



La composición léxica del vocabulario infantil en niños de cuatro años de distintos grupos sociales y su relación con el entorno lingüístico

Lexical Composition of the Vocabulary of Four-Year-Old Children From Different Social Groups and its Relationship to the Linguistic Environment

Macarena Sol Quiroga

Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental
– Consejo Nacional de Investigaciones Científicas y Técnicas (CIIPME-CONICET) /
Universidad Nacional de Hurlingham (UNaHur), Buenos Aires / Argentina
macarenasolquiroga@gmail.com

<https://orcid.org/0000-0002-9414-7454>

Celia Renata Rosemberg

Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental
– Consejo Nacional de Investigaciones Científicas y Técnicas (CIIPME-CONICET),
Buenos Aires / Argentina

crrosem@hotmail.com

<http://orcid.org/0000-0002-5343-5652>

Florencia Alam

Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental
– Consejo Nacional de Investigaciones Científicas y Técnicas (CIIPME-CONICET),
Buenos Aires / Argentina

florenciaalam@gmail.com

<http://orcid.org/0000-0002-3893-7873>

Resumen: El objetivo de este trabajo es analizar la composición léxica del vocabulario productivo infantil y su relación con el *input* lingüístico al que los niños de cuatro años de distintos grupos sociales están expuestos. A partir de un corpus de habla espontánea, se seleccionaron las transcripciones de 19 niños de dos grupos sociales distintos (n

= 38), correspondientes a 456 horas de grabación, y se analizó la cantidad total y los tipos distintos de sustantivos, adjetivos y verbos, tanto en la producción infantil como en el *input* lingüístico. El análisis de la varianza mostró que en ambos casos hubo una mayor cantidad de tipos distintos de sustantivos respecto a las demás palabras, pero una mayor cantidad total de verbos. Las regresiones beta con el vocabulario infantil como variable dependiente indicaron que la pertenencia al grupo socioeconómico fue un predictor de tipos distintos de todas las clases de palabras y de cantidad solo de adjetivos; la presencia de cada clase de palabra en el *input* fue predictor de la cantidad de las tres clases de palabras, y de mayor cantidad de tipos distintos de adjetivos y verbos. Estos resultados muestran la complejidad de las experiencias tempranas infantiles y la relevancia de atender a los distintos contextos en los que los niños adquieren el lenguaje.

Palabras clave: vocabulario; composición léxica; producción infantil; entorno lingüístico.

Abstract: The aim of this study was to analyze the lexical composition of child expressive vocabulary and its relationship with linguistic input of four-year-old children from different social backgrounds. From a spontaneous speech corpus, we analyzed the transcriptions of 19 children from two different social backgrounds ($n = 38$), corresponding to 456 hours of recordings, and we analyzed the lexical diversity and quantity of nouns, adjectives, and verbs, both in child vocabulary and input they were exposed to. The analysis of variance showed that in both cases there was a greater noun type quantity compared to the other word classes, but a greater token quantity of verbs. Beta regressions with child vocabulary as a dependent variable showed that belonging to a socioeconomic group was a predictor of the three lexical classes' types, but only of the adjective's tokens; the presence of each word class in the linguistic input was a predictor of all three word classes, and only of adjective and verb types. These results manifest the complexity of early childhood experiences and the need to pay attention to the various contexts where children acquire language.

Keywords: vocabulary; lexical composition; child production; linguistic input.

Recebido em 08 de fevereiro de 2022

Aceito em 23 de maio de 2022

1 Introducción

El aprendizaje de vocabulario es una de las dimensiones del desarrollo lingüístico que mayor atención ha atraído en las últimas tres décadas. En el

vocabulario se muestra con claridad el carácter cognitivo y representacional del lenguaje (NELSON, 1996): un vocabulario amplio y diverso incide no solo en las posibilidades de comunicación del niño, sino también en sus habilidades para comprender y producir discurso extendido (BEALS, 1997; ROSEMBERG *et al.*, 2011; ROSEMBERG, 2021). Además, el vocabulario tiene un impacto directo en la adquisición de la lectura y la escritura, y en el desempeño académico general (HOFF, 2013; SNOW; MATTHEWS, 2016). La amplitud del vocabulario da lugar a representaciones robustas, flexibles, precisas, interrelacionadas y de fácil acceso, necesarias para el acceso al sistema de escritura y para la comprensión y producción de textos orales y escritos (BIEMILLER; BOOTE, 2006; PERFETTI, 2007; SÉNECHAL; OUELETTE; RODNEY, 2006).

Dada la relevancia que el vocabulario posee en el desarrollo lingüístico y en la trayectoria académica de los niños, un número importante de estudios ha abordado la variabilidad individual y social en la comprensión y producción de léxico en la infancia (ver por ejemplo FERNALD; MARCHMAN; WEISLEDER, 2013; FRIEND; DEANDA; ARIAS-TREJO; POULIN-DUBOIS; ZESIGER, 2017; HART; RISLEY, 1995; HOFF, 2003; PACE; LUO; HIRSH-PASEK; GOLINKOFF, 2017; ROWE, 2008; SCHWAB; LEW-WILLIAMS, 2016; ROSEMBERG; ALAM, 2021). En estos estudios se asume que las disparidades encontradas en la comprensión y producción del vocabulario emergen de diferencias en las experiencias infantiles tempranas (MACWHINNEY, 1987; NELSON, 1996; TOMASELLO, 2003, entre otros), en particular, asociadas a las características generales del lenguaje al que los niños se hallan expuestos (*input* lingüístico) en el entorno del hogar desde su nacimiento. De hecho, varios de ellos han proporcionado evidencias de la relación entre las características del *input* lingüístico en hogares en distinta condición socioeconómica y el vocabulario que los niños producen (HART; RISLEY, 1995; HOFF, 2003) y comprenden (FERNALD *et al.*, 2013; ROWE, 2008, 2012). Cabe señalar que la mayoría de estos estudios han sido realizados en América del Norte, se han centrado en las experiencias infantiles durante los primeros dos años de vida, y, en la mayor parte de los casos, han definido como *input* lingüístico solo el habla dirigida al niño y no el resto del habla que el niño escucha en su entorno, pero que no está específicamente dirigida a él.

En todos los estudios mencionados se ha analizado el impacto del *input* sobre el desarrollo del vocabulario infantil en niños de distintas

condiciones socioeconómicas considerando la cantidad de *input* al que ha estado expuesto el niño (la cantidad total de palabras que el niño escucha) y, en algunos de ellos, se han considerado, asimismo, estimaciones generales de las características del *input* lingüístico, tales como la complejidad sintáctica, medida en la longitud promedio de enunciados, o la diversidad léxica global. Sin embargo, en ellos no se ha atendido a los tipos de palabras a las que los niños tienen acceso en su entorno lingüístico, a su relación con la composición del vocabulario infantil y a eventuales diferencias en función de las condiciones socioeconómicas en las que se encuentran las familias. Los tipos de palabras que transmiten la mayor parte del contenido semántico y/o conceptual de la emisión, los sustantivos, verbos y adjetivos, difieren en su referencia a objetos, personas, fenómenos, acciones y estados y muestran, cada uno, un comportamiento morfológico, léxico y sintáctico particular, generando demandas específicas al procesamiento lingüístico (BŁASZCZAK; KLIMEK-JANKOWSKA, 2015).

Es por ello que en este trabajo se analiza, en una población socioeconómicamente diversa de niños argentinos escasamente estudiada, la relación entre la producción infantil de sustantivos, verbos y adjetivos y el *input* lingüístico al que los niños están expuestos en el entorno natural de su hogar. Con este objeto, se revisan primero las investigaciones previas sobre la relación entre el *input* lingüístico y el habla infantil, en las que se consideran diferencias según las condiciones socioeconómicas, y en segundo lugar las destinadas a dar cuenta de las diferencias en el procesamiento de sustantivos adjetivos y verbos, para finalmente presentar los objetivos y preguntas específicas de este trabajo. A continuación, se presenta la metodología empleada en el estudio, los resultados obtenidos y su discusión en base a las investigaciones previas.

1.1 El *input* lingüístico y su relación con el habla infantil en niños de distintos grupos sociales

La evidencia actual ha puesto de relieve la importancia del entorno lingüístico en el desarrollo del lenguaje infantil y ha señalado que el vocabulario es uno de los aspectos del desarrollo más sensible al *input* y a las experiencias infantiles (HOFF, 2003; HOFF; TIAN, 2005; LIEVEN, 2010).

Hart y Risley (1995) fueron de los primeros en sostener que las diferencias en vocabulario entre niños de distintos grupos sociales podían atribuirse a diferencias en el *input* lingüístico al que se veían expuestos en sus experiencias tempranas. A partir de grabaciones realizadas en los hogares de niños de entre 7 y 9 meses durante dos años y medio, confeccionaron inventarios de las palabras producidas por los niños y aquellas que conformaban su *input*: exclusivamente las palabras producidas por su cuidador principal, es decir, sin tener en cuenta el habla producida por otros participantes, y sin considerar palabras familiares, onomatopeyas, palabras extranjeras, insultos y otros vocablos no incluidos en el diccionario. Este análisis los condujo a concluir que en los hogares de grupo socioeconómico bajo (que vivían en situaciones de pobreza o que recibían asistencia gubernamental [*welfare*]) los niños escuchaban sistemáticamente menos palabras que sus pares pertenecientes a la clase trabajadora o aquellos que eran hijos de profesores de la Universidad de Kansas, y que estas diferencias podrían ser lo suficientemente consistentes como para permitir una proyección en el tiempo. Sus resultados además mostraron que la cantidad de habla producida por los cuidadores correlacionaba con el incremento del vocabulario infantil.

Diversos estudios replicaron estos análisis y encontraron diferencias en el vocabulario entre niños de hogares que vivían en distintas circunstancias socioeconómicas. Así, por ejemplo, los resultados de Hoff (2003) indican que las diferencias en el vocabulario productivo de niños de distinto grupo social podrían explicarse por diferencias en el habla materna. En una línea similar, Rowe (2008) encontró que el *input* lingüístico al que acceden niños de grupos socioeconómicos más favorecidos, en las interacciones con su madre, se caracteriza por un vocabulario más amplio, con mayor diversidad léxica y enunciados más extensos, que cumplen en menor medida función de directivas, y que estas diferencias se relacionan con la comprensión del vocabulario infantil.

Pero, ¿cuáles son los aspectos del *input* lingüístico que inciden en el vocabulario infantil? Las investigaciones han identificado asociaciones entre el vocabulario de los niños y distintos aspectos del *input* lingüístico al que están expuestos en sus interacciones tempranas: así, se ha mostrado una relación entre la cantidad y diversidad de vocabulario (FERNALD *et al.*, 2013; HART; RISLEY, 1995; HOFF, 2003; SCHNEIDMAN; GOLDIN-MEADOW, 2012; ROWE, 2012), la estructura sintáctica

(HOFF, 2003; HUTTENLOCHER *et al.*, 2002, 2007) y la función pragmática de los enunciados (ROWE, 2008, 2012) que componen el lenguaje que los niños escuchan y distintos aspectos del vocabulario infantil. Los resultados del estudio longitudinal de Rowe (2012) indican, además, que los aspectos del *input* materno que impactan en mayor medida en el vocabulario infantil varían a lo largo del desarrollo: mientras que la cantidad de palabras distintas que los niños escuchan a los 2;5 años predice el vocabulario a los 3;5 años, a los 4;5 años el vocabulario infantil se relaciona con la cantidad de enunciados descontextualizados escuchados a los 4 años.

A pesar de que, como lo muestran los resultados de Rowe (2012), el *input* lingüístico que el niño escucha continúa impactando en el desarrollo del lenguaje infantil en los años preescolares, la amplia mayoría de los estudios que analizan la adquisición del vocabulario se focalizan en los primeros dos años y son escasos los trabajos que abordan la relación entre el vocabulario infantil y el *input* lingüístico en momentos posteriores del desarrollo. En este sentido, es importante destacar que los cuatro años de edad constituye un punto de inflexión en el desarrollo lingüístico y cognitivo: el lenguaje se convierte en una herramienta discursiva para construir e integrar modelos mentales (BORZONE; ROSEMBERG, 2000; NELSON, 1996) y los niños comienzan a reconstruir su experiencia en géneros discursivos más complejos, como las narraciones. En la interacción social, con el apoyo de las estrategias de los adultos, que buscan precisar y reconceptualizar las palabras empleadas por los niños (ROSEMBERG; SILVA, 2009), estos amplían su vocabulario. Los niños se vuelven capaces de establecer redes léxicas y jerarquías entre conceptos, así como relaciones de sinonimia, antonimia e hiponimia; gracias a ello logran independizar el uso de estos elementos léxicos de los contextos situacionales en los que surgieron (NELSON, 1985; TOMASELLO, 2002). De este modo, el vocabulario se complejiza para dar lugar a los nuevos usos lingüísticos, pero también como resultado de ellos (NELSON, 1996; ROSEMBERG *et al.*, 2016).

Aún cuando la evidencia de los estudios previos indica que en los dos primeros años de vida es el habla específicamente dirigida al niño, y no el resto del habla que el niño escucha en su entorno, aquello que predice el desarrollo del vocabulario (FERNALD *et al.*, 2013; SCHNEIDMAN; ARROYO; LEVINE; GOLDIN-MEADOW, 2013), cabe suponer que a los 4 años de edad los niños pueden hacer uso de toda

el habla que los circunda para aprender nuevas palabras. Los estudios de Akhtar (2005) con niños de dos años y los de Foushee *et al.* (2021) con niños de entre tres y seis años señalan que los niños son capaces de aprender palabras a partir del habla no dirigida a ellos.

Apoyándose en la evidencia de aprendizajes lingüísticos a partir de habla escuchada entre otros participantes, Sperry *et al.* (2018) discuten los resultados de Hart y Risley (1995) y argumentan la necesidad de redefinir aquello que se considera *input* lingüístico. En su estudio, con una muestra caracterizada por una mayor presencia de niños que vivían en condiciones socioeconómicas desfavorecidas, encuentran que la variabilidad en la cantidad de palabras en el *input* no se hallaba relacionada con el grupo socioeconómico de las familias, sino con la comunidad (es decir, con una estructura social más extensa, que integraban las familias). De ahí que los autores llamen la atención sobre el hecho de que la interacción diádica donde un cuidador principal se dirige exclusivamente al niño constituye una construcción social que no es uniforme a través de las culturas y grupos sociales, por lo que el habla dirigida al niño de parte del cuidador primario no podría asumirse como una condición excluyente para el desarrollo del vocabulario.

En una línea similar, el estudio de STEIN *et al.* (2021) en hogares de niños argentinos menores de 2 años mostró una amplia variabilidad intragrupo en la cantidad total de palabras en el habla dirigida al niño, así como diferencias cuantitativas y cualitativas en el entorno lingüístico de hogares en distintas condiciones socioeconómicas. Los niños que vivían en condiciones socioeconómicas vulneradas escucharon más palabras totales, una mayor diversidad léxica y una mayor cantidad de enunciados en el habla dirigida a otros participantes que en el habla específicamente dirigida a ellos, mientras que este patrón se invirtió para los niños que vivían en condiciones socioeconómicas no vulneradas.

Los resultados de otros trabajos realizados en distintas culturas (STOLL *et al.*, 2012; CASILLAS *et al.*, 2017, 2020) coinciden con los de Sperry y los de Stein en mostrar que el habla dirigida al niño no se produce con igual frecuencia en todos los grupos sociales y culturales, y que la matriz de interacción de la cual surge la distinción entre el habla dirigida al niño y el habla dirigida a otros participantes constituye una construcción cultural. Ello sugiere la necesidad de expandir la definición de entorno lingüístico e *input* considerando toda el habla disponible para el niño.

Las distintas culturas y grupos sociales suelen, asimismo, diferir en la cantidad de participantes que interactúan con los niños (CASILLAS *et al.*, 2020; SCHNEIDEMAN; GOLDIN-MEADOW, 2012). Schneidman y Goldin-Meadow (2012) encuentran, por ejemplo, que los niños mayas escuchaban solo un 31% del habla por parte de adultos, mientras que el 69% del *input* restante provenía de otros niños; en cambio, los niños norteamericanos escucharon un 85% de *input* por parte de adultos y solo 15% por parte de otros niños. Consecuentemente, puede también variar la cantidad total de habla que el niño escucha en su entorno (CASILLAS *et al.*, 2017, STEIN *et al.* 2021). De ahí que se haya planteado la relevancia de analizar la proporción de habla que el niño escucha dirigida a él y aquella entre otros participantes como una forma de controlar la cantidad total de habla. Los resultados de Casillas *et al.* (2017) ponen de manifiesto la relevancia de esta distinción. En su análisis del entorno lingüístico en los hogares socioeconómicamente diversos de niños norteamericanos durante los primeros dos años de vida, considerando tanto el habla dirigida a los niños como dirigida a los adultos, encontraron que el grupo socioeconómico no influyó en la proporción de habla dirigida al niño, pero sí en la cantidad total de esta habla y del habla total. Sin embargo, esta diferencia decrecía con la edad: al final del segundo año, todos los niños indistintamente del grupo socioeconómico escucharon la misma cantidad de habla dirigida a ellos y dirigida a adultos. Los autores encontraron también que progresivamente, durante los primeros dos años de vida, las madres incrementaban la cantidad de habla dirigida a los niños varones.

La mayoría de las investigaciones presentadas en este apartado se ha focalizado en aspectos cuantitativos como la cantidad total y diversidad de habla (HART; RISLEY, 1995; HOFF, 2003; HOFF; TIAN, 2005; ROWE, 2008, 2012; HUTTENLOCHER *et al.*, 2007; SCHNEIDMAN; ARROYO; LEVINE; GOLDIN-MEADOW, 2013; SPERRY *et al.*, 2018; STEIN *et al.*, 2021; STOLL *et al.*, 2012), y en indicadores generales de ciertas características *input* tales como la cantidad (SCHNEIDMAN; GOLDIN-MEADOW, 2012; HUTTENLOCHER *et al.*, 2002, 2007; CASILLAS *et al.*, 2017, 2020) o el largo promedio de las emisiones (MLU) (HOFF, 2003; HOFF; TIAN, 2005) e índices de diversidad léxica (VOCD) (ROWE, 2008). Sin embargo, como se señaló, en los años preescolares el vocabulario infantil puede verse afectado por otros aspectos que configuran la composición del *input* lingüístico.

1.2 La composición del vocabulario infantil y su relación con el *input* lingüístico

Las investigaciones sobre aprendizaje de vocabulario se han centrado mayormente en sustantivos y verbos (TOMASELLO, 2003), en parte porque estas categorías léxicas transmiten la mayor parte de la información semántica y conceptual de los enunciados (CROFT, 2000; DIXON, 2004). Asimismo, se han preguntado acerca de los factores que condicionan el aprendizaje de estas categorías léxicas y que inciden en el orden en el que se aprenden unas y otras. La hipótesis de la partición natural de Gentner (1982) postula que la diferencia entre sustantivos y verbos está basada en predisposiciones que condicionan el establecimiento de distinciones conceptuales entre elementos concretos (personas, objetos), y conceptos predicativos referidos a actividades, cambios de estado o relaciones causales. Gentner y Boroditsky (2001), además, sostienen que las distintas categorías léxicas imponen distintas demandas a los niños, porque la comprensión de los conceptos que denotan es de distinta índole: pueden requerir solo una identificación cognitivo-perceptual (dominancia cognitiva), o bien requerir una fusión de fragmentos perceptuales cuya agrupación no está predeterminada (dominancia lingüística). Argumentan que la facilidad con la que las personas pueden individualizar un elemento predeciría la denominación temprana; ello brindaría una ventaja a los sustantivos, y en especial a los sustantivos concretos, por sobre los verbos.

El sesgo en favor de los sustantivos ha sido identificado en inglés (BORNSTEIN *et al.*, 2004; FENSON *et al.*, 1994; GOLDFIELD, 2000), en italiano (BORNSTEIN *et al.*, 2004; CASELLO *et al.*, 1995), en griego (KAMBANAROS *et al.*, 2014); en hebreo (BORNSTEIN *et al.*, 2004; MAITAL; DROMI; SAGI; BORNSTEIN, 2000), en español (BORNSTEIN *et al.*, 2004; JACKSON-MALDONADO *et al.*, 1993), en francés (BASSANO, 2000; BORNSTEIN *et al.* 2004), en alemán (GENTNER, 1982), en japonés (GENTNER, 1982; YAMASHITA, 1999) y en holandés (BORNSTEIN *et al.*, 2004; DE HOUWER; GILLIS, 1998). Sin embargo, investigaciones en lenguas menos estudiadas proporcionan cierta evidencia de que el sesgo no sería universal, sino que la predominancia de una determinada categoría léxica, sustantivo o verbo, dependería de las características tipológicas de la lengua en cuestión, así como de factores contextuales tales como las actividades o las interacciones predominantes en una comunidad o grupo determinado que

configuran el *input* lingüístico al que accede el niño (CHOI; GOPNIK, 1995; DE LEÓN, 2001; STOLL *et al.*, 2012; TARDIF, 1996).

Así, por ejemplo, Stoll *et al.* (2012) encontraron, en una población hablante de chintang, que la proporción de sustantivos era mayor que la de verbos en el habla de los niños y ello se invertía en el habla de los adultos que los rodeaban. Los autores atribuyen estas diferencias a las características tipológicas del chintang: la morfología verbal de esta lengua polisintética puede resultar particularmente compleja para los niños, quienes sólo progresivamente entre los 2;6 y los 4 años acercan su desempeño productivo al de los adultos. Asimismo, señalan que las diferencias en la cantidad de ambas clases de palabras podrían también explicarse por factores contextuales, ya que las actividades en las que participaban los niños y los adultos son diferentes.

Con la intención de corroborar si el sesgo hacia los sustantivos identificado en niños hablantes de hebreo era el resultado del *input* lingüístico, Adi-Bensaid *et al.* (2015) analizaron el habla dirigida al niño de diadas de niños con sus cuidadores y también el habla dirigida a adultos. En el habla dirigida a adultos encontraron una mayor cantidad de verbos, lo cual se diferencia del sesgo hacia sustantivos del habla infantil. Los autores explican la prevalencia de verbos por la sintaxis del hebreo que, al ser una lengua pro-drop, tiende a enfatizar esta clase de palabras y a requerir una menor cantidad de frases nominales (TARDIF *et al.*, 1997). Por otro lado, explican la presencia del sesgo hacia sustantivos reportado en la producción infantil y su ausencia en el habla de los adultos por la hipótesis de la partición natural (GENTNER, 1982). No encontraron diferencias entre sustantivos y verbos en el habla dirigida a niños, palabras que fueron significativamente más frecuentes que adjetivos y adverbios, probablemente porque los adultos tienden a ajustar su habla en las interacciones con los niños.

En el caso del español, la evidencia del sesgo de preferencia por los sustantivos no resulta concluyente. Si bien el sesgo ha sido registrado en varios trabajos (DHILLON, 2010; JACKSON-MALDONADO *et al.*, 1993; WEISLEDER; WAXMAN, 2010), tanto para el habla infantil como para el *input* adulto, Dhillon (2010) encuentra que en español la preferencia por sustantivos no es tan marcada como en inglés, pero sí respecto al mandarín. La adquisición de sustantivos en español puede resultar más fácil porque su morfología involucra menos variaciones que la de los verbos: mientras que los sustantivos pueden variar solo

en género y número, habilitando como máximo cuatro variaciones para cada lema, los verbos flexionan en persona, número, tiempo y modo, habilitando una cantidad de variaciones mayor para cada lema, además de la presencia de irregularidades tanto en la raíz como en la desinencia verbal (ROJAS-NIETO, 2003). Casart e Iribarren Pérez (2007) analizaron datos de tres díadas hispanohablantes cuando los niños tenían uno, dos y tres años de edad. Si bien en un principio los niños mostraron un claro sesgo hacia sustantivos a diferencia del *input*, que mostraba lo opuesto, a los dos años los niños comienzan a mostrar una preferencia por verbos similar a la de su entorno. A los tres años, si bien se evidencia también un sesgo hacia verbos, se observa una mayor variabilidad entre los hablantes, que las autoras atribuyen a los contextos de interacción de cada familia, que actuarían en interjuego con la tendencia madurativa identificada.

En un estudio en hogares de niños hablantes de español en Argentina de entre 8 y 14 meses, Rosemberg *et al.* (2020) encontraron que la proporción de sustantivos en relación con los verbos difería en función del grupo socioeconómico de la familia y de si se trataba de habla específicamente dirigida al niño o habla entre otros participantes. En efecto, en los hogares de sectores medios, a diferencia de los hogares que residían en circunstancias socioeconómicas vulneradas, los adultos tendían a emplear una mayor proporción de tipos distintos de sustantivos cuando les hablaban a los niños; en este grupo, la proporción de verbos era mayor en el habla entre otros participantes. En el estudio se señala que ello puede resultar de los contextos de interacción particulares que caracterizan los hogares de ambos grupos de niños. En los hogares de sectores medios, el habla dirigida al niño tiene lugar en mayor medida durante actividades sociales de lectura de cuentos o juego, en las que los adultos incluyen una mayor proporción de emisiones referenciales con sustantivos que nombran objetos. En los hogares en circunstancias socioeconómicas vulneradas, en cambio, los adultos regulan en mayor medida la acción de los niños por medio de directivas que incluyen verbos. La mayor cantidad de participantes que caracteriza la vida cotidiana en estos hogares puede contribuir a explicar la mayor proporción de sustantivos en el habla entre otros participantes. En efecto, un mayor número de participantes puede implicar la introducción de una mayor cantidad de tópicos en conversaciones que se solapan y, por lo tanto, una mayor cantidad de sustantivos. Cabe señalar que en ese trabajo no se analizó la relación entre la proporción de sustantivos y verbos en el *input* y en el habla infantil.

Si bien las investigaciones previas se han centrado principalmente en el aprendizaje de sustantivos y verbos, los niños comienzan a aprender adjetivos a partir de los dos años (AKHTAR; MONTAGUE, 1999), y los utilizan con distintas funciones (ver NELSON, 1976, para una clasificación de usos de adjetivos en inglés). Los adjetivos están, de cierta forma, relacionados con el desarrollo conceptual o semántico (NELSON, 1975), ya que le permiten al niño/a ampliar su vocabulario y referirse a conceptos que no pueden ser resumidos con un solo término. Sin embargo, estudios previos han indicado que recién a los cuatro años los niños son capaces de utilizar adjetivos de forma flexible (KLIBANOFF; WAXMAN, 2000), probablemente debido a su naturaleza relacional respecto del sustantivo al que modifica (GENTNER, 1982). En español, los adjetivos pueden ser utilizados para atribuir características a un elemento (*una casa **amplia***), para circunscribirlos a un determinado ámbito (*las deudas **fiscales***) o para incluir una modalización o referencia deíctica (*la **última** cena*). Estas diferentes funciones imponen restricciones a la posición del adjetivo respecto del sustantivo (DI TULLIO, 1997), lo cual podría afectar la frecuencia de aparición de estas construcciones en el *input*, impactando a su vez en la adquisición.

Una de las formas más frecuentes con las que los adultos utilizan adjetivos al dirigirse a los niños involucra el contraste (*esta remera es **amarilla** y esa otra remera es **blanca***), ya que le facilita al niño concentrarse en los aspectos de los elemento en los que difieren (DAVIES; LINGWOOD; ARUNACHALAM, 2020); Tribushinina *et al.* (2013) analizaron el patrón de uso de esta estrategia en hogares de familias con distintas lenguas y encontraron que su aparición en el habla adulta se replica en el habla infantil a los tres años. Sin embargo, Blackwell (2005) encuentra que este uso de adjetivos parece ser poco frecuente en el *input* de niños angloparlantes.

La baja frecuencia de uso de los adjetivos podría redundar en una baja exposición de los niños a las distintas estructuras sintácticas en las que esta clase de palabras puede aparecer, y así dificultar su adquisición. Davies *et al.* (2020) analizaron la presencia de adjetivos en el habla dirigida a niños de entre tres y cuatro años durante actividades interactivas en familias de distintos grupos socioeconómicos, y no encontraron diferencias en función del grupo social en cuanto al uso de esta clase de palabras.

1.3 El presente estudio

Este estudio se fundamenta en los trabajos previos que proporcionaron evidencia sobre la importancia del *input* lingüístico para el desarrollo del lenguaje infantil en los años preescolares (HUTTENLOCHER; HAIGHT; BRYK; SELTZER; LYONS, 1991; ROWE, 2012; WEISLEDER; FERNALD, 2013) y busca, en particular, dar cuenta de la relación entre la composición léxica -sustantivos, adjetivos y verbos- en el *input* y en el habla infantil en una población escasamente estudiada de niños argentinos de 4 años de edad hablantes de español cuyas familias residen en distintas circunstancias socioeconómicas.

Cabe señalar que en Argentina, al igual que en muchos otros países latinoamericanos, la fragmentación social es muy amplia, con una parte significativa de su población viviendo en condiciones habitacionales y educacionales que vulneran sus derechos. Datos recientes (Encuesta Permanente de Hogares sobre la Incidencia de la pobreza y la indigencia, INDEC, 2021) indican que 31,2% de los hogares, donde reside el 40,6% de las personas, se encuentran por debajo de la línea de la pobreza; debajo de la línea de la indigencia se encuentra el 8,2% de los hogares que abarca el 10,7% de las personas. Estos porcentajes se incrementan cuando se considera la población de 0 y 14 años: el 54,3% se encuentran en la pobreza (Encuesta Permanente de Hogares sobre la Incidencia de la pobreza y la indigencia, INDEC, 2021). A diferencia de lo que sucede en otras áreas de Argentina y de América, la población que reside en la Ciudad de la Buenos Aires y el cordón periurbano que la circunda es mayormente hablante de español, por lo que resulta posible estudiar el impacto de las diferencias socioeconómicas sin que se confunda con una situación de bilingüismo.

En tanto los estudios previos indican que la cantidad de habla específicamente dirigida al niño y aquella dirigida a otros participantes puede variar en función de la cultura y/o del grupo social (HOFF, 2003; ROWE, 2008) y que los niños preescolares pueden aprender palabras del lenguaje que escuchan no específicamente dirigida a ellos (AKHTAR *et al.*, 2001; FLOOR; AKHTAR, 2006), en este trabajo se analiza la relación entre la composición léxica en el habla infantil y aquella que caracteriza al conjunto del habla que configura el entorno lingüístico en los hogares de los niños. Específicamente nos preguntamos:

- a) ¿Existen diferencias entre la proporción de sustantivos, adjetivos y verbos que configuran el *input* lingüístico en los hogares? ¿Existen estas diferencias en la producción infantil?
- b) ¿Cuál es el efecto de la condición socioeconómica en el *input* lingüístico al que están expuestos los niños cotidianamente en sus hogares, en particular en la proporción de sustantivos, adjetivos y verbos, tanto en cantidades totales como en palabras distintas? ¿Cuál es el efecto de la condición socioeconómica y de las proporciones de sustantivos, adjetivos y verbos, que configuran el *input* lingüístico al que los niños están expuestos en los hogares, en los sustantivos, verbos y adjetivos que los niños producen, tanto en cantidades totales como en palabras distintas?

Teniendo en cuenta los resultados obtenidos en otras poblaciones y con niños más pequeños, esperamos encontrar una mayor proporción de sustantivos que de verbos, y una menor cantidad de adjetivos, tanto en cantidades totales como en palabras distintas. Además, cabe suponer que la circunstancia socioeconómica en la que vivan las familias será un predictor significativo de la proporción de cada categoría léxica en el *input* lingüístico en los hogares. A su vez, cabe esperar que tanto la circunstancia socioeconómica como la proporción de cada una de estas clases de palabras en el *input* que los niños escuchan predecirán la proporción de sustantivos, verbos y adjetivos que los niños producen.

2 Metodología

2.1 Recolección de datos

En este estudio se analizó un corpus de habla espontánea (ROSEMBERG; ARRÚE; ALAM, 2005-2012) que consiste en audiograbaciones de interacciones cotidianas en hogares de niños de cuatro años (4:1 - 4:11) pertenecientes a distintos grupos socioeconómicos, étnicamente caracterizados como no indígenas ni afrodescendientes, según las categorías informadas en el censo 2010 (INDEC; 2010). El grupo 1 estuvo conformado por familias que vivían en barrios residenciales de Buenos Aires y el Gran Buenos Aires. Al menos uno de los adultos del hogar (madre y/o padre) tenía educación universitaria completa, lo que

en Argentina implica como mínimo 16 años de escolaridad. El grupo 2 estaba compuesto por familias que vivían en comunidades urbano-marginadas con condiciones socioeconómicas vulneradas, como villas de emergencia, de Buenos Aires y del Gran Buenos Aires, y cuyos padres o madres habían alcanzado como máximo el nivel secundario completo, es decir, tenían no más de 12 años de escolaridad. En la tabla 1 se presenta la composición del *input* para cada grupo; se consideran adultos a partir de los 12 años.

Tabla 1 - Medias de participantes del input en las grabaciones y del núcleo familiar sin contar al niño foco

Grupo	Franja etaria	Composición del <i>input</i> de las grabaciones		Composición del núcleo familiar	
		Media	DE	Media	DE
Grupo 1	Adultos	4.84	2.79	2.26	0.56
	Niños	2.10	1.85	1.11	0.57
Grupo 2	Adultos	10.4	4.47	2.68	1.00
	Niños	9.16	6.81	2.00	1.16

La muestra fue no probabilística¹ o dirigida del tipo de participantes voluntarios (HERNÁNDEZ SAMPIERI; FERNÁNDEZ COLLADO; BAPTISTA LUCIO, 2014). Las familias pertenecientes al grupo 1 fueron contactadas a través de las redes de contactos; las familias del grupo 2 fueron contactadas a través de centros comunitarios y comedores en los cuales las investigadoras habían realizado actividades de promoción de la alfabetización familiar.

Observadores especialmente entrenados para la tarea acompañaron al niño o niña de 4 años durante 12 horas distribuidas en varias visitas al hogar, en distintos momentos del día. Se les solicitó a las familias que realizaran sus actividades cotidianas tal como lo hacían habitualmente;

¹ Las muestras no probabilísticas se seleccionan a partir de criterios relacionados con la investigación y no necesariamente generalizables. El foco no está en la representatividad de la muestra, sino en la precisión de los parámetros de inclusión (HERNÁNDEZ SAMPIERI; FERNÁNDEZ COLLADO; BAPTISTA LUCIO, 2014).

los observadores no iniciaron actividades ni conversaciones, pero sí respondían y participaban de las actividades cuando así se lo solicitaban.

Las audiograbaciones fueron transcritas en formato CHAT (Codes for the Human Analysis of Transcription (SOKOLOV; SNOW, 1994) para ser procesadas por el software CLAN (MACWHINNEY, 2000). De ese corpus, se tomaron 456 horas de grabación, correspondientes a 19 niños de cada grupo social (n = 38).

Retomando la definición de entorno lingüístico propuesta por Sperry *et al.* (2018), clasificamos como *input* todas las emisiones disponibles al niño, con excepción de los enunciados emitidos por el observador. Respecto a los ítems léxicos, solo se excluyeron onomatopeyas y nombres propios; en el caso de las perífrasis verbales se contabilizó solamente el verboide, que constituye el elemento que expresa el significado, frente al verbo auxiliar, que expresa la información gramatical (ROSEMBERG *et al.*, 2020).

2.2 Procedimientos de análisis

Se utilizó la función *freq* del CLAN² para calcular la cantidad de lexemas (*types*) y la cantidad palabras totales (*tokens*) producidas por: (a) los niños de cuatro años y (b) el resto de los participantes (el *input*), excluyendo a los observadores. Tanto para (a) como para (b) se contabilizaron las cantidades totales de palabras de contenido: la suma de sustantivos, adjetivos y verbos. Se calculó la proporción de cada clase de palabra respecto al total de las clases de palabra de contenido, para evitar un sesgo producto de la diferencia en cantidades totales de habla entre las observaciones. Estos cálculos se realizaron de forma separada para los tipos de palabras distintas (*types*) y para las cantidades de palabras (*tokens*). Se calculó un índice de cantidad de sustantivos por verbos, siguiendo a Stoll *et al.* (2012), estableciendo la cantidad de sustantivos en relación con la suma de las cantidades totales de sustantivos y verbos, tanto para *types* como para *tokens*.

En línea con el planteo de Stoll *et al.* (2012) y de Sperry *et al.* (2018), los datos que comprende el corpus no se restringen al habla dirigida al niño, sino que incluyen el conjunto del habla que circunda al

² El programa *freq* construye listas de frecuencias de palabras o instancias de ítems léxicos presentes en la transcripción, según la especificación indicada en cada caso.

niño y que está accesible en su entorno. Dos argumentos sostienen esta decisión: por un lado, como fue mencionado anteriormente, la literatura indica que a los cuatro años de edad los niños son capaces de prestar atención y aprender a partir de conversaciones que no lo involucren directamente; por el otro, los niños que se encuentran en circunstancias socioeconómicas vulneradas suelen vivir en comunidades conformadas por familias extendidas y por grupos de vecinos, y ello da lugar a que los niños interactúen con una gran cantidad de personas diariamente, tanto adultos como niños de todas las edades. Por lo tanto, no considerar el habla total que se encuentra disponible para el niño podría subestimar la riqueza de las interacciones de las cuales participa.

Se llevó a cabo un análisis descriptivo de las variables de la cantidad total (*tokens*) y de tipos distintos de palabras (*types*) totales y por clase de palabra en la producción infantil y en el *input* por grupo socioeconómico del hogar. Se realizó una transformación arcoseno³ para el trabajo con proporciones y luego se realizó un análisis de la varianza de una vía para identificar si las diferencias entre las proporciones de sustantivos, adjetivos y verbos eran estadísticamente significativas, tanto en la producción infantil como en el *input*, al interior de cada grupo social, en *types* y en *tokens*. De esta forma, se realizaron cuatro ANOVA de una vía (producción infantil de grupo 1, *input* de grupo 1, producción infantil de grupo 2, *input* de grupo 2) con la cantidad de *types* y *tokens* como variables de respuesta y su clase de palabra como variable de agrupación. Luego, se utilizó el test post-hoc de Tukey para identificar las diferencias entre las clases de palabras. Solamente en el caso de los *tokens* del grupo 1 se utilizó la prueba no paramétrica de Kruskal-Wallis por presentar una distribución no normal. Los resultados fueron analizados a la luz de la significatividad de los factores de agrupación de las variables (en este caso, las distintas clases de palabras), y la magnitud de ese efecto a partir del valor del *Eta-cuadrado* (η^2)⁴.

³ La transformación arcoseno es una combinación de la función arcoseno y de la raíz cuadrada cuya sintaxis en el lenguaje de programación R es `asin(sqrt(x))`. Esta función permite expandir la varianza y por este motivo es adecuada para el trabajo con proporciones.

⁴ El *Eta-cuadrado* está basado en criterios similares al estadístico *F* que aporta la significación, es decir, en la suma de cuadrados en los cuales se divide la variabilidad total de la variable dependiente. Toma valores entre 0 y 1: mientras mayor sea el valor del η^2 , mayor será el efecto de la variable independiente sobre la dependiente.

Para responder a la pregunta de si el *input* al que el niño está expuesto en su hogar varía en función de la pertenencia al grupo social, y para analizar si el *input* o el grupo social al que pertenece el niño impacta en su producción, se realizaron regresiones beta (FERRARI; CRIBARI-NETO, 2004), cuya distribución es óptima para el trabajo con proporciones⁵. Se controló por género del niño foco, ya que en estudios anteriores se documentó que las niñas suelen tener mejores puntajes que los niños en pruebas de vocabulario como el CDI (STOKES; KLEE, 2009), y que el en el *input* también se identifican diferencias según el género del niño (Casillas *et al.*, 2017).

3 Resultados

3.1 Una descripción del léxico (palabras de contenido) que los niños producen y al que están expuestos en su hogar

La tabla 2 resume las cantidades totales (*tokens*) y los tipos distintos (*types*) de las tres clases de palabras analizadas, emitidos tanto por los niños foco (CHI) como por el *input*. Se encontró que en ambos grupos las personas que conforman el *input* de los niños producen verbos de forma más frecuente que sustantivos y adjetivos. En la producción infantil, si bien los niños producen una mayor cantidad de verbos que de sustantivos, la diferencia entre ambos tipos de palabras es menor que en el *input*. En el caso de los tipos distintos de palabras, en ambos grupos los hablantes emiten una mayor cantidad de tipos distintos de sustantivos que de verbos.

⁵ Las regresiones beta son óptimas para modelar proporciones porque están acotadas entre 0 y 1. Utiliza la distribución beta, que permite adecuarse ampliamente a las distintas formas que pueda tomar la distribución original.

Tabla 2 - Estadísticos descriptivos. Medias de cantidades totales (*tokens*) y de tipos de palabras distintas (*types*) y del total de las palabras de contenido, por grupo, del niño foco y del *input*

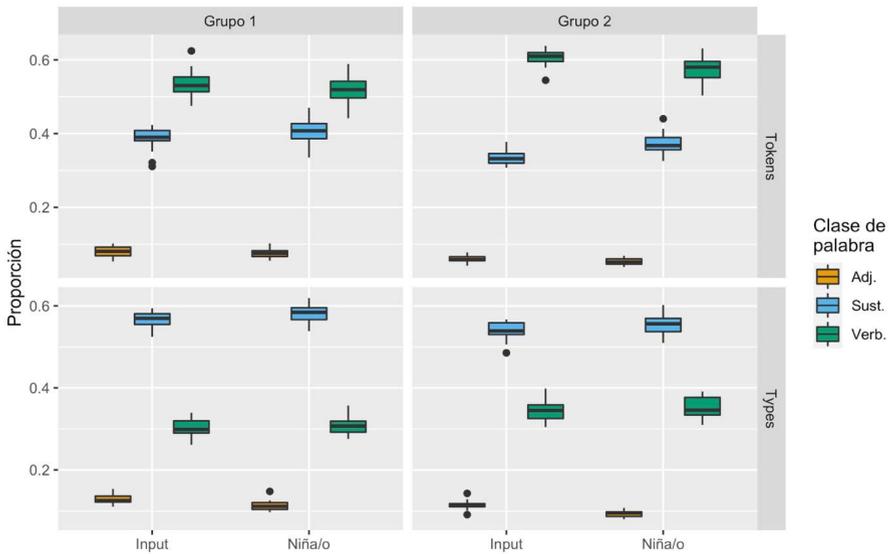
	Grupo	Tipos de palabras distintas (<i>types</i>)				Cantidades de palabras (<i>tokens</i>)			
		Sust.	Verb.	Adj.	Total	Sust.	Verb.	Adj.	Total
Niños foco	Grupo 1	384.5	200.3	74.5	659.32	1575.4	2021.7	297.1	3894.21
	Grupo 2	290.3	180.5	13.4	519.53	1329.6	2044.9	186.3	3560.68
<i>Input</i>	Grupo 1	534.9	279.0	124.2	938.11	2248.3	3075.3	482.6	5806.21
	Grupo 2	400.5	248.9	27.8	734.26	1543.4	2770.5	283.5	4597.42

Tabla 3 - Estadísticos descriptivos. Medias de porcentajes de cantidades (*tokens*) y tipos de palabras distintas (*types*), según categoría léxica, para niños foco e *input*

	Grupo	Tipos de palabras distintas (<i>types</i>)			Cantidades de palabras (<i>tokens</i>)		
		Sust.	Verb.	Adj.	Sust.	Verb.	Adj.
Niños foco	Grupo 1	58%	31%	11%	41%	52%	8%
	Grupo 2	56%	35%	9%	37%	58%	5%
<i>Input</i>	Grupo 1	57%	30%	13%	39%	53%	8%
	Grupo 2	54%	35%	11%	33%	61%	6%

El análisis de las proporciones, en la tabla 3 y en el gráfico 1, demostró que los niños producen una mayor proporción de tipos distintos de sustantivos (*types*), pero una mayor proporción de verbos totales (*tokens*); este patrón se repite también en el *input* lingüístico de ambos grupos. Los adjetivos, en todos los casos, representaron un porcentaje marcadamente menor.

Gráfico 1 - Proporción de sustantivos, adjetivos y verbos frente al total de palabras de contenido analizadas, en tipos distintos (*types*) y cantidades totales (*tokens*), producidas por los niños y por su entorno lingüístico, en ambos grupos



Se realizó un análisis de la varianza para cada grupo, tanto para el *input* como para la producción infantil, con el objetivo de identificar diferencias entre las proporciones de cada clase de palabra. Se encontraron diferencias estadísticamente significativas entre las clases de palabras en el *input* del grupo 1 tanto en *types* ($F(2,54)=2500.215$, $p < 0.01$) como en *tokens* ($H(2)=49.8$, $p < 0.01$); en ambos casos, el test post-hoc de Tukey mostró diferencias entre todas las clases de palabras ($p < 0.01$) y el tamaño del efecto fue grande ($\eta^2=0.99$ y $\eta^2=0.89$, respectivamente). En este mismo grupo social también se encontraron diferencias entre las proporciones de las tres clases de palabras en la producción infantil tanto en *types* ($F(2,54)=2539.39$, $p < 0.01$) como en *tokens* ($F(2, 54) = 1365.509$, $p < 0.01$); en ambos casos, el test post-hoc de Tukey mostró diferencias entre todas las clases de palabras ($p < 0.01$) y el tamaño del efecto fue grande ($\eta^2=0.99$ y $\eta^2=0.98$, respectivamente).

En cuanto al grupo 2, se encontraron diferencias estadísticamente significativas entre las proporciones de las tres clases de palabras en

el *input* tanto en *types* ($F(2,54)=2253.072$, $p < 0.01$) como en *tokens* ($F(2,54)=4722.74$, $p < 0.01$); en ambos casos, el test post-hoc de Tukey mostró diferencias entre todas las clases de palabras ($p < 0.01$) y el tamaño del efecto fue grande ($\eta^2=0.99$ y $\eta^2=0.99$, respectivamente). Finalmente, también se encontraron diferencias estadísticamente significativas en la producción infantil, tanto en *types* ($F(2,33)=3439$, $p < 0.01$, con el ajuste de Welch⁶) como en *tokens* ($F(2,54)=2492.931$, $p < 0.01$), con un tamaño de efecto grande ($\eta^2=0.99$); en ambos casos, el test post-hoc de Tukey mostró diferencias entre todas las clases de palabras ($p < 0.01$).

3.2 ¿Cuál es el impacto del grupo socioeconómico en los sustantivos, verbos y adjetivos a los que los niños están expuestos?

Se realizaron dos series de 4 regresiones beta cada una con el género del niño foco y la pertenencia al grupo socioeconómico como variables predictoras, y la proporción de cada clase de palabra (sustantivos, verbos y adjetivos) y la proporción de sustantivos en relación con los verbos como variables dependientes (*tokens* en una de las series y *types*, en la otra). En relación con los *tokens*, como muestra la tabla 4, la pertenencia al grupo socioeconómico resultó un predictor significativo de la media de proporciones en todos los casos: los niños del grupo 1 escuchan significativamente más sustantivos y adjetivos, y menos verbos. La proporción sustantivo-verbo también es más alta en el grupo 1 respecto del grupo 2. El género del niño foco también resultó ser un predictor significativo: mientras que los varones escucharon significativamente más *tokens* de verbos, las mujeres escucharon significativamente más *tokens* de sustantivos y la proporción de sustantivos en relación con los verbos también fue más elevada. Al analizar los resultados del modelo de precisión, se observa que hay una menor precisión -es decir, mayor variabilidad- en los *tokens* que forman parte del *input* del grupo 2, respecto del grupo 1.

⁶ El ajuste de Welch se utiliza en los casos en que no se cumple el supuesto de igualdad de varianzas y corrige el estadístico de contraste *F* del ANOVA de forma tal que considere esta disparidad. En este caso, las cantidades de *types* de sustantivos, adjetivos y verbos en el habla infantil del grupo 2 poseen varianzas desiguales, indicado por el test de Levene ($p < 0.05$).

Tabla 4 - Regresiones beta para estimar los *tokens* de sustantivos, verbos y adjetivos presentes en el *input*

Predictores	Variables dependientes – <i>Tokens</i> β (SE) / exp (β)			
	Sust.	Verbos	Adj.	Ratio
Coeficientes del modelo de medias con logit-link				
Intercepto	-0.65(-0.02)/0.34***	0.39(-0.02)/0.60***	-2.72(-0.04)/0.06***	-0.55(-0.02)/0.37***
Grupo (1)	0.23(-0.03)/0.57***	-0.30(-0.04)/0.43***	0.31(-0.06)/0.58***	0.28(-0.04)/0.57***
Género (varón)	-0.08(-0.03)/0.48**	0.10(-0.03)/0.52**	-0.08(-0.05)/0.48	-0.10(-0.03)/0.48**
Coeficientes del modelo de precisión con log-link				
Precisión: Intercepto	6.88(-0.39)/1.00***	6.48(-0.39)/1.00***	6.85(-0.39)/1.00***	6.63(-0.39)/1.00***
Precisión: Grupo (1)	-1.29(-0.46)/0.22**	-1.28(-0.46)/0.22**	-0.92(-0.46)/0.28*	-1.34(-0.46)/0.21**
Precisión: Género (varón)	-0.06(-0.46)/0.49	0.29(-0.46)/0.57	-0.17(-0.46)/0.46	0.12(-0.46)/0.53
Pseudo R ²	0.58	0.67	0.44	0.63
Log Likelihood	92.15	86.44	118.31	87.87
Num. obs.	38	38	38	38

***p < 0.001; **p < 0.01; *p ≤ 0.05

Tabla 5 - Regresiones beta para estimar los *types* de sustantivos, verbos y adjetivos presentes en el *input*

Predictores	Variables dependientes – <i>Types</i> β (SE) / exp (β)			
	Sust.	Verbos	Adj.	Ratio
Coeficientes del modelo de medias con logit-link				
Intercepto	0.18(-0.02)/0.55***	-0.66(-0.03)/0.34***	-2.04(-0.03)/0.16***	0.47(-0.03)/0.62***
Grupo (1)	0.10(-0.02)/0.53***	-0.19(-0.03)/0.45***	0.13(-0.03)/0.533***	0.18(-0.03)/0.54***
Género (varón)	-0.03(-0.02)/0.49	0.04(-0.03)/0.51	-0.01(-0.04)/0.50	-0.04(-0.03)/0.50
Coeficientes del modelo de precisión con log-link				
Precisión: Intercepto	5.99(-0.39)/1.00***	5.79(-0.39)/1.00***	7.16(-0.39)/1.00***	5.70(-0.39)/1.00***
Precisión: Grupo (1)	0.48(-0.46)/0.62	0.02(-0.46)/0.50	-0.37(-0.46)/0.41	0.08(-0.46)/0.52
Precisión: Género (varón)	0.91(-0.46)/0.71*	0.4(-0.46)/0.60	-0.52(-0.46)/0.37	0.59(-0.46)/0.64
Pseudo R ²	0.35	0.45	0.3	0.43
Log Likelihood	99.25	88.92	116.58	88.32
Num. obs.	38	38	38	38

*** p < 0.001; ** p < 0.01; * p ≤ 0.05

Al analizar los *types* que conforman el *input*, como se ve en la tabla 5, se encontró un patrón similar: en los hogares del grupo 1 los niños estaban expuestos a una mayor proporción de *types* de sustantivos (57%) y adjetivos (13%), y menos verbos (30%) que en los hogares del grupo 2 (54% sustantivos, 11% adjetivos, 35% verbos). Consecuentemente,

la proporción de sustantivo-verbo también fue más alta en el grupo 1. Al analizar el modelo de precisión, se observa que la variabilidad en la proporción de *types* de sustantivos que los varones escucharon fue mayor que en el caso de las mujeres.

3.3 ¿Cuál es el impacto del grupo socioeconómico y de los sustantivos, verbos y adjetivos presentes en el *input* lingüístico en la composición lexical de la producción infantil?

Se realizaron dos series de regresiones beta con el grupo, el género del niño y el *input* como variables predictoras, y la producción infantil de cada tipo de palabra (sustantivos, verbos y adjetivos) y la proporción de sustantivos en relación con los verbos (*tokens* en una de las series y *types* en la otra) como variables dependientes. La proporción de *tokens* en el *input*, como muestra la tabla 6, fue un predictor significativo en los cuatro casos: en ambos grupos, a mayor proporción de palabras presentes en el *input*, mayor proporción de palabras en el habla infantil. La pertenencia al grupo resultó un predictor estadísticamente significativo en el caso de los adjetivos: la pertenencia al grupo 1 aumentó significativamente la proporción de adjetivos en el habla infantil. Finalmente, el género no resultó un predictor significativo en ninguno de esos casos. Al analizar los resultados del modelo de precisión, se observa que la pertenencia al grupo 1 disminuyó significativamente la precisión de la proporción de cantidades totales (*tokens*) de verbos producidos por los niños; esto significa, en otras palabras, que las proporciones de *tokens* de verbos producidas por los niños pertenecientes al grupo 1 fueron significativamente más variables entre sí que aquellas de los niños del grupo 2.

Tabla 6 - Regresiones beta para estimar los *tokens* de sustantivos, verbos y adjetivos en la producción infantil

Predictores	Variables dependientes – <i>Tokens</i> β (SE) / exp (β)			
	Sust.	Verbos	Adj.	Ratio
Coeficientes del modelo de medias con logit-link				
Intercepto	-1.53 (0.29)/1.78***	-1.87 (0.39)/1.33***	-3.30 (0.13)/0.04***	-1.64 (0.28)/1.63***
Input	3.07 (0.83)/0.96***	3.57 (0.66)/0.98***	6.71 (1.90)/1.00***	3.42 (0.77)/0.97***
Grupo(1)	-0.03 (0.06)/0.49	0.05 (0.06)/0.51	0.25 (0.06)/0.56***	-0.05 (0.06)/0.49
Género (varón)	-0.02 (0.04)/0.50	0.01 (0.04)/0.50	0.01 (0.05)/0.50	-0.01 (0.04)/0.50
Coeficientes del modelo de precisión con log-link				
Precisión: Intercepto	7.56 (3.33)/1.00 *	11.65 (5.06)/1.00*	6.15 (1.30)/1.00***	5.69 (3.18)/1.00
Precisión: Input	-3.14 (9.67)/0.04	-8.51(8.47)/0.00	7.35 (19.70)/1.00	2.19 (8.65)/0.90
Precisión: Grupo (1)	-0.76 (0.69)/0.32	-1.53 (0.77)/0.18*	0.08 (0.60)/0.52	-1.13 (0.73)/0.24
Precisión: Género (varón)	-0.48 (0.49)/0.38	-0.39 (0.50)/0.40	-0.07 (0.47)/0.48	-0.50 (0.49)/0.38
Pseudo R ²	0.39	0.56	0.66	0.47
Log Likelihood	84.02	84.01	126.83	82.00
Num. obs.	38	38	38	38

***p < 0.001; **p < 0.01; *p < 0.05

Tabla 7 - Regresiones beta para estimar los *types* de sustantivos, verbos y adjetivos en la producción infantil

Predictores	Variables dependientes – <i>T</i> ypes β (SE) / exp (β)			
	Sust.	Verbos	Adj.	Ratio
Coeficientes del modelo de medias con logit-link				
Intercepto	-0.39 (0.51)/0.40	-1.22 (0.22)/0.23***	-2.79 (0.12)/0.06***	-0.45 (0.44)/0.39
Input	1.16 (0.93)/0.76	1.70 (0.65)/0.85**	5.00 (1.12)/1.00***	1.52 (0.70)/0.82*
Grupo(1)	0.07 (0.04)/0.52*	-0.13 (0.04)/0.47**	0.10 (0.03)/0.53***	0.11 (0.04)/0.53**
Género (varón)	-0.03 (0.03)/0.49	0.04 (0.03)/0.51	-0.06 (0.03)/0.48*	-0.04 (0.03)/0.93
Coeficientes del modelo de precisión con log-link				
Precisión: Intercepto	-6.57 (6.60)/0.00	12.41 (3.33)/1.00***	15.85 (2.29)/1.00***	-6.14 (5.80)/0.00
Precisión: Input	23.33 (12.10)/1.00	-18.05 (9.69)/0.00	-79.53 (19.73)/0.00***	19.90 (9.41)/1.00*
Precisión: Grupo (1)	-0.49 (0.56)/0.38	-0.60 (0.61)/0.35	0.89 (0.55)/0.71	-0.68 (0.60)/0.34
Precisión: Género (varón)	0.34 (0.47)/0.58	0.16 (0.47)/0.54	1.19 (0.46)/0.77**	0.21 (0.47)/0.55
Pseudo R ²	0.30	0.60	0.45	0.51
Log Likelihood	92.16	95.49	127.72	91.62
Num. obs.	38	38	38	38

***p < 0.001; **p < 0.01; *p ≤ 0.05

La segunda serie de regresiones beta tuvo a la proporción de *types* de cada una de las clases de palabras producidas por el niño como variable dependiente, y la proporción de estas clases de palabras en el *input* (sustantivos, verbos y adjetivos), el grupo social y el sexo del niño foco como variables predictoras; los resultados se muestran en la tabla 7. El grupo resultó un predictor significativo en algunos casos: mientras que la pertenencia al grupo 1 aumenta la proporción de *types* de adjetivos y marginalmente de sustantivos ($p = 0.0506$), la pertenencia al grupo 2 aumenta la proporción de *types* de verbos en la producción infantil. El género resultó un predictor marginalmente significativo solamente en el caso de los adjetivos: cuando el niño era varón, la proporción de esta clase de palabra disminuía. La proporción de *types* en el *input* resultó un predictor significativo que incrementaba la proporción de *types* en la producción infantil en todos los casos, excepto en los sustantivos.

Al analizar los resultados del modelo de precisión, se observa que, en el caso de los adjetivos, la proporción de *types* de adjetivos en el *input* fue un predictor significativo de la variabilidad de las proporciones de *types* de adjetivos en la producción infantil. Esto significa que a medida que aumenta la proporción de *types* de adjetivos en el *input*, se incrementa la variabilidad en las proporciones de *types* de adjetivos en la producción infantil.

4 Discusión

El objetivo de este trabajo fue estudiar la relación entre la composición del vocabulario productivo infantil -sustantivos, adjetivos y verbos- y el *input* lingüístico al cual están expuestos niños que crecen en hogares en distintas condiciones de vulnerabilidad socioeconómica.

Los resultados mostraron que, independientemente de la condición socioeconómica, tanto en el *input* lingüístico como en la producción infantil, se registró una mayor proporción de verbos que de sustantivos en las palabras totales (*tokens*), pero una mayor proporción de sustantivos que de verbos al tener en cuenta los lemas, es decir, en tipos de palabras distintas (*types*). Ello puede atribuirse a las características tipológicas del español que, al ser una lengua *pro-drop*, permite elidir las frases nominales con función de sujeto (ADI-BENSAID *et al.*, 2015; TARDIF *et al.*, 1997) y ello conduciría a que en sus emisiones los hablantes empleen más verbos que sustantivos. El hecho de que

la morfología verbal presente un sistema de inflexiones complejo de hasta 43 formas (ROJAS-NIETO, 2003) podría dar lugar a una baja probabilidad de aparición de un mismo verbo con la misma flexión, lo cual llevaría a que los niños aprendan y empleen menos tempranamente distintas formas de verbos que de sustantivos. Sin embargo, como esta discrepancia entre *types* y *tokens* -en promedio los hablantes de ambos grupos utilizaron más de diez veces de *tokens* que de *types* de verbos y solo cuatro veces más *tokens* que de *types* de sustantivos- se dio tanto en el habla infantil como en el *input* lingüístico, cabe suponer que este fenómeno resulta en mayor medida de una característica de la lengua, en lugar de un fenómeno ligado meramente a la adquisición, a pesar de que puede identificarse más marcadamente en la infancia.

Una interpretación alternativa o complementaria a la diferencia entre los *types* y los *tokens* de los sustantivos y verbos resulta de considerar el flujo de las actividades cotidianas registradas en las observaciones. Como lo mostraron estudios anteriores, determinadas actividades, tales como la lectura de cuentos, pueden dar lugar a patrones de uso lingüístico distintos en términos de *types* y *tokens* (MONTAG; JONES; SMITH, 2018; ROWE; WEISLEDER, 2020), o de proporción de sustantivos con relación a verbos (OGURA *et al.*, 2006; TARDIF *et al.*, 1999). Asimismo, Goldfield (1993) encuentra que durante juegos con objetos las madres de niños de un año utilizaban más sustantivos, mientras que durante juegos sin objetos utilizaban más verbos; Hoff-Ginsberg (1991), por otro lado, encuentra en el *input* lingüístico de niños entre 18 y 29 meses una mayor cantidad de directivas (que suelen incluir una mayor cantidad de verbos) durante juegos con juguetes, y la mayor diferencia entre grupos sociales en el uso de directivas se dio durante situaciones de comida; por último, Rosemberg *et al.* (2020) encuentran diferencias en función de las actividades en la proporción de sustantivos y verbos que ocurren en los hogares de distintas condiciones socioeconómicas. Ello indica, para próximos trabajos, la importancia de incorporar al análisis el tipo de actividad como fuente de variabilidad en el uso del lenguaje.

Los resultados también mostraron el impacto de las condiciones socioeconómicas en la composición léxica del *input* al cual los niños están expuestos en sus hogares: el efecto se identificó en todas las clases de palabras, tanto en *types* como en *tokens*. Este resultado coincide con los trabajos que han encontrado diferencias entre grupos en cuanto a la composición del vocabulario presente en el *input* (HOFF,

2013; HUTTENLOCHER *et al.*, 2002; ROWE, 2008; ROSEMBERG *et al.*, 2020). En este estudio, las condiciones socioeconómicas se operacionalizaron considerando dos indicadores: la educación materna y el lugar de residencia (en tanto subsume condiciones particulares, ligados a las condiciones de la vivienda, y la facilidad de acceso a centros educativos, recreacionales y culturales). Cabe tener en cuenta que el micro-contexto de las interacciones en el cual los niños desarrollan el lenguaje se ve afectado por el macro-contexto, producto del interjuego de múltiples dinámicas sociales, económicas y culturales, y que se materializan en cada país y comunidad de hablantes de forma única (ROSEMBERG, 2021; ROWE; WEISLEDER, 2020). Por lo tanto, dado que se ha identificado que, en esta población, las condiciones socioeconómicas constituyen un predictor de la composición del vocabulario en el *input* lingüístico en los hogares, resulta relevante estudiar el interjuego de aspectos del macro y micro-contexto que contribuyen a configurar estas diferencias. Uno de esos aspectos podría ser la cantidad de personas con las cuales interactúan los niños de los distintos grupos: las dinámicas familiares y sociocomunitarias son diversas y están relacionadas a factores tales como los tamaños de las viviendas, las redes de cuidado, la distribución de las áreas comunes dentro de los barrios y los niveles de autonomía que despliegan los niños en los espacios por donde circulan. Como se ve en la tabla 1, si bien la cantidad de personas que componen el núcleo familiar tiende a ser mayor en el grupo 2 que en el grupo 1, estas diferencias se amplían drásticamente al tener en cuenta las personas que no son parte de la familia pero que interactúan cotidianamente con los niños.

Además, se ha observado que los niños pertenecientes al grupo 1 estuvieron expuestos a una mayor variabilidad de proporciones de *tokens* de las tres clases de palabras. Esto evidenciaría una falta de homogeneidad al interior de este grupo. Si bien la heterogeneidad en el *input* es algo que ha sido señalado por diversos estudios (ALAM *et al.*, 2021), ello suele ser mencionado en relación con los hogares que viven en condiciones socioeconómicas vulnerabilizadas. En este estudio se aporta evidencia que indica que esta variabilidad también caracterizaría a los grupos de sectores medios.

Los niños que crecen en hogares en condiciones socioeconómicas vulnerabilizadas estuvieron expuestos a una mayor proporción de cantidades totales de verbos (*tokens*). Tanto los tipos particulares de actividades que caracterizan la vida cotidiana en estos hogares como

la necesidad de regular la acción de los participantes de la escena, especialmente los niños pequeños, pueden impactar en la disponibilidad de estas distintas clases de palabras en el entorno infantil. Es posible que tanto las condiciones de las viviendas como el hecho de que las familias que viven en condiciones socioeconómicas vulneradas suelen ser más extensas den lugar a que los niños y los adultos interactúen de forma compleja en un espacio pequeño. Esto podría dar lugar a la necesidad de un uso mayor de directivas y por lo tanto de verbos.

Los resultados del efecto de las condiciones socioeconómicas de los hogares y de la composición léxica del *input* en la composición léxica de la producción infantil de los niños de cuatro años en el flujo de la vida cotidiana son de relevancia en tanto proporcionan un panorama no abordado en los estudios previos, que se centraron en el *input* lingüístico en el habla dirigida al niño, a edades más tempranas y en muchos casos en situaciones cuasiexperimentales, o situaciones breves inducidas en los hogares (FERNALD *et al.*, 2013; HART; RISLEY, 1995; HOFF, 2003; ROWE, 2012; SCHNEIDMAN; GOLDIN-MEADOW, 2012). La proporción de cada una de las clases de palabras, sustantivos, verbos y adjetivos tanto en cantidades totales (*tokens*) como en tipos distintos (*types*), constituyó un predictor significativo en todos los casos -excepto en el caso de los *types* de sustantivos. En este sentido, los resultados aportan evidencia a los trabajos que enfatizan el hecho de que el vocabulario es en efecto un aspecto sensible al *input* y a la experiencia temprana del niño (HOFF; TIAN, 2005; LIEVEN, 2010).

De las tres clases de palabras, los sustantivos fueron la única cuyos *types* en el *input* no predijeron los *types* en el habla infantil. Cabe suponer que ello se deba, por una parte, a diferencias en las actividades de las que niños y adultos participan cotidianamente, que, como se ha mencionado, elicitan distintas clases de palabras. El eventual impacto del tipo de actividad y de los participantes involucrados será considerado en próximas investigaciones.

Las condiciones socioeconómicas de los hogares son un predictor significativo del habla infantil tanto de *types* como de *tokens* de adjetivos, pero solo de *types* de verbos y de sustantivos. El hecho de que las condiciones socioeconómicas constituyan un predictor significativo de los tipos distintos de palabras (*types*) al interior de cada una de estas categorías léxicas incluso controlando el efecto del *input*, podría estar relacionado con la diferencia en el acceso al sistema educativo tanto de

los niños como de los adultos con los que interactúan en los hogares y/o con desigualdades en el tipo de actividades y la disponibilidad de objetos en los hogares.

En relación con los adjetivos, dos aspectos llaman la atención: en primer lugar, su baja frecuencia en todas las condiciones de este estudio, tanto en los distintos hablantes como en los dos grupos; en segundo lugar, el hecho de que tanto en cantidades totales de palabras como en tipos distintos, los hablantes del grupo 1 produjeron casi un 50% más que los hablantes del grupo 2, la diferencia más marcada de todo el estudio. Esta situación se vuelve relevante considerando que recién a esta edad los niños comienzan a utilizar los adjetivos de forma flexible (KLIBANOFF; WAXMAN, 2003), por lo cual las diferencias contextuales en las oportunidades de aprendizaje de adjetivos podrían tener un impacto diferenciado al de los sustantivos o verbos. Weisleder y Waxman (2010) observaron que los marcos de aparición frecuentes para los adjetivos fueron menos precisos que los marcos para sustantivos y verbos, y Shlavik *et al.* (2020) encontraron una amplia variabilidad en distintas habilidades lingüísticas de aprendizaje de palabras que impactarían de forma distinta en los sustantivos y los adjetivos. Los resultados de nuestro estudio conducen a suponer que los adjetivos podrían tener una trayectoria de aprendizaje distinta que ameritaría un estudio específico.

Este estudio presenta, de todos modos, algunas limitaciones. Se trata de un corpus transversal que implicó la toma de datos en un momento dado. Sería productivo complementar estos datos con un estudio longitudinal, que permita observar cambios en el desarrollo, y con medidas directas de evaluación de la comprensión y producción de vocabulario. Por otro lado, y de modo muy importante, el análisis no incluyó el tipo de actividad que se estaba llevando a cabo en las situaciones registradas. Dado que los distintos tipos de actividades que realizan las familias en su vida cotidiana pueden dar lugar a usos del lenguaje distintos, así como a un vocabulario distinto, los próximos estudios tendrán en cuenta este análisis. Otro aspecto a tener en cuenta es la cantidad de interlocutores presentes en las situaciones: en tanto que cabe pensar que las diferencias reportadas en el *input* estuvieran relacionadas con diferencias en la cantidad de niños o adultos que lo componen.

Finalmente, consideramos relevante especificar que el hecho de que los niños de ambos grupos sociales utilizaran las categorías léxicas

con distinta frecuencia no significa que conozcan una mayor o menor cantidad de palabras, sino que las interacciones de las cuales participan brindan distintas oportunidades para producir (y escuchar) esas distintas clases de palabras. El conocimiento de las particularidades de los entornos lingüísticos que configuran la especificidad de las experiencias de alfabetización temprana de los niños de distintos grupos sociales constituye un insumo valioso para el diseño de políticas educativas que, atentas a la diversidad de realidades, eviten generalizaciones acerca del uso de lenguaje que puedan dificultar el acceso de grupos de niños a la escolarización.

5 Consideraciones finales

En resumen, los resultados de este trabajo aportan a la descripción de la complejidad de las experiencias lingüísticas infantiles. Se ha presentado evidencia a favor de las asociaciones entre la producción infantil y el entorno lingüístico del cual participa, tanto el *input* lingüístico que escucha, como las condiciones socioeconómicas en las que se encuentra su familia. Comprender en profundidad las experiencias lingüísticas tempranas es un aspecto clave del estudio de la alfabetización temprana y de la escolarización, y un eslabón fundamental para el diseño de intervenciones y materiales educativos.

Declaración de autorías

Macarena Quiroga: idea original, revisión y edición del corpus, construcción de la base de datos, análisis e interpretación de los resultados, escritura del artículo y edición general.

Celia Rosemberg: recolección de datos y conformación del corpus, dirección del diseño del artículo, interpretación y discusión de resultados, edición general.

Florencia Alam: recolección de datos y conformación del corpus, interpretación y discusión de resultados, edición general.

Agradecimientos

Agradecemos a las familias y a los niños por haber abierto las puertas de su hogar y por haber colaborado con generosidad.

Referencias

- ADI-BENSAID, L.; BEN-DAVID, A.; TUBUL-LAVY, G. Content words in Hebrew child-directed speech. *Infant Behavior and Development*, v. 40, p. 231-241, 2015. DOI: <https://doi.org/10.1016/j.infbeh.2015.06.012>
- AKHTAR, N. The robustness of learning through overhearing. *Developmental Science*, v. 8, n. 2, p. 199-209, 2005. DOI: <https://doi.org/10.1111/j.1467-7687.2005.00406.x>
- AKHTAR, N.; JIPSON, J.; CALLANAN, M. A. Learning words through overhearing. *Child development*, Michigan, v. 72, n. 2, p. 416-430, 2001.
- AKHTAR, N.; MONTAGUE, L. Early lexical acquisition: The role of cross-situational learning. *First Language*, London, v. 19, n. 57, p. 347-358, 1999. DOI: <https://doi.org/10.1177/014272379901905703>
- ALAM, F.; RAMÍREZ, M.L.; MIGDALEK, M. Other children's words in the linguistic environment of infants and young children from distinct social groups in Argentina. *Infancia y Aprendizaje*, v. 44, n. 2, p. 269-302, 2021.
- BASSANO, D. Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, Cambridge, v. 27, n. 3, p. 521-529, 2000. DOI: <https://doi.org/10.1017/S0305000900004396>
- BEALS, D. E. Sources of support for learning words in conversation: Evidence from mealtimes. *Journal of Child Language*, Cambridge, v. 24, n. 3, p. 673-694, 1997.
- BIEMILLER, A.; BOOTE, C. An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, [s.l.] v. 98, n. 1, p. 44-62, 2006. DOI: <https://doi.org/10.1037/0022-0663.98.1.44>
- BLACKWELL, A. Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology. *Journal of Child Language*, Cambridge, v. 32, n. 3, p. 535-562, 2005.
- BŁASZCZAK, J.; KLIMEK-JANKOWSKA, D. Noun and verb in the mind. An interdisciplinary approach. In: BŁASZCZAK, J.; KLIMEK-JANKOWSKA, D.; MIGDALSKI, K. (eds.) *How Categorical are Categories?* Berlin, München, Boston: Dr Gruyter Mouton, 2015, p. 75-112.

BORNSTEIN, M. H.; COTE, L. R.; MAITAL, S.; PAINTER, K.; PARK, S. Y.; PASCUAL, L.; ... ; VYT, A. Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child development*, Michigan, v. 75, n. 4, p. 1115-1139, 2004.

BORZONE, A.; ROSEMBERG, C. *¿Qué aprenden los niños cuando aprenden a hablar?*, Buenos Aires: Aique, 2000.

CASART QUINTERO, Y.; IRIBARREN PÉREZ, C. Proporción de sustantivos y verbos en el habla del cuidador y en el léxico temprano en Español. *Boletín de Lingüística*, Caracas, v. 19, n. 27, p. 42-69, 2007.

CASELLI, M. C.; BATES, E.; CASADIO, P.; FENSON, J.; FENSON, L.; SANDERL, L.; WEIR, J. A cross-linguistic study of early lexical development. *Cognitive Development*, [s. l.], v. 10, n. 2, p. 159-199, 1995.

CASILLAS, M.; AMATUNI, A.; SEIDL, A.; SODERSTROM, M.; WARLAUMONT, A. S.; BERGELSON, E. What do babies hear? Analyses of child-and adult-directed speech. *Interspeech 2017*, Stockholm, p. 2093-2097, 2017. DOI: <https://doi.org/10.21437/Interspeech.2017-1409>

CASILLAS, M.; BROWN, P.; LEVINSON, S. C. Early Language Experience in a Tzeltal Mayan Village. *Child Development*, Michigan, v. 91, n. 5, p. 1819-1835, 2020. DOI: <https://doi.org/10.1111/cdev.13349>

CHOI, S.; GOPNIK, A. Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, Cambridge, v. 22, n. 3, p. 497-529, 1995. DOI: <https://doi.org/10.1017/S0305000900009934>

CROFT, W. *Explaining language change: An evolutionary approach*. London: Pearson Education, 2000.

DAVIES, C.; LINGWOOD, J.; ARUNACHALAM, S. Adjective forms and functions in British English child-directed speech. *Journal of child language*, Cambridge, v. 47, n. 1, p. 159-185, 2020.

DE HOUWER, A.; GILLIS, S. *The acquisition of Dutch*. Amsterdam: John Benjamins, 1998.

DE LEÓN, L. Finding the richest path: language and cognition in the acquisition of verticality in Tzotzil (Mayan). In: BOWERMAN, M.; LEVINSON, S. (eds.). *Language Acquisition and Conceptual Development*. Cambridge: MIT-Press, 2001, p. 544-565. DOI: <https://doi.org/10.1017/cbo9780511620669.020>

- DHILLON, R. Examining the “noun bias”: A structural approach. *Proceedings of the 33rd Annual Penn Linguistics Colloquium*, Philadelphia, v. 16, n. 1, p. 51-60, 2010.
- DITULLIO, A. *Manual de gramática del español*. Buenos Aires: Edicial, 1997.
- DIXON, R. M. Adjective classes in typological perspective. In: DIXON, W.M.; AIKHENVALD, A. (eds.). *Adjective classes: A cross-linguistic typology*. Oxford: Oxford University Press, 2004. p. 1-49.
- FENSON, L.; DALE, P. S.; REZNICK, J. S.; BATES, E.; THAL, D. J.; PETHICK, S. J. Variability in early communicative development. *Monographs of the Society for Research in Child Development*, Chicago, v. 59, p. 1-185, 1994.
- FERNALD, A.; MARCHMAN, V. A.; WEISLEDER, A. SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, v. 16, n. 2, p. 234–248, 2013.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, v. 31, n. 7, p. 799-815, 2004. DOI: <https://doi.org/10.1080/0266476042000214501>
- FLOOR, P.; AKHTAR, N. Can 18-month-old infants learn words by listening in on conversations? *Infancy*, v. 9, n. 3, p. 327-339, 2006.
- FOUSHEE, R.; SRINIVASAN, M.; XU, F. Self-directed learning by preschoolers in a naturalistic overhearing context. *Cognition*, Amsterdam, v. 206, p. 1-18, 2021. DOI: <https://doi.org/10.1016/j.cognition.2020.104415>
- FRIEND, M.; DEANDA, S.; ARIAS-TREJO, N.; POULIN-DUBOIS, D.; ZESIGER, P. Developmental changes in maternal education and minimal exposure effects on vocabulary in English-and Spanish-learning toddlers. *Journal of experimental child psychology*, v. 164, p. 250-259, 2017.
- GENTNER, D. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In: KUCZAJ, S.A. (ed.). *Language Development: Vol. 2. Language, Thought, and Culture*. Hillsdale: Erlbaum, 1982. p. 301-334.
- GENTNER, D.; BORODITSKY, L. Individuation, relativity, and early word learning. In: BOWERMAN, M.; LEVINSON, S. (eds.). *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 2001. p. 215-256. <https://doi.org/10.1017/cbo9780511620669.010>

GOLDFIELD, B. A. Noun bias in maternal speech to one-year-olds. *Journal of child language*, Cambridge, v. 20, n. 1, p. 85-99, 1993.

GOLDFIELD, B. A. Nouns before verbs in comprehension vs. production: the view from pragmatics. *Journal of Child Language*, Cambridge, v. 27, n. 3, p. 501-520, 2000.

HART, B.; RISLEY, T. R. *Meaningful differences in the everyday experience of young American children*. Baltimore-London-Sidney: Paul H Brookes Publishing, 1995.

HERNÁNDEZ SAMPIERI, R.; FERNÁNDEZ COLLADO, C.; BAPTISTA LUCIO, P. *Metodología de la investigación*. México D.F.: McGraw-Hill, 2014.

HOFF-GINSBERG, E. Mother-child conversation in different social classes and communicative settings. *Child development*, Michigan, v. 62, n. 4, p. 782-796, 1991.

HOFF, E. The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, Michigan, v. 74, n. 5, p. 1386–1378, 2003. DOI: <https://doi.org/10.1111/1467-8624.00612>

HOFF, E. Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, Washington, v. 49, n. 1, p. 4–14, 2013. DOI: <https://doi.org/10.1037/a0027238>

HOFF, E.; TIAN, C. Socioeconomic status and cultural influences on language. *Journal of Communication Disorders*, v. 30, p. 271–278, 2005. DOI: <https://doi.org/10.1016/j.jcomdis.2005.02.003>

HUTTENLOCHER, J.; HAIGHT, W.; BRYK, A.; SELTZER, M.; LYONS, T. Early vocabulary growth: relation to language input and gender. *Developmental psychology*, Washington, v. 27, n. 2, p. 236-248, 1991.

HUTTENLOCHER, J.; VASILYEVA, M.; CYMERMAN, E.; LEVINE, S. Language input and child syntax. *Cognitive Psychology*, Cambridge, v. 45, n. 3, p. 337-274, 2002. DOI: [https://doi.org/10.1016/S0010-0285\(02\)00500-5](https://doi.org/10.1016/S0010-0285(02)00500-5)

HUTTENLOCHER, J.; VASILYEVA, M.; WATERFALL, H. R.; VEVEA, J. L.; HEDGES, L. V. The Varieties of Speech to Young Children. *Developmental Psychology*, Washington, v. 43, n. 5, p. 1062-1083, 2007. DOI: <https://doi.org/10.1037/0012-1649.43.5.1062>

INDEC. Encuesta Permanente de Hogares sobre la Incidencia de la pobreza y la indigencia en 31 aglomerados urbanos. Primer semestre de 2021. *Condiciones de Vida*, v. 5, n. 13, p. 1-19, 2021.

JACKSON-MALDONADO, D.; THAL, D.; MARCHMAN, V.; BATES, E.; GUTIERREZ-CLELLEN, V. Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language*, Cambridge, v. 20, n. 3, p. 523-549, 1993. DOI: <https://doi.org/10.1017/S0305000900008461>

KAMBANAROS, M.; GROHMANN, K.; MICHAELIDES, M.; THEODOROU, E. On the nature of verb–noun dissociations in bilingual SLI: A psycholinguistic perspective from Greek. *Bilingualism: Language and Cognition*, Cambridge, v. 17, n. 1, p. 169-188, 2014.

KLIBANOFF, R. S.; WAXMAN, S. R. Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child development*, Michigan, v. 71, n. 3, p. 649-659, 2000.

LIEVEN, E. Input and first language acquisition: Evaluating the role of frequency. *Lingua*, v. 120, n. 11, p. 2546–2556, 2010. DOI: <https://doi.org/10.1016/j.lingua.2010.06.005>

MACWHINNEY, B. *The competition model. Mechanisms of language acquisition*, New York: Routledge, 1987.

MACWHINNEY, B. *The CHILDES Project: Tools for analyzing talk*. London: Psychology Press, 2000.

MAITAL, S. L.; DROMI, E.; SAGI, A.; BORNSTEIN, M. H. The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, Cambridge, v. 27, n. 1, p. 43-67, 2000.

MONTAG, J. L.; JONES, M. N.; SMITH, L. B. Quantity and diversity: Simulating early word learning environments. *Cognitive science*, v. 42, n. 2, p. 375-412, 2018.

NAIGLES, L. R.; HOFF-GINSBERG, E. Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, Cambridge, v. 25, n. 1, p. 95-120, 1998. DOI: <https://doi.org/10.1017/S0305000997003358>

NELSON, K. The nominal shift in semantic-syntactic development. *Cognitive Psychology*, Cambridge, v. 7, n. 4, p. 461-479, 1975. DOI: [https://doi.org/10.1016/0010-0285\(75\)90018-3](https://doi.org/10.1016/0010-0285(75)90018-3)

NELSON, K. Some attributes of adjectives used by young children. *Cognition*, v. 4, n. 1, p. 13-30, 1976. DOI: [https://doi.org/10.1016/0010-0277\(76\)90008-1](https://doi.org/10.1016/0010-0277(76)90008-1)

NELSON, K. *Making sense: The acquisition of shared meaning*. New York: Academic. 1975.

NELSON, K. *Language in Cognitive Development*. Cambridge: Cambridge University Press. 1996. DOI: <https://doi.org/10.1017/cbo9781139174619>

OGURA, T.; DALE, P. S.; YAMASHITA, Y.; MURASE, T.; MAHIEU, A. The use of nouns and verbs by Japanese children and their caregivers in book-reading and toy-playing contexts. *Journal of Child Language*, Cambridge, v. 33, n. 1, p. 1-29, 2006.

PACE, A.; LUO, R.; HIRSH-PASEK, K.; GOLINKOFF, R. M. Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, v. 3, p. 285-308, 2007.

PERFETTI, C. Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, Philadelphia, v. 11, n. 4, p. 357-383, 2007. <https://doi.org/10.1080/10888430701530730>

ROJAS-NIETO, C. Early acquisition of Spanish verb inflexion: A usage-based account. *Psychology of Language and Communication*, Warsaw, v. 7, n. 2, p. 17-36, 2003.

ROSEMBERG, C.R. Hogar Dulce hogar: oportunidades para el desarrollo del vocabulario y la alfabetización temprana en el contexto lingüístico del hogar. In: ABUSAMRA, V.; CHIMENTI, A.; TISCORNIA, S. (orgs.). *La ciencia de la lectura*. Buenos Aires: Tilde Editorial, 2021. p. 85-105.

ROSEMBERG, C.R.; ALAM, F. Socioeconomic disparities in the comprehension of lexical categories. A study with Spanish-speaking Argentinian toddlers. *European Journal of Psychology of Education*, v. 36, n. 4, 2021. DOI: <https://doi.org/10.1007/s10212-020-00522-0>

ROSEMBERG, C.R.; ALAM, F.; AUDISIO, C.P.; RAMÍREZ, M.L.; GARBER, L.; MIGDALEK, M.J. Nouns and verbs in the linguistic environment of Argentinian toddlers: socioeconomic and context-related differences. *First Language*, London, v. 40, n. 2, p. 192-217, 2020. DOI: <https://doi.org/10.1177/0142723719901226>

ROSEMBERG, C.R.; ARRÚE, J.; ALAM, F. Corpus: Home language environments of 4-year-old Argentinean children from different socio-cultural groups. CONICET (doi en trámite)

ROSEMBERG, C.R.; MENTI, A.; STEIN, A.; ALAM, F.; MIGDALEK, M. Vocabulario, narración y argumentación en los primeros años de la infancia y la niñez. Una revisión de investigaciones. *Revista Costarricense de Psicología*, San José, v. 35, n. 2, p. 139-158, 2016.

ROSEMBERG, C.R.; SILVA, M.L. Teacher-children interaction and concept development in kindergarten. *Discourse processes*, v. 46, n. 6, p. 572-591, 2009.

ROSEMBERG, C.R.; SILVA, M.L.; STEIN, A. Narrativas infantiles en contexto: un estudio en hogares urbano marginados de Buenos Aires. *Revista del Instituto de Ciencias de la Educación*, Buenos Aires, n. 28, p.135-154, 2011.

ROWE, M. L. Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, Cambridge, v. 35, n. 1, p. 185–205, 2008. DOI: <https://doi.org/10.1017/S0305000907008343>

ROWE, M. L. A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, Michigan, v. 83, n. 5, p. 1762-1774, 2012. DOI: <https://doi.org/10.1111/j.1467-8624.2012.01805.x>

ROWE, M. L.; WEISLEDER, A. Language development in context. *Annual Review of Developmental Psychology*, v. 2, p. 201-223, 2020.

SÉNÉCHAL, M.; OUELLETTE, G.; RODNEY, D. The misunderstood giant: On the predictive role of early vocabulary to future reading. In: Dickinson, D.; Neuman, S. (eds.). *Handbook of early literacy research*. New York: Guilford Press, 2006. p. 173-182.

SCHWAB, J. F.; LEW-WILLIAMS, C. Language learning, socioeconomic status, and child-directed speech. *Cognitive Science*, v. 7, n. 4, p. 264-275, 2016.

SHAVLIK, M.; DAVIS-KEAN, P. E.; SCHWAB, J. F.; BOOTH, A. E. Early word-learning skills: A missing link in understanding the vocabulary gap? *Developmental Science*, v. 24, n. 2, e13034, 2021. DOI: <https://doi.org/10.1111/desc.13034>

SHNEIDMAN, L. A.; ARROYO, M. E.; LEVINE, S. C.; GOLDIN-MEADOW, S. What counts as effective input for word learning? *Journal of Child Language*, Cambridge, v. 40, n. 3, p. 672-686, 2013. DOI: <https://doi.org/10.1017/S0305000912000141>

SHNEIDMAN, L. A.; GOLDIN-MEADOW, S. Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, v. 15, n. 5, p. 659-673, 2012. DOI: <https://doi.org/10.1111/j.1467-7687.2012.01168.x>

SNOW, C. E.; MATTHEWS, T. J. Reading and language in the early grades. *Future of Children*, Princeton, v. 26, n. 2, p. 57-74, 2016. DOI: <https://doi.org/10.1353/foc.2016.0012>

SOKOLOV, J.L.; SNOW, C.E. (eds). *Handbook of research in language development using CHILDES*. Hillsdale: Lawrence Erlbaum, 1994.

SPERRY, D. E.; SPERRY, L. L.; MILLER, P. J. Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*, Michigan, v. 90, n. 4, p. 1303-1318, 2018. DOI: <https://doi.org/10.1111/cdev.13072>

STEIN, A.; MENTI, A.; ROSEMBERG, C.R. Socioeconomic status differences in the linguistic environment: a study with Spanish-speaking populations in Argentina. *Early Years*, p. 1-15, 2021. DOI: <https://doi.org/10.1080/09575146.2021.1904383>

STOKES, S. F.; KLEE, T. Factors that influence vocabulary development in two-year-old children. *Journal of Child Psychology and Psychiatry*, v. 50, n. 4, p. 498-505, 2009.

STOLL, S.; BICKEL, B.; LIEVEN, E.; PAUDYAL, N. P.; BANJADE, G.; BHATTA, T. N.; GAENSZLE, M.; PETTIGREW, J.; RAI, I. P.; RAI, M.; RAI, N. K. Nouns and verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language*, Cambridge, v. 39, n. 2, p. 284-321, 2012. DOI: <https://doi.org/10.1017/S0305000911000080>

TARDIF, T. Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, Washington, v. 32, n. 3, p. 492-504, 1996. DOI: <https://doi.org/10.1037/0012-1649.32.3.492>

TARDIF, T.; SHATZ, M.; NAIGLES, L. Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, Cambridge, v. 24, n. 3, p. 535-565, 1997. DOI: <https://doi.org/10.1017/S030500099700319X>

TARDIF, T.; GELMAN, S. A.; XU, F. Putting the “noun bias” in context: A comparison of English and Mandarin. *Child development*, Michigan, v. 70, n. 3, p. 620-635, 1999.

TOMASELLO, M. Things are what they do: Katherine Nelson’s functional approach to language and cognition. *Journal of Cognition and Development*, v. 3, n. 1, p. 5-19, 2002. DOI: https://doi.org/10.1207/S15327647JCD0301_2

TOMASELLO, M. *Constructing a language*. Cambridge: Harvard University Press, 2003.

TRIBUSHININA, E.; VAN DEN BERGH, H.; KILANI-SCHOCH, M.; AKSU-KOÇ, A.; DABAŠINSKIENĖ, I.; HRZICA, G.; ...; DRESSLER, W. The role of explicit contrast in adjective acquisition: A cross-linguistic longitudinal study of adjective production in spontaneous child speech and parental input. *First Language*, London, v. 33, n. 6, p. 594-616, 2013.

VAN DRUTEN-FRIETMAN, L.; DENESSEN, E.; GIJSEL, M.; VERHOEVEN, L. Child, home and institutional predictors of preschool vocabulary growth. *Learning and Individual Differences*, v. 43, p. 92-99, 2015. DOI: <https://doi.org/10.1016/j.lindif.2015.08.032>

WACHS, T. D. Environmental considerations in studies with nonextreme groups. In: WACHS, T.D.; PLOMIN, R. (eds.). *Conceptualization and measurement of organism-environment interaction*. Washington: American Psychological Association. p. 44–67, 1991. DOI: <https://doi.org/10.1037/10100-003>

WEISLEDER, A.; FERNALD, A. Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological science*, New York, v. 24, n. 11, p. 2143-2152, 2013.

WEISLEDER, A.; WAXMAN, S. R. What’s in the input? Frequent frames in child-directed speech offer distributional cues to grammatical categories in Spanish and English. *Journal of Child Language*, Cambridge, v. 37, n. 5, p. 1089-1108, 2010. DOI: <https://doi.org/10.1017/S0305000909990067>

YAMASHITA, Y. The acquisition of nouns and verbs in young Japanese children: why do verbal nouns emerge early. *Proceedings of the 23rd Annual Boston University Conference on Language Development*, p. 741-752, 1999.