían entenderse como bloques sin analizar (“no-hay”, “un-montón”, etcétera). En esta subsección, la más amplia del cuestionario, los ítems están organizados en veinte áreas o campos temáticos.

<sup>1</sup>Si bien no existe un español del Perú, y como los datos provienen de la ciudad de Lima, se adoptó este término por tratarse de la primera adaptación para ese país y por ser lo tradicional en los nombres de los otros CDI. Se discute este tema en Blume et al. (2019).

La sección Gramática comprende cinco subsecciones: Terminaciones de palabras, Verbos difíciles, Palabras sorprendentes, Combinación de palabras y Complejidad morfosintáctica. En este estudio se emplearon las dos últimas. En Combinación de palabras se piden tres ejemplos de las frases más largas producidas por el niño que los padres recuerden, a partir de las cuales se obtiene una longitud media. En Complejidad morfosintáctica, se indaga sobre el tipo de estructuras sintácticas y los morfemas que produce el niño. Al final del cuestionario, se pidió información sobre la salud del niño y su escolarización, así como sobre la de los padres, y el contacto con otras lenguas y variedades del castellano.

Las diferentes secciones y subsecciones de la Forma II de los CDI originales en inglés, así como las versiones en español peninsular y mexicano, han servido para estudiar diversos aspectos del desarrollo lingüístico y comunicativo infantil. Los datos provenientes de las subsecciones empleadas en este estudio (Vocabulario, Combinación de palabras y Complejidad morfosintáctica) han mostrado importantes correlaciones con medidas obtenidas directamente del habla espontánea (español mexicano: Thal, Jackson-Maldonado y Acosta, 2000) y han sido utilizadas para medir la relación entre vocabulario y gramática (inglés: McGregor, Sheng y Smith, 2005; español peninsular: Mariscal y Gallego, 2012).

Ninguna de las varias formas disponibles para evaluar el lenguaje en niños pequeños es suficiente por sí sola para capturar la multidimensionalidad del lenguaje (Dockrell, 2001; Feldman et al., 2005; Law y Roy, 2008). Los CDI, como todo reporte parental, carecen de las ventajas propias de las medidas directas del lenguaje. A pesar de ello, se ha empleado la Forma II de los CDI, por un lado, por la ausencia de medidas referenciales para el

español infantil del Perú y, por otro, por las ventajas que ofrece como herramienta práctica, poco costosa, eficiente y no invasiva (Feldman et al., 2005; Law y Roy, 2008; Pérez-Pereira y Resches, 2011) que aprovecha el conocimiento amplio y representativo que los padres tienen del lenguaje de sus hijos en contextos naturales (Dale y Goodman, 2005; Pérez-Pereira y Resches, 2011).<sup>2</sup> La medición de la validez concurrente de estas herramientas –que en este caso permite saber qué tan buenos informantes del desarrollo lingüístico temprano de sus hijos son los padres– es costosa y compleja porque se realiza comparando los resultados de la prueba con datos obtenidos de forma independiente por medio de otras metodologías. La evaluación de la Forma II en la versión original en inglés de los CDI (Law y Roy, 2008) y la versión en español peninsular han arrojado buenos resultados de validez concurrente (López-Ornat et al., 2005; Mariscal, Nieva y López-Ornat, 2010). La validez de los CDI peruanos aún no ha sido estudiada con una prueba ya validada para el español del Perú porque no existen pruebas ya validadas para niños tan pequeños. La evaluación está prevista y se realizará con las muestras de lenguaje espontáneo recogidas de forma paralela a la aplicación de los cuestionarios y posteriormente codificadas, pero que aún deben ser empleadas para la validación final de la prueba. Por las razones presentadas, aunque se consideren confiables los datos empleados en este estudio, es necesario ser cautelosos al momento de interpretarlos.

### **La composición del vocabulario temprano y la preferencia por los sustantivos**

Respecto de algunas lenguas como el inglés, se ha encontrado que los sustantivos son la categoría predominante en el vocabulario infantil temprano y la que antes se

2 Aunque los CDI son apropiados para investigar el desarrollo típico del lenguaje porque son confiables, válidos y sensibles a la edad, los datos acerca de su sensibilidad para detectar casos de desarrollo atípico no son alentadores (sobre los originales en inglés, ver Law y Roy, 2008 y Fenson et al., 2007).

adquiere (Gentner, 1982), por lo que se planteó la hipótesis de que fuera un sesgo universal. Para el idioma español, Jackson-Maldonado, Thal, Marchman, Bates y Gutiérrez-Clellen (1993), usando los CDI, encontraron que los sustantivos se adquieren antes que otras categorías, representan el mayor porcentaje del vocabulario y son la única categoría cuya producción aumenta con el tiempo en el rango de edad de 15 a 30 meses evaluado por los autores. Actualmente, se considera que la preferencia por los sustantivos no es universal (Dhillon, 2010), sino más bien caracteriza a algunas lenguas como el italiano, el hebreo, el finlandés, el francés, el japonés y el holandés, por ejemplo (en D'Odorico y Fasolo, 2007); pero no a otras, como el mandarín, el cantonés y el tzeltal (Brown, 1998; Tardif, Gelman y Xu, 1999; Tardif, Shatz y Naigles, 1997). Las explicaciones ofrecidas para dar cuenta de las diferencias son variadas, pero coinciden en atribuir el contraste a características lingüísticas y culturales que se reflejan en el *input*. Este estudio se propone, además de ofrecer una descripción del vocabulario infantil reportado, brindar información sobre el lugar que tienen los sustantivos en el lenguaje de los niños peruanos de edades entre 16 y 30 meses.

## **La relación entre el vocabulario y la gramática en el desarrollo temprano**

### ***El poder predictivo del vocabulario***

El vocabulario y la gramática son considerados los principales indicadores de la adquisición del lenguaje en edades tempranas (Silva et al., 2017). Además, la estrecha relación entre ambos ha sido demostrada en diferentes estudios. Por un lado, ya Bates y Goodman (1997) presentaron una serie de evidencias –provenientes del desarrollo típico como atípico, de las afasias y del procesamiento lingüístico en tiempo real– a favor de la existencia de relaciones tempranas y complejas entre el conocimiento léxico y el gramatical. Algunos estudios basados en

versiones del CDI para niños de 15 o 16 meses a 30 meses de edad han mostrado la relación vocabulario-gramática, tanto en inglés (Dale, Dionne, Eley y Plomin, 2000; Marchman y Bates, 1994) como en lenguas romances (italiano: Devescovi et al., 2005; catalán: Serrat et al., 2010; portugués europeo: Silva et al., 2017; gallego: Pérez-Pereira y Resches, 2011). Con respecto al español, Mariscal y Gallego (2012) hallaron una fuerte relación entre vocabulario (amplitud léxica, número de sustantivos y número de verbos) y gramática (secciones Terminaciones de palabras y Complejidad morfosintáctica), medidos con el CDI en niños hablantes de español peninsular. Por otro lado, Thal et al. (2000) examinaron la producción de niños hablantes de español mexicano: emplearon el número de palabras producidas, la Longitud media de los tres enunciados más largos y la Complejidad de frases (equivalente a nuestra Complejidad morfosintáctica) del CDI, así como medidas de vocabulario y gramática obtenidas directamente del habla espontánea. Estas autoras encontraron correlaciones significativas entre la medida de vocabulario del CDI y la medida directa de gramática, así como entre las medidas directas de vocabulario y la medida de gramática del CDI.

A pesar de la fuerte evidencia a favor del poder predictivo del vocabulario para la gramática, análisis más finos de la relación entre los tipos de palabras, por un lado, y la morfología y la sintaxis, por otro, han demostrado la complejidad de dicha interacción. Así, Braginsky, Yurovsky, Marchman y Frank (2015) examinaron los datos de cuatro lenguas –inglés, español, noruego y danés– disponibles en *Wordbank* y demostraron que, a pesar de la fuerte relación entre el vocabulario y la gramática, el desarrollo de esta última se puede explicar también por la edad, independientemente del crecimiento del vocabulario.

En este estudio, a partir de la información reportada, se evaluó el poder predictivo del vocabulario en relación con la gramática en niños hablantes de español del Perú, buscando

contribuir con información proveniente de una variedad no trabajada anteriormente.

### ***La edad, el nivel socioeconómico y el desarrollo gramatical***

Para medir la relación entre el vocabulario y la gramática, se han considerado como variables de control la edad y el nivel socioeconómico por su conocida relación con el desarrollo lingüístico. Con respecto a la edad, diferentes estudios que emplearon los CDI han encontrado que el poder explicativo del vocabulario es tan fuerte que supera al de la edad (Devescovi et al., 2005; Marchman y Bates, 1994; Marjanovič-Umek, Fekonja-Peklaj y Podlesek, 2013). En relación con el nivel socioeconómico, la evidencia apunta en distintas direcciones. Algunos estudios encontraron que este indicador explica el desarrollo lingüístico desde edades muy tempranas, en particular el desarrollo del vocabulario (Hart y Risley, 1992) y de la gramática (Huttenlocher, Vasilyeva, Cymerman y Levine, 2002), probablemente con la mediación del *input* recibido (Huttenlocher, Waterfall, Vasilyeva, Vevea y Hedges, 2010). Con respecto a la relación entre el nivel socioeconómico por un lado, y el vocabulario y la gramática medidos por el CDI por el otro, la evidencia resulta difícil de interpretar. En las versiones en inglés se ha encontrado un efecto del nivel socioeconómico en las habilidades lingüísticas, pero la dirección del efecto, en algunos casos, es a favor de los niveles más altos y, en otros, en contra –lo que suele explicarse como producto de la sobreestimación de los padres de familias de niveles socioeconómicos más bajos respecto del lenguaje de sus hijos (Law y Roy, 2008). En las versiones en español, en cambio, no se ha hallado evidencia del efecto del nivel socioeconómico en las medidas de vocabulario y gramática de la Forma II del CDI (español mexicano: Jackson-Maldonado et al., 1993; español peninsular: Mariscal et al., 2007). En el presente estudio, se consideró el nivel de escolaridad más alto alcanzado por la madre como indicador de nivel socioeconómico en

línea con estudios precedentes (Dollaghan et al., 1999; Hart y Risley, 1995; Hoff, 2003).

### **Objetivos**

A partir de estos antecedentes, se analizaron los datos de niños hablantes de español del Perú con edades entre 16 y 30 meses con los siguientes objetivos:

- 1) Describir la composición de su vocabulario productivo sobre la base de la clasificación de las palabras registradas por la Forma II de los CDI.
- 2) Identificar si los sustantivos son la clase de palabras más frecuente en el vocabulario infantil reportado y la que más aumenta con la edad, como ocurre con niños hablantes de italiano (Caselli et al., 1995; D’Odorico y Fasolo, 2007) y de español mexicano (Jackson-Maldonado et al., 1993) de edades equivalentes.
- 3) Evaluar si, en los datos reportados, el vocabulario productivo explica el nivel de desarrollo gramatical, como en el español peninsular (Mariscal y Gallego, 2012) y el mexicano (Thal et al., 2000) de niños de edades equivalentes, controlados los efectos de la edad y el nivel socioeconómico.

## **Metodología**

### **Participantes**

El grupo de participantes estuvo compuesto por 104 niños residentes en Lima (45 niñas y 59 niños), de edades entre los 16 y 30 meses, cuyos datos fueron obtenidos a partir de la aplicación de una versión preliminar de la Forma II de los CDI. Según lo declarado por los padres, ningún niño había tenido problemas de audición o lenguaje. No hubo diferencias significativas entre niños y niñas en ninguna de las medidas de vocabulario o gramática, a pesar de que la evidencia acerca de la adquisición del lenguaje en general ha mostrado una ventaja de las niñas (Alexander y Wilcox,

2012; Bornstein, Haynes, Painter y Genevro, 2000; Murray, Johnson y Peters, 1990). En el caso de los CDI, sin embargo, las diferencias suelen ser muy pequeñas o inexistentes, tanto para la versión original en inglés (Law y Roy, 2008) como para las versiones del español de México (Jackson-Maldonado et al., 1993) y España (López-Ornat et al., 2005).

## **Materiales y procedimiento: CDI peruano**

La administración del cuestionario fue llevada a cabo en la casa o el lugar de trabajo de los cuidadores principales, quienes respondían a las preguntas del cuestionario ayudados por los asistentes de investigación del proyecto.

Para este trabajo, se tomó en cuenta la edad de los niños (en meses), el nivel de escolarización de la madre y la información proveniente de las subsecciones Vocabulario (sección Palabras), Combinación de palabras y Complejidad morfosintáctica (ambas de la sección Gramática) del cuestionario.

Vocabulario contiene 616 ítems; Combinación de palabras, 3 ítems, de los que se obtiene como única medida la Longitud Media del Enunciado de los tres enunciados más largos o LME3; y Complejidad morfosintáctica, 34 ítems.

Las subsecciones seleccionadas de Gramática fueron las consideradas más sensibles para medir habilidades morfológicas y sintácticas, en línea con estudios anteriores en español (Mariscal y Gallego, 2012; Thal et al., 2000). En Combinación de palabras, se pidió a los padres ejemplos de las tres construcciones más largas producidas por el niño. Por otro lado, Complejidad morfosintáctica mide la producción de ciertos tipos de construcciones con cuatro niveles de complejidad que incluyen estructuras con sustantivos y modificadores, así como oraciones simples y complejas, y estructuras interrogativas e imperativas. En la versión empleada en esta investigación se mantuvieron los cuatro niveles de comple-

jidad de respuestas usados en la versión de los CDI para el español peninsular, ya utilizados en otros estudios (Mariscal y Gallego, 2012; Mariscal, Nieva y López-Ornat, 2010).

## **Análisis de los datos**

Todos los datos fueron analizados empleando el paquete estadístico SPSS versión 21. En los análisis, se incluyeron las variables Edad, Escolarización materna, Vocabulario, LME3 y Complejidad morfosintáctica.

## **Descripción de las variables**

La información sobre la edad (en meses) y la escolarización materna, como la información lingüística, provino del reporte de los padres. Al nivel de escolarización materna se le asignaron los siguientes puntajes según el grado de instrucción completado: 1 (primaria), 2 (secundaria), 3 (estudios técnicos o en instituto) y 4 (estudios universitarios). Todas las madres habían concluido, al menos, la primaria. En 55 de los 104 casos, la madre (sola o con otras personas) era la principal cuidadora del niño; en los otros casos se trataba de la niñera, el padre o algún abuelo y en dos casos no se obtuvo información al respecto.

En todas las subsecciones del CDI empleadas en este estudio, se siguieron las pautas de calificación ofrecidas por el manual de la versión en español peninsular (López-Ornat et al., 2005). A los ítems de la sección Vocabulario se les asignaron los siguientes valores: 0 cuando el niño no producía el ítem en cuestión y 1 cuando sí lo producía. La variable Vocabulario se calculó sumando los puntajes de los 616 ítems de la sección. La variable LME3 se obtuvo promediando el número de palabras de los ejemplos de las tres frases más largas reportadas. A los participantes que no producían aún combinaciones se les asignó un 1 (en tanto todos producían al menos palabras sueltas). En los ítems de la subsección Complejidad morfosintáctica, se emplearon cuatro valores según el nivel

de desarrollo: 0 si el niño aún no producía nada parecido al ejemplo de las estructuras gramaticales mostradas en el cuestionario; y 1, 2 o 3, según produjera las formas de menos compleja a más compleja. La variable Complejidad morfosintáctica se obtuvo sumando los puntajes de los 34 ítems de la sección.

Las subsecciones Vocabulario y Complejidad morfosintáctica evidenciaron una consistencia interna muy alta (alfas de Cronbach de .99 y .97, respectivamente;  $p < .01$ ). Los puntajes obtenidos para Vocabulario y Complejidad morfosintáctica, así como la variable LME3, resultaron sensibles a la edad (índices de correlación de Spearman de .73, .64 y .61, respectivamente;  $p < .01$ ).

Luego se examinaron las distribuciones de las siguientes variables: Edad, Escolarización materna, Vocabulario, LME3 y Complejidad morfosintáctica, que se presentan en la sección de resultados.

### **Descripción del vocabulario y preferencia por los sustantivos**

Para conocer la composición del vocabulario e identificar si existía una mayor presencia de sustantivos, se realizaron los siguientes análisis. Los ítems de la subsección Vocabulario se clasificaron del modo siguiente: (i) Sustantivos, (ii) Verbos, (iii) Adjetivos, (iv) Adverbios, (v) Pronombres, determinantes e interrogativas, (vi) Preposiciones, (vii) Conjunciones y subordinantes, (viii) Onomatopeyas, y (ix) Interjecciones. Luego, con el objetivo de describir la distribución de las clases anteriores y evaluar si los sustantivos eran la clase más producida, se realizaron los siguientes análisis.

#### ***Número de ítems producidos por clase***

Para ofrecer una descripción completa de la producción de los niños, se contó el número de ítems que cada uno producía por clase de palabras y en total.

#### ***Relación entre el número de ítems producidos por clase y la edad***

Con el fin de evaluar si había algún cambio evolutivo, se calcularon las correlaciones entre el número de ítems producidos por cada clase y la edad empleando el coeficiente de Spearman.

#### ***Comparación entre sustantivos, predicados y palabras funcionales producidos***

Con el objetivo de describir la producción de sustantivos, predicados y palabras funcionales y evaluar si los sustantivos eran la clase de palabras más producida, se crearon los grupos de palabras Predicados y Palabras funcionales, siguiendo a Jackson-Maldonado et al. (1993). Para la descripción, se calculó el porcentaje promedio que representaba cada uno de los grupos de palabras respecto del total de ítems clasificados en la producción de los niños. Luego, con la frecuencia de Sustantivos, Predicados y Palabras funcionales, se estableció la comparación entre estos grupos de palabras, empleando un ANOVA de Welch (porque las variables tenían varianzas desiguales) y la prueba de Games-Howell para el examen post-hoc.

#### **Poder explicativo del vocabulario**

En primer lugar, se realizaron regresiones múltiples jerárquicas para evaluar si el vocabulario predecía el nivel de desarrollo gramatical. Se llevaron a cabo dos regresiones equivalentes, con distinta variable dependiente: en el primer caso, con LME3 y, en el segundo, con Complejidad morfosintáctica. Ambas regresiones se realizaron en dos pasos. Como se buscaba controlar el efecto de Edad y Escolarización materna, estas variables se introdujeron en el primer paso. En el segundo paso, se añadió Vocabulario, porque se buscaba identificar el poder explicativo de dicha variable.

## Resultados

### Descripción de las variables

La edad de los niños que participaron variaba entre 16 y 30 meses, con un promedio de 22.34 ( $N = 104$ ,  $DT = 4.38$ ). La variable Escolarización materna presentó las siguientes frecuencias para 101 participantes (3 no respondieron; se considera el máximo nivel alcanzado): 4 tenían completados estudios primarios; 7, secundarios; 18, técnicos o en instituto; y 72, universitarios. Como se puede observar, la gran mayoría de las madres (89 %) contaban con estudios superiores (técnicos o universitarios), lo que refleja que el grupo de familias a las que se aplicó el CDI era muy homogéneo.

Las demás medidas se distribuyeron de la siguiente forma (en todos los casos:  $N = 104$ ). En primer lugar, Vocabulario presentó un puntaje promedio de 153.71 (mínimo: 3, máximo: 569 y  $DT = 149.73$ ). Esto significa que los niños conocían en promedio 154 palabras de las 616 por las que se preguntó, y había niños que conocían solo 3 y otros hasta 569, lo que mostró una producción baja y muy variada.

En segundo lugar, LME3 tuvo como promedio 2.25 (mínimo: 1, máximo: 6.33 y  $DT = 1.41$ ). La medida representa el número promedio de palabras de los tres enunciados más largos, que correspondió a poco más de dos palabras como media del grupo.

Por último, las respuestas a los ítems de Complejidad morfosintáctica -que podían ser 0 (cuando el niño aún no produce ni siquiera la forma más simple presentada), 1, 2 o 3, según el nivel de complejidad de la construcción- se distribuyeron de la siguiente forma. El total de respuestas a los 34 ítems con los cuales se evaluó a los 104 niños fue de 3522 (14 ítems no fueron contestados). De este total, la mayoría, 2029, correspondió al nivel 0, seguido por 961 respuestas correspondientes al nivel 1; 228, al nivel 2 y 304, al nivel 3. Esto significa que gran parte de las respuestas de los niños (el 85 %) se ubicó en los dos niveles más bajos de

la escala. El puntaje total por niño, empleado como variable Complejidad morfosintáctica, arrojó un promedio de 22.39 (mínimo: 0, máximo: 90 y  $DT = 22.07$ ).

### Descripción del vocabulario y preferencia por los sustantivos

Se siguió una clasificación de las palabras similar a las de Caselli, Casadio y Bates (1999) y Jackson-Maldonado et al. (1993), pero en lugar de agrupar los ítems de algunas áreas temáticas de la subsección Vocabulario del CDI dentro de una misma clase, se clasificaron los ítems uno por uno, siguiendo la categorización del Diccionario de la lengua española de la Real Academia Española (2014); además, a diferencia de los estudios anteriores, se consideró también los nombres propios como sustantivos. A continuación, se describe nuestra clasificación.

Para empezar, de los 616 ítems de la subsección Vocabulario, se excluyeron aquellos formados por una o más palabras que no podían ser clasificadas en una única categoría gramatical (como “mucho”, que puede ser adjetivo o adverbio; o “a ver”, que está formada por una preposición y un verbo). Los 585 ítems restantes fueron clasificados en (i) Sustantivos; (ii) Verbos; (iii) Adjetivos; (iv) Adverbios; (v) Pronombres, determinantes e interrogativas; (vi) Preposiciones; (vii) Conjunciones y subordinantes; (viii) Onomatopeyas y (ix) Interjecciones.

La clase Sustantivos contenía los ítems de las áreas temáticas Animales, de verdad o de juguete, Partes del cuerpo, Juguetes, Vehículos, de verdad o de juguete, Alimentos y bebidas, Ropa, Muebles y cuartos, y Utensilios de la casa del CDI (como en Caselli et al., 1999 y Jackson-Maldonado et al., 1993), y también ítems de áreas que dichas autoras no incluyeron: todos los de Personas, y Objetos y lugares fuera de la casa, así como algunos de Juegos, rutinas y fórmulas sociales (“aplausos” y “besitos”), Cualidades (“hambre” y “miedo”) y Tiempo (“día” y “(por la) tarde”). La clase

Sustantivos resultante incluía 342 ítems. La clase Verbos constaba de 103 ítems, todos de las áreas temáticas Acciones, y Auxiliares y perífrasis. La clase Adjetivos estaba compuesta por 40 ítems e incluía un ítem de Juegos, rutinas y fórmulas sociales (“tonto”) y todos los de Cualidades (“alto” y “difícil”), una vez excluidos los sustantivos. Los 28 ítems de la clase Adverbios provienen de Juegos, rutinas y fórmulas sociales (“sí”, “no” y “ya/ok”), Tiempo (“ahora/ahorita” y “pronto”), Preguntas (“cómo” y “cuándo”), Preposiciones y locativos (“arriba” y “ahí”), Cuantificadores y artículos (“también” y “tampoco”) y Conectivas (“entonces”). La clase Pronombres, determinantes e interrogativas estaba formada por un ítem de Juegos, rutinas y fórmulas sociales (“¿por qué?”), así como por ítems del área de Pronombres, determinantes y cuantificadores (“algo” y “conmigo”) y Preguntas (“¿qué?” y “¿quién?”), con un total de 38 ítems. La clase Preposiciones estaba compuesta por 9 ítems, de Preposiciones y locativos (“a” y “con”); y Conjunciones y subordinantes, por 6 ítems de Conectivas (“o” y “pero”). La clase Onomato-

peyas estaba compuesta por 15 ítems de Interjecciones y sonidos de animales y cosas (“bee [oveja]”, “miau [gato]” y “tic-tac [reloj]”). Por último, la clase Interjecciones estaba compuesta por 4 ítems de Interjecciones y sonidos de animales y cosas (“¿ah?/¿eh?” y “pucha/uy”). Las 585 palabras clasificadas corresponden al 95 % de las palabras del CDI.

### *Número de ítems producidos por clase*

El número promedio de ítems de cada clase que produjeron los niños ( $N = 104$ ) fue el siguiente: Sustantivos, 86.12 ( $DT = 85.64$ ); Verbos, 21.45 ( $DT = 29.28$ ); Adjetivos, 7.15 ( $DT = 9.74$ ); Adverbios, 7.75 ( $DT = 5.91$ ); Pronombres, determinantes e interrogativas, 9.14 ( $DT = 9.11$ ); Preposiciones, 1.18 ( $DT = 1.99$ ); Conjunciones y subordinantes, .79 ( $DT = 1.31$ ); Onomatopeyas, 8.17 ( $DT = 4.06$ ); e Interjecciones, 2.60 ( $DT = 1.29$ ). Además, el promedio total de ítems de las clases evaluadas que produjeron los niños fue de 150.70 ( $DT = 150.82$ ). La Tabla 1 ofrece más información sobre la distribución del vocabulario.

Tabla 1  
*Número de ítems producidos según clase de palabras.*

<b>Clase (número de ítems en el CDI)</b>	<b>Rango</b>	<b>M (DT)</b>
Sustantivos (342)	2 - 330	86.12 (85.64)
Verbos (103)	0 - 101	21.45 (29.28)
Adjetivos (40)	0 - 39	7.15 (9.74)
Adverbios (28)	0 - 27	7.75 (5.91)
Pronombres, determinantes e interrogativas (38)	0 - 35	9.14 (9.11)
Preposiciones (9)	0 - 8	1.18 (1.99)
Conjunciones y subordinantes (6)	0 - 6	0.79 (1.31)
Onomatopeyas (15)	0 - 15	8.17 (4.06)
Interjecciones (4)	0 - 4	2.60 (1.29)
<b>Total (585)</b>	<b>3 - 577</b>	<b>150.70 (150.82)</b>
Predicados (143)	0 - 138	28.61 (38.35)
Palabras funcionales (53)	0 - 47	11.12 (11.80)

$N = 104$

### ***Relación entre el número de ítems producidos por clase y la edad***

Se calcularon las correlaciones entre el número de ítems producidos por clase de palabras y la variable Edad, con el objetivo de ver si la producción aumentaba con ella. Todas las correlaciones fueron positivas y significativas: altas para Sustantivos y Adjetivos ( $r = .72$  y  $.70$ , respectivamente); moderadas para Verbos; Pronombres, determinantes e interrogativas; Adverbios; y Preposiciones ( $r = .69$ ,  $.66$ ,  $.58$  y  $.54$ , respectivamente); y bajas para Onomatopeyas; Conjunciones y subordinantes; e Interjecciones ( $r = .44$ ,  $.43$  y  $.40$ , respectivamente); para todas las correlaciones,  $r(102)$  y  $p < .001$ . Este resultado muestra que, para todas las clases, la producción aumenta con la edad, aunque en diferente medida; los sustantivos son los que más aumentan, seguidos por los adjetivos y los verbos.

### ***Comparación entre sustantivos, predicados y palabras funcionales producidos***

Siguiendo a Jackson-Maldonado et al. (1993), para cada niño se calculó el número Sustantivos, el de Predicados (sumando Verbos y Adjetivos) y el de Palabras funcionales (sumando Pronombres y determinantes; Preposiciones; y Conjunciones y subordinantes). Los promedios de las medidas consideradas fueron los siguientes (en todos los casos,  $N = 104$ ): Sustantivos (342 ítems), 86.12 ( $DT = 85.64$ ); Predicados (143 ítems), 28.6 ( $DT = 38.35$ ) y Palabras funcionales (53 ítems), 11.12 ( $DT = 11.80$ ).

Los resultados de los análisis realizados con estas medidas fueron los siguientes. Primero, con respecto al porcentaje del total de ítems clasificados (585) que representaba cada clase, a Sustantivos correspondió, en promedio, el 56.64 % ( $DT = 12.14$ , rango: 17.81 % - 100 %); a Predicados, el 11.92 % ( $DT = 9.19$ , rango: 0 % - 27.61 %); y a Palabras funcionales, el 6.95 % ( $DT = 4.33$ , rango: 0 % - 21.43 %); el restante 6.68 % correspondió

a otras clases de palabras. Estos resultados muestran que los sustantivos representan más de la mitad del vocabulario productivo clasificado. Segundo, la diferencia entre la producción de Sustantivos, Predicados y Palabras funcionales resultó significativa. El ANOVA de Welch entre el número de ítems producidos de cada una de estas tres clases mostró diferencias significativas entre ellas:  $F$  de Welch (2,309) = 53.70,  $p < .001$ . La prueba post hoc de Games-Howell reveló que la producción de Sustantivos era significativamente mayor que la de Predicados ( $p < .001$ ) y la de Predicados significativamente mayor que la de Palabras funcionales ( $p < .001$ ). Esto evidencia que la producción total de sustantivos, según resultados de nuestro CDI, es mayor que la de las otras dos categorías evaluadas.

### **Poder explicativo del vocabulario**

Con respecto a los análisis de regresión, el primer modelo, con LME3 como variable dependiente, explicó un porcentaje importante de la variabilidad: 56 %. El primer paso explicó un 37 %, debido a las contribuciones de las dos variables introducidas: Escolarización materna ( $\beta = .17$ ,  $p = .04$ ) y Edad ( $\beta = .57$ ,  $p < .001$ ), mientras que el segundo paso explicó un 20 %, con la contribución del Vocabulario ( $\beta = .60$ ,  $p < .001$ ), como se muestra en la Tabla 2. El segundo modelo, ahora con Complejidad morfosintáctica como variable dependiente, explicó una parte más importante de la variabilidad que el anterior, 75 %. El primer paso dio cuenta de un 38 %, debido exclusivamente a la contribución de la Edad ( $\beta = .61$ ,  $p < .001$ ), mientras que el segundo paso explicó un 37 % más de la variabilidad, con la contribución del Vocabulario ( $\beta = .82$ ,  $p < .001$ ), como se muestra en la Tabla 2. Ambos resultados revelan un patrón equivalente y muestran que una parte de las diferencias en el resultado de los niños en las medidas de gramática depende de sus niveles de vocabulario, independientemente de la edad que tengan y del nivel educativo de sus madres.

Tabla 2

Regresiones con LME3 y Complejidad Morfosintáctica (CM) como variables dependientes.

	R <sup>2</sup>		ΔR <sup>2</sup>		B		SEB		β	
	LME3	CM	LME3	CM	LME3	CM	LME3	CM	LME3	CM
Paso 1	.37	.38	.37*	.38*						
Escolarización materna					.31	1.31	.15	2.25	.17**	.05
Edad					.18	3.09	.03	.40	.57*	.61*
Paso 2	.56	.75	.20*	.37*						
Vocabulario					.01	.12	.001	.10	.60*	.82*

\* $p < .001$ ; \*\* $p = .04$

## Discusión y conclusiones

En este trabajo, se han empleado los datos obtenidos a partir de la aplicación de una versión preliminar del CDI a 104 niños hablantes de español del Perú con edades entre los 16 y 30 meses para examinar, a partir del reporte parental, la composición de su vocabulario productivo, así como la relación entre éste y la gramática. Se buscó identificar si los sustantivos son la clase de palabras más producida y la que más aumenta con la edad. Además, se evaluó si el nivel de desarrollo del vocabulario explica el nivel de desarrollo gramatical.

### Sobre el vocabulario temprano y la preferencia por los sustantivos

Como se mencionó en la introducción, el vocabulario es una medida importante para entender y predecir el desarrollo lingüístico (Lee, 2011; Marchman y Fernald, 2008). Por esta razón, resulta imprescindible conocer la composición del vocabulario en edades muy tempranas como las que evalúan los CDI, en particular en el caso del español del Perú, sobre el cual la información es escasa. Los objetivos de este trabajo, en relación con el vocabulario

reportado, fueron describir su composición e identificar si los sustantivos son la clase de palabras más frecuente y la que más aumenta con la edad, como ocurre con niños hablantes de italiano (Caselli et al., 1995; D'Odorico y Fasolo, 2007) y de español mexicano (Jackson-Maldonado et al., 1993) de edades equivalentes.

En relación con el primer objetivo, se halló que los Sustantivos representan el 57 % de las palabras clasificadas (que son el 94 % del total de palabras del CDI peruano), y su producción es mayor que la de Predicados y ésta, a su vez, mayor que la de Palabras funcionales. Estos resultados mostraron que, según el reporte, en el español infantil del Perú, hay una clara preferencia por los sustantivos frente a los predicados, tal como sucede en el italiano y en el español mexicano.

Con respecto al segundo objetivo, la relación entre las medidas de cada clase de palabras y la edad, se encontró que la producción de Sustantivos es la que más aumenta ( $r = .72$ ), seguida por la de Adjetivos y Verbos ( $r = .70$  y  $.69$ , respectivamente), clases que conforman el grupo de Predicados. Las clases que conforman el grupo de Palabras funcionales aumentan en menor medida con la edad, por lo menos en el periodo aquí estudiado,

con valores para  $r$  entre .43 (Conjunciones y subordinantes) y .66 (Pronombres, determinantes e interrogativas). Los datos sugieren que, en este periodo, existe un aumento en la producción de sustantivos más importante que en otras categorías gramaticales, como sucede en el español mexicano (Jackson-Maldonado et al., 1993), aunque para confirmarlo serían útiles más estudios, en particular análisis finos de producción espontánea.

El aporte de este estudio a la investigación sobre el vocabulario infantil, a pesar de basarse en datos obtenidos de forma indirecta por medio de un reporte parental, ha permitido (1) proporcionar nueva información sobre la composición del vocabulario productivo temprano del español del Perú, sobre la cual no existen más datos accesibles registrados, y (2) comprobar que, en esta variedad, el vocabulario temprano está compuesto principalmente por sustantivos, como sugieren estudios previos con otras variedades.

### **Sobre la relación entre el vocabulario y la gramática en el desarrollo temprano**

La relación entre vocabulario y gramática en el lenguaje infantil, medidos por el CDI, ha sido fuertemente demostrada en diferentes lenguas, incluso una vez controlada la edad y el nivel socioeconómico (Devescovi et al., 2005; Mariscal y Gallego, 2012; Serrat et al., 2010; Silva et al., 2017; Pérez-Pereira y Resches, 2011). En este estudio, un tercer objetivo era evaluar si el vocabulario productivo de los niños peruanos entre 16 y 30 meses explicaba su nivel de desarrollo gramatical, como ocurre en el español peninsular (Mariscal y Gallego, 2012) y en el mexicano (Thal et al., 2000) de niños de edades equivalentes, controlados los efectos de la edad y el nivel socioeconómico. Para ello, se realizaron análisis de regresión que midieron el poder explicativo del Vocabulario en relación con la Longitud media de los tres enunciados más largos (LME3) y la Complejidad morfosintáctica. Los resultados muestran que, una vez controlados la edad

y el nivel socioeconómico, el vocabulario explica un 20 % de la LME3 y un 37 % de la Complejidad morfosintáctica, lo que se interpreta como evidencia del fuerte poder explicativo del vocabulario en la gramática infantil del español del Perú. De esta manera, se halla una importante relación entre vocabulario y gramática en una variedad infantil antes no estudiada, a partir de los datos registrados con el CDI. Se espera que estudios posteriores puedan confirmar este hallazgo empleando medidas directas de lenguaje infantil.

Por último, se incluye una reflexión con respecto a la interpretación de la relación vocabulario-gramática. En esta investigación se asumió una relación unidireccional (desde el vocabulario hacia la gramática); sin embargo, la relación es compleja y no se limita necesariamente a una única dirección (más bien, podría ser bidireccional o recíproca). Más aún, es posible preguntarse si el *input* no es el factor que explique simultáneamente el desarrollo del vocabulario y el de la gramática, como parece ocurrir en relación con la predominancia, según la lengua, de algunos tipos de palabras, como los sustantivos, en el lenguaje infantil. Hay abundante evidencia de que las características del *input* explican las diferencias en el desarrollo lingüístico infantil: desde Hart y Risley (1995), quienes lo asociaron con el número de palabras que los niños escuchaban y éste, a su vez, con el nivel socioeconómico de la familia –y que calculaban en 30 millones de palabras la diferencia entre niños de familias acomodadas y pobres–, hasta estudios más recientes que son críticos de la asociación entre *input* y nivel socioeconómico, pero que confirman la importancia de la calidad y cantidad del *input* de los padres. En esta línea, Gilkerson y colegas (2017), por ejemplo, ratificaron la relación entre el número de palabras escuchadas y el desarrollo lingüístico, aunque no la linealidad propuesta por Hart y Risley (1995) de la relación nivel socioeconómico-número de palabras. Hurtado, Marchman y Fernald (2008) y Weisleder y Fernald (2013) también mostraron la importancia del *input*

de los padres para el desarrollo lingüístico en español. Los resultados presentados aquí podrán explicarse de manera más profunda con estudios que incluyan también como variables las medidas provenientes del *input* de los cuidadores con el objetivo de identificar diferencias en la producción infantil basadas en el *input* al que están expuestos los niños.

## Referencias bibliográficas

- Alexander, G.M. y Wilcox, T. (2012). Sex differences in early infancy. *Child Development Perspectives*, 6(4), 400-406. <https://doi.org/10.1111/j.1750-8606.2012.00247.x>
- Bates, E. y Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes* 12, 507-586. <https://doi.org/10.1080/016909697386628>
- Blume, M., Fernández-Flecha, M., Junyent, A. y Tijero, T. (2019). Proceso de adaptación de los Inventarios de Desarrollo Comunicativo MacArthur-Bates al castellano estándar del Perú. *Lexis*, 43(2), 369-420. <https://doi.org/10.18800/lexis.201902.003>
- Bornstein, M.H., Haynes, O.M., Painter, K.M. y Genevro, J.L. (2000). Child language with mother and with stranger at home and in the laboratory: A methodological study. *Journal of Child Language*, 27(2), 407-420. <https://doi.org/10.1017/s0305000900004165>
- Braginsky, M., Yurovsky, D., Marchman, V.A. y Frank, M.C. (2015). Developmental changes in the relationship between grammar and the lexicon. En D.C. Noelle, R. Dale, A.S. Warlaumont, J. Yoshimi, T. Matlock, C.D. Jennings y P.P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, 256-261. Austin, TX: Cognitive Science Society. Recuperado de <https://mindmodeling.org/cogsci2015/papers/0054/paper0054.pdf>
- Brown, P. (1998). Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics*, 36(4), 713-754. <https://doi.org/10.1515/ling.1998.36.4.713>
- Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L. y Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10(2), 159-199. [https://doi.org/10.1016/0885-2014\(95\)90008-X](https://doi.org/10.1016/0885-2014(95)90008-X)
- Caselli, C., Casadio, P. y Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, 26(1), 69-111. <https://doi.org/10.1017/S0305000998003687>
- D'Odorico, L. y Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language*, 34(4), 891-907. <https://doi.org/10.1017/S0305000907008240>
- Dale, P., Dionne, G., Eley, T.C. y Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language* 27, 619-42. <https://doi.org/10.1017/S0305000900004281>
- Dale, P.S. y Goodman, J.C. (2005). Commonality and individual differences in vocabulary growth. In M. Tomasello y D. I. Slobin (eds.), *Beyond nature-nurture. Essays in honor of Elizabeth Bates*, 41-78. Mahwah, NJ: Lawrence Erlbaum.
- Devescovi, A., Caselli, M.C., Marchione, D., Pasqualetti, P., Reilly, J. y Bates, E. (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, 32(4), 759-786. <https://doi.org/10.1017/S0305000905007105>
- Dhillon, R. (2010). Examining the 'Noun Bias': A structural approach. *University of Pennsylvania Working Papers in Linguistics*, 16(1), 7. Recuperado de <https://repository.upenn.edu/pwpl/vol16/iss1/7/>
- Dockrell, J.E. (2001). Assessing language skills in preschool children. *Child Psychology and Psychiatry Review*, 6(2), 74-85. <https://doi.org/10.1017/s1360641701002532>
- Dollaghan, C.A., Campbell, T.F., Paradise, J.L., Feldman, H.M., Janosky, J.E., Pitcairn, D. N. y Kurs-Lasky, M. (1999). Maternal education and measures of early speech and

- language. *Journal of Speech, Language, and Hearing Research*, 42(6), 1432-1443. <https://doi.org/10.1044/jslhr.4206.1432>
- Feldman, H.M., Dale, P.S., Campbell, T.F., Colborn, D.K., Kurs-Lasky, M., Rockette, H. E. y Paradise, J. L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development*, 76(4), 856-868. <https://doi.org/10.1111/j.1467-8624.2005.00882.x>
- Fenson, L., Marchman, V., Thal, D., Dale, P., Bates, E. y Reznick, S. (2007). *The MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual* (2nd ed.). Baltimore (EEUU): Brookes Publishing.
- Fenson, L., Dale, P., Reznick, J.S., Thal, D., Bates, E., Hartung, J., Pethink, S. y Reilly, J. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA (EEUU): Singular Publishing Group.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report*, 257. Recuperado de [https://www.ideals.illinois.edu/bitstream/handle/2142/17514/ctrstreadtechrepv01982i00257\\_opt.pdf](https://www.ideals.illinois.edu/bitstream/handle/2142/17514/ctrstreadtechrepv01982i00257_opt.pdf)
- Gilkerson, J., Richards, J.A., Warren, S.F., Montgomery, J.K., Greenwood, C.R., Oller, D.K., Hansen, J. y Paul, T.D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, 26(2), 248-265. [https://doi.org/10.1044/2016\\_AJSLP-15-0169](https://doi.org/10.1044/2016_AJSLP-15-0169)
- Hart, B. y Risley, T.R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, 28(6), 1096. <https://doi.org/10.1037/0012-1649.28.6.1096>
- Hart, B. y Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD (EEUU): Paul H Brookes.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368-1378. <https://doi.org/10.1111/1467-8624.00612>
- Hurtado, N., Marchman, V.A. y Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11(6), F31-F39. <https://doi.org/10.1111/j.1467-7687.2008.00768.x>
- Huttenlocher, J., Vasilyeva, M., Cymerman, E. y Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337-374. [https://doi.org/10.1016/S0010-0285\(02\)00500-5](https://doi.org/10.1016/S0010-0285(02)00500-5)
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J. y Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61(4), 343-365. <https://doi.org/10.1016/j.cogpsych.2010.08.002>
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E. y Gutiérrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language*, 20(3), 523-549. <https://doi.org/10.1017/S0305000900008461>
- Law, J. y Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, 13(4), 198-206. <https://doi.org/10.1111/j.1475-3588.2008.00503.x>
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(1), 69-92. <https://doi.org/10.1017/S0142716410000299>
- López-Ornat, S., Gallego, C., Gallo, P., Karousou, A., Mariscal, S. y Martínez, M. (2005). *Inventarios de Desarrollo Comunicativo MacArthur Manual*. Madrid: TEA Ediciones.
- Marchman, V. y Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21, 339-366. <https://doi.org/10.1017/S0305000900009302>
- Marchman, V.A. y Fernald, A. (2008). Speed of word recognition and vocabulary knowledge

- in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11(3), F9–F16. <https://doi.org/10.1111/j.1467-7687.2008.00671.x>
- Mariscal, S. y Gallego, C. (2012). The relationship between early lexical and grammatical development in Spanish: Evidence in children with different linguistic levels. *The Spanish Journal of Psychology*, 15(1), 112-123. [https://doi.org/10.5209/rev\\_sjop.2012.v15.n1.37293](https://doi.org/10.5209/rev_sjop.2012.v15.n1.37293)
- Mariscal, S., López-Ornat, S., Gallego, C., Gallo, P., Karousou, A. y Martínez, M. (2007). La evaluación del desarrollo comunicativo y lingüístico mediante la versión española de los inventarios MacArthur-Bates. *Psicothema*, 19(2), 190-197. <https://doi.org/10.1174/021093906777571664>
- Mariscal, S., Nieva, S. y López-Ornat, S. (2010). Observar y medir el desarrollo gramatical temprano en español. *Psicothema*, 22, 51-56. Recuperado de <http://www.redalyc.org/articulo.oa?id=72712699009>
- Marjanovič-Umek, L., Fekonja-Peklaj, U. y Podlesek, A. (2013). Characteristics of early vocabulary and grammar development in Slovenian-speaking infants and toddlers: A CDI-adaptation study. *Journal of Child Language*, 40(4), 779-798. <https://doi.org/10.1017/S0305000912000244>
- McGregor, K.K., Sheng, L.I. y Smith, B. (2005). The precocious two-year-old: Status of the lexicon and links to the grammar. *Journal of Child Language*, 32(3), 563-585. <https://doi.org/10.1017/s0305000905006926>
- Murray, A.D., Johnson, J. y Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language*, 17(3), 511-525. <https://doi.org/10.1017/S0305000900010862>
- Pérez-Pereira, M. y Resches, M. (2011). Concurrent and predictive validity of the Galician CDI. *Journal of Child Language*, 38(1), 121-140. <https://doi.org/10.1017/s0305000909990262>
- Real Academia Española (2014). *Diccionario de la lengua española* (23ª ed.). Consultado en <http://www.rae.es/>
- Serrat, E., Sanz-Torrent, M., Badia, I., Aguilar, E., Olmo, R., Lara, F., Llorenç, A. y Serra, M. (2010). La relación entre el aprendizaje léxico y el desarrollo gramatical. *Infancia y Aprendizaje*, 33(4), 435-448. <https://doi.org/10.1174/021037010793139590>
- Silva, C., Cadime, I., Ribeiro, I., Santos, S., Santos, A.L. y Viana, F.L. (2017). Parents' reports of lexical and grammatical aspects of toddlers' language in European Portuguese: Developmental trends, age and gender differences. *First Language*, 37(3), 267-284. <https://doi.org/10.1177/0142723716689274>
- Tardif, T., Gelman, S.A. y Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, 70, 620–635. <https://doi.org/10.1111/1467-8624.00045>
- Tardif, T., Shatz, M. y Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, 24, 535–565. <https://doi.org/10.1017/s030500099700319x>
- Thal, D., Jackson-Maldonado, D. y Acosta, D. (2000). Validity of a parent-report measure of vocabulary and grammar for Spanish-speaking toddlers. *Journal of Speech, Language, and Hearing Research*, 43(5), 1087-1100. <https://doi.org/10.1044/jslhr.4305.1087>
- Weisleder, A. y Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143-2152. <https://doi.org/10.1177/0956797613488145>
- Wordbank: <http://wordbank.stanford.edu>

Recibido: 2 de noviembre de 2018

Aceptado: 20 de agosto de 2020