



Canada

**FORGE**

FORTALECIMIENTO  
DE LA GESTIÓN DE LA  
EDUCACIÓN EN EL PERÚ

 **GRADE**



**CIES**  
consorcio de investigación  
económica y social

## INFORME FINAL

### ***Identificación de predictores del desempeño en la comprensión de textos orales y escritos.***

Andrea Anahí Junyent Moreno

INSTITUTO DE ESTUDIOS PERUANOS

Setiembre de 2015

Este documento fue producido gracias al apoyo del Proyecto Fortalecimiento de la Gestión de la Educación en el Perú - FORGE que es implementado por el Grupo de Análisis para el Desarrollo- GRADE, con fondos otorgados por el gobierno de Canadá a través de su Ministerio de Asuntos Exteriores, Comercio y Desarrollo. (Proyecto No. A034597)

## ÍNDICE

<b>RESEÑA</b>	4
<b>ABSTRACT</b>	4
<b>INTRODUCCIÓN</b>	5
<b>EL PROBLEMA DE ESTUDIO</b>	7
1. El escenario actual: los datos de las pruebas PISA, TERCE y ECE.	7
2. Política pública a favor de la mejora en comprensión textual	12
<b>EL MARCO TEÓRICO</b>	16
1. El desarrollo lingüístico	16
1.1. La comprensión de textos	17
1.2. El vocabulario	18
2. Aspectos que influyen el desarrollo de la comprensión textual	19
2.1. El desarrollo del niño	19
2.1.1. La edad	20
2.1.2. El desarrollo cognitivo	20
2.1.3. El desarrollo físico	20
<i>El crecimiento</i>	21
<i>Las discapacidades</i>	22
2.2. El sexo y el género	22
2.2.1. Relación entre desarrollo lingüístico y sexo	23
2.2.2. Relación entre desarrollo lingüístico y género	25
2.2.3. Relación entre comprensión lectora y género en el Perú	25
2.3. Las oportunidades del niño	27
2.3.1. Las prácticas de literacidad	28
2.3.2. Estatus socioeconómico	29
2.3.3. Educación	30
<b>LA METODOLOGÍA</b>	32
1. La base de datos	32
2. Los participantes	34
3. Las variables y el diseño	34
3.1. Las variables	35
3.1.1. Género	35
3.1.2. Variables relacionadas con el desarrollo del niño	35
<i>“Edad”</i>	35
<i>“Crecimiento”</i>	35
<i>“Vocabulario”</i>	36
<i>“Texto oral” y “texto escrito”</i>	36
3.1.3. Variables relacionadas con las oportunidades del niño	37
<i>“Español oral de la madre”</i>	37
<i>“Español escrito de la madre”</i>	37
<i>“Educación de la madre”</i>	38
<i>“Años de educación”</i>	38
<i>“Tipo de institución educativa”</i>	38
<i>“Estatus socioeconómico”</i>	39
<i>“Lugar de residencia”</i>	39
3.2. El diseño	39

4. Los análisis	42
4.1. Estadísticos Descriptivos	42
4.2. Correlaciones	42
4.3. Comparación de medias	43
4.4. Modelo de ecuación estructural (o SEM)	43
<b>LOS RESULTADOS</b>	<b>45</b>
1. Estadísticos Descriptivos	45
1.1. Lugar de residencia y tipo de institución educativa.	45
1.2. Español oral y escrito de la madre, y su nivel de educación.	45
1.3. Desarrollo cognitivo, crecimiento, educación, estatus socioeconómico, vocabulario texto oral y escrito.	46
2. Correlaciones	46
3. Comparación de medias	46
4. Modelo de ecuaciones estructurales	47
4.1. Regresiones para “vocabulario”	47
4.2. Regresiones para “texto oral”	48
4.3. Regresiones para “texto escrito”	48
4.4. Modelo completo	51
<b>DISCUSIÓN</b>	<b>53</b>
1. Comparación entre grupos niño-niña, urbano-rural, institución pública-privada con respecto al vocabulario y la comprensión de textos orales y escritos	53
2. El vocabulario y sus variables predictoras	54
3. El texto oral y sus variables predictoras	56
4. El texto escrito y sus variables predictoras	57
<b>CONCLUSIONES Y RECOMENDACIONES DE POLÍTICA</b>	<b>60</b>
1. Identificación de grupos de riesgo	60
2. Propuestas de intervención	61
2.1. Intervención a través de los padres o cuidadores	62
2.2. Intervención a través de los docentes de Educación Inicial	63
<b>PLAN DE INCIDENCIA</b>	<b>65</b>
<b>BIBLIOGRAFÍA</b>	<b>69</b>
<b>ANEXO 1</b>	<b>88</b>
<b>ANEXO 2</b>	<b>89</b>
<b>ANEXO 3</b>	<b>91</b>
<b>ANEXO 4</b>	<b>92</b>
<b>ANEXO 5</b>	<b>93</b>
<b>ANEXO 6</b>	<b>94</b>
<b>ANEXO 7</b>	<b>95</b>
<b>ANEXO 8</b>	<b>96</b>
<b>ANEXO 9</b>	<b>97</b>
<b>ANEXO 10</b>	<b>98</b>

## **RESEÑA**

Este estudio buscó identificar los predictores de la comprensión lectora para promover políticas públicas a favor de intervenciones tempranas, utilizando la base de datos Niños del Milenio.

Se evaluaron las relaciones entre las características de los niños y sus oportunidades; y el desarrollo lingüístico. Se encontró que (1) las características de la madre, el desarrollo del niño, el lugar de residencia, y el estatus socioeconómico predijeron el vocabulario a los 5 años; (2) el vocabulario a los 5 y la comprensión de textos orales predijeron la comprensión lectora a los 7; (3) los niños rurales componen el grupo a riesgo de desarrollar un vocabulario insuficiente. Proponemos entrenar a cuidadores y docentes de Educación Inicial en prácticas de literacidad tempranas.

## **ABSTRACT**

The aim of the study was to identify predictors of reading comprehension in order to promote early interventions. We examined data from Young Lives study, using structural equation modeling, for assessing relationships between child and environmental features, and language development. Main findings were that (1) mothers' features – which we relate to means to employ literacy practices -; child development, place of residence and socio-economic status predicted vocabulary at five; (2) vocabulary at five, and later oral text comprehension predicted reading comprehension at eight; and (3) rural children are at risk of insufficient vocabulary. Based on these results, interventions are proposed to improve vocabulary, and therefore, later reading comprehension, through literacy practices employed by caregivers and preschool teachers.

## INTRODUCCIÓN

En términos generales, este estudio busca identificar los predictores de la variabilidad en la habilidad de comprensión textual de los niños, utilizando la base de datos Niños del Milenio, con la finalidad de promover políticas públicas a favor de estrategias de intervención temprana para mejorar dicha habilidad.

La comprensión de textos es una competencia imprescindible para el desarrollo personal y social de los niños, así como para su futuro éxito académico y laboral. La situación de los estudiantes peruanos con respecto a la adquisición de esta habilidad es conocida: por ejemplo, en las renombradas pruebas de lectura PISA 2012, el Perú quedó en el último lugar de los 65 países participantes. El Estado afronta el desafío de revertir esta situación y mejorar la capacidad lectora de los niños y adolescentes peruanos. La principal dificultad consiste en cómo y cuándo intervenir, pues el bajo rendimiento en comprensión de lectura se hace evidente ya desde la primaria: según la Evaluación Censal de Estudiantes 2014, el 56,5% de los niños de 2º grado no alcanza los logros esperados.

La comprensión de textos es una habilidad que se basa en el desarrollo de varios dominios de la cognición y el lenguaje, que inicia desde el nacimiento. Así, cuando un niño empieza la educación formal, ya están sentadas las bases para la adquisición de la competencia textual. Por ello, es posible indagar sobre los predictores de la competencia textual que emergen en etapas anteriores, por ejemplo, durante la etapa preescolar. Identificar los predictores tempranos permitirá diseñar estrategias de intervención que verán sus frutos años después y durante todo el período de educación formal.

Nuestro estudio tiene como objetivo principal identificar los predictores de la comprensión de textos escritos en niños de 7 y 8 años entre variables relacionadas a las características de los niños – incluidas la amplitud de su vocabulario y su capacidad de entender textos orales – y las oportunidades que le ofrece su entorno. Inicialmente, partimos con la hipótesis de que el vocabulario y la comprensión de textos orales predecirían la comprensión de textos escritos. Esta hipótesis fue

confirmada por nuestros resultados. Además, elaboramos un modelo de relaciones entre las características del niño (género; desarrollo cognitivo y lingüístico) y las oportunidades que tuvo (relacionadas a su educación; las características de su familia y su ambiente); y lo evaluamos – a través de un análisis de ecuaciones estructurales – para identificar el poder de predicción de las variables independientes y mediadoras. Como hallazgos principales tuvimos que (1) las características de la madre – las cuales relacionamos con la posibilidad que esta tiene de emplear prácticas de literacidad con su hijo –; así como el desarrollo del niño; el lugar de residencia y el estatus socioeconómico de su familia predijeron el desempeño de los niños de 4-5 años en las pruebas de vocabulario; (2) la amplitud del vocabulario a los 4-5 años y la posterior habilidad de comprender textos orales predijeron la comprensión de textos escritos a los 7-8 años.

Como objetivos secundarios, nuestro estudio buscó, a partir de los resultados obtenidos, (1) identificar los grupos que corren el riesgo de no alcanzar un adecuado desarrollo lingüístico, y (2) proponer intervenciones para mejorar directa e indirectamente el desarrollo lingüístico. Identificamos como principal grupo vulnerable a no alcanzar un buen vocabulario, a los niños residentes en zonas rurales. Además, proponemos que se intervenga entrenando a las madres, padres o cuidadores principales, así como a los docentes de Educación Inicial para emplear prácticas de literacidad que, directamente, mejorarán el vocabulario de los niños, e indirectamente, su posterior comprensión lectora.

No tenemos conocimiento de ningún estudio que haya examinado las relaciones que se evaluaron. En particular, no conocemos trabajos con el objetivo de identificar el poder predictivo en el tiempo del vocabulario sobre la comprensión lectora<sup>1</sup>.

---

<sup>1</sup> Para una breve descripción de los resultados de trabajos sobre determinantes de la comprensión de lectura en el Perú, véase el anexo 1.

## EL PROBLEMA DE ESTUDIO

Como hemos mencionado, la situación de los estudiantes peruanos en comprensión de texto está claramente por debajo de lo deseado, a pesar del interés por parte del Estado en mejorarla. En las siguientes líneas, describimos con más detalles esta situación.

### 1. El escenario actual: los datos de las pruebas PISA, TERCE y ECE<sup>2</sup>.

La Evaluación Internacional de Estudiantes PISA, elaborada por la Organización para la Cooperación Económica y el Desarrollo OCDE, tiene por objeto evaluar el nivel de adquisición de “conocimientos y habilidades necesarios para la participación plena en la sociedad del saber” que tienen los alumnos al final de la educación obligatoria. Evalúa el rendimiento de los estudiantes de 15 años en lectura, matemática y ciencia que hayan concluido al menos seis grados de escolaridad. La prueba de lectura, aplicada a aproximadamente 6000 estudiantes, sitúa al Perú en la cola de los 65 países participantes, con un puntaje medio entre los niveles de desempeño 1 y 2, donde 1 corresponde al más bajo y 6, al más alto. Aún así, al comparar los resultados en las pruebas de comprensión lectora de PISA en las que participó el Perú, 2000 (administrada en Perú y algunos otros países en 2001), 2009 y 2012, se observa una mejora en el desempeño: se ha pasado de un rendimiento mayoritario por debajo del nivel 1 al límite con el nivel 2. En la tabla 1 puede verse la distribución en porcentajes para cada nivel.

---

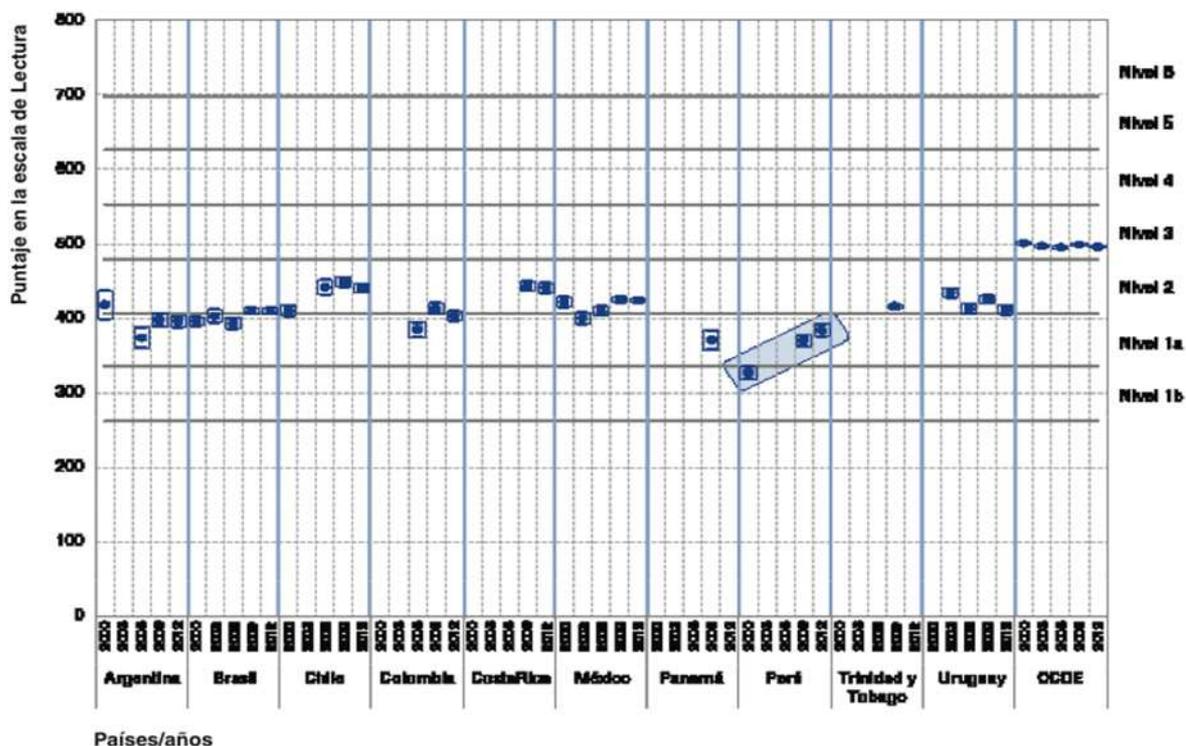
<sup>2</sup> Toda la información presentada en esta sección, incluidos los cuadros, fue obtenida de la web del MINEDU <<http://minedu.gob.pe>> y de la Unidad de Medición de la Calidad de la Enseñanza del MINEDU, en la sección relativa a las evaluaciones <<http://umc.minedu.gob.pe>>

Tabla 1 Resultados por niveles de desempeño y porcentajes de alumnos peruanos en las pruebas de lectura de PISA 2001, 2009 y 2012.

PISA 2001	% alumnos	PISA 2009	% alumnos	PISA 2012	% alumnos
5	0,1	6	0,0	6	0,0
4	1,0	5	0,4	5	0,5
3	4,9	4	2,6	4	3,3
2	14,5	3	10,1	3	11,4
1	25,5	2	22,1	2	24,9
< 1	54,1	1a	28,7	1a	29,5
		1b	22,0	1b	20,6
		< 1b	14,1	< 1b	9,8

La mejora en el rendimiento medio del Perú se destaca en la región. Según el informe de la OCDE (2010b), aunque el promedio general en las pruebas de comprensión lectora no ha variado de 2000 a 2010, Perú ha sido el país latinoamericano que más avanzó desde el inicio de su participación y este avance se mantuvo con las pruebas de 2012. En el gráfico 1, proveniente del Informe Nacional sobre los resultados de PISA elaborado por el Ministerio de Educación (de ahora en adelante: MINEDU; 2012) se pueden observar los puntajes medios para los países latinoamericanos en las pruebas de 2000-2001, 2009 y 2012 y el puntaje general.

Gráfico 2. Nivel promedio en las pruebas de lectura en los países latinoamericanos que participaron en las pruebas PISA de 2000 a 2012 y nivel general de todos los países participantes. Fuente: MINEDU 2013c



En el Perú, con respecto al desempeño promedio, tenemos brechas por género, por área geográfica y por gestión de las instituciones educativas. Por un lado, las niñas tuvieron, en 2001, un desempeño mejor por 7 puntos que los niños; mientras que en las pruebas de 2009 y 2012, la diferencia fue mayor: de 22 puntos. Es decir, las diferencias entre niños y niñas, a favor de estas últimas, aumentaron inicialmente y parecen haberse estabilizado. Por otro lado, entre poblaciones urbanas y rurales hay mayores diferencias aún, las cuales, a su vez, han aumentado hasta, aparentemente, estabilizarse. En 2001, el promedio en las ciudades fue de 83 puntos más que en las zonas rurales; esta diferencia fue de 91 puntos en 2009 y 90 puntos en 2012. Finalmente, hay una brecha, que parecer estar cerrándose, entre colegios públicos y privados. En 2001, los primeros tenían un promedio mayor por 122 puntos; pero en 2009 y 2012 su ventaja se redujo a 89 y 74 puntos, respectivamente, mostrando un estancamiento, pues la diferencia entre estos puntajes no es estadísticamente significativa (MINEDU 2012). Estos últimos datos deben ser interpretados cuidadosamente ya que, como bien advierte el mismo informe, las estimaciones tienen

un margen de error mayor debido a la heterogeneidad entre los colegios privados y la mayor atracción de alumnos por parte de este tipo de instituciones en los últimos años.

El Tercer Estudio Regional Comparativo y Explicativo (TERCE), llevado a cabo en 2013 por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la UNESCO, comparó el desempeño de estudiantes de tercero y sexto grado de primaria de 15 países latinoamericanos en lectura, matemática y ciencias (UNESCO 2014). Aquí presentamos los resultados relativos a la comprensión lectora en Perú en tercer grado, en comparación con los otros países y en relación al desempeño de la prueba anterior, el Segundo Estudio Regional Comparativo y Explicativo (SERCE), de 2006 (UNESCO 2008). En TERCE 2013, el desempeño de los niños peruanos los colocó en la octava posición, con un puntaje de 505,44 (el puntaje máximo, de Chile, fue 557,01; y el mínimo, de República Dominicana, 495,94); hubo un aumento estadísticamente significativo con respecto al puntaje en SERCE 2006, que fue de 476,29 (el puntaje máximo, de Costa Rica, fue 563,19; y el mínimo, de República Dominicana, 421,47), pero la posición con respecto a los demás países se mantuvo<sup>3</sup>. En la Tabla 2 se presentan los puntajes de todos los países participantes a SERCE 2006 y TERCE 2013. Con respecto a posibles brechas entre grupos, tenemos información solo respecto al género: no hubo diferencias significativas entre niños y niñas de 3º y 6º en las evaluaciones SERCE 2006 ni TERCE 2013 realizadas en Perú (UNESCO 2014).

---

<sup>3</sup> Todos estos datos provienen de la primera entrega de resultados (UNESCO 2015), la entrega final aún no está disponible.

Tabla 2. Países participantes y puntajes promedios en SERCE 2006 y TERCE 2013, para las pruebas de lectura en tercer grado, ordenados por desempeño.<sup>4</sup>

		SERCE			TERCE
1	Costa Rica	563,19	1	Chile	557,01
2	Chile	546,07	2	Costa Rica	545,50
3	Uruguay	542,15	3	Uruguay	531,79
4	México	529,92	4	México	528,77
5	Brasil	520,32	5	Colombia	525,57
6	Colombia	514,94	6	Brasil	523,93
7	Argentina	506,45	7	Argentina	508,58
<b>8</b>	<b>Perú</b>	<b>476,29</b>	<b>8</b>	<b>Perú</b>	<b>505,44</b>
9	Nicaragua	472,92	9	Ecuador	490,70
10	Panamá	472,05	10	Guatemala	489,03
11	Paraguay	455,24	11	Panamá	482,63
12	Guatemala	451,46	12	Honduras	479,19
13	Ecuador	447,44	13	Nicaragua	478,96
14	Rep.Dominican.	421,47	14	Paraguay	469,14
	Honduras	-	15	Rep.Dominicana	455,94

La Evaluación Censal de Estudiantes (ECE) 2014, análogamente a lo que evidencian las pruebas PISA y TERCE, sugiere una mejora en la comprensión lectora de los estudiantes peruanos. Los datos de ECE 2014 que presentamos provienen de una muestra censal de comprensión de lectura en niños de segundo grado de primaria, la población de nuestro interés. Los resultados, en puntajes, son clasificados en tres niveles: satisfactorio, en proceso y en inicio. En el nivel satisfactorio, se encuentra solo el 43,5% de los niños, según ECE 2014; sin embargo, esto indica una mejora en el tiempo, pues en 2013 en este nivel estaba el 33% y en 2012, el 30,9% (ver tabla 3 para una comparación más completa).

Tabla 3: Resultados por niveles de logro de ECE 2012, ECE 2013 y ECE 2014.

Nivel	ECE 2012	ECE 2013	ECE 2014
Satisfactorio	30,9%	33,0%	43,5%
En proceso	49,3%	51,3%	44,0%
En inicio	19,8%	15,8%	12,5%

Examinando los resultados más recientes, de 2013 y 2014, encontramos brechas de equidad con respecto al género, el área geográfica y la gestión de las

<sup>4</sup> Cuadro elaborado a partir de la información en UNESCO (2014).

instituciones educativas. En primer lugar, las niñas tienen un desempeño superior a los niños en 2013 y 2014, de 4,2 puntos en ambas mediciones. En segundo lugar, los estudiantes de colegios privados obtuvieron resultados mejores que los de colegios privados por 19,7 y 19,3, respectivamente en 2013 y 2014. Finalmente, los niños provenientes de zonas urbanas tuvieron promedios mayores que los niños de zonas rurales por 28,1 puntos en 2013 y 33 puntos en 2014.

Los datos proporcionados por las pruebas PISA, SERCE y TERCE, y ECE dibujan un escenario en el que hay ligeras mejoras en la comprensión de lectura, seguramente fruto de los esfuerzos del Estado, que aún así, no llegan a tener resultados completamente satisfactorios. Es evidente también que hay diferencias de desempeño que colocan a los varones, a los estudiantes de zonas rurales y de instituciones educativas públicas en una posición de desventaja. A continuación, mostramos las acciones que ha realizado el Estado y los compromisos que ha tomado para la mejora de la educación pública en general, y de la comprensión de textos, en particular.

## 2. Política pública a favor de la mejora en comprensión textual

El Estado, a través del Ministerio de Educación, ha mostrado interés en conocer y mejorar la situación educativa en el Perú, en general, y el desarrollo de habilidades comunicativas, como la comprensión de textos, en particular. El interés en conocer la situación de la comprensión lectora se hace evidente principalmente con las pruebas ECE, administradas cada año y que se describen en el apartado anterior. En este apartado vamos a describir brevemente los intentos, en parte satisfactorios, para revertir la situación en que se encuentran los estudiantes peruanos en comprensión de lectura.

Desde 2007, el Perú cuenta con el “Proyecto Educativo Nacional al 2021” (PEN), exigido por mandato de la Ley General de Educación (artículo 7), que define los objetivos que el Estado quiere alcanzar en materia de educación y traza el camino a seguir hasta el año 2021 (MINEDU 2007). En el PEN, se establecen seis objetivos estratégicos. Uno de estos objetivos es que los estudiantes e instituciones educativas logren “aprendizajes pertinentes y de calidad”, cuyo primer resultado consistiría en que “todos logran competencias fundamentales para su desarrollo personal y el progreso

e integración nacional” (MINEDU 2007). Desde 2007 hasta hoy ha habido mejoras y logros en la educación, y, específicamente, en la comprensión de lectura; sin embargo, el desempeño sigue siendo muy bajo, recordemos que incluso para el 2014 más de la mitad de los niños de segundo grado de primaria no logran el nivel satisfactorio en las pruebas ECE del Ministerio de Educación. El balance y las recomendaciones sobre el PEN, elaboradas para el 2014, claramente describen esta situación (MINEDU 2014). El aspecto más interesante de este informe es la atribución de las mejoras a ciertos factores: “la oportunidad de los procesos de contratos docentes y de entrega de material de apoyo pedagógico; el programa de reforzamiento para estudiantes con mayores dificultades de aprendizaje, la expansión de la cobertura del programa de acompañamiento; así como el compromiso y acciones desarrolladas por los Gobiernos regionales, las organizaciones de la sociedad civil y las empresas” (MINEDU 2014: 27). Esto sugiere que se ha conseguido salir de una condición en la cual habría resultado imposible que nuestros estudiantes tuvieran una educación adecuada. Creemos que ahora, sin embargo, es necesario implementar actividades específicas para mejorar la comprensión textual, competencia básica e imprescindible para un aprendizaje óptimo. Estas actividades podrán ser planeadas solo sobre la base de la relación entre habilidades que ofrece este estudio. Como adelantamos en la introducción, nuestros resultados muestran que la habilidad de comprender un texto escrito, depende de la habilidad de comprender un texto oral, la cual se desarrolla desde mucho antes de aprender a leer o iniciar la primaria. Además, la comprensión de textos escritos a los 8, 9 años de edad se basa en el léxico que tuvo un niño antes de empezar la educación formal, a los 4, 5 años. Sobre la base de esta información se podrán planear actividades e implementarlas junto a las actividades ya en vigor en inicial y primaria, integradas en las estrategias puestas en marcha para mejorar la educación de los niños peruanos.

En los últimos años, se han elaborado diferentes estrategias para mejorar la educación en general, y la comunicación en particular, de los niños en los niveles inicial y básico. Por un lado, el Ministerio de Educación hace explícito su interés por “mejorar los niveles de logros de aprendizaje combinando tres resultados intermedios: el desarrollo de una gestión orientada por resultados de logros de aprendizaje, el desarrollo de una educación inicial articulada con la primaria y el desarrollo de un nuevo sistema de capacitación y acompañamiento pedagógico”. Los dos primeros resultados son de nuestro interés para el tipo de actividades que proponemos. El primer resultado implica una gestión caracterizada por un sistema de rendición de

cuentas ligado a la difusión de evaluaciones estandarizadas que midan la mejora en el rendimiento educativo (MINEDU 2011). Las pruebas ECE, administradas anualmente, son un claro ejemplo de este tipo de evaluaciones. Además, el primer resultado se apoya en el segundo: desarrollar una educación inicial que logre “competencias específicas y necesarias para contribuir a lograr las metas de logros en lecto-escritura y razonamiento matemático en los primeros grados de primaria” (MINEDU 2011: p 27). Es en este segundo punto donde mejor se inserta nuestra propuesta. Los resultados de nuestro estudio indican que la comprensión de textos en niños en edad correspondiente a segundo grado de primaria depende de su comprensión de textos orales y, también del léxico que tuvieron cuando estaban en edad de recibir educación inicial. Creemos, sobre la base de nuestros resultados, que mejorar el léxico en educación inicial es una forma de aumentar las probabilidades de que los niños tengan una mejor comprensión de textos escritos durante los primeros años de primaria.

Por otro lado, en la misma línea, el Ministerio de Educación ha diseñado el Programa Estratégico Logros de Aprendizaje Ciclo III (PELA), en el cual se define como problema central el “bajo nivel de logro en comprensión lectora y pensamiento lógico-matemático en alumnos de segundo grado de instituciones educativas públicas de nivel primario”<sup>5</sup>. La forma de superar este problema, así como de alcanzar otros objetivos que también se integran en el objetivo general de mejorar la calidad del servicio que ofrecen las instituciones educativas en el Perú, tanto públicas como privadas, se ha implementado a través de un sistema curricular. Este sistema asegurará los aprendizajes requeridos y orientará la labor de los docentes. Está compuesto por el Marco Curricular, los Mapas de Progreso y las Rutas de Aprendizaje, y se constituye en el orientador y articulador de los Currículos Regionales. El Marco Curricular define el total de los aprendizajes fundamentales que tienen que alcanzar los estudiantes durante la educación básica; los Mapas de Progreso describen qué deben conocer, saber hacer y valorar los estudiantes en cada ciclo de la educación básica y proporciona criterios para controlar y evaluar dichos aprendizajes; y las Rutas del Aprendizaje sirven de apoyo al trabajo de los docentes, orientando sus estrategias de enseñanza para favorecer el aprendizaje (MINEDU 2013b). La última versión del Marco Curricular (MINEDU 2014) incluye, entre las ocho competencias que debe haber adquirido un alumno al finalizar la educación básica, la habilidad de comunicarse; la cual entiende como la habilidad de utilizar “el lenguaje de

---

<sup>5</sup> Esta información proviene de la página web del Ministerio de Educación: <http://www.minedu.gob.pe/>

manera eficaz para participar en diversas prácticas sociales e interculturales” (MINEDU 2014: 11). En los Mapas de Progreso y las Rutas de Aprendizaje se definen con detalles las competencias que comprenden la comunicación, incluyendo explícitamente el vocabulario y las diferentes habilidades relacionadas a la comprensión de textos orales y escritos. Ambos indican, con diferente grado de detalle, a qué nivel debe encontrarse el desempeño de los niños en cada competencia desde educación inicial (MINEDU 2013a, 2013b, 2015a y 2015b).

Como se ha mostrado, el Ministerio de Educación ha puesto en marcha un sistema coordinado que prevé el desarrollo de competencias lingüísticas, principalmente la comprensión de textos, cuya adquisición inicia en los primeros años de educación preescolar. Sin embargo, en ninguno de estos documentos se ha encontrado referencia a la necesidad de intervenir para aumentar el vocabulario durante el periodo inicial con el objetivo de mejorar la comprensión de textos escritos a lo largo de la primaria. Creemos que este tipo de intervención, la que nosotros sugerimos sobre la base de resultados, podría perfectamente integrarse en el sistema educativo nacional.

## **EL MARCO TEÓRICO**

En esta sección, presentamos los conceptos que utilizaremos para la selección de los datos de la base, el análisis y la interpretación de los resultados, con referencia al marco teórico y la literatura que nos sirven de referencia. Presentamos en primer lugar, el concepto de desarrollo lingüístico, que incluye la comprensión textual y el vocabulario; en segundo lugar, otros aspectos que influyen la comprensión del texto, como el nivel desarrollo del niño y las oportunidades que tiene para mejorarla.

### **1. El desarrollo lingüístico**

El lenguaje de un niño empieza a desarrollarse desde el inicio de su vida. Desde que es capaz de escuchar las vocalizaciones de su entorno, el ser humano está construyendo las bases de lo que en el futuro será su lengua o sus lenguas. Adquirirá el lenguaje en todos sus dominios (fonético-fonológico, morfológico, sintáctico, léxico) y aprenderá a usarlo para comunicarse (pragmática). Aprenderá a entender y a producir palabras, y a lo largo de su vida continuará a adquirir nuevas palabras y olvidará aquellas cuyo uso no le resulta útil. En las primeras etapas de su vida ampliará significativamente su vocabulario y este proceso será más evidente cuando, con el contacto con la educación formal, incorpore palabras de contextos distintos al de su entorno familiar. Cuanto más amplio, de nivel más alto y, sobre todo, más acorde al ámbito académico sea su vocabulario, más importante será para la adquisición de conocimientos en la escuela. Por otro lado, las palabras que escuchará estarán organizadas en frases, oraciones y en estructuras mayores, según la función comunicativa. Los cuentos, las historias, las clases respetarán la forma de un texto, sea este oral o escrito. La habilidad de decodificar e interpretar estos textos será imprescindible en la construcción de nuevos conocimientos y predecirá el futuro éxito académico y laboral (OCDE 2011).

A continuación, presentamos los conceptos de comprensión de textos orales y escritos y de vocabulario.

## 1.1. La comprensión de textos

Los años inmediatamente anteriores al inicio de la educación formal, la etapa preescolar, constituyen una fase importante para la adquisición del lenguaje y, en particular, para la comprensión de textos orales (Florit et al. 2009, 2011; Kendeou et al. 2008). La evidencia científica muestra que la comprensión de textos orales es un proceso que comparte muchas características con la comprensión de textos escritos, más estudiada en la literatura (Hoover y Gough 1990; y para un meta análisis reciente, Florit y Cain 2011), por lo que aquí describimos conjuntamente ambos procesos.

La comprensión de un texto, oral o escrito, implica llevar a cabo un proceso complejo y multidimensional que permita construir una representación coherente del significado del texto (Hannon y Daneman 2001; Kintsch 1998). En la construcción de esta representación se pueden identificar dos niveles. El nivel básico corresponde a la construcción del significado explícito a nivel local y global del texto (Florit, Roch y Levorato 2011). Sobre este nivel, se construye el nivel más alto, donde se integra la información con conocimientos previos (Kintsch 1998; Kintsch y Kintsch 2005). Para extraer la información explícita, el oyente/lector necesita utilizar habilidades lingüísticas adquiridas previamente. Si el texto es oral, necesita identificar las palabras, entender las frases y recordar lo comprendido anteriormente. Si es escrito, necesita, además, decodificar los símbolos. Es decir, en ambos casos entran en juego el vocabulario receptivo, la comprensión morfo-sintáctica y la memoria. Diferentes estudios han demostrado que la adquisición de estas habilidades predice, en diferente grado, la comprensión textual (Florit et al. 2011, 2013, 2014; Muter et al. 2004; Ouellette 2006; Roth et al. 2002; Seigneuric y Ehrlich, 2005). Sin embargo, el poder de predicción del vocabulario es tan alto en comparación a las demás habilidades que, al evaluarlas en un mismo grupo, la variabilidad explicada depende exclusivamente del vocabulario (Junyent 2011).

La comprensión del texto oral, no solo es un proceso equivalente al de la comprensión de texto escrito, sino que, además, es una habilidad crítica para el desarrollo de la literacidad (McCardle et al. 2001). Este patrón es evidente tanto en las primeras etapas de adquisición de la lectura (Dickinson et al. 2003; National Institute of Child Health and Human Development 2005), como en los últimos años de primaria, cuando tanto la amplitud del vocabulario como el conocimiento de la estructura de los textos orales ayudan en la comprensión de textos escritos (Dickinson

y Tabors 2001; Griffin et al. 2004; Reese et al. 2010; Sénéchal y LeFevre 2002; Storch y Whitehurst 2002). Para que un niño desarrolle una buena habilidad de comprensión lectora es necesario que antes haya desarrollado un vocabulario adecuado y una capacidad de comprensión de textos orales apropiada (Snow et al. 1998; Whitehurst y Lonigan 1998). Investigaciones recientes sugieren que los problemas que muchos niños experimentan al aprender a leer durante los primeros años de educación pueden deberse a deficiencias en las habilidades de base que generalmente se desarrollan durante la etapa preescolar (para una revisión, ver Caesar y Nelson 2013).

Con respecto a la comprensión de textos escritos en los niños del Perú, sabemos que su desarrollo es pobre desde el inicio de su escolaridad (para una descripción de los últimos hallazgos ver Thorne 2013) como muestran los datos presentados en la sección relativa al problema de estudio.

## 1.2. El vocabulario

Las investigaciones han demostrado repetidamente que el vocabulario es una base fundamental en la adquisición del lenguaje y de la literacidad, aunque no la única (August et al. 2005). Con respecto a la relación entre vocabulario y comprensión textual, los resultados de numerosos estudios evidencian una relación causal entre ellos (Freebody y Anderson 1983; Daneman 1988; de Jong y van der Leij 2002; Roth et al. 2002; Oakhill y Cain 2012; Florit y Levorato 2012). Esta relación puede explicar casos de bajo desempeño en la comprensión de textos: los niños que no conocen el significado de muchas palabras o no pueden acceder a ellos rápida y eficientemente, encuentran, debido a esta razón, dificultad en entender los textos (Perfetti 1985). Algunos de los estudios mencionados encontraron la relación entre vocabulario y comprensión textual en niños con las edades de nuestro interés. Roth y colegas (2002) hallaron que el vocabulario receptivo de niños de 5 y 6 años predecía su comprensión de textos escritos a los 7 y 8 años; y Oakhill y Cain (2012) encontraron la misma relación entre los 7 y 8 años, por un lado; y los 8 y 9, por otro lado. Además, de Jong y van der Leij (2002) descubrieron que el léxico en comprensión y producción a los 7 y 8 años predecía la competencia lectora a los 9 y 10 años. Análogamente, Florit y Levorato (2012) mostraron una relación equivalente entre léxico y comprensión del texto oral, en niños que tenían 4 años en la primera medición y 5, en la segunda. Además de la relación entre vocabulario, como causa, y comprensión textual, como

efecto, hay evidencias de una relación causal también en la dirección opuesta (Bishop 1997; Cunningham 2005; Eldredge et al. 1990), aunque de manera mediada (Florit y Levorato 2012). Así, es posible interpretar la relación entre vocabulario y comprensión textual como una relación bidireccional (Seigneuric y Ehrlich 2005; Verhoeven y van Leeuwe 2008; Bast y Reitsma 1998), con respecto a la cual, sin embargo, queda más claro el rol del vocabulario como causa que como efecto (Florit y Levorato 2012).

La relación entre vocabulario y comprensión textual se confirma en estudios recientes realizados con niños hablantes de español. En primer lugar, se han encontrado correlaciones entre el vocabulario receptivo y la comprensión de textos orales en niños chilenos de 3 años y medio a 5 años (Strasser et al 2010). En segundo lugar, se han hallado importantes relaciones entre el vocabulario y la comprensión de textos escritos en niños de 8 a 11 años en Argentina (Canet-Juric 2013; Villalonga et al. 2014) y Perú (Morales 2009 y 2011).

## 2. Aspectos que influyen en el desarrollo de la comprensión textual

La comprensión textual, adecuada o inadecuada, puede depender – además del desarrollo lingüístico, como vimos en el apartado anterior – de una serie de factores relacionados tanto con las características del niño como con las oportunidades que su entorno le ofrece. En esta sección presentamos aquellas que consideramos más importantes y acerca de las cuales podemos extraer información de la base de datos que analizamos.

### 2.1. El desarrollo del niño

La comprensión de textos tiende a mejorar con el tiempo transcurrido, no solo por la experiencia, sino también por la maduración que ocurre en los niños con la edad. En esta sección, presentamos, en primer lugar la edad y su efecto en la comprensión textual; en segundo lugar, separadamente, el desarrollo cognitivo y físico asociado a dicha habilidad; y por último, el género.

### 2.1.1. La edad

Como todas las habilidades cognitivas, las competencias lingüísticas aumentan con la edad durante el periodo infantil; al respecto hay abundante evidencia científica (para una revisión, ver Singleton y Ryan 2004).

### 2.1.2. El desarrollo cognitivo

Asociadas a la edad, tenemos mejoras en las competencias lingüísticas debido al desarrollo cognitivo de los niños. La perspectiva piagetiana colocaba al lenguaje en una posición de dependencia frente al pensamiento, cuyo desarrollo dependía de este último (Piaget 1979). En ese contexto, el lenguaje no era sino una expresión de la más general capacidad simbólica (Zampini y D'Odorico 2007). Desde esta perspectiva, el desarrollo de las habilidades lingüísticas depende del desarrollo cognitivo (Bates y Dick 2002). Sin embargo, marcos teóricos más recientes, como en el que se inserta este trabajo, interpretan la dependencia que tiene el lenguaje del desarrollo cognitivo de maneras más cauta. La relación entre cognición y lenguaje desde esta perspectiva es múltiple, variada y ligada a estadios específicos de del desarrollo (Gopnik y Meltzoff 1986). No entraremos en detalles sobre la compleja relación cognición-lenguaje; para nuestros fines, consideraremos que la comprensión textual, habilidad lingüística que implica procesos cognitivos básico y de alto nivel, depende del desarrollo cognitivo.

En este estudio no se utilizará una medida directa del desarrollo cognitivo, sino solo una medida indirecta, el crecimiento, que se describe en la sección siguiente. La base de datos Niños del Milenio ofrece una medida directa de desarrollo cognitivo que excluimos por ser inadecuada para nuestros fines; la descripción de la medida y las razones de la exclusión están en el anexo 2.

### 2.1.3. El desarrollo físico

Nuestro interés en el desarrollo físico está basado en la necesidad de considerar indicadores del desarrollo cognitivo, a falta de una medida directa y adecuada, para evaluar el efecto (sin influencia del lenguaje) del desarrollo en las habilidades lingüísticas, como la comprensión del texto. Consideramos el crecimiento un indicador indirecto del desarrollo cognitivo y aprovechamos la información sobre

discapacidad para seleccionar a los niños que no muestran ningún signo de retraso en el desarrollo lingüístico.

### *El crecimiento*

El crecimiento infantil, indicador de la salud y el estado nutricional de un niño, ha demostrado ser de gran importancia para el desarrollo cognitivo (Grantham-McGregor 2002; Behrman et al. 2004 y 2008; Crookston et al. 2011). Muchas investigaciones se han centrado principalmente en los dos primeros años de vida, pues generalmente se asume que los déficits cognitivos asociados a un inadecuado desarrollo, producto de un pobre estado nutricional, en esta etapa son irreversibles posteriormente (Martorell et al. 1994; Walker et al. 1996; Checkley et al. 2003). Otros estudios, sin embargo, han demostrado, en cambio, que sí es posible revertir los déficits de desnutrición temprana si se interviene durante los primeros cinco años de vida (Morgane et al. 1993; Strauss y Dietz 1998; Pollitt 1993; Benjamin et al. 2013); además se ha hallado mejoras en las habilidades cognitivas, asociadas a una mejor nutrición (Cheung y Ashorn 2010; Crookston et al. 2010, 2013; Yang et al. 2011; Fink y Rockers 2014). Todos estos estudios utilizan el índice “puntaje z de la altura para la edad” o HAZ (por sus siglas en inglés), que relaciona la altura del niño y su edad con la población de referencia, ofreciendo información sobre el crecimiento infantil (World Health Organization 1997). Un crecimiento insuficiente representa la historia de salud y nutrición de un niño y es un signo de malnutrición crónica (Victora et al. 2008). A nivel poblacional, niveles altos de crecimiento insuficiente se asocia a pobres condiciones socioeconómicas y alto riesgo de exposición temprana y frecuente a condiciones adversas, como enfermedad o prácticas alimenticias inapropiadas (World Health Organization 1997). El nivel de insuficiencia del crecimiento se calcula utilizando la desviación estándar (sd) de la media poblacional. La insuficiencia severa se encuentra por debajo de 3 sd; la media, entre 2 sd y 3 sd; y la leve, por debajo de 2 sd.

En la base de datos tenemos el puntaje HAZ de los niños; indicador del crecimiento de cada niño hasta la edad de la medición – debido a factores como malnutrición crónica o exposición a condiciones adversas – y lo consideramos indicador indirecto del desarrollo cognitivo.

## *Las discapacidades*

Las discapacidades pueden afectar de múltiples maneras el desarrollo de un niño. Al referirnos a discapacidades estamos considerando una parte de aquellas sobre las que la base de datos nos ofrece información. Esta información incluye datos sobre discapacidades físicas como mentales que puedan traer como consecuencias impedimentos en las actividades cotidianas. Hemos utilizado la información exclusivamente sobre las discapacidades que tienen un efecto claro y evidente en el desarrollo lingüístico con el objetivo de excluir a aquellos niños que puedan tener un trastorno del lenguaje. Esta información está relacionada a preguntas sobre la capacidad auditiva del niño, su capacidad de comprensión del habla de otros, de producir lenguaje y de ser comprendido. Según diferentes estudios, la capacidad de entender y ser entendidos durante la etapa preescolar está relacionada con la identificación de déficits lingüísticos en etapas posteriores (Shriber et al. 1999; Fish y Pinkerman 2003).

### 2.2. El sexo y el género

El concepto de sexo y el de género, aún en mayor medida, han sido reelaborados a lo largo del tiempo. Por mucho tiempo, y quizás aún hasta hoy, ha prevalecido la idea del sexo, masculino o femenino, como característica profunda compartida por un grupo de individuos y representada por los genitales, a partir de la cual se desarrollan específicas respuestas sociales, similares en todas las culturas, lo que Nicholson (1994) llama “fundamentalismo biológico”. Este paradigma, que da por sentado el sexo como base inmutable sobre la que se construye socialmente el género, ha sido desafiado por el paradigma construccionista social; desde esta nueva perspectiva, el género se convierte en la base, la institución social que determina cómo vemos los cuerpos y las diferencias sexuales (Knaak 2004). Para Butler (1990/2007), el sexo entendido como la base natural del género es una concepción que tiene lugar en un sistema social ya marcado por la normativa del género. Para la autora, “el sistema binario de géneros sostiene de manera implícita la idea de una relación mimética entre género y sexo, en la cual el género refleja al sexo o, de lo contrario, está limitado por él” (Butler 1990/2007: 54). Butler no solo ve el género como una construcción, sino que lo considera “el medio discursivo/cultural a través del cual [...] ‘un sexo natural’ se forma y establece como ‘prediscursivo’, anterior a la cultura,

una superficie políticamente neutral *sobre la cual* actúa la cultura” (Butler 1990/2007: 56). En pocas palabras, para las nuevas perspectivas, el género no es más una respuesta social a una clara división sexual, sino una construcción que es, a su vez, la base de la distinción sexual.

El debate sobre el tema de sexo y género es amplio y una descripción profunda excede los límites de este trabajo; sin embargo, consideramos importante estar alerta a los conceptos que están detrás de las etiquetas “hombre” y “mujer”, cuando se examinan evidencias provenientes de otras investigaciones y cuando se interpretan los resultados de los propios estudios.

Para empezar, es necesario tener claros los conceptos de sexo y género que se están utilizando para interpretar los resultados de la evidencia científica. En toda investigación sobre el desarrollo de los seres humanos, en la que se haga distinción entre hombre y mujer, niño y niña, está implícita una concepción sobre la que se fundamenta la pertenencia a una u otra categoría. Diversos autores han señalado la importancia de reconceptualizar el sexo y el género en el diseño e interpretación de estudios de investigación (Johnson y Repta 2012; Knaak 2004; Lorber 1996 y 2005). Sin embargo, hasta el momento, en la mayoría de estudios se identifica sexo con género en un sistema binario. Por esa razón, vamos a interpretar los resultados provenientes de estudios sobre el desarrollo lingüístico, como información sobre “hombres” y “mujeres”, entendidos desde lo que suponemos es la perspectiva de los autores: una distinción en dos géneros sobre la que se interpretan las diferencias sexuales. Señalaremos aquellos casos en los que parece interpretarse una distinción más cercana a las diferencias sexuales. A continuación, describiremos (1) la relación entre el desarrollo lingüístico y el sexo; (2) la relación entre el desarrollo de la lengua oral y escrita y el género; y (3) la relación entre el desarrollo de la lengua escrita y el género para el caso peruano.

#### 2.2.1. Relación entre desarrollo lingüístico y sexo

La relación entre el género, el cerebro y las funciones cognitivas ha sido tema de interés científico desde hace mucho (Kimura 1996 y 1996; Witelson 1976). En lo que respecta al lenguaje, la investigación ha demostrado una clara ventaja a favor de las mujeres. El desarrollo del lenguaje es, generalmente, más rápido en las niñas cuando se les compara con los niños (Alexander y Wilcox 2012). Hay razones para pensar que estas diferencias puedan tener bases biológicas: son evidentes en

estudios con niños muy pequeños; investigaciones sobre el funcionamiento cerebral y su relación con las hormonas sexuales; así como el examen de incidencias de trastornos lingüísticos. En primer lugar las niñas, como grupo, empiezan a hablar antes que los niños (Murray et al. 1990), adquieren el vocabulario más rápidamente (Roulstone et al. 2002); y a los 2 y 3 años de edad tienen un desempeño lingüístico superior al de los niños (Bornstein et al, 2000; Dionne et al, 2003). Por un lado, no es posible atribuir estos resultados a diferencias puramente sexuales; pues aunque pequeños, estos niños y niñas ya han sido inmersos en la cultura que los recibió, y han recibido de los adultos un trato, quizás distinto, dependiente del género que se les atribuye. Por otro lado, sin embargo, sí se puede hipotizar que la influencia cultural tendrá un menor efecto en niños pequeños que en niños más grandes y en adultos, y que, por lo tanto, los resultados distintos podrían ser, al menos en parte, consecuencia de una diferencia biológica. En segundo lugar, tenemos investigaciones sobre la relación entre funcionamiento cerebral y género/sexo, como el estudio de Burman y colegas (2008). Estos autores demostraron que el desempeño superior de las niñas en pruebas de lenguaje oral y escrito se asocia a un patrón de activación del cerebro diferente al de los niños e independiente de la modalidad (auditiva o visual). Aparte de evidencia de una diferente activación cerebral en los niños y las niñas, se ha demostrado que hay diferencias grupales en la estructura cerebral que se explican con la mayor o menor presencia de ciertas hormonas sexuales. La teoría de Geschwind (Geschwind y Galaburda, 1987; Geschwind y Galaburda, 1985a, 1985b y 1985c), basada en datos empíricos, sostiene que la superioridad del hemisferio derecho en los varones se debe a los altos niveles de testosterona, lo que ocurre en detrimento del desarrollo del hemisferio izquierdo, que suele ser el dominante para las funciones lingüísticas. Esta es una posible explicación a la ventaja, aparentemente biológica, de las niñas frente a los niños con respecto al lenguaje. La relación entre hormonas, estructura del cerebro y competencia lingüística se ve, además, confirmada con otros resultados, como los obtenidos por Lutchmaya et al. (2002): entre los 18 y 24 meses, las niñas de su estudio tenían un vocabulario más amplio que los niños; y esta diferencia se relacionaba directamente con el nivel de testosterona fetal medida previamente (a menor testosterona, mayor amplitud léxica). En último lugar, hay abundante evidencia que los trastornos de lenguaje en el desarrollo son menos frecuentes en las niñas que en los niños (Tomblin et al. 1997; Whitehouse 2010). El conjunto de resultado apoya fuertemente la hipótesis una diferencia biológica entre niños y niñas, a la base de la ventaja de estas últimas en el desarrollo lingüístico.

### 2.2.2. Relación entre desarrollo lingüístico y género

Además de los estudios presentados, hay muchos otros que también muestran la ventaja de las niñas frente a los niños, aunque en estos casos el concepto que subyace a la división es de género y no de sexo. Por ejemplo, Bauer y otros (2002) evalúan la amplitud léxica del habla espontánea en niños y niñas de 8 a 14 meses, basándose en el reporte de los padres. Por lo tanto, es evidente que aquí se mide la relación entre el género que los padres atribuyen al hijo y la percepción que tienen acerca de la producción lingüística del niño. Con respecto a la ventaja de las chicas frente a los chicos durante la educación formal, la situación se complica y la interpretación de la evidencia científica se hace más difícil. Por un lado, el impacto de la diferencia de género sobre el desarrollo de la habilidad lingüística de los niños y niñas – que probablemente muestra ventajas biológicas en las niñas – se acentuará con el ingreso al sistema educativo. Por otro lado, el aprendizaje de la lengua escrita y, en general, el desempeño académico, se verán fuertemente influenciados por los roles de género, lo que los padres, los maestros y los propios niños y niñas consideran que se espera de ellos. La evidencia muestra que tanto en primaria (Martin y Hoover 1987; Strand et al. 2006) como en secundaria (Mann et al. 1999; Undheim y Norvic 1992), así como en diferentes culturas<sup>6</sup>, las niñas presentan una ventaja en el desarrollo oral y escrito del lenguaje. Como dijimos, interpretar estos datos es complejo; aunque evidentemente la relación que entra en juego vincula género y desempeño lingüístico. En este trabajo, no analizaremos esta relación en general, sino que nos limitaremos a interpretar los datos de niños y niñas peruanas sobre la relación género-desempeño lingüístico.

### 2.2.3. Relación entre comprensión lectora y género en el Perú

En esta sección examinamos los datos sobre el desempeño de los niños y niñas peruanas en la lectura, desde la perspectiva de las diferencias género. Para empezar, recordamos los datos presentados en la sección “el problema de estudio” sobre los diferentes desempeños entre niños y niñas en pruebas de comprensión lectora en el Perú. Para interpretar estos resultados, nos servimos del trabajo de

---

<sup>6</sup> Aunque la mayoría de investigación en estas áreas proviene de los Estados Unidos y el Reino Unido, hay también estudios con otras poblaciones, como la japonesa (Martin y Hoover 1987) y la noruega (Undheim y Norvic 1992).

McGeown y otros (2012), que, hasta donde sabemos, es el único estudio que evalúa el desempeño en la lectura de niños y niñas, examinando separadamente las relaciones con el sexo y el género. Este estudio, realizado con población británica, además nos servirá para, junto con literatura local, interpretar nuestros propios resultados.

Las pruebas que evalúan la comprensión lectora de niños y niñas peruanos evidencian dos tipos de resultados. Por un lado, las pruebas PISA (2001, 2009 y 2012) y las evaluaciones ECE (2013 y 2014) muestran una ventaja de las estudiantes frente a los estudiantes. Por otro lado, en las evaluaciones SERCE 2006 y TERCE 2013, niños y niñas tuvieron resultados equivalentes. Es interesante notar que la ventaja de las niñas no se limita a los resultados del Perú o de la región: “las chicas superan a los chicos en lectura prácticamente en todos los países y economías” (OCDE 2014). Así, la ventaja de las niñas frente a los niños, se verifica en poblaciones de diferentes países y entre ellos, el nuestro.

El estudio de McGeown et al. (2012) señala no solo las conocidas ventajas de las niñas frente a los niños en la habilidad de lectura; sino también el hecho de que las niñas leen con mayor frecuencia, tienen actitudes más positivas hacia la lectura, mayor motivación a leer, mayor confianza en las propias habilidades de lectura, y otorgan mayor valor a las actividades de lectura, en comparación con los niños<sup>7</sup>. A partir de esta información, McGeown y colegas (2012) evalúa a niños y niñas británicos, para identificar diferencias en la habilidad de leer y la motivación; y medir el poder predictivo del sexo y la identidad de género. La habilidad de leer en los niños o niñas, como otras actividades que requieren esfuerzo e interés para llevarse a cabo, ha sido fuertemente asociada con la motivación. Los niños que se sienten intrínsecamente motivados a leer – es decir, cuya motivación proviene de factores internos – seguirán leyendo porque lo disfrutan; lo harán aunque no haya presiones externas o recompensas por hacerlo. En cambio, los niños cuya motivación es extrínseca – proviene de factores externos, como valores o demandas – realizarán la actividad solo para conseguir un resultado, como conseguir una buena nota o ganar el reconocimiento de sus padres o maestros, y no por el placer de leer (McGeown et al.

---

<sup>7</sup> En McGeown (2012) se encuentran las siguientes referencias: sobre la frecuencia de lectura, Coles y Hall 2002; las actitudes, Coles y Hall 2002, Kush y Watkins 1996, Logan y Johnston 2009; McKenna y otros 1995, Sainsbury y Schagen 2004; la motivación Baker y Wigfield 1999, Marinak y Gambrell 2010, Wigfield y Guthrie 1997; la confianza en las propias habilidades Wigfield et al., 1997; y el mayor valor otorgado a la lectura; Durik et al. 2006; Eccles et al. 1993; Marinak y Gambrell 2010; Wigfield et al. 1997.

2012: 329). El resultado más importante de este estudio fue que los niños y niñas se identificaron con los rasgos estereotípicos de su género y la identidad de género tuvo una fuerza predictiva más fuerte que el sexo, con respecto a la motivación intrínseca a leer (McGeown et al. 2012: 333). La importancia de estos resultados es que nos permiten interpretar los datos sobre los niños peruanos desde una perspectiva de género. La diferencia entre niños y niñas en la comprensión lectora puede estar asociada a una mayor dedicación y empeño de estas últimas a la actividad de lectura, producto de su identificación como actividad “femenina”.

Con respecto al contexto peruano, hay datos que demuestran que la actitud hacia la lectura – probablemente determinada por la motivación – hace que “las estudiantes mujeres tengan un rendimiento en comunicación por encima del que obtienen los estudiantes varones” (MINEDU 2006).

Además, hay razones para hipotizar una relación entre estereotipos de género, y dedicación – y, por lo tanto, desempeño – a ciertas actividades. Benavides (2006: 9) señala que “el mayor rendimiento de los varones en las matemáticas y las ciencias”, y, agregamos nosotros, el mayor rendimiento de las mujeres en la comprensión lectora, “podría estar escondiendo un estereotipo social sobre lo que le ‘corresponde’ aprender a cada género”. Así, el mejor desempeño de las niñas en comprensión lectora – incluso basada en una ventaja biológica en el desarrollo lingüístico – puede ser interpretado como producto de la influencia de estereotipos de género en el aprendizaje de niños y niñas, que restringen el libre desarrollo de las habilidades de unos y otros.

### 2.3. Las oportunidades del niño

Como mencionamos anteriormente, las habilidades lingüísticas que adquiere un niño dependen no solo de su desarrollo general, sino también de las oportunidades que le ofrece su entorno para que estas habilidades puedan florecer. Las oportunidades que describimos están relacionadas con las prácticas de literacidad, el acceso a la educación inicial y primaria; y el status socio-económico de la familia.

### 2.3.1. Las prácticas de literacidad<sup>8</sup>

Los problemas de comprensión de textos escritos que afrontan muchos niños en los primeros años de educación formal pueden estar relacionados con las habilidades de literacidad tempranas, que se adquieren generalmente durante la etapa preescolar (Caesar y Nelson 2013). Las investigaciones señalan tres habilidades, entre las relacionadas con las prácticas de literacidad, como los mejores predictores de la comprensión del texto escrito: la aptitud oral, el procesamiento fonológico y la familiaridad con el material impreso (August y Shanahan 2006). Podemos suponer que las dos primeras, además, tendrán efecto también en la comprensión del texto oral. Estas habilidades afloran en un entorno favorable, como el de las familias que ofrecen prácticas de literacidad a los niños. Con prácticas de literacidad se entienden, según cada estudio, situaciones en las que se promueve un uso complejo del lenguaje y se recurre a textos escritos. La oportunidad de experimentar estas prácticas en casa ha demostrado desempeñar un rol importante en el desarrollo de lenguaje de los niños y en la comprensión de los textos (Hart y Risley 1995). Los niños provenientes de hogares que ofrecen prácticas de literacidad ricas, como la lectura compartida, e involucran al niño en conversaciones que van más allá de los temas básicos cotidianos suelen mostrar un lenguaje más desarrollado que sus pares en condiciones menos favorables (Hart and Risley 1995; Wells 1986; Bennet et al. 2002). Para poder llevar a cabo las prácticas de literacidad en el hogar, la familia, y en modo particular la madre o cuidador principal, debe poseer los recursos necesarios: suficiente fluencia oral y escrita en la lengua en cuestión; y un nivel de educación que haya estimulado su contacto con material escrito. Diversos estudios han encontrado que los hogares con madres con un nivel medio o bajo en la lengua que utilizan (su segunda lengua), con niveles bajos de educación y con pocos ingresos emplean menos prácticas que favorecen la literacidad (Snow et al. 1998; González y Uhing 2008). Además, los hogares con madres menos educadas y con menor dominio de la lengua escrita ofrecen situaciones de lectura compartida menos buenas (Caspé 2009) y proporcionan menos oportunidades a los niños para interactuar utilizando el lenguaje oral y escrito (Lynch 2008).

Las prácticas de literacidad no solo se asocian con la comprensión de textos, sino también con la amplitud del vocabulario. Hay estudios que muestran que el tipo

---

<sup>8</sup> Utilizamos un concepto amplio de literacidad que puede definirse como el conjunto de “conocimientos y actitudes necesarios para el uso eficaz en una comunidad de los géneros escritos” (Cassany, 2005).

de lenguaje usado en casa y la presencia de recursos textuales escritos (libros) se correlacionan con un vocabulario más amplio en los niños (Stanovich 1986; Juel 1988; Snow et al. 1998; De Temple y Snow 2003; Hoff 2006; Huttenlocher et al. 2002). Además, otros estudios han evidenciado que la amplitud del vocabulario no solo se relaciona con las prácticas de literacidad – como un lenguaje materno dirigido al niño rico en palabras distintas, frecuencia de conversaciones referidas a contextos distintos al aquí y ahora – sino también, de modo directo, con características de la madre, como sus capacidades de lectura (Pan et al. 2005).

La base de datos Niños del Milenio nos ofrece información sobre el nivel de español oral y escrito de las madres, así como su grado de educación. Utilizaremos estos datos como evidencia de los recursos que poseen para generar prácticas de literacidad adecuadas.

### 2.3.2. Estatus socioeconómico

Sobre la relación entre estatus socioeconómico (SES, por su siglas en inglés) y desarrollo cognitivo existe abundante investigación (Bradley y Corwyn 2002). La relación empieza desde la infancia (McCall 1981) y se mantiene durante toda la etapa escolar (Alexander et al. 1993, Duncan et al. 1994; Zill et al. 1995). Los niños de familias con bajo SES tienen más probabilidades de presentar problemas en el lenguaje oral y escrito (Snow et al. 1998; Whitehurst y Lonigan 1998). En particular, se ha encontrado relación entre SES y amplitud del léxico (Hart y Risley 1995) y entre SES y comprensión de textos escritos en la etapa preescolar (Aikens y Barbarin 2008) y durante la primaria (Stanovich 1986). La explicación a las diferencias en el desarrollo cognitivo, en general, y lingüístico, en particular, debido al SES puede ser muy compleja. En primer lugar, se debe considerar que el tipo de lenguaje utilizado en la escuela (preescolar y escolar) es una de las variedades que ofrece la lengua en la que se instruye a los niños. Esta es una posible explicación a los resultados, como los presentados, que muestran que los niños de SES bajos tienen desempeños inferiores en pruebas de lenguaje; la desventaja de estos niños aparece solo con respecto al ambiente escolar y no al ambiente familiar o de su comunidad (Walker et al. 1994). Es decir, el desempeño es bajo cuando se les evalúa en la variedad utilizada en la escuela, pero no cuando se les evalúa en la variedad utilizada en casa o con su entorno más cercano. No conocemos estudios que confirmen estos resultados para

poblaciones de la región. Sin embargo, creemos en nuestra realidad también se pueda explicar pobre desempeño de los niños con bajo SES en pruebas lingüísticas con la diferencia entre el tipo del lenguaje utilizado en su comunidad y el tipo de lenguaje utilizado en la escuela (Feagans 1982).

Además, hay otras posibles explicaciones que considerar. Las familias con menores medios económicos pueden ser familias con menores recursos de literacidad y este puede ser el factor determinante de la diferencia. También es posible concebir situaciones en las que un conjunto de aspectos relacionados con la vida diaria de familias en desventaja obstaculicen el desarrollo de un niño, como el poco tiempo dedicado a él debido a una gran carga laboral de los padres o incluso su propia participación del mismo niño en actividades laborales.

### 2.3.3. Educación

La influencia de la educación en el desarrollo cognitivo y lingüístico depende principalmente de las actividades realizadas en este espacio, en particular en inicial y en los primeros años de primaria. Existe gran variabilidad entre las actividades realizadas en las escuelas de los diferentes países, por lo que para este acápite no revisaremos información proveniente de la literatura internacional. Preferimos remitirnos a las competencias que oficialmente deberían ser alcanzadas en la escuela inicial y los dos primeros años de primaria, basándonos en los lineamientos de las Rutas del Aprendizaje.

En las Rutas del Aprendizaje (2015a) se indica que los niños de 4 y 5 años deben tener las capacidades que describimos a continuación. Con respecto a los textos orales, deben poder “escucharlos activamente” “inferir su significado” y “reflexionar sobre la forma, contenido y contexto” de ellos. Con respecto a los textos escritos, se plantea una relación con ellos a través de la escucha de textos leídos por otros y la introducción a la propia lectura. Así, los niños deben poder “apropiarse del sistema de escritura” – es decir, conocer el uso social de los textos, identificar textos escritos de su entorno y saber que se lee de izquierda a derecha y de arriba hacia abajo, diferenciar palabras de imágenes –, “recuperar información de diversos textos escritos” – o sea, localizar donde hay información escrita –, “reorganizar la información e inferir el significado” de textos que le leen y “reflexionar sobre la forma, contenido y

contexto de los textos escritos”. Si estos objetivos se cumplen, esperamos que el recibir educación inicial tenga un fuerte y claro efecto en el nivel de comprensión de textos.

A los 6, 7 y 8 años (primero y segundo de primaria), según las Rutas del Aprendizaje (2015b), los niños deben ser capaces de comprender y producir textos orales y escritos. Con respecto a la comprensión, los niños deben ser capaces de escuchar activamente los textos orales; inferir su significado; reflexionar sobre su forma, contenido y contexto; y recuperar y organizar la información de diversas fuentes. Además, deben poder apropiarse del sistema de escritura; recuperar información de diversos textos escritos; inferir su significado; reflexionar sobre su forma, contenido y contexto; y reorganizar la información proveniente de diferentes fuentes. Si estos objetivos se cumplen, la asistencia a la escuela debería tener un impacto evidente en la habilidad de comprender textos, ya en los primeros años de la primaria.

En la descripción del problema de estudio, hemos visto que para el caso específico de los niños peruanos, la gestión de la institución educativa tiene efecto en el desempeño de la comprensión lectora. Los resultados de PISA 2001, 2009 y 2012, así como ECE 2013 y 2014 muestran brechas importantes que perjudican a los niños de escuelas públicas.

Otra diferencia importante que corresponde a las oportunidades del niño, y que está relacionada con el tipo de educación que reciben, es el lugar de residencia, que coincide con la ubicación de la escuela. Cuando se describió el problema de estudio se mencionó que PISA 2001, 2009 y 2012, así como ECE 2013 y 2014 señalan fuente brechas que desfavorecen a los niños de zonas rurales.

## LA METODOLOGÍA

Muchos niños desarrollan su lenguaje y, en particular, las habilidades de comprensión de textos sin mucho esfuerzo, incluso con poca instrucción; mientras muchos otros se enfrentan a grandes dificultades. La variabilidad es la característica principal del desarrollo lingüístico infantil. Si queremos designar recursos a mejorar la comprensión de textos, es importante encontrar la fuente de la variabilidad entre niños. Si, además, el desempeño de gran parte de los niños es bajo, como en el caso del Perú, con mayor razón necesitamos encontrar las variables relacionadas con dicho desempeño. Conocer las fuentes de variación permitirá la implementación de políticas públicas de intervención con altas expectativas de éxito.

Este estudio está diseñado para, a partir de los datos ya recogidos y organizados en la base Niños del Milenio, identificar fuentes de variación de la comprensión del texto en español. Hemos decidido utilizar solo la información del español, la lengua que fue elegida por más niños para tomar las pruebas de comprensión. Para identificar las fuentes de variación en el desempeño de estos niños, hemos extraído información sobre su nivel de desarrollo y sobre las oportunidades que tuvieron para desplegar sus habilidades de comprensión de textos; hemos clasificado esta información en variables. Evaluaremos cuáles de estas variables – algunas medidas cuando los niños eran preescolares, otras en simultánea con la variable dependiente – son buenos predictores de su desempeño en la comprensión de textos cuando su edad correspondía al segundo año de primaria.

A continuación describimos la base de datos, el grupo de participantes, las variables utilizadas y el tipo de análisis aplicado.

### 1. La base de datos

Para el análisis, se ha utilizado la base de datos Niños del Milenio, la cual contiene información obtenida a través de la administración de cuestionarios a niños, a sus padres y a las comunidades de residencia de las familias. El estudio Niños del

Milenio ha realizado cuatro rondas de encuesta, en los años 2002, 2006, 2009 y 2014 (ver tabla 4). La muestra fue seleccionada de manera aleatoria y se realizó, para la primera ronda, en 20 localidades de 14 regiones del Perú; pero, debido a la importante movilización de las familias durante los años, el número de localidades y regiones se amplió en las rondas siguientes. En comparación con otras encuestas nacionales, se estima que la representatividad de la encuesta Niños del Milenio es muy alta: oscila alrededor del 95% de la población infantil del Perú (excluyendo al 5% con más ingresos).

Los datos de esta base incluyen características demográficas, socioeconómicas y culturales de los hogares; servicios en el hogar, ingresos y gastos, disponibilidad de activos y capital social; estrategias para enfrentar adversidades; acceso, uso y calidad de servicios públicos y privados; percepciones sobre el bienestar psicosocial del individuo y su familia; así como datos antropométricos e indicadores del desarrollo cognitivo y del rendimiento escolar de los niños. Se han empleado cuestionarios distintos para los grupos entrevistados: la cohorte menor, con una edad entre 6 y 18 meses en la primera ronda; y la cohorte mayor con una edad entre 7 años y medio y 8 años y medio en la primera ronda; los padres o cuidadores; y los líderes de la comunidad (ver tabla 4).

Para nuestro estudio, se han utilizado los datos de la cohorte menor en las rondas 2 y 3 (en negritas en la tabla 4); se ha utilizado solo una parte de la enorme información de la base, descrita en la sección relativa a los datos. A continuación se describe cómo se hizo la selección de la información.

Tabla 4: Rondas y cohortes del estudio Niños del Milenio

Encuestas	Año	Cohorte menor	Cohorte mayor
Ronda 1	2002	6 a 18 meses	7 a 8 años
Ronda 2	2006-7	<b>4 a 5 años</b>	11 a 12 años
Ronda 3	2009	<b>7 a 8 años</b>	14 a 15 años
Ronda 4	2013	11 a 12 años	18 a 19 años
Ronda 5	2016	14 a 15 años	21 a 22 años

## 2. Los participantes

Los datos utilizados corresponden a los participantes de las rondas 2 y 3. Los participantes de la ronda 3 son un subgrupo de los participantes de la ronda 2, pues el objetivo del estudio Niños del Milenio era continuar con los mismos niños desde la primera ronda (la diferencia se debe solo a la pérdida de participantes). Nosotros utilizaremos los datos provenientes de las rondas 2 y 3. De este grupo, hemos seleccionado los datos de los niños que optaron por tomar las pruebas PPVT (Test de vocabulario en imágenes) y EGRA (Prueba diagnóstica de lectura inicial) en español y, además, eran hablantes nativos de español, pues nuestro estudio está centrado en el desarrollo lingüístico de los hispanohablantes. Por otro lado, hemos excluido a todos los niños cuyos datos sugieren un posible trastorno del lenguaje o, al menos, un retraso en su desarrollo. Nos hemos basado en las preguntas que se hicieron a los padres respecto a la habilidad de cada niño de comprender y ser comprendidos cuando habla: ¿puede oír tan bien como los otros niños de su edad? ¿Entiende lo que se le dice tan bien como los otros niños de su edad? ¿Puede hablar tan bien como los otros niños de su edad? ¿Se le entiende cuando habla? Se excluyeron los datos correspondientes a las respuestas: “absolutamente no” y “con dificultad”.

A continuación presentamos los estadísticos descriptivos del grupo definitivo de participantes en nuestro estudio.

El grupo está compuesto por 1583 participantes, de los cuales 798 (50.4%) son niños y 785 son niñas (49.6%). La edad del grupo varía, en la ronda 2, entre 53 y 74 meses cumplidos; correspondientes a 4 años y 5 meses; y 6 años y 6 meses, respectivamente. En la ronda 3, la edad varía entre 68 meses y 100 meses cumplidos; correspondientes a 6 años y 8 meses, y 8 años y 4 meses, respectivamente.

## 3. Diseño y materiales

Para este estudio, ha sido seleccionada y organizada parte de la información de la base de datos Niños del Milenio. Se emplean variables obtenidas, algunas veces, de manera directa de la base de datos y otras, producto de transformaciones. A continuación se describen las variables utilizadas y, en los casos que aplica, cuál fue la transformación que se realizó.

### 3.1. Las variables

#### 3.1.1. Género

La variable “género” indica si se trata de un niño o una niña. Los niños están identificados con el valor 0 y las niñas, con el valor 1. Para hacer más sencilla la interpretación de las correlaciones y regresiones, para esos análisis la variable está etiquetada como “Género: niña”.

#### 3.1.2. Variables relacionadas con el desarrollo del niño.

##### “Edad”

La variable edad ha sido tomada directamente de la base de datos de Niños del Milenio. Se ha utilizado la edad, calculada en meses, contabilizada en la ronda 2. Nos interesa el efecto de esta variable y de la variable “crecimiento” en la prueba de vocabulario, administrada en la misma ronda 2, pues nos da información sobre el desarrollo del niño.

##### “Crecimiento”

La variable que aquí llamamos “crecimiento”<sup>9</sup> equivale al puntaje z correspondiente a la relación entre la talla y la edad del niño (*height-for-age z-score*, HAZ), con respecto a la población de referencia. El cálculo del puntaje HAZ fue obtenido, por el estudio Niños del Milenio, basándose en las propias mediciones en la ronda 2 y siguiendo las indicaciones de *WHO International Growth Reference* (World Health Organization, 1997).

Esta es la segunda variable, después de la edad, que nos da información, aunque de manera indirecta, sobre el desarrollo cognitivo.

---

<sup>9</sup> Esta variable habría podido llamarse “estado nutricional” o “proxy de estado nutricional”, pues permite inferir el estado nutricional, y efectivamente, así ha sido utilizada en muchos estudios (ver sección 2.1.3., acápite “El crecimiento”). Sin embargo, hemos querido utilizar el nombre “crecimiento” para designar a esta variable porque nos ofrece información directa solo del crecimiento, es decir, si el niño ha alcanzado la altura esperada al momento de la medición.

### “Vocabulario”

Esta variable proviene del resultado del Test de vocabulario por imágenes Peabody (*Peabody Picture Vocabulary Test*, PPVT), en su versión para el español (Dunn et al., 1986). Este test mide la amplitud del vocabulario en comprensión: el niño escucha una palabra y tiene que indicar qué figura, de las cuatro que se le muestran, representa la palabra. Los ítems están ordenados por dificultad. El test prevé que se administre solo una parte de los ítems: la administración empieza con el ítem asignado como inicio para cada edad, a partir del cual se identifica el límite inferior o base (se asume que el niño conoce todas las palabras de los ítems anteriores) y el límite más alto o techo para cada niño (se asume que el niño desconoce todas las palabras de los ítems posteriores). El test ofrece puntajes brutos que se calculan restando el número de errores del ítem techo. La variable que utilizamos es el puntaje bruto y fue recogida en la ronda 2. Los puntajes posibles varían entre 0 y 125.

### “Texto oral” y “texto escrito”

Estas dos variables provienen del *Early Grade Reading Assessment* (EGRA), una evaluación desarrollada por USAID para medir los fundamentos básicos de la adquisición de la literacidad en los primeros años de formación escolar. EGRA está compuesto por varios subtests, de los cuales tres fueron traducidos y adaptados para el estudio Niños del Milenio: identificación de palabras familiares, lectura de pasaje y comprensión auditiva (Cueto et al., 2009). De los tres subtest administrados por Niños del Milenio en la ronda 3, hemos elegido solo los ítems correspondientes al segundo y tercero porque son aquellos que miden la comprensión de los textos leídos y oídos. En el segundo subtest, entre otras cosas, se pide al niño que lea un texto primero en voz alta y luego en voz baja; y que conteste a ocho preguntas sobre su contenido, formuladas oralmente por el entrevistador. Estas respuestas han sido clasificadas por nosotros como ítems de comprensión del texto escrito. En el tercer subtest, el entrevistador lee al niño un pequeño texto para que conteste a seis preguntas sobre su contenido; estos son nuestros ítems de comprensión del texto oral. Para cada tipo de comprensión hemos obtenido un promedio de respuesta sumando los puntajes y dividiéndolos entre el número de ítems; así, los puntajes pueden variar entre 0 (ninguna pregunta contestada correctamente) a 1 (todas las preguntas contestadas correctamente) para cada una de las variables.

### 3.1.3. Variables relacionadas con las oportunidades del niño.

En esta sección presentamos las variables relacionadas con las características maternas, del entorno y de la familia. Las primeras, para nosotros, son índices de los recursos de la madre para emplear prácticas de literacidad en casa. Hemos creado las variables “español oral de la madre”, “español escrito de la madre” y “educación de la madre” exclusivamente para los casos de los niños cuyo cuidador principal es la propia madre. La razón es que, si consideramos que estas variables nos dan información sobre los recursos de la madre para emplear prácticas de literacidad, tenemos que estar seguros de que la madre desempeña el rol principal en la crianza del niño.

#### *“Español oral de la madre”*

Esta variable refleja la competencia en español oral que tiene la madre del niño y corresponde a la información obtenida directamente de la madre con respecto a su propia competencia en español. La variable es una combinación de tres variables de la ronda 2 presentes en la base. La primera variable indica la lengua materna de la madre. Distingue entre español, quechua, aymara, lengua de la selva y combinaciones español-quechua y español-aymara. La segunda variable corresponde a la respuesta a la pregunta de si habla español. La tercera variable, en cambio, proviene de la pregunta sobre cuán bien habla español; la respuesta consistía en elegir entre: “fluentemente, como mi lengua materna”, “bien, entiendo sin problemas”, “básico, me cuesta entender”. La variable “español oral”, creada por nosotros, recoge la información de estas tres variables y tiene 4 posibles valores: 0 (no habla español), 1 (habla un español básico), 2 (habla un español bueno), 3 (habla un español fluido) y 4 (es hablante nativa de español).

#### *“Español escrito de la madre”*

Corresponde a la variable de la ronda 2 relativa a la pregunta que se les hizo a las madres de si podían leer y entender una carta o un periódico en español. Las posibles respuestas eran “absolutamente nada”, “con dificultad” , y “fácilmente” que corresponden a los valores 0, 1 y 2, respectivamente.

### *“Educación de la madre”*

Se recodificó la variable relativa a la educación de la madre que ofrece la base de datos. La nueva variable responde al nivel educativo más alto que haya alcanzado la madre y puede tomar los siguientes valores:

- 0 = Corresponde a ningún nivel educativo; no se considera ningún tipo de educación anterior a la primaria.
- 1 a 11 = Corresponden a los años de escuela frecuentados: 1 para primero de primaria, 2 para segundo, 3 para tercero y así, sucesivamente, hasta el final de la secundaria.
- 13 y 14 = Corresponden a estudios incompletos o completos de escuela técnica, respectivamente.
- 15 y 16 = Corresponden a estudios incompletos o completos de universidad, respectivamente.

En la nueva variable han sido excluidos los programas adultos de literacidad y los estudios calificados como “otros”.

### *“Años de educación”*

Esta variable hace referencia a la asistencia del niño a alguna institución educativa desde los tres años hasta el momento de la entrevista de la ronda 3, es decir, incluyendo el año en curso. Las instituciones pueden ser nido, jardín de infancia, preescolar o escuela primaria; pero no guarderías o cunas. El puntaje indica el número de años y puede variar de 0 a 7 años.

### *“Tipo de institución educativa”*

La base de datos de Niños del Milenio nos ofrece información, recogida en la ronda 3 sobre el tipo de institución educativa a la que asistieron los niños durante los años anteriores. A partir de esa información, hemos creado una nueva variable con dos posibles valores. El valor “privada” indica que el tipo de institución a la que el niño asistió durante más tiempo desde que empezó a su formación (preescolar y/o escolar)

era privada; el valor “pública”, indica que se trataba de una institución pública. Para hacer más sencilla la interpretación de las correlaciones y regresiones, para esos análisis la variable está etiquetada como “Tipo de institución: privada”.

#### *“Estatus socioeconómico” (SES)*

Esta variable corresponde al índice de riqueza calculado por el estudio Niños del Milenio. Este índice tiene la intención de ser el instrumento principal para medir el estatus socioeconómico de la familia. Fue construido sobre la base de la calidad de la vivienda donde residía el niño, el acceso a los servicios y la posesión de bienes de consumo de larga duración que tenía la familia. Está basado en los trabajos del Banco Mundial y Macro Internacional usados para desarrollar el índice de riqueza citado en las Encuestas de Indicadores Múltiples por Conglomerados (MICS) (Young Lives 2001 y Espinoza 2014). Utilizamos la información del estatus socioeconómico medido en la ronda 2.

#### *“Lugar de residencia”*

Esta variable hace referencia al lugar de residencia de los niños cuando fueron encuestados en la ronda 2. La variable tiene dos valores posibles: 0, que hace referencia a las zonas urbanas; y 1, a las zonas rurales. Para hacer más sencilla la interpretación de las correlaciones y regresiones, para esos análisis la variable está etiquetada como “Lugar de residencia: rural”.

### 3.2. El diseño

El objetivo de este estudio es identificar los predictores del texto escrito entre tres tipos de variables. En primer lugar, tenemos las variables relacionadas a las características del niño: género; edad, crecimiento y desarrollo cognitivo (índices de desarrollo, de la ronda 2). La información sobre el poder predictivo de estas variables nos permitirá identificar a los niños que están sujetos a un riesgo mayor de no aprender a comprender textos escritos.

En segundo lugar, están las variables relativas a las oportunidades del niño y que se subdividen en tres tipos. El primer tipo está compuesto por las características de la madre – español oral, español escrito y educación – y se relaciona con la capacidad del principal cuidador – en todos los casos, la madre – de emplear prácticas de literacidad con el niño. El segundo tipo está formado por el estatus socioeconómico y el lugar de residencia. El último tipo constituye la educación que ha recibido el niño y está compuesto por la variable que contabiliza los años de educación recibida y la que define el tipo de institución educativa a la que asistió, hasta el momento de la entrevista en la ronda 3. La información sobre el poder predictivo de estas variables nos permitirá, en algunos casos, identificar a los niños a riesgo; en otros, identificar factores sobre los que se podrá intervenir directa o indirectamente. Con respecto a la identificación de grupos de riesgo, podemos tomar como ejemplo los resultados sobre las variables estatus socioeconómico y lugar de residencia: saber que ser más pobre que otros o residir en una zona rural disminuye las probabilidades de un niño de tener un vocabulario amplio<sup>10</sup>, nos llevará a proponer intervenciones focalizadas en los grupos rurales y de menos recursos. En relación a la identificación de factores modificables, tenemos que es posible proponer intervenciones directas o indirectas. Encontrar que el tiempo que un niño recibió educación en una institución y que el grado de educación que consiguió su madre predicen, aunque sea en parte, la futura comprensión del texto oral del niño, nos llevará a sugerir intervenciones en estos factores. Por ejemplo, de manera directa se puede intervenir para garantizar que los niños reciban educación por los años necesarios; de manera indirecta, se puede ofrecer a las madres entrenamiento para emplear prácticas de literacidad aunque no posean de antemano las herramientas necesarias.

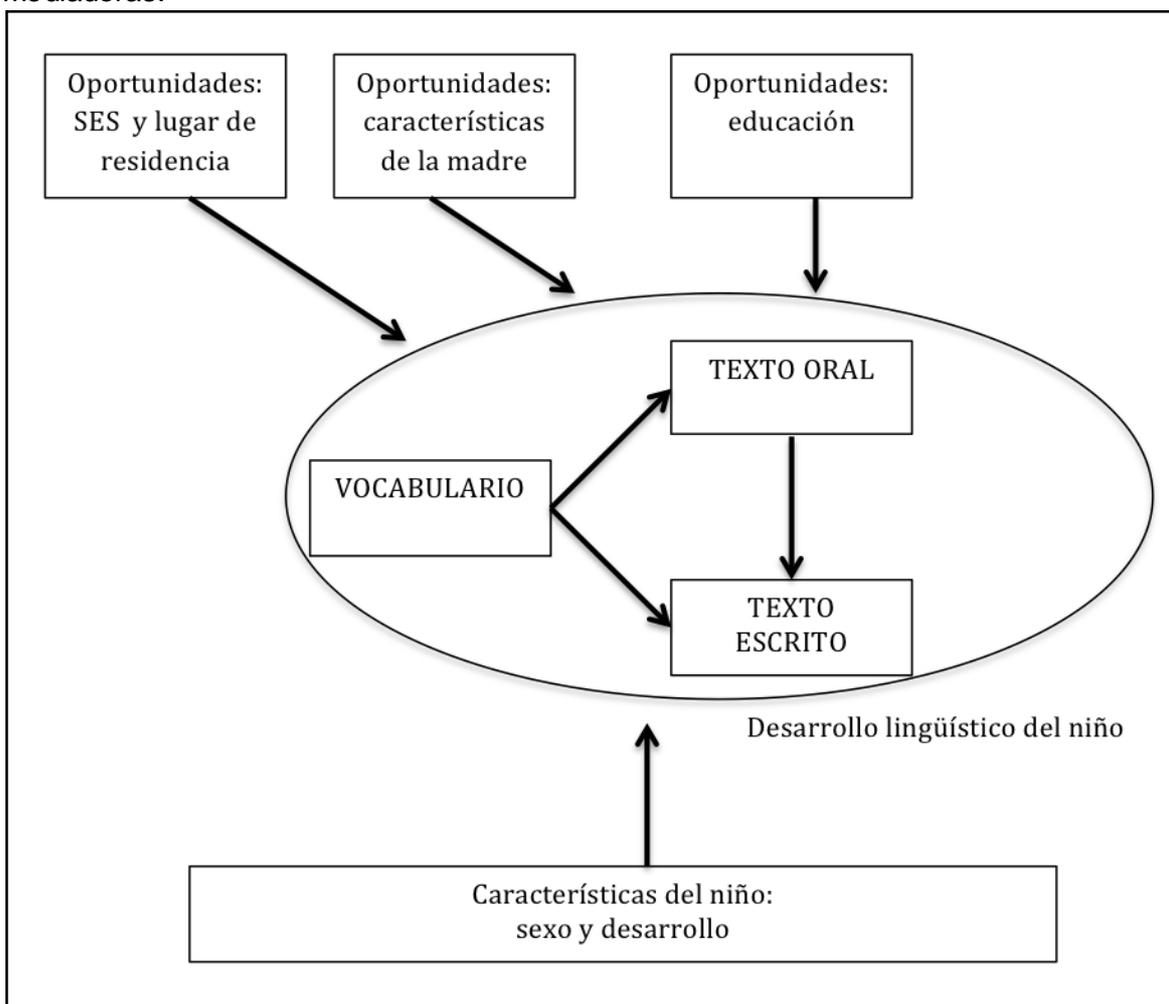
En tercer lugar, están las variables de desarrollo lingüístico – vocabulario y texto oral – que, en nuestro modelo, desempeñan el rol tanto de variables independientes como de variables dependientes. La información sobre el poder predictivo de estas variables nos permitirá identificar los principales factores modificables y sobre los que se propondrá intervenir de forma directa y específica.

La relación entre las variables que proponemos y que evaluamos con el modelo de ecuaciones estructurales se puede describir como en el gráfico 2.

---

<sup>10</sup> Nos referimos, exclusivamente, al tipo de vocabulario medido por la prueba.

Gráfico 2: Modelo de relación entre variables independientes, dependientes y mediadoras.



#### 4. Los análisis

Con el objetivo de evaluar el modelo que proponemos y de ofrecer información suficiente para el lector sobre los datos, hemos realizado los siguientes análisis. En primer lugar, describimos la distribución de las variables (estadísticos descriptivos). En segundo lugar, presentamos las correlaciones existentes entre todas las variables utilizadas. En tercer lugar, presentamos la comparación de medias. Hemos dividido el grupo total por género (niño-niña), por lugar de residencia (urbano-rural) y por tipo de institución a la que mayormente han asistido hasta el momento de la entrevista de la ronda 3 (pública-privada) y hemos comparado el desempeño de cada par de grupos en el vocabulario, la comprensión de texto oral y de texto escrito. Hemos realizado estas comparaciones porque los antecedentes sobre el desempeño de los niños peruanos muestran brechas en el desarrollo y la adquisición de habilidades lingüísticas, en particular, en la comprensión de textos escritos. Las diferencias juegan a favor de las niñas, de los residentes en zonas urbanas y de los estudiantes de escuelas privadas, en cada grupo. En cuarto lugar, presentamos las regresiones que nos permitirán identificar el poder predictivo de las variables seleccionadas sobre vocabulario, texto oral y texto escrito para finalmente, presentar el modelo completo y las relaciones entre todas las variables relevantes.

##### 4.1. Estadísticos Descriptivos

En la sección relativa a los resultados, se presentan los estadísticos descriptivos según el tipo de variables. Para las variables nominales, las cuales además son dicotómicas, se ofrecen las frecuencias y los porcentajes; para las ordinales, el rango y la moda; y para las cuantitativas, el rango, la media y la desviación estándar (sd). Se ofrece, además, información sobre el número de niños de los cuales se tiene datos válidos para cada variable.

##### 4.2. Correlaciones

Para mostrar las correlaciones entre las variables utilizadas, presentamos diferentes índices, según el tipo de variables que se relacionen:

- R de Pearson: entre variables cuantitativas.
- Rho de Spearman: variables cuantitativas con ordinales; y entre variables ordinales.
- Punto biserial rb: variables cuantitativas con nominales.
- Lambda: variables ordinales con nominales.
- Phi: entre variables nominales.

#### 4.3. Comparación de medias

Se comparó las medias de las variables dependientes (vocabulario, texto oral y texto escrito) – dividiendo el grupo por género, lugar de residencia y tipo de institución – a través de un análisis multivariado de covarianza (MANCOVA) que permitió controlar por edad.

#### 4.4. Modelo de ecuación estructural (o SEM)

Para evaluar el modelo que postulamos, hemos utilizado un tipo de análisis (SEM) que nos permite medir la relación entre las variables, a través de un sistema de ecuaciones con variables independientes y dependientes; y en el cual las dependientes pueden ser también predictores de otras variables (Schumacker y Lomax 2004). En particular, evaluamos el poder de predicción de las variables que presentamos en la tabla 5 sobre las variables vocabulario, texto oral y texto escrito. Se indica con un “SÍ” las variables que son consideradas posibles predictores, en la columna respectiva; y con un “NO”, las que no lo son.

Tabla 5: Variables predictoras (en la primera columna) y variables dependientes (en la primera fila).

	Vocabulario	Texto oral	Texto escrito
Edad	SÍ	SÍ	SÍ
Género (o “niña”)	SÍ	SÍ	SÍ
Crecimiento	SÍ	SÍ	SÍ
Español oral de la madre	SÍ	SÍ	SÍ
Español escrito de la madre	SÍ	SÍ	SÍ
Educación de la madre	SÍ	SÍ	SÍ
Años de educación	NO	SÍ	SÍ
Tipo de institución educativa (o “privada”)	NO	SÍ	SÍ
Estatus socioeconómico (SES)	SÍ	SÍ	SÍ
Lugar de residencia (o “rural”)	SÍ	SÍ	SÍ
Vocabulario	NO	SÍ	SÍ
Texto oral	NO	NO	SÍ
Texto escrito	NO	NO	NO

En la sección resultados, se presentan los resultados para cada ecuación (una para vocabulario, otra para texto oral y otra para texto escrito), separados en grupo completo, niños y niñas; y luego, el modelo total. Además de presentar tablas con los principales coeficientes, en el texto se describen los resultados. En particular, se presenta la bondad del ajuste; los coeficientes  $\beta$  no estandarizado y  $\beta$  estandarizado ( $\beta_e$ ); y el error estándar (EE). El primero indica cuán explicativo es el modelo, es un factor que muestra la reducción de la incertidumbre gracias al modelo y las variables independientes incluidas en él. El segundo indica la relación entre la variable independiente y la variable dependiente consideradas, el signo de esta relación, y su fuerza.

## LOS RESULTADOS

### 1. Estadísticos Descriptivos

#### 1.1. Lugar de residencia y tipo de institución educativa.

La información sobre el género ha sido presentada en la descripción de los participantes.

Con respecto al lugar de residencia, tenemos que gran parte de los niños provienen de zonas urbanas: el 78% de los niños (1202) provienen de ciudades y el 21,24% (327) de zonas rurales.

Con respecto a la gestión de las instituciones educativas (preescolares y escolares) a las que asistieron los niños, tenemos la siguiente información. Del total de niños que asistieron a instituciones públicas o privadas, tenemos que el 81,2% (1279) asistió principalmente a instituciones públicas y el 18,8% (297), a instituciones privadas.

#### 1.2. Español oral y escrito de la madre, y su nivel de educación.

La información sobre las variables español oral de la madre, español escrito de la madre y educación de la madre se presentan en la tabla 6.

Tabla 6: Rango y moda para las variables español oral, español escrito y educación de la madre.

	N	Rango	Moda
Español oral de la madre	1529	0 – 4	4
Español escrito de la madre	1451	0 – 2	2
Educación de la madre	1447	0 – 16	11

### 1.3. Desarrollo cognitivo, crecimiento, educación, estatus socioeconómico, vocabulario texto oral y escrito.

En relación a estas variables, presentamos la información en la tabla 7

Tabla 7: Rango, media y desviación típica para las variables desarrollo cognitivo, crecimiento, años de educación, SES, vocabulario, texto oral y texto escrito

	N	Rango	Media	Desviación típica
Crecimiento	1527	-1,39 – 4,50 <sup>11</sup>	-1,40	1,09
Años de educación	1576	0 – 8	3,99	1,35
SES	1529	0 – 0,92	0,50	0,23
Vocabulario	1529	0 – 87	30,88	18,00
Texto oral	1449	0,17 – 1,00	0,75	0,16
Texto escrito	1165	0 – 1	0,69	0,21

## 2. Correlaciones

Por comodidad de lectura, los índices de correlación entre las variables se presentan en los anexos 3, 4, 5 y 6.

## 3. Comparación de medias

Los resultados de MANCOVA mostraron que hubo una diferencia estadísticamente significativa en las pruebas de vocabulario, comprensión de texto oral y comprensión de texto escrito en relación a la variable género,  $F(3, 1071) = 2,965$ ,  $p < .05$ ; Wilk's  $\Lambda = 0.992$ ,  $\eta^2$  parcial = .008 y a la variable lugar de residencia,  $F(6, 2142) = 5,864$ ,  $p < .0001$ ; Wilk's  $\Lambda = 0.968$ ,  $\eta^2$  parcial = .016; pero no en relación a la variable tipo de institución educativa.

Con respecto al género, tenemos que, a pesar de que los resultados sugieren diferencias en las evaluaciones consideradas; las posteriores pruebas inter-sujetos no muestran que el género tenga algún efecto en las evaluaciones, cuando utilizamos un  $p < .017$ , producto de la corrección de Bonferroni para comparaciones múltiples.

---

<sup>11</sup> Los valores fuera del rango -5 a 5 han sido considerados errores de medición. Esto ha llevado a excluir a 3 niños; sus puntajes fueron -9,87, -6,35 y -5,38.

Además, tenemos que las diferencias debido al lugar de residencia se encuentran solamente en las pruebas de vocabulario ( $F(2, 1073) = 16,285; p < .0001; \eta^2 \text{ parcial} = .029$ ) y no en las pruebas de comprensión de textos.

#### 4. Modelo de ecuaciones estructurales

##### 4.1. Regresiones para “vocabulario”

A continuación, se presentan los resultados de las regresiones con variable dependiente vocabulario para el grupo total de niños (los resultados separados por niños y niñas aparecen en el anexo 7). El modelo explicó un porcentaje muy importante de la variabilidad, el 51%, debido a las contribución de las variables edad, crecimiento, español escrito de la madre, educación de la madre, estatus socioeconómico de la familia y lugar de residencia, como se muestra en la tabla 8. Con respecto a las variables que favorecen el vocabulario, tenemos que, manteniendo constante los otros factores, el aumento de una desviación estándar (sd) en la medida de crecimiento, aumenta de 0,10 sd el vocabulario. Análogamente, el aumento de una desviación estándar en el nivel de educación de la madre genera un aumento de 0,29 sd en el vocabulario; y en el aumento en el estatus socioeconómico, un 0,28 sd. En cambio, las variables que desfavorecen el desarrollo son el español escrito de la madre y el vivir en una zona rural. Por un lado, un aumento en una desviación estándar en el nivel del español escrito de la madre significa una disminución de 0,09 sd en el vocabulario. Por otro lado, ante un cambio discreto de 0 a 1 en la variable “lugar de residencia: rural” (de vivir en una zona urbana a vivir en una zona rural) se genera una disminución de 0,10 sd en el vocabulario.

Tabla 8: Regresión para toda la muestra con vocabulario como variable dependiente.

	R <sup>2</sup>	$\beta$	$\beta^e$	EE
<b>Vocabulario</b>	0,510			
Edad		1,180	0,308***	0,021
Género: niña		-0,413	-0,012	0,022
Crecimiento		1,714	0,102***	0,025
Español oral de la madre		-0,071	-0,004	0,026
Español escrito de la madre		-3,808	-0,089**	0,027
Educación de la madre		1,255	0,291***	0,031
Estatus socioeconómico		22,860	0,276***	0,032
Lugar de residencia: rural		-4,500	-0,091***	0,025

\*  $p < ,05$ ; \*\*  $p < ,01$  ; \*\*\*  $p < ,001$

#### 4.2. Regresiones para “texto oral”

Ahora presentamos los resultados de las regresiones con variable dependiente texto oral para el grupo total de niños, en la tabla 9 (los resultados separados por género aparecen en el anexo 8). La variabilidad explicada por el modelo fue muy baja, de solo 2%. Contribuyeron variables que favorecen el texto oral, educación de la madre y número de años que el niño recibió educación; y la única variable del modelo que desfavorece el desempeño lingüístico, el español escrito de la madre. Si se mantienen constantes los otros factores, el aumento en una desviación estándar del nivel de educación de la madre ocasiona un aumento de 0,10 sd en el texto oral; y, de manera análoga, un aumento de una desviación estándar en años de educación del niño produce un incremento de 0,08 sd en su vocabulario. De manera inversa, un aumento de una desviación estándar del español escrito de la madre significa una disminución de 0,09 sd en el vocabulario del niño.

El vocabulario no contribuyó significativamente a la explicación del modelo; no hubo ningún efecto indirecto de las variables predictoras del vocabulario.

Tabla 9: Regresión para toda la muestra con texto oral como variable dependiente.

	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Texto oral</b>	0,02			
Género: niña		<-0,001	-0,001	0,032
Crecimiento		0,002	0,010	0,035
Español oral de la madre		0,008	0,049	0,042
Español escrito de la madre		-0,036	-0,092*	0,037
Educación de la madre		0,004	0,098*	0,046
Años de educación		0,009	0,080*	0,037
Tipo de institución: privada		0,028	0,075	0,041
Estatus socioeconómico		-0,035	-0,047	0,046
Lugar de residencia: rural		0,027	0,061	0,036
Vocabulario		<-0,001	-0,006	0,043

\*  $p < ,05$ ; \*\*  $p < ,01$ ; \*\*\*  $p < ,001$

#### 4.3. Regresiones para “texto escrito”

A continuación se muestran los resultados de las regresiones para texto escrito en el grupo total de niños. El modelo explicó el 20% de la variabilidad, debido a la contribución del género y de los puntajes en vocabulario y texto oral. Por un lado, un cambio discreto en el género (de ser niño a ser niña) significa un incremento en 0,09

sd el puntaje en texto escrito. Por otro lado, un aumento de una desviación estándar en vocabulario y texto oral significa un incremento de 0,32 y 0,25, respectivamente en texto escrito, como se muestra en la tabla 10.

Tabla 10: Regresión para toda la muestra con texto escrito como variable dependiente.

	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Texto escrito</b>	0,196			
Género: niña		0,036	0,089**	0,028
Crecimiento		<-0,001	<-0,001	0,031
Español oral de la madre		-0,006	-0,030	0,039
Español escrito de la madre		0,012	0,025	0,043
Educación de la madre		0,001	0,022	0,041
Años de educación		<-0,001	-0,002	0,033
Tipo de institución: privada		0,003	0,007	0,035
Estatus socioeconómico		0,006	0,006	0,041
Lugar de residencia: rural		-0,029	-0,050	0,037
Vocabulario		0,004	0,319***	0,036
Texto oral		0,318	0,246***	0,029

\*  $p < ,05$ ; \*\*  $p < ,01$ ; \*\*\*  $p < ,001$

Además, los análisis mostraron un efecto indirecto positivo de las variables edad ( $\beta = 0,004$ ; EE = < 0,001); crecimiento ( $\beta = 0,007$ ; EE = 0,002); educación de la madre ( $\beta = 0,006$ ; EE < 0,001); años de educación que recibió el niño ( $\beta = 0,003$ ; EE = 0,001); estatus socioeconómico de la familia ( $\beta = 0,0074$ ; EE = 0,017); y un efecto indirecto negativo del español escrito de la madre ( $\beta = -0,025$ ; EE = 0,007). La tabla completa está en el anexo 9.

Ya que en la regresión para el grupo total de niños, la variable género contribuye a la variabilidad, presentamos también las regresiones separadas para niños y niñas. Como se ve en las tablas 11 y 12, los modelos para varones y mujeres explican el 20% y el 18% de la variabilidad, respectivamente. En el caso de los niños, las contribuciones importantes son nuevamente el vocabulario y el texto oral. Un incremento de una desviación estándar en vocabulario significa un aumento de 0,39 en texto escrito y un incremento de 1 sd en texto escrito, un aumento de 0,26 sd en texto escrito. El análisis nos muestra, además, efectos indirectos de la edad ( $\beta = 0,005$ ; EE < 0,001); el crecimiento ( $\beta = 0,014$ ; EE = 0,004); el nivel de educación de la madre ( $\beta = 0,005$ ; EE = 0,001); y el estatus socioeconómico de la familia ( $\beta = 0,980$ ; EE = 0,028). La tabla completa está en el anexo 10.

Para las niñas la regresión muestra que como para los niños, las contribuciones importantes provienen del vocabulario y el texto oral. Análogamente un aumento de 1 sd en vocabulario y en texto oral, generan, respectivamente, un incremento de 0,26 sd y 0,23 sd en texto escrito.

En el caso de las niñas, también encontramos efectos indirectos positivos de la edad ( $\beta = 0,005$ ; EE = 0,001), la educación de la madre ( $\beta = 0,007$ ; EE = 0,001), el estatus socioeconómico de la familia ( $\beta = 0,054$ ; EE = 0,020) y efectos negativos del español escrito de la madre ( $\beta = -0,003$ ; EE = 0,008). La tabla completa está en el anexo 10.

Tabla 11: Regresión para los niños con texto escrito como variable dependiente.

NIÑOS	R <sup>2</sup>	$\beta$	$\beta^e$	EE
<b>Texto escrito</b>	0,196			
Crecimiento		0,003	0,015	0,045
Español oral de la madre		0,001	0,005	0,057
Español escrito de la madre		-0,012	-0,024	0,060
Educación de la madre		<0,001	0,011	0,057
Años de educación		-0,004	-0,028	0,048
Tipo de institución: privada		-0,015	-0,029	0,053
Estatus socioeconómico		-0,036	-0,036	0,059
Lugar de residencia: rural		-0,030	-0,053	0,051
Vocabulario		0,005	0,389***	0,050
Texto oral		0,361	0,264***	0,040

\*  $p < ,05$ ; \*\*  $p < ,01$ ; \*\*\*  $p < ,001$

Tabla 12: Regresión para las niñas con texto escrito como variable dependiente.

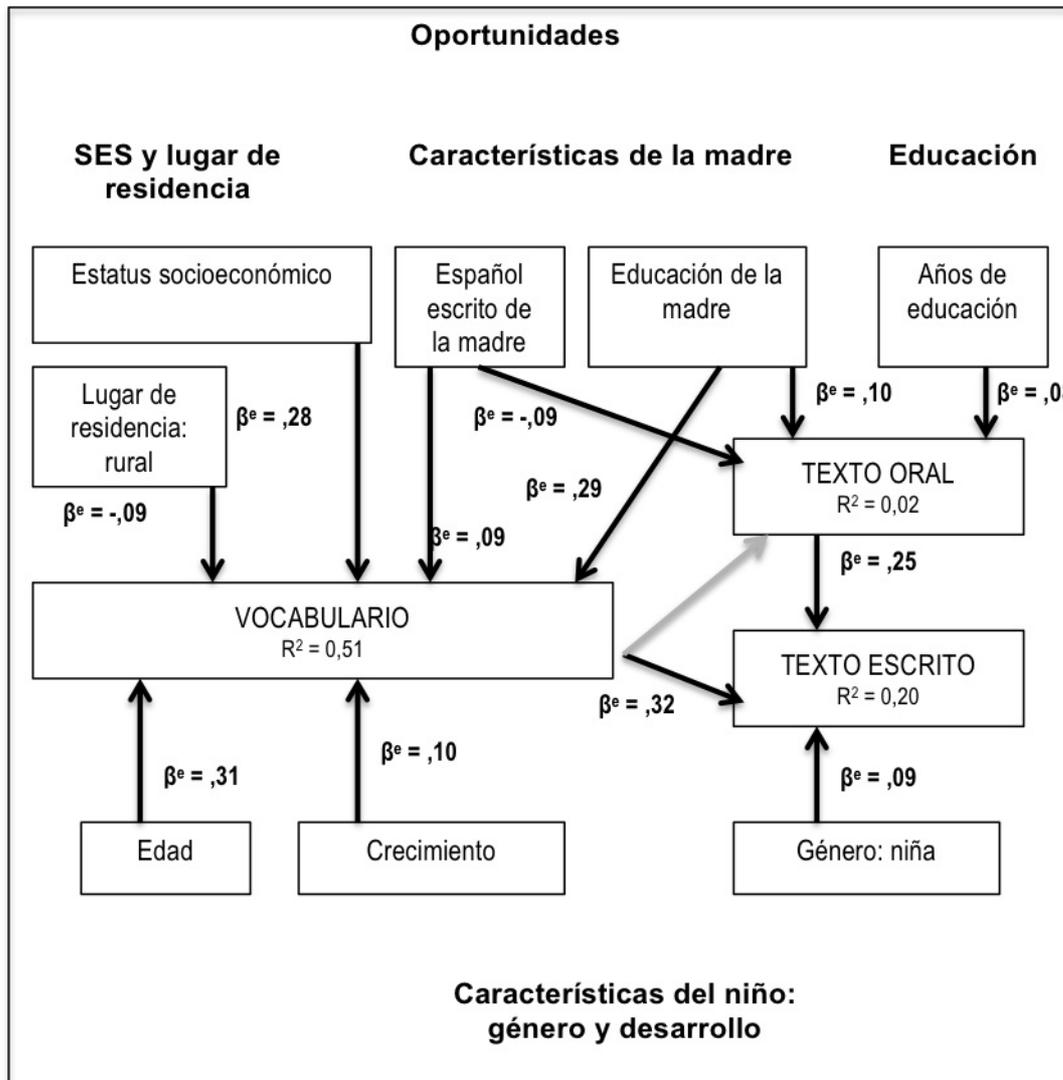
NIÑAS	R <sup>2</sup>	$\beta$	$\beta^e$	EE
<b>Texto escrito</b>	0,178			
Crecimiento		-0,005	-0,025	0,043
Español oral de la madre		-0,014	-0,067	0,053
Español escrito de la madre		0,030	0,063	0,060
Educación de la madre		0,003	0,052	0,059
Años de educación		0,003	0,022	0,047
Tipo de institución: privada		0,018	0,039	0,046
Estatus socioeconómico		0,046	0,054	0,057
Lugar de residencia: rural		-0,025	-0,049	0,055
Vocabulario		0,003	0,255***	0,050
Texto oral		0,277	0,227***	0,042

\*  $p < ,05$ ; \*\*  $p < ,01$ ; \*\*\*  $p < ,001$

#### 4.4. Modelo completo

En el gráfico 3, se presentan las relaciones entre las variables que resultaron ser predictivas de vocabulario, texto oral y texto escrito y estas últimas. Se ofrece información sobre el nivel de ajuste de cada modelo ( $R^2$ ) y la variación que aporta el aumento en una desviación estándar de la variable en cuestión sobre la variable dependiente ( $\beta^e$ ).

Gráfico 3: Modelo de ecuaciones estructurales para las variables vocabulario, texto oral y texto escrito para el grupo total de niños.



## DISCUSIÓN

Los resultados de los análisis presentan un panorama complejo para el que ofreceremos una interpretación en esta sección. Se empezará por la discusión de los resultados de la comparación de medias para seguir con los resultados de las regresiones.

1. Comparación entre grupos niño-niña, urbano-rural, institución pública-privada con respecto al vocabulario; y la comprensión de textos orales y escritos

Los resultados mostraron que la única diferencia en las pruebas de vocabulario, texto oral y texto escrito, entre grupos diferenciados por género, lugar de residencia y tipo de institución educativa, se encuentra en el vocabulario: los niños de zonas urbanas tienen un desempeño significativamente mejor que los niños de zonas rurales. Estos resultados coinciden en parte con los antecedentes al respecto. Como mencionamos en la sección correspondiente, con respecto al desarrollo lingüístico, en particular a la comprensión de textos, el desempeño de los estudiantes peruanos presenta importantes brechas entre niños y niñas, a favor de las niñas (pruebas PISA 2001, 2009 y 2012; y ECE 2013 y 2014) o performances equivalentes (pruebas SERCE y TERCE); brechas entre estudiantes de zonas rurales y zonas urbanas, a favor de estos últimos y entre inscritos a escuelas públicas y privadas, a favor de los alumnos de instituciones privadas (pruebas PISA 2001, 2009 y 2012; y ECE 2013 y 2014; Guerrero et al. 2012). Nuestros resultados, en cambio, no muestran ninguna de estas diferencias con respecto al texto escrito ni oral, solo en relación al vocabulario y para la variable zona de residencia. A continuación se comenta, en primer lugar, la diferencia entre niños urbanos y rurales en el desempeño de la prueba de vocabulario, y en segundo lugar, la ausencia de diferencias en las otras comparaciones.

En primer lugar, la prueba con la que se midió el léxico, el Test de vocabulario por imágenes Peabody, es muy útil porque presenta palabras de variedades estándar de español y conceptos, incluso abstractos, muy utilizados en ámbitos como el escolar. Así, nos da información sobre el conocimiento que tiene el niño del tipo de léxico que necesita en la escuela. Sin embargo, esta prueba no mide el conocimiento

de palabras de variedades regionales ni conceptos de ámbitos específicos. Es posible que este tipo de prueba mida solo una parte del vocabulario de los niños rurales, dejando de lado la parte del léxico característico de la variedad de la zona y los conceptos que no se consideran en ámbitos como el escolar, pero que están presentes cuando se habla del entorno natural y social del niño. Naturalmente, los niños urbanos también tienen una variedad regional y acceso a conceptos diferentes a los del ámbito escolar, pero, probablemente, estos niños tienen más contacto con la variedad estándar y los conceptos de la escuela que los otros niños debido a la mayor apertura y las mayores oportunidades que les ofrecen las ciudades.

La brecha entre niños urbanos y rurales se encontró solamente en la prueba de vocabulario. En nuestro estudio, los niños y las niñas, así como los estudiantes de escuelas privadas y escuelas públicas tuvieron un desempeño equivalente en las tres medidas de desarrollo lingüístico; además, con respecto a las pruebas de comprensión textual, los niños urbanos y rurales tuvieron puntajes equivalentes. La ausencia de brechas entre niños y niñas está en línea con resultados de estudios precedentes, los cuales, dependiendo de la medida encuentran equivalencias o ventajas de las niñas. La ausencia de brechas entre niños que asisten a escuelas privadas y públicas – en línea con la tendencia hacia la reducción – debe ser interpretada con cautela, pues ya el Ministerio de Educación (MINEDU 2012) advierte que las estimaciones están muy influenciadas por la heterogeneidad entre colegios privados y el aumento de matrículas en ese tipo de instituciones. Por otro lado, aunque no se hayan identificado brechas entre los grupos comparados, es posible que estas aparezcan al combinar los factores. Los factores considerados suelen estar asociados, y además relacionados con un factor que no hemos considerado en este estudio: el bilingüismo. Por diseño, hemos excluido a los niños bilingües, población que frecuentemente agrupa características asociadas a un pobre desempeño en el desarrollo lingüístico y académico: un bajo estatus socioeconómico, residencia en una zona rural, madres con niveles bajos de educación formal y, hablantes nativas de lenguas originarias. Guerrero et al. (2012) demostró que el grupo de niños que comparte todas estas características lleva una vida más difícil que los demás niños en el Perú. Entre estos niños hay mayor desnutrición, menor acceso a educación inicial, mayor extra-edad en primaria y mayor deserción escolar en secundaria; todos factores asociados a un menor desarrollo lingüístico y a un menor aprendizaje (Guerrero et al. 2012). Futuros estudios deberían incorporar los datos de niños bilingües, incluyendo

medidas de lenguaje en ambas lenguas, de tal forma que se logre distinguir la influencia de los factores mencionados, de la primera lengua y de la segunda lengua.

## 2. El vocabulario y sus variables predictoras

El modelo explicativo propuesto tuvo un excelente nivel de predicción (51%) para el desempeño de los niños en la prueba de vocabulario, debido a la contribución de variables relacionadas a las características del niño y a sus oportunidades.

En primer lugar, como era de esperarse, las variables edad, crecimiento, que definimos indicadores indirectos del desarrollo cognitivo, predijeron un cambio importante ( $\beta^e = 0,31$  y  $\beta^e = 0,10$ , respectivamente). Esto indica, simplemente, que con el desarrollo los niños mejoran su vocabulario.

En segundo lugar, tenemos las variables que más nos interesan, aquellas relacionadas con las oportunidades del niño. Con respecto a las características de la madre, vemos que su nivel de educación predice un cambio importante en el vocabulario del niño ( $\beta^e = 0,29$ ). Interpretamos estos datos, proponiendo que las madres con mayor educación tendrán más herramientas para emplear prácticas de literacidad con sus hijos, y esto tendrá un efecto en el desarrollo de su vocabulario. Con respecto al español de la madre, en cambio, tenemos resultados inesperados. Por un lado, el español oral no contribuye a predecir la variabilidad en las pruebas de vocabulario; por otro, el español escrito predice un cambio, aunque muy pequeño ( $\beta^e = 0,01$ ), en el sentido contrario de lo que esperábamos. Es decir, las madres con un nivel de español más bajo tienen más probabilidades de tener hijos con buen vocabulario que las madres con un nivel de español más alto. Nos planteamos que esto se deba a la combinación de dos razones. La primera es el limitado número de niveles de esta variable (3) y su distribución en la muestra: el 81% de las madres contestó que leía español sin dificultad, mientras que solo un 6,2% y un 4,8% respondieron que lo hacían con dificultad o que no sabían leer en absoluto (en español), respectivamente. La segunda razón, que se combina con la primera, es que es posible que haya otras características asociadas a este 11% de madres, principales cuidadores<sup>12</sup> que no han sido medidas durante las entrevistas o que no hemos incluido

---

<sup>12</sup> Recordemos que la información sobre las madre fue utilizada solo para los casos en los que la esta era la cuidadora principal del niño.

aquí. Podemos imaginar, por ejemplo, que en este grupo de madres haya muchas que pasan gran parte del tiempo en casa con sus hijos. El no leer español o leerlo con dificultad reduce fuertemente el campo laboral al que puede acceder una mujer joven y, posiblemente, quedarse en casa con los hijos es una de las pocas opciones restantes. Las madres que dedican la mayor parte de su tiempo a sus hijos tienen más momentos para emplear prácticas de literacidad que las madres que pasan menos tiempo en casa. Es posible que, aunque estas madres no sean las que tienen los mayores recursos para emplear este tipo de prácticas, sí sean las que tienen más oportunidades para hacerlo, mejorando así el vocabulario de sus niños.

En tercer lugar, el estatus socioeconómico (SES) de la familia y el lugar de residencia implican cambios en el vocabulario. El estatus socioeconómico genera un cambio muy importante ( $\beta^e = 0,29$ ) en las pruebas de vocabulario, mientras que la el lugar de proveniencia, uno pequeño ( $\beta^e = -0,09$ ). El primer resultado es consistente con los estudios que demuestran que niños de familias con SES bajos tienen un desempeño inferior al de niños de familias con SES más alto, cuando se les evalúa en la variedad utilizada en la escuela (Walker et al. 1994; Feagans 1982; Hart y Risley 1995). Además, para el caso peruano, muchos estudios han confirmado la relación entre SES y comprensión lectora (Cueto et al. 2005; MINEDU 2004, 006 y 2015; Morales 2009 y 2011).

Análogamente, el lugar de residencia ha mostrado ser determinante en la comprensión lectora de los niños del Perú (Cueto et al. 2005; MINEDU 2004, 2006 y 2015). Además, como mencionamos en la sección anterior, la combinación SES bajo-residencia rural, unida a los factores con los que suele estar asociado (madre con baja educación formal, etc.) aumenta las probabilidades de que un niño presente un pobre desarrollo lingüístico y un bajo desempeño académico.

### 3. El texto oral y sus variables predictoras

El modelo que propusimos explicó solo un bajísimo porcentaje de la variabilidad del texto oral (2%). Por lo tanto, la información que ofrece sobre las variables contribuyentes debe ser tomada solo de manera indicativa. Aun así, quisiéramos dedicar unas líneas a comentar esta información. En primer lugar, ninguna de las variables relacionadas con el desarrollo del niño contribuyó a explicar

la variabilidad, ni siquiera el vocabulario, la variable de desarrollo lingüístico. Esto significa que la prueba no midió ninguna mejora en la comprensión del texto oral con el mayor desarrollo lingüístico ni cognitivo. En segundo lugar, con respecto a las características de la madre, tenemos resultados que reflejan, de manera más tenue, los del modelo de explicación del vocabulario: a mayor educación de la madre, mayor comprensión del texto oral del hijo ( $\beta^e = 0,10$ ) y a mejor español escrito de la madre, menor comprensión del texto oral del hijo ( $\beta^e = -0,09$ ). En tercer lugar, el estatus socioeconómico y el lugar de residencia no tuvieron ningún poder de predicción. En cuarto lugar, los años de educación del niño contribuyeron, de manera leve, a un mejor desempeño en el texto oral ( $\beta^e = 0,08$ ).

Sobre el bajo poder explicativo de este modelo, podríamos comentar que, aparentemente, el desarrollo del texto oral se debe a factores que no han sido incluidos en el modelo. Estas variables pueden no haber sido medidas por las entrevistas de Niños del Milenio o, simplemente, no haber sido consideradas por nosotros. Creemos que es importante desarrollar análisis diferentes de esta habilidad pues su relevancia para la comprensión de textos escritos se ha visto confirmada por los resultados del modelo que comentamos a continuación.

#### 4. El texto escrito y sus variables predictoras

El modelo propuesto explicó una parte importante (20%) de la variabilidad en la prueba de comprensión de textos escritos.

En primer lugar, con respecto a las características del niño, tenemos que el vocabulario a los 5 años, el texto oral, a los 8, y el género contribuyeron de manera distinta a la explicación de la variabilidad. El vocabulario y el texto oral significaron un cambio muy importante en la comprensión del texto ( $\beta^e = 0,32$  y  $\beta^e = 0,25$ , respectivamente). Esto confirma nuestra hipótesis más importante: el vocabulario y la comprensión del texto oral, habilidades que se desarrollan en etapas tempranas, predicen la comprensión de texto escrito. En particular, el vocabulario, medido tres años antes que la comprensión del texto escrito (en preescolar), predice fuertemente el desempeño de los niños en esta prueba (aplicada en primaria). Además, el género también representó un cambio, aunque más pequeño, en el desempeño de la prueba de comprensión de texto escrito ( $\beta^e = 0,09$ ). Veremos la interpretación de esta

información, más adelante, cuando comentamos las regresiones separadas por género.

En segundo lugar, las variables relacionadas a las oportunidades del niño contribuyen a la explicación del modelo, de manera indirecta, a través del desarrollo lingüístico, en particular del vocabulario. El crecimiento y la edad, indicadores indirectos del desarrollo cognitivo favorecieron la comprensión del texto escrito, así como la educación de la madre, el estatus socioeconómico de la familia y los años de educación del niño. Solo el español escrito tuvo un efecto indirecto negativo en la comprensión lectora. Así, vemos que todos estos factores contribuirán al desarrollo de la habilidad de comprensión de textos escritos además del nivel previo del vocabulario y la habilidad de comprensión de textos orales. Esto significa que los factores mencionados determinan el desarrollo de la comprensión lectora: su efecto en el vocabulario a los 5 años tiene efectos mediados en la comprensión textual a los 8 años.

Con respecto a los resultados de las regresiones separadas para niños y para niñas, se encontró que para ambos grupos el vocabulario y el texto oral predicen la comprensión del texto escrito (para los niños  $\beta^e = 0,39$  y  $\beta^e = 0,26$ ; para las niñas  $\beta^e = 0,26$  y  $\beta^e = 0,23$ ); y los factores estatus socioeconómico de la familia y educación de la madre tienen efectos indirectos. Por otro lado, sin embargo, los grupos tienen un perfil opuesto con respecto al efecto indirecto del crecimiento (medida de desarrollo cognitivo) y el español escrito de la madre. El crecimiento tuvo un efecto indirecto positivo en el grupo de niños – el español escrito de la madre no tuvo ningún efecto –; y el español escrito de la madre tuvo un efecto indirecto negativo en el grupo de las niñas – el crecimiento no tuvo ningún efecto. Estos resultados sugieren que la comprensión de textos escritos en los niños depende de su desarrollo cognitivo; mientras que en las niñas, de la literacidad de la madre. Estos perfiles especulares entre niños y niñas pueden interpretarse a la luz de las diferencias de sexo y género. A pesar de que las pruebas de vocabulario, comprensión de textos orales y escritos de esta muestra no evidencian diferencias significativas entre niños y niñas, sabemos por otros estudios que generalmente las niñas son mejores que los niños en estas habilidades. La ventaja de las niñas parece tener tanto una base biológica como un origen en los rasgos asociados a los estereotipos femeninos y masculinos. Es posible que la mayoría de las niñas aprenda a entender los textos escritos de manera adecuada, más allá de su nivel de inteligencia, gracias a la ventaja biológica por un

lado, y, por otro, debido a que actividades como la lectura – y no, en cambio, las matemáticas o las ciencias – son consideradas las apropiadas para las niñas y, como causa y/o consecuencia de ello, aquellas en las que las niñas sobresalen. En los niños, en cambio, la comprensión lectora se desarrollará de manera diversificada y será buena dependiendo principalmente del nivel de inteligencia de cada niño. Con respecto a la interpretación del efecto indirecto y negativo de la literacidad en la madre en las niñas y no en los niños, es necesaria mayor cautela. Como comentamos en la discusión sobre el vocabulario y sus variables predictoras, podría haber un efecto de la escasa variabilidad en los niveles de la variable español escrito de la madre. La segunda razón que propusimos está relacionada con el tiempo que las madres pueden poner a disposición de sus hijos. Es posible imaginar un escenario en que las madres con un nivel escrito de español bajo o muy bajo tengan menos oportunidades de trabajo y, por lo tanto, pasen más tiempo en casa con sus hijos. Así, aunque no sean las madres con mayores recursos para emplear las prácticas de literacidad, pueden ser las madres con más oportunidades para hacerlo; con lo que mejorarían el vocabulario de sus hijos e hijas; así como la comprensión del texto escrito de sus hijas. La razón por la cual sean las mujeres y no los varones los que se beneficien de esta situación puede deberse a que son justamente las niñas las que comparten más tiempo en casa con la madre, envueltas en situaciones que promueven la literacidad (narración de historias, cuentos, anécdotas) y en las que las madres encuentran buenos interlocutores, con un alto desarrollo lingüístico.

## **CONCLUSIONES Y RECOMENDACIONES DE POLÍTICA**

Los resultados de este estudio confirman que algunos factores relacionados a las características del niño y, lo más importante, a las oportunidades que su entorno le ofrece predicen el desarrollo de su vocabulario y su capacidad de entender textos escritos. Además, confirma que el desarrollo del vocabulario y la comprensión de textos orales predice la comprensión de textos escritos, incluso con años de anticipación.

Respecto a los objetivos que tenemos, la información que ofrece este estudio nos permite identificar grupos a riesgo de no desarrollar su lenguaje de manera adecuada, y programar intervenciones directas e indirectas para mejorar el lenguaje de los niños.

### **1. Identificación de grupos de riesgo**

Nuestros resultados señalan principalmente al grupo de niños residentes en zonas rurales como grupo a riesgo de no desarrollar un vocabulario adecuado para el ámbito escolar, en línea con los hallazgos de estudios anteriores. Además, sobre la base de la literatura, el riesgo será mayor si estos niños provienen de familias con estatus socioeconómico bajo, asisten a una escuela pública y tienen madres hablantes nativas de una lengua originaria. Esto significa que cualquier tipo de intervención programada debería tener como objetivo prioritario los niños con las características descritas.

Los resultados de este estudio también sugieren que cambios positivos en los factores que, directa o indirectamente, favorecen el desarrollo del vocabulario y la comprensión de textos generará mejoras en estas dos habilidades. Aunque las recomendaciones de política para mejorar las condiciones de vida de los niños exceden los objetivos de este estudio, es importante señalar lo siguiente. Sobre la base de nuestros resultados, programas que mejoran la nutrición de los niños – y que les ofrecen la oportunidad de un desarrollo físico y cognitivo óptimo (como Qali

Warma<sup>13</sup>) – así como aquellos que mejoran la situación económica de las familias más pobre (como JUNTOS<sup>14</sup>) contribuirá a un mejor desarrollo del vocabulario y la comprensión lectora en los niños.

## 2. Propuestas de intervención

El conjunto de resultados, interpretados a la luz de la evidencia científica sobre el desarrollo de las habilidades lingüísticas, sugiere que la exposición temprana de los niños a prácticas de literacidad tendrá un efecto dominó sobre el vocabulario y la futura comprensión lectora. Esto significa que un programa enfocado a mejorar el vocabulario en edad preescolar servirá de intervención directa para ampliar el léxico, pero también como intervención indirecta para mejorar la comprensión lectora en edad escolar. Por esa razón, nuestras recomendaciones de política pueden sintetizarse en una única propuesta: intervenir tempranamente para mejorar el vocabulario y, por lo tanto, la comprensión de textos escritos, dando prioridad al grupo vulnerable (niños de zonas rurales). Los siguientes párrafos están dedicados al modo de intervenir.

La necesidad de mejorar el lenguaje, en particular, ampliar el vocabulario de los niños no es un problema que se presente exclusivamente en el Perú. Muchos otros países se han encontrado, o se encuentran, frente a esta situación. Las recomendaciones y sugerencias que se ofrecen a continuación están basadas en evidencia científica actual y en programas de intervención cuya efectividad ha sido confirmada<sup>15</sup>. Además, se describe un tipo de intervención de efectividad comprobada experimentalmente, implementado en la región.

---

<sup>13</sup> El Programa Nacional de Alimentación Escolar Qali Warma, implementado por el Ministerio de Desarrollo e Inclusión Social ofrece, entre otros, servicios alimentarios y promueve mejores hábitos de alimentación en los usuarios del programa.

<sup>14</sup> El Programa Nacional de Apoyo Directo a los más Pobres –"JUNTOS", es un Programa de Transferencias Monetarias Condicionadas enfocadas en hogares en condición de pobreza prioritariamente de las zonas rurales, integrados por gestantes, niños, adolescentes y jóvenes.

<sup>15</sup> Esta información proviene principalmente de las conclusiones extraídas por el National Early Literacy Panel (2008), a través de un meta-análisis de los estudios y programas de desarrollo de la literacidad conocidos hasta el momento y de los principios extraídos, a partir de una revisión de los resultados científicos más importantes sobre el desarrollo del lenguaje de Odom et al. 2012, en particular el capítulo de Hirsh-Pasek).

## 2.1. Intervención a través de los padres o cuidadores

La intervención a través de los padres o cuidadores puede consistir en recomendaciones concretas sobre cómo interactuar con el niño, empleando prácticas de literacidad. Estas recomendaciones son las siguientes:

- (1) Ofrecer un *input* abundante y variado: Por un lado, un niño necesita recibir de los adultos un lenguaje abundante, en el sentido de cantidad de palabras, pues “los niños aprenden las palabras que más escuchan” (Hirsh-Pasek et al. 2012: 101). El lenguaje que recibe de los adultos, además, debe ser variado, rico en palabras distintas pues eso ampliará el vocabulario.
- (2) Ofrecer un *input* elaborado: El lenguaje que reciben los niños no solo debe ser abundante y variado, sino también debe ser organizado según la situación comunicativa. Los niños no aprenden nuevas palabras si se le enseña el significado de una lista de términos, sino que necesitan que las palabras estén organizadas en oraciones con contenido significativo según la situación comunicativa. Así, pueden extraer el significado de las palabras y, al mismo tiempo, identificar la estructura sintáctica correspondiente, el campo semántico y el contexto comunicativo en que su uso es apropiado. Por ejemplo, el niño aprenderá el significado de la palabra “freír” cuando esta aparezca dentro de una explicación sobre cómo se cocina. Gracias al contexto extraerá su significado; aprenderá que requiere un agente (alguien fríe) un paciente con función de objeto directo (alguien fríe algo); que se utiliza cuando se habla de preparar alimentos; y que su uso es apropiado en un contexto comunicativo de explicación, en un registro informal.
- (3) Crear contextos comunicativos interactivos: Los niños aprenden en conversaciones en las que hay respuestas a sus preguntas y pedidos; están enfocadas en sus intereses y consideran su nivel de atención. Está demostrado que los niños pequeños no aprenden el lenguaje en contextos pasivos como estar solos frente a la televisión<sup>16</sup>.

---

<sup>16</sup> Hay evidencia de esto en los estudios de Kuhl y otros (2003); Roseberry y otros (2009); y DeLoache y otros (2010).

- (4) Realizar lecturas compartidas: Las situaciones en las que los padres o cuidadores leen a sus hijos promueven el aprendizaje de nuevas palabras, así como mejora la comprensión de textos orales, cuyo efecto positivo se verá también en la comprensión lectora posterior.

Estas recomendaciones son válidas no solo para los padres hispanohablantes, sino también para los hablantes de lenguas originarias. Los padres que realicen prácticas de literacidad con sus hijos en una lengua distinta del español sentarán las bases de la literacidad en dicha lengua y esto promoverá el desarrollo del lenguaje y la literacidad también en la segunda lengua de sus hijos (Konishi et al. 2014).

## 2.2. Intervención a través de los docentes de Educación Inicial

La intervención a través de los maestros, puede llevarse a cabo implementando dentro de sus currículos o programas momentos de lectura al grupo de niños, cuya efectividad como práctica de literacidad ha sido ampliamente comprobada para mejorar el vocabulario y el texto oral bajo ciertas condiciones. La lectura debe estar acompañada de estrategias que favorezcan el aprendizaje. Como ejemplo de intervención exitosa de este tipo, tenemos el programa “Un Buen Comienzo”. Este es un proyecto que busca mejorar la calidad de la Educación Inicial en Chile que reciben los niños en situación de vulnerabilidad social, a través de un trabajo integrado en distintas áreas, con foco en lenguaje. Presentamos la evaluación, realizada por Gómez (2014), del efecto de las estrategias empleadas por las maestras durante la lectura a los niños. La autora examinó el efecto del uso de estrategias como la discusión de aspectos importantes de la narración, la definición de nuevas palabras y el análisis de eventos descritos en la ampliación del vocabulario de niños de 4 y 5 años de escuelas públicas de Santiago de Chile. La variabilidad de este tipo de prácticas predijo el crecimiento del vocabulario de los niños. Este tipo de resultados sugiere que capacitar a las maestras en el uso de estrategias de alto nivel durante la lectura a los niños tendrá efectos positivos en su vocabulario, y en la comprensión textual.

Sobre la implementación de las intervenciones, sugerimos instruir a los padres sobre el tipo de interacción que favorece el desarrollo lingüístico de sus hijos, a través de programas dedicados a la orientación sobre prácticas de cuidado y aprendizaje. El

programa Cunamás, por ejemplo, asesora en la crianza a las familias rurales con niños menores de tres años; podría ser un excelente medio para informar a los padres sobre las prácticas que favorecen el desarrollo del vocabulario. Estas intervenciones, que requieren continuidad, pueden llevarse a cabo desde muy temprana edad, por lo que informar a los padres de niños menores de 3 años sería conveniente. Por otro lado, sugerimos que se capacite a las maestras de Educación Inicial en la selección de narraciones que enriquezcan el vocabulario de los niños y en el uso de estrategias de alto nivel durante la lectura de las narraciones. Las situaciones en las que se pueden llevar a cabo estas prácticas ya están previstas en las Rutas del Aprendizaje, las cuales incluyen situaciones “donde los niños tuvieran oportunidades genuinas de [...] relatar y escuchar historias de la vida cotidiana o de temas específicos”. (Rutas del Aprendizaje versión 2015, área de Comunicación para los 3, 4 y 5 años de Educación Inicial).

## **PLAN DE INCIDENCIA**

Con el objetivo de difundir los resultados de este estudio y de tener impacto en las políticas públicas, se ha diseñado el siguiente plan de incidencia.

En primer lugar, se elaborará breves memorias que sinteticen los resultados y hagan énfasis en las conclusiones y recomendaciones de política, que serán distribuidas a los hacedores de política con el poder de poner en práctica las recomendaciones generadas a partir de los resultados del estudio.

Cada memoria tendrá la misma estructura, pero se hará énfasis en ciertos resultados y recomendaciones, según el público objetivo. A continuación describimos para cada uno de los dos resultados principales, las relativas recomendaciones y las instituciones a las que irían dirigidas.

Resultado 1: Los niños de escuelas rurales son un grupo a riesgo de desarrollar un vocabulario reducido y, por lo tanto, una pobre comprensión lectora.

Recomendación 1: Dar prioridad a este grupo como beneficiario de los programas para mejorar el vocabulario y la comprensión lectora.

Instituciones a las que va dirigida la recomendación:

- Ministerio de Desarrollo e Inclusión Social: en particular, el programa Cunamás.
- Ministerio de Educación:
  - Dirección General de Educación Básica Regular, en particular a la Dirección de Educación Inicial y a la Dirección de Educación Primaria;
  - Dirección General de Educación Básica Alternativa, Intercultural Bilingüe y de Servicios Educativos en el Ámbito Rural, en particular a la Dirección de Educación Intercultural Bilingüe y la Dirección de Servicios Educativos en el Ámbito Rural;

- Dirección General de Desarrollo Docente, en particular a la Dirección de Formación Inicial Docente y la Dirección de Formación Docente en Servicio.

Resultado 2: La comprensión lectora de los niños de 8 años depende directamente de su comprensión de textos orales y del vocabulario que tuvieron a los 5 años e, indirectamente, del estatus socioeconómico de la familia y del número de años de educación del niño y de su madre.

Recomendación 2a: Fortalecer programas que mejoren las condiciones económicas de las familias, especialmente rurales.

- Ministerio de Desarrollo e Inclusión Social: en particular el programa JUNTOS.

Recomendación 2b: Garantizar la asistencia de los niños a la escuela.

- Ministerio de Desarrollo e Inclusión Social.
- Ministerio de Educación.

Recomendación 2c: Mejorar el vocabulario en edad preescolar, a través de programas dirigidos a los padres y cuidadores.

- Ministerio de Desarrollo e Inclusión Social: en particular, el programa Cunamás.
- Ministerio de Educación:
  - Dirección General de Educación Básica Regular, en particular a la Dirección de Educación Inicial y a la Dirección de Educación Primaria;
  - Dirección General de Educación Básica Alternativa, Intercultural Bilingüe y de Servicios Educativos en el Ámbito Rural, en particular a la Dirección de Educación Intercultural Bilingüe y la Dirección de Servicios Educativos en el Ámbito Rural;
  - Dirección General de Desarrollo Docente, en particular a la Dirección de Formación Inicial Docente y la Dirección de Formación Docente en Servicio.

Recomendación 2d: Mejorar el vocabulario en edad preescolar, a través de programas dirigidos a los educadores.

- Ministerio de Educación:
  - Dirección General de Educación Básica Regular, en particular a la Dirección de Educación Inicial y a la Dirección de Educación Primaria;
  - Dirección General de Educación Básica Alternativa, Intercultural Bilingüe y de Servicios Educativos en el Ámbito Rural, en particular a la Dirección de Educación Intercultural Bilingüe y la Dirección de Servicios Educativos en el Ámbito Rural;
  - Dirección General de Desarrollo Docente, en particular a la Dirección de Formación Inicial Docente y la Dirección de Formación Docente en Servicio.

En segundo lugar, se difundirá los resultados de este estudio entre los especialistas del tema a través de una publicación en una revista especializada y con presentaciones en los principales centros de investigación.

“Los datos usados en esta publicación provienen del estudio Niños del Milenio, conocido internacionalmente como Young Lives, una investigación longitudinal de 15 años que analiza la naturaleza cambiante de la pobreza infantil en Etiopía, India (estado de Andhra Pradesh), Perú y Vietnam ([www.ninosdelmilenio.org](http://www.ninosdelmilenio.org) / [www.younglives.org.uk](http://www.younglives.org.uk)). Niños del Milenio es co-financiado por UK aid del Departamento de Desarrollo Internacional (DFID, por sus siglas en inglés) y por el Ministerio de Relaciones Exteriores de los Países Bajos (2010-2014). Las opiniones aquí expresadas son de los autores y no necesariamente compartidas por el estudio Niños del Milenio / Young Lives, la Universidad de Oxford, DFID ni otros donantes.”

## BIBLIOGRAFÍA

Agüero, J. y Cueto, S.

2004 Dime con quién andas y te diré cómo rindes. Peer-effects como determinantes del rendimiento escolar. Lima: GRADE y CIES.

Aikens, N. L., y Barbarin, O.

2008 Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, n° 100, pp. 235–251.

Alcázar, L.

2009 *El gasto público social frente a la infancia. Análisis del programa Juntos y de la oferta y demanda de servicios asociadas a sus condiciones*. Niños del Milenio/GRADE

Alexander K.L., Entwisle D.R. y Dauber S.L.

1993 “First-grade behavior: its short- and long-term consequences for school performance”. *Child Development*, n° 64, pp. 801–814.

Alexander, G. M. y Wilcox, T.

2012 “Sex Differences in Early Infancy”. *Child Development*, n° 6, vol. 4, pp. 801–814.

August, D., Carlo, M., Dressler, C. y Snow, C.

2005 “The critical role of vocabulary development for English language learners”. *Learning Disabilities Research and Practice*, n° 20, vol. 1, pp. 50–57.

Bast, J. y P. Reitsma

1998 “Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study”. *Developmental Psychology*, n° 34, pp. 1373-1399.

Bates, E. y F. Dick

2002 “Language, gesture, and the developing brain”. *Developmental Psychobiology*, n° 40, vol. 3, pp. 293–310.

Bauer, D., B. Goldfield y J. Reznick

2002 "Alternative approaches to analyzing individual differences in the rate of early vocabulary development". *Applied Psycholinguistics*, n° 23, pp. 313–335.

Behrman, J., H. Alderman y J. Hodinott

2004 *Hunger and Malnutrition. In Global Crises, Global Solutions*. Edited by Lomborg B. Cambridge, UK: Cambridge University Press, pp. 363 - 420.

Behrman, J., H. Hodinott, J. Maluccio, E. Soler-Hampejsek, E. Behrman, R. Martorell , M. Ramirez-Zea y A. D. Stein.

2008 *What Determines Adult Cognitive Skills? Impacts of Pre-Schooling Schooling and Post-Schooling Experiences In Guatemala*. Washington DC, USA: IFPRI; Discussion Paper No. 826.

Benavides, M.

2006 "Educación de las niñas: Lecciones del proceso peruano". *Informe final de consultoría*. Lima: GRADE.

Bennett K. K., D. J. Weigel y S. S. Martin

2002 "Children's acquisitions of early literacy skills: Examining family contributions". *Early Childhood Research Quarterly*, n° 17, pp. 297–317.

Bishop, D. V. M.

1997 *Uncommon understanding: Development and disorders of language comprehension in children*, UK, Psychology Press.

Bornstein, M. , O. Haynes, M. Painter y J. Genevro

2000 "Child language with mother and with stranger at home and in the laboratory: A methodological study". *Journal of Child Language*, vol. 27, pp. 407–420.

Bradley, R. H., y R. F. Corwyn

2002 "Socioeconomic status and child development". *Annual review of psychology*, n 53, vol. 1, pp. 371-399.

Burman, D., T. Bitan y J. Booth

2008 "Sex differences in neural processing of language among children". *Neuropsychologia*, vol. 46, pp. 1349–1362.

Butler, J.

2007 *El género en disputa: el feminismo y la subversión de la identidad*. (M. A. Muñoz, traductora). Barcelona: Ediciones Padiós Ibérica. (Trabajo original publicado en 1990).

Caesar, L. G., y N. W. Nelson

2013 "Parental involvement in language and literacy acquisition: A bilingual journaling approach". *Child Language Teaching and Therapy*, 0265659013513028

Campbell, J.M., S. K. Bell y L.K. Keith

2001 "Concurrent Validity of the Peabody Picture Vocabulary Test - Third Edition as an Intelligence and Achievement Screener for Low SES African American Children", *Assessment* n° 8, vol. 1, pp. 85-94

Campbell, J.M.

1998 "Review of the Peabody Picture Vocabulary Test - Third Edition", *Journal of Psychoeducational Assessment*, n° 16, vol. 4, pp. 334-8

Canet-Juric, L.; D. Burin; M. Andrés y S. Urquijo

2013 "Perfil cognitivo de niños con rendimientos bajos en comprensión lectora". *Anales de Psicología*, 29(39), 996-1005.  
<http://dx.doi.org/10.6018/analesps.29.3.138221>

Caspe, M.

2009 "Low-income Latino mothers' booksharing styles and children's emergent literacy development". *Early Childhood Research Quarterly*, n° 24, pp. 306-24.

Cassany, D.

2005 "Investigaciones y propuestas sobre literacidad actual: multiliteracidad, internet y criticidad". En *Conferencia presentada en Congreso Nacional Cátedra UNESCO para la lectura y la escritura, Universidad de Concepción*, vol. 24.

Caulfield L.; S. Richard; J. Rivera; P. Musgrove y R. Black

2006 "Stunting, wasting, and micronutrient deficiency disorders". En Jamison D, Breman J, Measham A, Alleyne G, Claeson M, Evans D, Jha P, Mills A, Musgrove P, et al. (eds.). *Disease control priorities in developing countries*. 2ª ed. Washington, DC: Oxford University Press, pp.551-67.

Checkley W.; L. Epstein; R. Gilman; L. Cabrera y R. Black

2003 "Effects of acute diarrhoea on linear growth in Peruvian children". *American Journal of Epidemiology*, vol. 157, pp. 166-175.

Cheung Y. B. y P. Ashorn

2010 "Continuation of linear growth failure and its association with cognitive ability are not dependent on initial length- for-age: a longitudinal study from 6 months to 11 years of age". *Acta Paediatrica*, n° 99, pp. 1719 – 1723.

Crookston B.T.; W. Schott; S. Cueto; K. A. Dearden; P. Engle; A. Georgiadis; E. A. Lundeen; M. E. Penny; A. D. Stein y J. R. Behrman

2013 "Postinfancy growth, schooling, and cognitive achievement: Young Lives". *American Journal of Clinical Nutrition* n° 98, pp. 1555 – 1563.

Crookston B.T.; K. A. Dearden; S. C. Alder; C. A. Porucznik; J. B. Stanford; R. M. Merrill; T. T. Dickerson y M. E. Penny

2011 "Impact of early and concurrent stunting on cognition". *Journal of Maternal and Child Nutrition*, n°7, pp. 397 – 409.

Crookston B.T.; M. E. Penny; S. C. Alder; T. T. Dickerson; R. M. Merrill; J. B. Stanford; C. A. Porucznik y K. A. Dearden

2010 "Children who recover from early stunting and children who are not stunted demonstrate similar levels of cognition". *Journal of Nutrition* n° 140, pp. 1996–2001.

Cueto, S; J. Escobal; M. Penny y P. Ames

2011 "Tracking Disparities: Who Gets Left Behind? Initial Findings from Peru". *Round 3. Survey Report*.

Cueto, S., G. Guerrero, J. León, M. De Silva, S. Huttly, M. Penny, C. Lanata, E. Villar

2005 "Social capital and education outcomes in urban and rural Peru." *Working Paper* 28.

Cueto, S.; G. Guerrero, J. León, E. Seguin e I. Muñoz

2009 "Explaining and overcoming marginalisation in education: a focus on ethnic/language minorities in Peru". *Background paper for the EFA Global Monitoring Report 2010*, Paris: UNESCO.

Cueto, S., G. Guerrero, J. León, M. Zapata y S. Freire

2013 "¿La cuna marca las oportunidades y el rendimiento educativo? Una mirada al caso peruano" *Grupo de Análisis para el Desarrollo (GRADE)*.

Cueto, S.; J. Leon; G. Guerrero e I. Muñoz

2009b "Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives". *Young Lives Technical Note*, n° 15.

Cueto, S., J. Leon e I. Muñoz.

2011 "Education Opportunities and Outcomes for Children in Peru: A Longitudinal Model, draft manuscript." *DEMUNA (Defensoría Municipal del Niño y del Adolescente) Manual de Gestión DEMUNA*, Lima: Acción por los Niños.

Cunningham, A. E.

2005 *Vocabulary growth through independent reading and reading aloud to children, in teaching and learning vocabulary: Bringing research to practice*, eds. E. H. Hiebert y M. L. Kamil, Mahwah, NJ, Erlbaum, pp. 45-48.

Daneman, M.

1988 *Word knowledge and reading skills, in Reading research: Advances in theory and practice*, eds. M. Daneman, G. MacKinnon y T. G. Waller, San Diego, Academic, 145-175.

de Jong, P. F., y van der Leij, A.

2002 "Effects of phonological abilities and linguistic comprehension on the development of reading". *Scientific Studies of Reading*, vol. 6, pp. 51-77

DeLoache J., C. Chiong; K. Sherman; N. Islam; M. Vanderborght1; G. Troseth; G. Strouse y K. O'Doherty

2010 "Do babies learn from baby media?". *Psychological Science.*, vol. 21, n o 11, pp. 1570–1574

De Temple, J.; y C. E. Snow

2003 "Learning words from books. On reading books to children" *Parents and teachers*, pp. 16-36.

Díaz, J.

2007 "Educación inicial y rendimiento en la Escuela". *Boletín Análisis y Propuestas*, vol. 12. GRADE.

Dickinson, D. K., A. McCabe; L. Anastasopoulos; E. S. Peisner-Feinberg, E. S., y M. D. Poe

2003 "The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children". *Journal of Educational Psychology*, n o 95, pp. 465 - 481.

Dickinson, D. K., y P. O. Tabors

2001 *Beginning literacy with language*. Baltimore: Brookes.

Dionne, G., P. Dale, M. Boivin y R. Plomin

2003 "Genetic evidence for bidirectional effects of early lexical and grammatical development". *Child Development*, n° 74, pp. 394–412.

Duncan G.; J. Brooks-Gunn y P. Klebanov

1994. "Economic deprivation and early childhood development". *Child Development*, n° 65, pp. 296 - 318

Dunn, L., E. Padilla, D. Lugo y L. Dunn

1986 *Manual del Examinador para el Test de Vocabulario en Imágenes Peabody (Peabody Picture Vocabulary Test) – Adaptación Hispanoamericana (Hispanic-American Adaptation)*, Minnesota: AGS

Eldredge, J.; B. Quinn y Butterfield, D.

1990 "Causal relationships between phonics, reading comprehension, and vocabulary achievement in the second grade". *Journal of Educational Research*, vol 83, pp. 201-214.

Espinoza, P.

2014 Young Lives Rounds 1 to 3. Constructed Files. *Young Lives Technical Note*, n° 30.

Feagans, L.

1982 "The development and importance of narratives for school adaptation". En L. Feagans, & D. C. Farran (eds.), *The language of children reared in poverty: implications for evaluation and intervention*, pp. 95 - 116. New York: Academic Press.

Fish, M., & Pinkerman, B.

2003 "Language skills in low-SES rural Appalachian children: Normative development and individual differences, infancy to preschool". *Journal of Applied Developmental Psychology*, no 23, vol. 5, 539-565.

Fink G. y P. Rockers

2014 "Childhood growth, schooling, and cognitive development: further evidence from the Young Lives study". *American Journal of Child Nutrition* n° 100, pp. 182 – 188.

Flanagan, D. y S. Dixon

2014 "The Cattell-Horn-Carroll theory of cognitive abilities". En Reynolds, C.; K. Vannest y E. Fletcher-Janzen. *Encyclopedia of Special Education*. Wiley Online Library. <http://onlinelibrary.wiley.com/book/10.1002/9780470373699>

Florit, E. y K. Cain, K.

2011 "*The simple view of reading: Is it valid for different types of alphabetic orthographies?*". *Educational Psychology Review*, 23(4), 553-576.

Florit, E.; M. Roch; G. Altoè, G., y M. C. Levorato

2009 "Listening comprehension in preschoolers: The role of memory". *British Journal of Developmental Psychology*, vol. 27, n° 4, pp. 935-951.

Florit, E., y M. C. Levorato

2012 "A longitudinal study on listening text comprehension and receptive vocabulary in preschoolers". *Rivista di psicolinguistica applicata*, n° 12, vol. 1/2, pp. 65 - 80.

Florit, E., Roch, M. y Levorato, M. C.

2011 "Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills". *Discourse Processes*, n° 48, vol. 2, pp. 119 - 138.

2013 "The relationship between listening comprehension of text and sentences in preschoolers: Specific or mediated by lower and higher level components?". *Applied Psycholinguistics*, n° 34, vol. 02, pp. 395 - 415.

2014 "Listening text comprehension in preschoolers: a longitudinal study on the role of semantic components". *Reading and Writing*, n° 27, vol. 5, pp. 793-817.

Freebody, P. y R. C. Anderson

1983 "Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension". *Reading Research Quarterly*, n° 18, vol. 3, 277-294.

Geschwind, N., y A. Galaburda

1987 *Cerebral lateralization*. Cambridge, MIT Press.

Geschwind, N., y A. Galaburda

1985<sup>a</sup> "Cerebral lateralisation: biological mechanisms, associations and pathology. II. A hypothesis and a program for research. *Archives of Neurology*, n° 42, pp. 521-552.

Geschwind, N., y A. Galaburda, A. M.

1985b "Cerebral lateralisation: biological mechanisms, associations and pathology. III. A hypothesis and a program for research". *Archives of Neurology*, n° 42, pp. 634-654.

Geschwind, N. y A. Galaburda

1985c "Cerebral lateralization. biological mechanisms, associations, and pathology: I. A hypothesis and a program for research". *Archive of Neurology*, n° 42, 428-459.

Gilman, R.; L. Cabrera y R. Black

2003 "Effects of acute diarrhea on linear growth in Peruvian children". *American Journal of Epidemiology*, n° 157, pp. 166 - 75.

Gómez, L.

2014 "Exploring teachers' read-aloud practices of children's language skills: the case of low-income Chilean preschool classrooms". (Tesis de doctorado no publicada). Boston College University. Boston, E.E.U.U.

Gonzalez, J. y B. Uhing

2008 "Home literacy environments and young Hispanic children's English and Spanish oral language: A community analysis." *Journal of Early Intervention* n° 30, pp. 116 - 39.

Gopnik, A., y A. Meltzoff

1986 "Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis". *Child Development*, n° 57, pp. 1040 - 1053.

Grantham-McGregor S.

2002 "Linear growth retardation and cognition", *The Lancet* n° 359, pp. 542.

Griffin, T.; L. Hemphill; L. Camp y D. Wolf

2004 "Oral discourse in the preschool years and later literacy skills". *First Language*, n° 24, pp. 123 - 147.

Guerrero, G; J. León, E. Rosales, M. Zapata, S. Freire, V. Saldarriaga y S. Cueto

2012 "Young Lives School Survey in Peru: Design and Initial Findings". Working Paper, n° 22.

Guerrero, G., Leon, J., Pells, K., y Woodhead, M.

2014 *Changing Children's Lives Risks and Opportunities*. Niños del Milenio.

Hannon, B. y M. Daneman

2001 "A new tool for measuring and understanding individual differences in the component process of reading comprehension". *Journal of Educational Psychology*, n° 93, pp. 103- 128.

Hart, B. y T. Risley

1995 *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.

Hirsh-Pasek, K. y R. Golinkoff

2012 "How babies talk: Six principles of early language development". En *Re-visioning the beginning: Developmental and Health Science Contributions to Infant/Toddler Programs for Children and Families Living in Poverty*. New York, Guilford Press, pp. 77-101.

Hoff, E.

2006 "How social contexts support and shape language development". *Developmental Review*, vol. 26, 55–88.

Hoover, W. y P. Gough

1990 "The simple view of reading". *Reading and writing*, vol. 2, n° 2, pp. 127-160.

Huttenlocher, J.; M. Vasilyeva; E. Cymerman y S. Levine

2002 "Language input and child syntax". *Cognitive Psychology*, vol. 45, pp. 337–374.

Johnson, J. y R. Repta

2012 Sex and gender. En Oliffe, J. y L. Greaves (eds.) *Designing and conducting gender, sex, and health research*. Sage Publications, pp. 17-37.

Juel, C.

1988 "Learning to read and write: A longitudinal study of 54 children from first through fourth grades". *Journal of educational Psychology*, n° 80, vol. 4, pp. 437.

Junyent, A.

2011 *Individual differences in Specific Language Impairment: profiles of preschoolers exposed to Italian* (Tesis de doctorado no publicada). Università degli Studi di Padova. Padua, Italia.

Kendeou, P. ; T. Papadopoulos, y M. Kotzapoulou

2013 "Evidence for the early emergence of the simple view of reading in a transparent orthography". *Reading and Writing*, vol. 26, n° 2, pp. 189-204.

Kimura, D.

1996 "Sex, sexual orientation and sex hormones influence human cognitive function". *Current opinion in neurobiology*, vol. 6, n° 2, pp. 259-263.

Kimura, D.

1999 *Sex and cognition*. London: MIT press.

Kintsch, W.

1998 *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University.

Kintsch,, W. y E. Kintsch

2005 Comprehension. En S. G. Paris, S. A. Stahl (eds.), *Children's reading comprehension and assessment*, pp. 71–92. Mahwah, NJ: Lawrence Erlbaum.

Knaak, S.

2004 "On the reconceptualizing of gender: Implications for research design". *Sociological Inquiry*, vol. 74, n° 3, pp. 302-317.

Konishi, H.; J. Kanero; M. Freeman; R. Golinkoffy K. Hirsh-Pasek

2014 "Six Principles of Language Development: Implications for Second Language Learners". *Developmental neuropsychology*, vol. 39, n° 5, pp. 404-420.

Kuhl, P.; F. Tsao y H.Liu

2003 "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning". *Proceedings of the National Academy of Sciences*, vol. 100, pp. 9096-9101.

León, J. A., y S. Cueto, S.

2004 "Dime con quién estudias y te diré cómo rindes: Peer-effects como determinantes del rendimiento escolar". *Informe Final para el Consorcio de Investigación Económica y Social*.

Lorber, J.

1996 "Beyond the binaries: Depolarizing the categories of sex, sexuality, and gender." *Sociological Inquiry*, vol. 66, n° 2, pp. 143-160.

Lutchmaya, S., S. Baron-Cohen, S. y R. Raggatt

2002 "Foetal testosterone and vocabulary size in 18- to 24-month-old infants". *Infant Behavior and Development*, n° 24, 418–424.

Lynch, A.

2008 "The linguistic similarities of Spanish heritage and second language learners". *Foreign Language Annals* n° 41, pp. 252–81.

Mann, V.; S. Sasanuma; N. Sakuma y S. Masaki

1990 "Sex differences in cognitive abilities: A cross-cultural perspective". *Neuropsychologia*, vol. 28, pp. 1063–1077.

Martin, D. y H. Hoover

1987 "Sex differences in educational achievement: A longitudinal study". *Journal of Early Adolescence*, vol. 7, pp. 65– 83.

Martorell R., L. Khan y D. Schroeder

1994 "Reversibility of stunting: epidemiologic findings from children in developing countries". *European Journal of Clinical Nutrition*, n° 48, vol. 1 pp.45–7.

McCall, R. B.

1981 "Nature-nurture and the two realms of development: a proposed integration with respect to mental development". *Child Development*, n° 52. pp. 1–12.

McCardle, P.; H. Scarborough y H. Catts

2001 "Predicting, explaining, and preventing children's reading difficulties". *Learning Disabilities Research and Practice*, vol. 16, n° 4, pp. 230-239.

McGeown, S.; H. Goodwin; N. Henderson y P. Wright

2012 "Gender differences in reading motivation: does sex or gender identity provide a better account?." *Journal of Research in Reading* vol. 35 n° 3, pp. 328- 336.

Ministerio de Educación (MINEDU)

2004 *Una aproximación a la alfabetización lectora de los estudiantes peruanos de 15 años. Resultados del Perú en la evaluación internacional PISA.*

2006 *Evaluación nacional del Rendimiento Estudiantil 2004. ¿Cómo disminuir la inequidad del sistema educativo peruano y mejorar el rendimiento de sus estudiantes? Factores explicativos más relevantes en la Evaluación Nacional 2004.*

2007 Consejo Nacional de Educación. *Proyecto Educativo Nacional al 2021.*

2010 Consejo Nacional de Educación. *Proyecto Educativo Nacional en el 2009 – Balance y recomendaciones.*

2012 *PISA 2012: primeros resultados. Informe Nacional del Perú.*

2013a Instituto Peruano de Evaluación, Acreditación y Certificación de la Calidad de la Educación Básica IPEBA – Programa Estándares de Aprendizaje. *Mapas de Progreso del Aprendizaje. Comunicación: oral.*

2013b Instituto Peruano de Evaluación, Acreditación y Certificación de la Calidad de la Educación Básica IPEBA – Programa Estándares de Aprendizaje. *Mapas de Progreso del Aprendizaje. Comunicación: lectura.*

2013c *PISA 2012: Primeros Resultados. Informe Nacional del Perú.*

2014 *Marco del sistema curricular nacional. Tercera Versión para el Diálogo.*

2015a *Rutas del Aprendizaje versión 2015. II ciclo. Área curricular: Comunicación.*

2015b *Rutas del Aprendizaje versión 2015. III ciclo. Área curricular: Comunicación.*

2015c *Aprendizajes de primero a sexto de primaria en lectura y matemática. Un estudio longitudinal en instituciones educativas estatales de Lima Metropolitana.*

Morales, S.

2009 *Socio-cultural variables and Reading literacy acquisition in Lima, Perú* (Tesis doctoral). Universidad de Nijmegen: Nijmegen.

Morales, S.

2011 *Variables socioculturales y cognitivas en el desarrollo de la comprensión de lectura en Lima, Perú. Revista Peruana de Investigación Educativa*, 3, 96-129.

Morgane P.; R. Austin-LaFrance; J. Bronzino; J. Tonkiss; S. Diaz-Cintra; L. Cintra; T. Kemper y J. Galler

1993 "Prenatal malnutrition and development of the brain". *Neuroscience and Behavior Review*, n° 17, pp. 91 – 128.

Morriset, C.

1995 "Toddlers' language development: Sex differences within social risk". *Developmental Psychology*, vol. 31, pp. 851–865.

Murray, A., J. Johnson y J. Peters, J.

1990 "Fine-tuning of utterance length to preverbal infants: Effects on later language development". *Journal of Child Language*, n° 17, pp. 511–525.

Muter, V.; C. Hulme; M. J. Snowling y J. Stevenson

2004 "Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study". *Developmental psychology*, n° 40, vol. 5, 665.

National Early Literacy Panel

2008 *Developing Early Literacy: Report of the National Early Literacy Panel*. Jessup, Maryland: National Institute for Literacy.

National Institute of Child Health and Human Development Early Child Care

2005 "Research Network. Pathways to reading: The role of oral language in the transition to reading". *Developmental Psychology*, n° 41, pp. 428 - 442.

Nicholson, L.

1994 "Interpeting gender". *Signs: journal of women in cultuer and society*, vol. 20, n° 1, pp. 79-105

Oakhill, J. V., y K. Cain

2012 "The precursors of reading ability in young readers: Evidence from a four-year longitudinal study". *Scientific Studies of Reading*, n° 16, vol. 2, pp. 91-121.

Odom, S.; E. Pungello y N. Gardner-Neblett

2012 *Re-visioning the beginning: Developmental and Health Science Contributions to Infant/Toddler Programs for Children and Families Living in Poverty*. New York, Guilford Press.

Organization for Economic Co-operation and Development

2001 Program for International Student Assessment (PISA). *Knowledge and skills for life. First results from PISA 2000*. Paris, France.

Ouellette, G. P.

2006 "What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension". *Journal of Educational Psychology*, n° 98, pp. 554 - 566.

Pan, B.; M. Rowe; J. Singer y C. Snow

2005 "Maternal correlates of growth in toddler vocabulary production in low-income families. *Child development*, vol. 76, n° 4, pp. 763-782.

Perova, E. y R. Vakis

2009 "Welfare impacts of the JUNTOS Program in Peru: evidence from a non-experimental evaluation". Washington DC: World Bank pp. 1-59

Perfetti, C. A.

1985 *Reading ability*, New York, Oxford University Press.

Piaget, J.

1979 *La formación del símbolo en el niño*. Fondo de cultura Económica. Publicación original de 1945 (en francés).

Pollitt E.; K. Gorman; P. Engle; R. Martorell; J. Rivera; T. Wachs y N. Scrimshaw

1993 "Early supplementary feeding and cognition: effects over two decades". *Monographs of the Society for Research in Child Development*, 108-118

Reese, E.; D. Leyva; A. Sparks y W. Grolnick

2010 "Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading. *Early Education and Development*", vol. 21, n° 3, pp. 318-342.

Roseberry, S.; K. Hirsh-Pasek; J. Parish-Morris y R. Golinkoff

2009 "Live action: Can young children learn verbs from video?". *Child Development*, vol. 80, n° 5, pp. 1360-1375.

Roth, F. P., D. Speece y D. Cooper

2002 "A longitudinal analysis of the connection between oral language and early reading". *Journal of Educational Research*, n° 95, pp. 259 - 272.

Roulstone, S., Law, J., Rush, R., Clegg, J., y Peters, T.

2011 *Investigating the role of language in children's early educational outcomes*.

Roulstone, S., S. Loader, K. Northstone y M. Beveridge

2002 "The speech and language of children aged 25 months: Descriptive data from the Avon Longitudinal Study of Parents and Children". *Early Child Development and Care*, n° 172, vol. 3, pp. 259-268.

Schumacker, R. y R. Lomax

2004 *A beginner's guide to structural equation modeling*. Psychology Press.

Seigneuric, A. y M. Ehrlich

2005 "Contribution of working memory capacity to children's reading comprehension. A longitudinal investigation". *Reading and Writing: An Interdisciplinary Journal*, n° 18, pp. 617 – 56

Sénéchal, M. y J. LeFevre

2002 "Parental involvement in the development of children's reading skill: A five-year longitudinal study". *Child development*, vol. 73, n° 2, pp. 445-460.

Shriberg, L. D., J. Tomblin y J. McSweeney

1999 "Prevalence of speech delay in 6-year-old children and comorbidity with language impairment". *Journal of Speech, Language, and Hearing Research*, 42, 1461 - 1481.

Snow, C; M. Burns y P. Griffin

1998 *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Sorsby, A.J. y M. Martlew

1991 "Representational demands in mothers' talk to preschool children in two contexts: Picture book reading and a modeling task". *Journal of Child Language* n° 18, pp. 373 - 95.

Stanovich, K. E.

1986 "Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy". *Reading Research Quarterly*, n° 21, 360–407.

Singleton, D. y L. Ryan

2004 *Language acquisition: The age factor*, vol. 9. Multilingual Matters.

Storch, S. y G. Whitehurst

2002 "Oral language and code-related precursors to reading: evidence from a longitudinal structural model". *Developmental psychology*, vol. 38, n° 6, 934.

Strand, S.; I. Deary y P. Smith.

2006 "Sex differences in cognitive abilities test scores: A UK national picture." *British Journal of Educational Psychology*, vol. 76, n° 3, pp. 463-480.

Strasser, K.; A. Larraín; S. López de Lérída, y M. R. Lissi

2010 "La comprensión narrativa en edad preescolar: Un instrumento para su medición". *Psykhé (Santiago)*, n° 19, vol. 1, pp. 75-87.

Strauss R. y W. Dietz

1998 "Growth and development of term children born with low birth weight: effects of genetic and environmental factors". *Journal of Pediatrics* n° 133, pp. 67–72.

Taub, G., R. Floyd, T. Keith, y K. McGrew

2008 "Effects of general and broad cognitive abilities on mathematics achievement from kindergarten through high school". *School Psychology Quarterly*, vol. 23, n° 2, pp.187-198.

Thorne, C.; K. Morla; P. Uccelli; T. Nakano; B. Mauchi; L. Landeo y R. Huerta

2013 "Efecto de una plataforma virtual en comprensión de lectura y vocabulario: una alternativa para mejorar las capacidades lectoras en primaria". *Revista de Psicología PUCP*, n° 31, vol. 1, pp. 3-35.

Tomblin, J. ; N. Records; P. Buckwalter; X. Zhang; E. Smith y M. O'Brien

1997 "Prevalence of specific language impairment in kindergarten children". *Journal of Speech, Language, and Hearing Research*, n° 40, pp. 1425–1460.

Ukrainetz, T.A. y D.S. Duncan

2000 "From Old to New: Examining Score Increases on the Peabody Picture Vocabulary Test–III", *Language, Speech, and Hearing Services in Schools* vol. 31, pp. 336 – 339

Undheim, J. y H. Nordvik.

1992 "Socioeconomic Factors and Sex Differences in an Egalitarian Educational System: academic achievement in 16-year-old Norwegian students." *Scandinavian Journal of Educational Research* vol. 36, n° 2, pp. 87-98.

UNESCO - Organización de las Naciones Unidas para la Educación, la Ciencia y la Lectura

2008 *Los aprendizajes de los estudiantes de América Latina y el Caribe. Primer reporte de los resultados del SERCE Segundo Estudio Regional Comparativo y Explicativo.*

2014 *Primera entrega de resultados TERCE Tercer Estudio Regional Comparativo y Educativo*.

Velarde, E., Canales, R., y Meléndez, M.

2010 Enfoque cognitivo y psicolingüístico de la lectura: diseño y validación de una prueba de habilidades prelectoras (THP) en niños y niñas de la Provincia Constitucional del Callao, Perú. *Revista de investigación psicológica*, n° 13, vol. 1, pp. 53-68.

Verhoeven, L. y J. Van Leeuwe

2008 "Prediction of the development of reading comprehension: A longitudinal study". *Applied Cognitive Psychology*, n° 22, pp. 407 - 423.

Victora, C.; L. Adair; C. Fall; P. Hallal; R. Martorell; L. Richter y H. Sachdev

2008 "Maternal and child undernutrition: consequences for adult health and human capital". *The Lancet* n° 371, pp. 340 –357.

Villalonga, M., Padilla, C. y D. Burin

2014 "Relaciones entre decodificación, conocimiento léxico-semántico e inferencias en niños de escolaridad primaria". *Interdisciplinaria*, n° 31, vol.2, pp. 259-274.

Walker, S.; S. Chang; C. Powell; S. Grantham-McGregor

2005 "Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study". *The Lancet* n° 366, pp 1804 – 1807.

Walker S.P., S. Grantham-McGregor; C. Powell y S. Chang

2000 "Effects of growth restriction in early childhood on growth, IQ, and cognition at age 11 to 12 years and the benefits of nutritional supplementation and psychosocial stimulation". *Journal of Pediatrics*, n° 137, pp. 36 – 41.

Walker S.P., S. Grantham-McGregor; Himes, J.; C. Powell y S. Chang

1996 "Early childhood supplementation does not benefit the long-term growth of stunted children in Jamaica". *Journal of Nutrition*, n°126, pp. 3017 – 3024.

Walker, D., Greenwood, C., Hart, B., & Carta, J.

1994 "Prediction of school outcomes based on early language production and socioeconomic factors". *Child Development*, n° 65, pp. 606–621.

Washington, J.A. y H.K. Craig

1999 *Performance of At-risk, African American Preschoolers on the Peabody Picture Vocabulary Test-III*, Language, Speech, and Hearing Services in Schools n° 30, pp. 75-82

Wells G.

1986 *The meaning makers: Children learning language and using language to learn*. 1st edition. Portsmouth, NH: Heinemann Educational Books.

Whitehouse, A. J.

2010 "Is there a sex ratio difference in the familial aggregation of specific language impairment? A meta-analysis". *Journal of Speech, Language, and Hearing Research*, n° 53, vol. 4, pp.1015-1025.

Whitehurst, G. J., y C. J. Lonigan

1998 "Child development and emergent literacy". *Child Development*, 69, 848–872.

Williams, K.T. y J. Wang

1997 *Technical References to the Peabody Picture Vocabulary Test - Third Edition*, Circle Pines, MN: American Guidance Service.

Witelson, D. F.

1976 "Sex and the single hemisphere: Specialization of the right hemisphere for spatial processing." *Science*, vol. 193, n° 4251, pp. 425-427.

Woldehanna, T. y L. Gebremedhin

2012 "The Effects of Pre-school Attendance on the Cognitive Development of Urban Children aged 5 and 8 Years. Evidence from Ethiopia". *Young Lives Working Papers*, n° 89.

World Health Organization

1997 *Global database on child growth and malnutrition*. Geneva.

Yang, S.; K. Tilling; R. Martin; N. Davies; Y. Ben-Shlomo y M Kramer

2011 "Pre-natal and post-natal growth trajectories and childhood cognitive ability and mental health". *International Journal of Epidemiology*, vol. 40, pp. 1215–1226.

Young Lives

2001 Summary of the Young Lives Conceptual Framework, en  
<[www.younglives.org.uk](http://www.younglives.org.uk)>

Zampini, L., y L. D'Odorico, L.

2011 "Lexical and syntactic development in Italian children with Down's syndrome." *International journal of language & communication disorders*, vol. 46, n° 4, pp. 386-396.

Zill, N.; K. Moore; E. Smith; T. Stief y M. Coiro

1995 "The life circumstances and development of children in welfare families: a profile based on national survey data". En *Escape From Poverty: What Makes a Difference For Children?* ed. PL Chase-Lansdale, J Brooks- Gunn, pp. 38–59. New York: Cambridge University Press.

Zill, N., y G. Resnick, G.

2006 "Emergent literacy of low-income children in Head Start: Relationships with child and family characteristics, program factors, and classroom quality". En D. K. Dickinson & S. B. Neuman (eds.), *Handbook of Early Literacy Research* (Vol. II, pp. 347 - 371). New York, NY: Guilford.

Zill, N.; A. Sorongon; K. Kim; C. Clark y M. Wolverton

2006 *FACES 2003: Research Brief*.

## ANEXO 1: Determinantes de la comprensión lectora en los niños hispanohablantes del Perú

Los trabajos más recientes sobre comprensión lectora – utilizando metodologías variadas – identifican a los siguientes factores como principales determinantes del desempeño de los niños hablantes de español. Con respecto a los factores individuales, se encuentran la conciencia fonológica y la identificación de las letras (Velarde et al. 2010); la habilidad de decodificación y el vocabulario (Morales 2009 y 2011); la memoria verbal y el lenguaje oral (Velarde et al. 2010); la disposición hacia la lectura (MINEDU 2004 y 2006). Con respecto al entorno y las oportunidades, están los factores relacionados a la vida escolar, como la educación inicial (Díaz 2007 y MINEDU 2015c), la relación con los pares en la escuela (Agüero y Cueto 2004; Guerrero et al 2012); el tipo de gestión (pública o privada) de las escuelas (Guerrero et al 2012); el clima y la infraestructura de las escuelas (MINEDU 2004 y 2015c); así como la calidad de la escuela y su capacidad de adaptación a los alumnos (Guerrero et al 2014); el grado de profesionalización, la autoeficacia y el trabajo colegiado de los docentes (MINEDU 2004 y 2015c). En relación con el entorno familiar, tenemos el nivel de educación de los padres, principalmente el de la madre, el número de hermanos y la comunicación en el hogar (MINEDU 2004); el capital cultural de la familia, la infraestructura del hogar y el grado de participación de los padres en la vida escolar de los niños (MINEDU 2015c) y, principalmente, el estatus socioeconómico (Cueto et al. 2005; MINEDU 2004, 2006 y 2015c; Morales 2009 y 2011). Además, el género tuvo un efecto independiente o en combinación con otros factores en prácticamente todos los estudios citados.

## ANEXO 2: La variable de desarrollo cognitivo en la base de datos Niños del Milenio y las razones de su exclusión en este estudio.

La base de datos Niños del Milenio ofrece una medida de desarrollo cognitivo que corresponde al resultado del Cognitive Development Assessment (CDA) creado por la International Evaluation Association (IEA) para evaluar el desarrollo cognitivo de niños de 4 años. El estudio Niños del Milenio utiliza una subescala de este instrumento, compuesta por 15 ítems que evalúan conceptos de cantidad, que fue previamente traducida y adaptada al español. En la administración de la subescala, se utiliza un juego de cartillas con figuras, en el cual cada cartilla representa un solo ítem. Antes de empezar, el examinador dice al niño que le mostrará algunas imágenes, y que harán diferentes cosas con las imágenes; le pide al niño que escuche y haga lo que se le dice. Para cada ítem, el examinador lee unas instrucciones específicas, haciendo énfasis en las palabras subrayadas. Los ítems son los siguientes:

1. Mira el plato con quequitos. Señala el plato que tenga pocos quequitos.
2. Mira el tazón de huevos. Señala el tazón que tenga más huevos.
3. Mira las botellas. Señala la que está casi vacía.
4. Mira las manzanas. Señala la manzana a la que le falta la mitad.
5. Mira la caja de bolitas y los grupos de bolitas. Señala el grupo de bolitas que tenga tantas bolitas como la caja.
6. Mira las figuras de los tazones y las cucharas. Señala la figura que muestre una cuchara en cada tazón.
7. Mira las figuras de los guantes. Señala la figura que muestre un par de guantes.
8. Mira al niño con los globos y estos grupos de globos. Señala el grupo que tiene igual número de globos que el niño.
9. Mira el grupo de estrellas. Señala el grupo que tenga menos estrellas.
10. Mira los tazones. Señala el tazón que está lleno.
11. Mira los bloques. Señala la figura que tenga muchos bloques.
12. Mira los círculos. Señala el círculo completo.
13. Mira los pájaros. Señala el pájaro que no tiene nada en su boca.
14. Mira los abrigos. Señala el abrigo que tenga todos sus botones.
15. Mira los perros y los gatos. Señala la figura en la que el perro tenga menos comida que el gato.

Como se puede observar, esta prueba requiere que un niño sea capaz de entender e interpretar correctamente las instrucciones – en particular el significado de las palabras subrayadas – y utilizar la información para seleccionar la figura adecuada a la descripción, en algunos casos, necesita procesar la información para hacer comparaciones (ítems 8, 9, 11 y 15). Sin duda, esta prueba está midiendo la inteligencia del niño – especialmente aquellos ítems que requieren un procesamiento mayor de la información – y, por lo tanto, su desarrollo cognitivo, por lo que ha sido

seleccionada por el estudio Niños del Milenio como indicador de desarrollo cognitivo y la validez y confiabilidad de su uso han sido medidas (Cueto et al. 2009b); así como ha sido empleada para medir el desarrollo cognitivo en estudios recientes que utilizan la base de datos (por ejemplo, Cueto et al. 2011, 2013; Woldehanna y Gebremedhin 2012). Por esas razones, inicialmente se incluyó esta variable en nuestro estudio. Sin embargo, en la etapa final de la elaboración del trabajo, se decidió excluir la variable por no ser un indicador de desarrollo cognitivo, independiente del lenguaje. Por el contrario, el desempeño en esta prueba depende fuertemente del lenguaje, pues, ante todo, el niño necesita un vocabulario y una sintaxis perfectos para poder iniciar el proceso de selección de la figura adecuada. La clara relación entre vocabulario y el resultado en esta prueba se hace evidente en la correlación entre ambas medidas en nuestra muestra: 0,582 ( $p = 0,01$ ;  $n = 1529$ ). La base no tiene ninguna medida del nivel de sintaxis por lo que no se puede ofrecer datos al respecto.

ANEXO 3: Correlaciones entre crecimiento, años de educación, estatus socioeconómico, vocabulario, texto oral y texto escrito.

Correlaciones de Pearson: r y número de participantes (entre paréntesis).

	Crecimiento	Años de educación	SES	Vocabulario	Texto oral
Crecimiento	-				
Años de educación	,019 (1527)	-			
SES	,424** (1527)	-,077** (1529)	-		
Vocabulario	,430** (1527)	,000 (1529)	,633** (1529)	-	
Texto oral	,070** (1406)	,045 (1449)	,049 (1408)	,083** (1408)	-
Texto escrito	,170** (1131)	-,017 (1165)	,254** (1133)	,367** (1133)	,265** (1133)

\*  $p < .05$ ; \*\* $p < ,01$

ANEXO 4: Correlaciones entre vocabulario, español oral y escrito de la madre, su educación, crecimiento del niño, sus años de educación, estado socioeconómico, vocabulario, texto oral y escrito,

Correlaciones de Spearman: rho y número de participantes (entre paréntesis),

	Vocabulario	1	2	3	4	5	6	7	8
(1) Español oral de la madre	,165** (1529)	-							
(2) Español escrito de la madre	,196** (1451)	,478** (1451)	-						
(3) Educación de la madre	,361** (1447)	,410** (1447)	,512** (1447)	-					
(4) Crecimiento	,261** (1527)	,270** (1527)	,280** (1449)	,411** (1445)	-				
(5) Años de educación	,037 (1529)	,065* (1529)	,082** (1451)	,028 (1447)	,030 (1527)	-			
(6) SES	,388** (1529)	,278** (1529)	,335** (1451)	,574** (1447)	,431** (1527)	-,069** (1529)	-		
(7) Vocabulario	,599** (1529)	,274** (1529)	,317** (1451)	,570** (1447)	,441** (1527)	,012 (1529)	,647** (1529)	-	
(8) Texto oral	,075** (1408)	,038 (1408)	,011 (1335)	,065* (1331)	,047 (1406)	,046 (1449)	,041 (1408)	,071** (1449)	-
(9) Texto escrito	,194** (1133)	,096** (1133)	,125** (1077)	,227** (1073)	,142** (1131)	-,016 (1165)	,224** (1133)	,350** (1133)	,247** (1133)

\*  $p < ,05$ ; \*\* $p < ,01$

ANEXO 5: Correlaciones entre sexo, lugar de residencia, tipo de institución educativa, crecimiento, años de educación, estatus socioeconómico, vocabulario, texto oral y escrito.

Correlaciones de punto bi-serial:  $r_b$  y número de participantes (entre paréntesis).

	Niña	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Residencia rural	.016 (1529)	-						
(2) Institución privada	.044 (1529)	-.206** (1529)	-					
(3) Crecimiento	-.012 (1527)	-.309** (1527)	.228** (1527)	-				
(4) Años de educación	.028 (1529)	.043 (1529)	-.501** (1576)	.019 (1527)	-			
(5) SES	-.009 (1529)	-.580** (1529)	.398** (1529)	.424** (1527)	-.077** (1529)	-		
(6) Vocabulario	-.032 (1529)	-.477** (1529)	.345** (1529)	.430** (1527)	.000 (1529)	.633** (1529)	-	
(7) Texto oral	-.028 (1408)	-.035 (1408)	.059* (1449)	.070** (1406)	.045 (1449)	.049 (1408)	.083** (1408)	-
(8) Texto escrito	.058 (1133)	-.194** (1133)	.147** (1165)	.170** (1131)	-.017 (1165)	.254** (1133)	.367** (1133)	.265** (1113)

\* $p < ,05$ ; \*\* $p < ,01$

ANEXO 6: Correlaciones entre sexo, lugar de residencia, tipo de institución educativa, español oral de la madre, español escrito de la madre y educación de la madre.

Tabla 1: Correlaciones entre español oral y escrito de la madre, y su educación; y sexo del niño, lugar de residencia y tipo de institución: lambda y número de participantes (entre paréntesis).

	Niña	Residencia rural	Institución privada
Español oral de la madre	,020 (1529)	<,0005 (1529)	<,0005 (1529)
Español escrito de la madre	<,0005 (1451)	,013 (1451)	<,0005 (1451)
Educación de la madre	,079* (1447)	,027 (1447)	,007 (1447)

\*  $p < ,05$ ; \*\* $p < .01$

Tabla 2: Correlaciones entre sexo, lugar de residencia y tipo de institución: phi y número de participantes (entre paréntesis).

	Niña	Residencia rural
Residencia rural	,016 (1529)	-
Institución privada	,044 (1529)	-,206* (1529)

\*  $p < ,05$ ; \*\* $p < .01$

ANEXO 7: Regresiones con variable dependiente vocabulario para los niños y para las niñas.

Tabla 1: Regresión para el grupo de niños, con vocabulario como variable dependiente.

NIÑOS	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Vocabulario</b>	0,521			
Edad		1,183	0,317***	0,030
Crecimiento		2,084	0,126***	0,035
Español oral de la madre		-0,180	-0,011	0,039
Español escrito de la madre		-2,044	-0,047	0,037
Educación de la madre		1,023	0,241***	0,045
Estatus socioeconómico		24,394	0,298***	0,046
Rural		-4,919	-0,103**	0,036

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

Tabla 2: Regresión para el grupo de niñas, con vocabulario como variable dependiente.

NIÑAS	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Vocabulario</b>	0,504			
Edad		1,164	0,296	0,030
Crecimiento		1,291	0,075	0,035
Español oral de la madre		0,109	0,006	0,034
Español escrito de la madre		-5,640	-0,134	0,039
Educación de la madre		1,520	0,347	0,041
Estatus socioeconómico		21,222	0,254	0,043
Rural		-3,966	-0,078	0,034

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

ANEXO 8: Regresiones con variable dependiente texto oral para los niños y para las niñas.

Tabla 1: Regresión para los niños con texto oral como variable dependiente.

NIÑOS	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Texto oral</b>	0,017			
Crecimiento		0,011	0,077	0,049
Español oral de la madre		0,011	0,077	0,060
Español escrito de la madre		-0,015	-0,040	0,053
Educación de la madre		<-0,001	-0,003	0,064
Años de educación		0,006	0,052	0,062
Tipo de institución: privada		0,020	0,051	0,069
Estatus socioeconómico		-0,041	-0,056	0,065
Lugar de residencia: rural		0,023	0,055	0,049
Vocabulario		<-0,001	-0,030	0,058

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

Tabla 2: Regresión para las niñas con texto oral como variable dependiente.

NIÑAS	R <sup>2</sup>	β	β <sup>e</sup>	EE
<b>Texto oral</b>	0,040			
Crecimiento		-0,008	-0,054	0,050
Español oral de la madre		0,004	0,022	0,059
Español escrito de la madre		-0,057	-0,148**	0,052
Educación de la madre		0,008	0,207**	0,067
Años de educación		0,011	0,90	0,047
Tipo de institución: privada		0,033	0,089	0,050
Estatus socioeconómico		-0,023	-0,030	0,064
Lugar de residencia: rural		0,030	0,066	0,053
Vocabulario		<-0,001	<-0,001	0,062

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

ANEXO 9: Efectos indirectos sobre la variables texto escrito para el grupo total de niños.

Efectos indirectos sobre texto escrito.

TODOS	$\beta$	EE
<b>Texto oral</b>		
Edad	0,004***	<0,001
Sexo	-0,002	<0,001
Crecimiento	0,007**	<0,001
Español oral de la madre	0,002	<0,001
Español escrito de la madre	-0,025***	0,001
Educación de la madre	0,006***	<0,001
Años de educación	0,003*	<0,001
Tipo de institución: privada	0,009	0,005
Estatus socioeconómico	0,003*	0,009
Lugar de residencia: rural	0,009	<0,001

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

ANEXO 10: Efectos indirectos sobre la variables texto escrito para los grupos de niños y niñas.

Tabla 1: Efectos indirectos sobre texto escrito para el grupo de niños.

NIÑOS	$\beta$	EE
<b>Texto escrito</b>		
Edad	0,005***	<0,001
Crecimiento	0,014***	0,004
Español oral de la madre	0,003	0,005
Español escrito de la madre	-0,015	0,011
Educación de la madre	0,005**	0,001
Años de educación	0,002	0,002
Tipo de institución: privada	0,007	0,010
Estatus socioeconómico	0,980***	0,028
Lugar de residencia: rural	-0,014	0,011
Vocabulario	<-0,001	<0,001

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$

Tabla 2: Efectos indirectos sobre texto escrito para el grupo de niñas.

NIÑAS	$\beta$	EE
<b>Texto escrito</b>		
Edad	0,0053***	0,001
Crecimiento	0,001	0,003
Español oral de la madre	0,001	0,004
Español escrito de la madre	-0,0032***	0,008
Educación de la madre	0,007***	0,001
Años de educación	0,003	0,002
Tipo de institución: privada	0,009	0,005
Estatus socioeconómico	0,054***	0,020
Lugar de residencia: rural	-0,003	0,009
Vocabulario	<-0,001	<0,001

\*  $p < ,05$ ; \*\*  $p < ,01$ ; ; \*\*\*  $p < ,001$