



**UNIVERSIDAD DE CONCEPCIÓN  
FACULTAD DE HUMANIDADES Y ARTE  
Programa de Doctorado en Lingüística**

# **MEDIACIÓN METACOGNITIVA DE LAS INTERACCIONES ENTRE LECTURA Y ESCRITURA**

Tesis presentada a la Facultad de Humanidades y Arte de la Universidad de Concepción para optar al grado académico de doctor en lingüística

**POR: PATRICIO ANDRÉS PINO CASTILLO**

Profesor Guía: Christian Soto Fajardo

Septiembre 2023  
Concepción, Chile

## **Derecho de Autor**

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

*A César, el amor de mi vida,  
por su compañía, apoyo,  
y paciencia constante.*

*A Cristina y Christian,  
mis hermanos y amigos  
de toda la vida, que me conocen  
y me comprenden  
quizá más que yo mismo.*

*A mi mamá, Cristina, y mi papá, Patricio,  
que siempre han apoyado mis estudios  
y celebrado mis logros.*

*A todos mis estudiantes, de pre y postgrado,  
quienes me enseñaron a enseñar cada vez mejor  
lo que es la lectura y la escritura académica.*

## **Agradecimientos**

Agradezco a mi profesor guía, Dr. Christian Soto Fajardo, por su acompañamiento y guía en este proceso; al Dr. Antonio Gutiérrez de Blume, por la generosidad al compartir todo su conocimiento, y su paciencia y compañía a lo largo de este proceso, y a la Coordinadora Nacional de Lenguaje de la Universidad Santo Tomás, Carolina Cereceda Triviño, por la gentileza y desinterés en su colaboración con este proyecto.

Este trabajo fue financiado por la Agencia Nacional de Investigación y Desarrollo (ANID) / Beca Doctorado Nacional 21201026, FONDECYT 1231433, “Metacognición de la lectura y escritura: dimensiones teóricas y aplicadas”, y FONDEF ID22110268, “ESCRIBE: un tutor inteligente para evaluar y entrenar estrategias de escritura en el sistema escolar chileno”.

## Tabla de Contenido

Derecho de Autor .....	ii
Agradecimientos .....	iv
Tabla de Contenido.....	v
Índice de Tablas.....	viii
Índice de Ilustraciones .....	ix
Resumen .....	x
Abstract.....	xii
Introducción .....	1
Capítulo I: Problema de Investigación y Marco Teórico .....	4
Comprensión y Producción del Discurso Escrito .....	11
Comprensión Lectora .....	11
Expresión Escrita .....	15
Interacciones entre la Comprensión y la Expresión.....	29
Competencias Lectoescriturales en el Discurso Especializado .....	34
Características de los Textos Especializados .....	34
Niveles de Competencia de los Sujetos y Metacognición .....	42
Procesos Mentales en la Comprensión y Producción Especializada.....	48
El sistema de Educación Superior en Chile .....	51
Síntesis y Propósito Investigativo .....	52
Objetivos.....	53

Objetivos Generales .....	53
Objetivos Específicos .....	54
Hipótesis.....	55
Hipótesis de Investigación.....	55
Hipótesis Nula .....	56
Capítulo II: Metodología.....	57
Características de la Investigación .....	57
Población y Muestra .....	58
Estudio 1.....	59
Corpus.....	61
Procedimiento de Análisis .....	61
Estudio 2.....	63
Instrumentos y Materiales .....	63
Procedimiento de Análisis .....	67
Capítulo III: Resultados.....	73
Resultados del Estudio 1 .....	73
Resultados de Análisis Factoriales.....	73
Diferencias entre Textos de Mayor y Menor Complejidad .....	78
Resultados del Estudio 2.....	80
Conclusiones .....	82
Discusión General .....	84
Implicancias para la Teoría.....	87

Implicancias para la Enseñanza y el Aprendizaje.....	93
Futuras Direcciones.....	98
Bibliografía.....	100
Anexos.....	117
Análisis de Juicios de Expertos para las Pruebas de Lectura .....	117
Ciencias Sociales.....	117
Derecho.....	118
Economía y Negocios .....	118
Ingeniería .....	119
Recursos Naturales y Medicina Veterinaria.....	120
Salud.....	120

## Índice de Tablas

Tabla 1 .....	21
Tabla 2 .....	64
Tabla 3 .....	74
Tabla 4 .....	78
Tabla 5 .....	79
Tabla 6 .....	79

## Índice de Ilustraciones

Ilustración 1.....	56
Ilustración 2.....	71
Ilustración 3.....	81

## Resumen

La lectura y la escritura son actividades fundamentales para la participación en la sociedad. Estas habilidades cobran especial relevancia cuando se trata de la inserción a círculos discursivos especializados, lo que es parte central de la formación universitaria. Existe abundante literatura sobre el proceso de comprensión lectora y producción escrita y sus interacciones, distinguiendo estadios de desarrollo iniciales y avanzados. En esta distinción, se identifica la metacognición como componente esencial de los procesos más avanzados y estratégicos, y la precisión metacognitiva como un factor fundamental para la efectividad de los procesos regulatorios que los componen. A través de un diseño transeccional, la presente investigación buscó, por una parte, determinar dimensiones de complejidad escrita con sustento teórico a partir de los índices de TRUNAJOD para las tareas de escritura académica de estudiantes de primer año de una universidad privada no selectiva y, por otra, determinar el potencial mediador de la precisión metacognitiva en la predicción de las habilidades de lectura desde las habilidades de escritura para estos estudiantes. Se recolectaron medidas de comprensión lectora, producción escrita y precisión metacognitiva de la lectura en 102 estudiantes de Arica a Puerto Montt. Los resultados muestran que es posible explicar significativamente un 52,3% de la varianza en la complejidad escrita a través de 4 dimensiones de índices TRUNAJOD: riqueza léxica, complejidad sintáctica, estructura informativa, y discurso especializado.

Por otra parte, se demostró que aspectos superficiales de la producción escrita, potenciados por la precisión metacognitiva, predicen significativamente la comprensión lectora. Estos hallazgos confirman la teoría existente sobre las relaciones entre lectura y escritura, añadiendo el poder mediador de la precisión metacognitiva en esta relación.

*Palabras clave:* complejidad escrita, literacidad, precisión metacognitiva, interacciones entre lectura y escritura, metacognición

## **Abstract**

Reading and writing are fundamental activities for participation in society. These skills become especially relevant when it comes to insertion into specialized discursive circles, which is a central part of university education. There is abundant literature on the process of reading comprehension and written production and their interactions, distinguishing initial and advanced stages of development. In this distinction, metacognition is identified as an essential component of the most advanced and strategic processes, and metacognitive accuracy as a fundamental factor for the effectiveness of the regulatory processes that compose them. Through a cross-sectional design, the present research sought, on the one hand, to determine dimensions of written complexity with theoretical support from the TRUNAJOD indexes for the academic writing tasks of first-year students of a non-selective private university and, on the other hand, to determine the mediating potential of metacognitive accuracy in the prediction of reading skills from writing skills. Results of reading comprehension, written production and metacomprehension accuracy were collected in 102 students from Arica to Puerto Montt. Results show that it is possible to significantly explain 52.3% of the variance in written complexity through 4 dimensions of TRUNAJOD indices: lexical richness, syntactic complexity, information structure, and specialized discourse. Additionally, results show that superficial aspects of written production, mediated by metacognitive accuracy, significantly predict reading

comprehension. These findings confirm the existing theory on the relationships between reading and writing, adding the mediating power of metacognitive accuracy in this relationship.

*Keywords:* writing complexity, literacy, metacognitive accuracy, interactions between reading and writing, metacognition

## **Introducción**

La comprensión y producción de la lengua son sin duda dos actividades centrales en la experiencia humana en sociedad. En el contexto universitario, la lectura y escritura académicas permiten la inserción efectiva en círculos especializados, lo que determina el éxito en la formación académica y la trayectoria profesional.

En torno a este importante tema, existe abundante literatura sobre el proceso de comprensión lectora y producción escrita y sus interacciones, distinguiendo estadios de desarrollo iniciales y avanzados para ambas habilidades. En esta distinción, se destaca el rol de la metacognición como componente esencial de los procesos más avanzados y estratégicos, y la precisión metacognitiva como un factor fundamental para la efectividad de los procesos regulatorios que los componen. Sin embargo, se desconoce cómo aporta la precisión metacognitiva en la interacción entre lectura y escritura.

Conocer la relación entre la precisión metacognitiva, la producción escrita y la comprensión lectora permitirá detectar un aspecto central que potencia la interacción entre las competencias de lectura y escritura en contextos académicos, y propiciar intervenciones educativas sobre estas competencias que

beneficien la transición hacia procesos lectoescriturales más eficientes y especializados.

Considerando estos antecedentes, la presente investigación buscó determinar dimensiones de complejidad escrita a partir de los índices de la herramienta TRUNAJOD y determinar el potencial mediador de la precisión metacognitiva en la interacción entre las habilidades de lectura y escritura para estudiantes de primer año de una universidad privada no selectiva. Con estos propósitos, fue necesario analizar la validez de constructo de modelos factoriales exploratorios que agrupan índices de TRUNAJOD; determinar las competencias en lectura académica, competencias en escritura académica, y precisión metacognitiva de la comprensión lectora; identificar las variables que predicen significativamente las competencias de lectura académica; y determinar el potencial mediador de la precisión metacognitiva en esta interacción.

La investigación se configuró como un estudio postpositivista transeccional (*single shot*) con énfasis en la lectura y escritura natural al interior de la universidad, potenciando la validez ecológica del estudio. Se recolectaron resultados de comprensión lectora, producción escrita y juicios de conocimiento sobre la comprensión lectora de 4804 estudiantes de Arica a Puerto Montt, de los cuales un número reducido (N = 102) completó todas las mediciones requeridas

por la investigación. Esta alta mortalidad muestral responde a que se trató de una investigación con participación voluntaria.

El texto se divide en tres capítulos principales: el problema de investigación y marco teórico, donde se presentan los antecedentes del problema y una revisión actualizada de las temáticas centrales; la metodología, que detalla las características de la investigación, la población, muestra, instrumentos y procedimientos de análisis; y los resultados, que aborda los principales hallazgos de la investigación. El trabajo finaliza con las conclusiones, que detallan el cumplimiento de los objetivos de la investigación y su relación con la teoría en el área de estudio, ofreciendo también proyecciones para la investigación en estas temáticas.

## **Capítulo I: Problema de Investigación y Marco Teórico**

La lectura y la escritura son actividades centrales para la participación de las personas en la sociedad. En la actualidad, la comprensión lectora no se concibe como una habilidad que se adquiere solamente en la escolaridad, sino que es “un conjunto creciente de conocimientos, destrezas y estrategias que los individuos van desarrollando a lo largo de la vida en distintos contextos, a través de la interacción con sus iguales y con la comunidad en general” (Organización para la Cooperación y Desarrollo Económicos [OCDE], 2017, p. 33). Por su parte, la escritura surge y contribuye a las circunstancias sociales, llevando consigo culturas e historias (Bazerman, 2016) y se considera cada vez más importante, ya que “las tareas que implican la producción de material escrito consideran a un porcentaje cada vez mayor de la fuerza laboral” (OCDE, 2018, p. 4). Esto no es sorprendente si consideramos que, además de construir representaciones personales, sociales y culturales, la escritura conlleva acciones tangibles como resultado de estas mismas (Bazerman, 2016).

En el contexto de la educación superior, la lectoescritura es especialmente relevante, ya que es el medio principal que cristaliza el proceso de aprendizaje del estudiantado y su acercamiento gradual a las comunidades académicas y científicas de su elección (Cassany & López, 2010). Paralelamente, los desarrollos en ciencia y tecnología han dado lugar a discursos cada vez más

especializados, conformando un continuo de subtipos lingüísticos orientados a un tema específico. Estos textos se configuran de forma especializada para facilitar una comunicación efectiva, y varían con el tiempo para adaptarse a los cambios en estas disciplinas (Schröder, 1991). Por consiguiente, al ingresar a la educación superior, cada estudiante debe emplear las habilidades de alfabetización que adquirió en el contexto escolar para someterse a un proceso de aprendizaje adicional que le permita acceder y participar en discursos especializados (Bazerman, 2016). El acompañamiento en esta trayectoria de especialización e inserción se denomina alfabetización académica (Carlino, 2003, 2005, 2013).

La necesidad de este aprendizaje adicional está respaldada por un reciente estudio etnográfico longitudinal que sigue las experiencias, prácticas y comprensión de las habilidades de lectura y escritura de estudiantes desde el final de la escolaridad hasta el comienzo de la educación superior (Baker, 2018). Los resultados demuestran que la escritura no solo se acomoda superficialmente en términos de formalidad y género, sino también en las prácticas de escritura del estudiantado. La escritura requiere descubrimiento, creación y construcción de la voz a nivel universitario, en oposición a su naturaleza más reproductiva en el nivel escolar. Del mismo modo, la lectura en la educación superior es drásticamente distinta, ya que este nuevo escenario requiere un enfoque autodirigido y más independiente que el requerido en la escuela. El estudio

también señala que la lectura está ausente de las prácticas de escritura en la escolarización, a diferencia del proceso de composición en la universidad que se basa en la investigación y la comprensión de diferentes fuentes sin mucha orientación por parte del profesorado, todo lo cual sugiere no solo una transición en las prácticas, sino también en la relación del estudiantado con el acceso y la creación del conocimiento.

El término “alfabetización académica” o “alfabetizaciones académicas” se utiliza generalmente para referirse a las prácticas de lectura y escritura en contextos académicos, a asignaturas relacionadas con el desarrollo de habilidades de lectura y escritura para estudios universitarios, a convenciones lingüísticas en las disciplinas, o a la investigación relacionada con cualquiera de estas áreas (Lillis & Scott, 2015). Sin embargo, postulan los autores, es un campo específico de investigación con su historia y epistemología: la concepción de las alfabetizaciones como prácticas sociales, desde un punto de vista transformador. La noción crítica de alfabetización como prácticas sociales hace alusión no sólo a las actividades sino también a las estructuras sociales que estas reproducen. Adicionalmente, la postura transformadora, a diferencia de una normativa, considera no solo a estudiantes como aprendices de una forma de escritura, sino que también en cuanto a su desarrollo de conciencia sobre cómo esta nueva práctica discursiva interactúa con la propia, y qué nuevas formas de dar sentido pueden surgir de esta interacción (Lillis & Scott, 2015).

La preocupación educativa por las alfabetizaciones académicas tiene sus orígenes en la década de 1970 en los Estados Unidos con la iniciativa *Writing Across the Curriculum* (WAC), inspirando movimientos similares como *Writing in the Disciplines* (WID) y *Academic Literacies* (ACLITS) en el Reino Unido. En la historia de estas iniciativas, se observa un cambio gradual de paradigma a lo largo del tiempo, desde una visión superficial, atomista y prescriptiva, basada en habilidades, a una concepción sociocultural y compleja del proceso de internalización hacia las disciplinas (Pino Castillo et al., 2019). Estas iniciativas surgieron como resultado de la diversificación del estudiantado, cuando las universidades de muchos países, como el Reino Unido, los Estados Unidos y Sudáfrica, pasaron de un sistema exclusivo que solo seleccionaba estudiantes de grupos sociales o culturales específicos, a sistemas más abiertos, aceptando poblaciones más diversas tanto a nivel nacional como internacional. Curiosamente, este movimiento hacia la apertura fue acompañado por discursos de estándares deficientes en alfabetización por parte del estudiantado, e inicialmente se dedicó un esfuerzo mínimo a la enseñanza de habilidades de lectura y escritura a nivel universitario (Lillis & Scott, 2015).

Se han identificado tres perspectivas sobre la investigación sobre el desarrollo de la alfabetización en la educación superior (Lea & Street, 1998, 2006): (1) la perspectiva de las habilidades de estudio, en la que la alfabetización

se conceptualiza como una habilidad cognitiva individual compuesta por el uso instrumental de aspectos superficiales específicos de la lengua; (2) la perspectiva de socialización académica, en la que se cree que cada estudiante experimenta un proceso de aculturación en discursos y géneros específicos de la materia, los que rigen las habilidades de estudio superficiales; y (3) la perspectiva de las alfabetizaciones académicas, que adiciona las relaciones de poder, autoridad e identidad a la perspectiva de socialización. Teniendo esto en cuenta, las necesidades de aprendizaje del estudiantado en la educación superior también pueden clasificarse como (1) un conjunto de normas lingüísticas que se deben identificar y aplicar en lectura y escritura; (2) un conjunto de normas sociales que es necesario comprender para participar en diferentes círculos especializados; y (3) la conciencia crítica de la propia identidad presente y futura, y las relaciones de poder existentes en la comunidad académica que eventualmente se convertirá en un nuevo escenario de participación.

Como puede inferirse de esta revisión, el desarrollo de estas habilidades sufre una transformación en el tránsito de un contexto a otro: desde una forma dirigida de lectura y escritura en la escuela, hacia una más estratégica y autodirigida en la universidad. En otras palabras, el desarrollo de la alfabetización en el contexto escolar depende de las habilidades cognitivas, mientras que, en la universidad, demanda conciencia y regulación metacognitivas. El entrenamiento de estrategias en el desarrollo de la alfabetización ha recibido gran atención en

la literatura, ya que se ha demostrado que las estrategias metacognitivas pueden predecir significativamente la competencia lectora, que también se aplica a estudiantes con trastorno del espectro autista (TEA; Rebolledo Luna et al., 2021) y a estudiantes con discapacidades cognitivas, en cuyo caso las estrategias metacognitivas pueden compensar eficazmente las desventajas cognitivas (Soto, Rodríguez Poblete, et al., 2018). Estos esfuerzos también se han combinado con desarrollos tecnológicos, como la creación de un tutor en lectura basado en la web en inglés y español (McCarthy et al., 2018, 2020; McNamara et al., 2004) y el entrenamiento metacognitivo por un agente educativo animado (Soto, Gutiérrez de Blume, Rodríguez, et al., 2019), que resultó especialmente provechoso en lectores con dificultades. También es importante considerar que el desarrollo de la alfabetización en los entornos escolares se concentra en gran medida en las habilidades de lectura, mientras que, en la universidad, hay un énfasis en el uso del trabajo escrito producido por el estudiantado como la principal forma de evaluación (Allen et al., 2019), no solo de sus propias habilidades lectoescriturales, sino que también de su aprendizaje de las disciplinas específicas de sus respectivas carreras universitarias.

En psicolingüística, existe abundante teoría y evidencia sobre los procesos de comprensión lectora y producción escrita, y existe también evidencia sobre distintos estadios de desarrollo o competencia en estas habilidades. En este contexto, la investigación más reciente se ha concentrado en el rol de la

metacognición en propiciar procesos lectoescriturales más estratégicos (Gutierrez De Blume et al., 2021; McCarthy et al., 2018, 2020; Soto, Gutiérrez De Blume, et al., 2018; Soto, Gutiérrez de Blume, Jacovina, et al., 2019; Soto, Rodríguez Poblete, et al., 2018) y en las interacciones entre lectura y escritura (Graham, Liu, Aitken, et al., 2018; Graham, Liu, Bartlett, et al., 2018; Graham & Hebert, 2011), pero se desconoce cómo interactúan estas variables, sobre todo considerando que el gatillamiento de procesos estratégicos o metacognitivos depende de la precisión al juzgar la efectividad de la lectura y la escritura.

Como puede desprenderse de la discusión, este objeto de estudio tiene diversas aristas: la comprensión y producción del discurso escrito, la interacción entre comprensión y producción del discurso, las competencias lectoescriturales en el discurso especializado, y los procesos cognitivos y metacognitivos que están en la base de la comprensión y producción en estos contextos. Habiendo establecido las principales problemáticas relacionadas, se presentará a continuación una revisión actualizada de la investigación en cada una de estas, culminando con el propósito de la presente investigación, sus preguntas, objetivos e hipótesis.

## **Comprensión y Producción del Discurso Escrito**

Según Scovel (1998), el campo de indagación de la psicolingüística es la naturaleza y estructura de la mente humana, que se desvela a través del estudio del lenguaje y del habla. Para el autor, esta área interdisciplinaria de estudio incluye los procesos de análisis y síntesis de la lengua de forma sincrónica y diacrónica; es decir, (1) la síntesis diacrónica o adquisición de la lengua, (2) la síntesis sincrónica o producción de la lengua, (3) el análisis sincrónico o comprensión de la lengua y (4) el análisis diacrónico o disolución y pérdida de la lengua. La psicolingüística, como área de desarrollo investigativo que establece redes entre la lingüística, la psicología, la neurociencia y la inteligencia artificial, ha provisto de abundante conocimiento para el entendimiento de los procesos mentales que subyacen a las actividades de comprensión y producción de la lengua, en particular sobre los modelos cognitivos que explican el proceso de comprensión del discurso escrito, arista que ha sido tradicionalmente la más productiva (De Vega & Cuetos, 1999).

### ***Comprensión Lectora***

Existen distintos modelos teóricos que buscan explicar la comprensión lectora, con distintas perspectivas y énfasis. Estos modelos tienen en común el afán por explicar la naturaleza de la representación mental que emerge del

procesamiento de estructuras discursivas (McNamara & Magliano, 2009). Asimismo, existe consenso entre las distintas perspectivas en concebir la representación mental como una amalgama entre la información del texto, la información relacionada al texto y las inferencias realizadas durante el proceso.

En su revisión de estos modelos teóricos para proponer una visión integradora de la comprensión, McNamara y Magliano (2009) identifican ocho dimensiones centrales para todos ellos:

1. arquitectura conexionista: los modelos asumen activación paralela del ambiente, significado y conocimientos previos;
2. despliegue de activaciones: la activación de conceptos se extiende y activa otros conceptos en la memoria;
3. procesamiento inconsciente automático: se asume que se dispone de cierta información de forma automática e inconsciente, aunque no existe consenso en qué información, la naturaleza de la automaticidad ni los efectos del procesamiento inconsciente;
4. enfoque discursivo: se asume que quien lee tiene un foco de atención, que cambia en el tiempo y con el input;
5. convergencia y satisfacción de límites: se asume que un concepto se activa cuando recibe activación de otros relacionados, siendo los

límites de la representación mental los conceptos activados, sus relaciones, y la información en la memoria de largo plazo;

6. mapeo: se refiere al establecimiento de relaciones entre el input lingüístico y el contexto anterior de conceptos activados, influenciado por la cohesión referencial y situacional, dando lugar a un sentido de continuidad;
7. inferencias basadas en el texto: se establecen conexiones entre constituyentes del texto cuando hay vacíos en la cohesión y se infiere la relación entre ideas. Las inferencias se consideran parte del modelo de situación, ya que reflejan relaciones causales, motivacionales, temporales y espaciales; y
8. límites de la memoria: se asume una memoria de trabajo limitada (2-4 proposiciones a la vez).

Algunos de estos modelos separan la memoria de trabajo limitada de la memoria de largo plazo; otros modelos proponen una memoria operativa de largo plazo. Si bien cada uno de los 7 modelos analizados por los autores aportan con distintos aspectos de la comprensión, el modelo más influyente hasta el día de hoy es el de Kintsch y Van Dijk (Kintsch, 1988, 1998; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983), particularmente por la introducción de la noción de modelo de situación, y la concepción de la comprensión en tres niveles interconectados: (1) la estructura superficial, o decodificación de palabras y

oraciones, y su significado literal; (2) la base textual, o significado del texto como red de proposiciones; y (3) el modelo de situación, o integración de la base textual con los conocimientos previos de quien lee.

En cuanto a la operacionalización y medición de la comprensión lectora, una de las formas más ampliamente utilizadas para constatarla es a través de pruebas de comprensión o competencia lectoras. La prueba del Programa Internacional de Evaluación de los Estudiantes (*Program for International Student Assessment* [PISA]), desarrollada por la OCDE (2017) es la prueba internacional de lectura más ampliamente utilizada por los gobiernos que pertenecen a esta organización. El programa PISA se define como “un recurso para ofrecer información abundante y detallada que permita a los países miembros adoptar las decisiones y políticas públicas necesarias para mejorar los niveles educativos” (OCDE, n.d., p. 3). Esta evaluación se realiza cada 3 años en las áreas de matemática, lectura y ciencias, en una muestra de alrededor de 600.000 estudiantes de 15 años de todo el mundo.

El programa define la comprensión lectora como “comprender, utilizar, reflexionar e interactuar con los textos escritos para alcanzar los propios objetivos, desarrollar el conocimiento y el potencial personal y participar en la sociedad” (OCDE, 2017, p. 34). Esta definición se operacionaliza en tres niveles de evaluación: (1) acceder y obtener, (2) integrar e interpretar y (3) reflexionar y

valorar. Los procesos de acceder y obtener evalúan la búsqueda y localización de información concreta. Por su parte, los procesos de integrar e interpretar evalúan la capacidad de conectar datos para elaborar significado e ir más allá del significado literal, identificando ideas o implicaciones latentes en el texto de forma general o local, siendo ambas necesarias para desarrollar una comprensión global del texto. Finalmente, los procesos de reflexionar y valorar evalúan la capacidad de acudir a ideas o actitudes externas al texto para relacionar la información textual con los conocimientos previos del lector. Como puede observarse, tanto los fundamentos de la prueba PISA de lectura como los niveles de comprensión que evalúa son consonantes con los modelos vigentes, representados principalmente en el trabajo de Kintsch y Van Dijk (Kintsch, 1988, 1998; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983).

### ***Expresión Escrita***

Se asume que los modelos de situación son el punto de partida para la expresión, por lo que la escritura se conceptualizó inicialmente como lectura inversa (Kintsch & van Dijk, 1978). La investigación sobre la escritura todavía depende en parte de una perspectiva sociocultural y una cognitiva (MacArthur & Graham, 2016), por lo que no existe un consenso completo sobre los procesos cognitivos involucrados en la expresión escrita. Una visión unificada sobre la escritura conduciría, como en el caso de la lectura, a una noción única sobre

cómo medirla objetivamente, a identificar las diferencias entre escritores novatos y avanzados, y a una perspectiva consensuada sobre la creación de actividades educativas necesarias para ayudar al estudiantado a transitar de escritores novatos a expertos. Aunque aún no existe esta visión única, la investigación y la teoría en los últimos 80 años pueden delinear ciertos procesos asociados con estos problemas no resueltos.

Desde la perspectiva cognitiva de la expresión escrita, que concierne a la psicolingüística, pueden detectarse cinco hitos en el modelamiento de la expresión. Un primer hito importante es el trabajo de Levelt (1938), que conceptualiza el proceso de producción de la lengua en tres módulos (conceptualizador, formulador y articulador), y un sistema de almacenamiento de conocimiento declarativo (lexicón). Según el autor, la expresión oral ocurre de la siguiente manera:

1. el conceptualizador toma las ideas y las estructura en un esquema gramatical, similar a la estructura profunda, de naturaleza pre-verbal;
2. el formulador —el único módulo que tiene acceso al lexicón, que almacena el conocimiento declarativo semántico y sintáctico (lemma), y también morfológico y fonológico (formas)— crea un patrón oracional, que corresponde a la estructura superficial, y un plan fonético, que es el output de este módulo;

3. finalmente, el articulador transforma el plan fonológico en habla audible y, podemos inferir, que el plan fonológico podría ser reemplazado por un plan grafológico en el caso de la producción escrita.

Dos características importantes de este modelo son el encapsulamiento (sólo el formulador tiene acceso al conocimiento gramatical y fonológico) y la incrementalidad, que supone la creación inmediata de output, que a su vez es utilizado inmediatamente por el próximo módulo.

Una segunda aproximación es el modelo de comprensión y producción de textos de Kintsch y Van Dijk (1978) que, según declaran, es aplicable a la lengua oral y escrita, ya que las diferencias entre estos medios ocurren en niveles inferiores a los que aborda el modelo. Es importante recordar que, en el tiempo en que fue concebido, aún no se manejaba la noción de modelo de situación, por lo que el modelo de comprensión involucra solo la extracción de proposiciones y la conformación de una estructura macroproposicional derivada a través de macrorreglas e inferencias. La producción, por su parte, es la aplicación inversa de macrorreglas desde la macroestructura para derivar (micro)proposiciones. Este modelo de producción comprende solo la escritura de resúmenes, distinguiendo en este proceso la reproducción de micro y macro proposiciones desde el texto original, la reconstrucción de proposiciones en base a inferencias sobre información incompleta usando macrorreglas de forma inversa, y los

metacomentarios que se presentan en un protocolo de producción en voz alta. El plan de producción es guiado por los esquemas de producción, representados por la tarea de escritura y el conocimiento previo, lo que explica que los resúmenes tengan un orden canónico independiente del orden de la lectura que los originan. Para los autores, este modelo puede eventualmente incorporarse a una teoría general de producción de textos, en cuyo caso las proposiciones centrales no son recordadas, y deben describirse nuevos mecanismos para generarlas.

Un tercer hito importante es el modelo de Hayes y Flower (1980), quienes a través de un análisis de dos años de protocolos de escritura en voz alta pudieron identificar tres espacios o dominios: (1) el ambiente de la tarea, que involucra las especificaciones de la tarea de escritura, el tema, la audiencia intencionada, las motivaciones del escritor y el texto escrito; (2) la memoria de largo plazo del escritor, que incluye conocimiento sobre distintos temas y audiencias, además de planes de escritura generales; y (3) el proceso de escritura, que contiene los subprocesos de planificación, traducción y revisión. La planificación, a su vez, se compone de la generación de ideas, la organización de ideas, y el establecimiento de metas; la traducción consiste en la producción de texto según el plan de escritura; y la revisión consiste en la lectura y edición para mejorar la calidad del texto producido.

Una cuarta aproximación fue el trabajo de Scardamalia y Bereiter (1992), que establece la distinción entre escritores novatos y escritores experimentados y propone un modelo llamado “decir el conocimiento” que explica la producción de textos para los primeros, y un modelo llamado “transformar el conocimiento”, para los segundos. Decir el conocimiento conceptualiza los procesos mentales en la escritura de textos sobre un tópico y género conocidos. El modelo opera a través de un proceso lineal y con mínima planificación, y sus etapas son la construcción de la representación sobre lo solicitado, la localización de indicadores de tópico y género y su activación en la memoria, la producción de un texto que sirve como indicador adicional de tópico y género —lo que, a su vez, propicia la coherencia— y la repetición del proceso. Transformar el conocimiento, por su parte, conceptualiza la escritura experimentada y estratégica. Implica tres etapas interconectadas: la construcción de la representación mental de la tarea, su análisis y el establecimiento de objetivos a partir de esta. Luego, estos objetivos se transfieren a una etapa que alterna la creación y resolución de problemas en dos espacios, retórico y de contenido, entre los cuales se traspasa. El espacio problematizador de contenido crea y resuelve problemas compuestos por creencias, a través de deducciones e hipótesis. El espacio problematizador retórico crea y resuelve problemas que consisten en representaciones del texto, a través de alteraciones del mismo, los objetivos de la escritura o relaciones entre ideas. Luego de la creación, solución y traspaso entre estos espacios, se pasa a

un proceso de decir el conocimiento, y este texto producido sirve como indicador adicional, para repetir el proceso.

Un último hito en el modelamiento de la producción de textos es el modelo revisado de Hayes y Flower (1980) a partir del trabajo de Scardamalia y Bereiter (1992), que da como resultado el modelo actual de Hayes (2012). Los cambios más significativos son la adición de la memoria operativa, la consideración de la tecnología en la transcripción, la inclusión de la motivación como un factor que explica la persistencia, y la consideración de la planificación y revisión como actividades especializadas de escritura. Además de este modelo, un aporte importante de este trabajo fue la identificación de sub-tipos de escritura dentro de decir el conocimiento, que fueron comprobados a través de modelamiento computacional. Se identificaron tres tipos de textos: (1) foco flexible, compuestos por cadenas de comentarios sobre el texto inmediatamente precedente; (2) tema fijo, compuesto por ruedas o listas centradas en un tema; y (3) elaboración de tema, que consisten en estructuras jerárquicas de temas, subtemas y elaboraciones. También se comprobó que los primeros dos tipos de texto son reemplazados gradualmente por el tercero en la trayectoria desde primero a noveno grado de educación.

**Tabla 1**

*Resumen de los Principales Modelos de Expresión*

	Levelt (1938)	Kintsch & van Dijk (1978)	Hayes & Flower (1980)	Scardamalia & Bereiter (1992)	Hayes (2012)
Generación de ideas y metas de la escritura.	Conceptualizador: toma ideas y las organiza en una estructura gramatical profunda.	Estructura macroproposicional. Plan de producción en base a esquema de producción: tarea y conocimiento previo.	Memoria de largo plazo y ambiente de tarea. Planificación: Generación de ideas, establecimiento de metas.	Construcción de representación mental de la tarea. Localización y activación de indicadores de tópico y género.	Memoria de largo plazo. Memoria operativa, atención, lectura. Motivación: determina la persistencia. Proponedor: genera ideas desde la tarea.
Planificación y estructuración del texto a producir.	Formulador: crea un patrón oracional de estructura superficial en base a lemmas y formas morfológicas y fonológicas.	Proposiciones: se derivan de la estructura macroproposicional a través de la aplicación inversa de macrorreglas.	Planificación: organización de ideas.	Espacio de contenido: resuelve problemas en creencias a través de deducciones e hipótesis. Espacio retórico: resuelve problemas sobre la representación del texto a través de modificaciones.	Planificación: parte del establecimiento de metas, depende de la motivación.
Producción del texto según el plan de escritura.	Articulador: Transforma la estructura superficial en habla audible*.	----	Traducción: producción de texto.	Producción de texto.	Foco flexible: comentarios sobre el texto precedente. Tema fijo: comentarios sobre un tema. Elaboración de tema: estructura jerárquica de temas, subtemas y elaboraciones.
Revisión del texto escrito y modificaciones.	----	----	Revisión y edición del texto.	El texto retroalimenta la escritura.	Revisión: parte del establecimiento de metas, depende de la motivación.

Como puede observarse en el resumen de la tabla 1, no existe consenso en cuanto a los modelos de escritura descritos en esta revisión. Sin embargo, pueden distinguirse cuatro componentes o etapas:

1. la generación de ideas y establecimiento de metas de la escritura, esquema mental almacenado en la memoria de largo plazo que, como esquema, filtra la información entrante sobre la tarea de escritura a través de los conocimientos previos sobre el tema, la audiencia, género y registro;
2. la planificación y estructuración del texto a producir, que depende de la motivación y genera un plan de escritura en base a resolución de problemas retóricos y de contenido;
3. la producción del texto según el plan de escritura o, en una versión más simple, la generación de texto con estructuras flexibles, fijas o jerárquicas; y
4. la revisión del texto escrito y modificaciones a este que, al igual que la planificación, depende de la motivación y del establecimiento de metas.

Es importante considerar que el planteamiento de Hayes (2012) de considerar la planificación y la revisión como aspectos que dependen de la motivación en su modelo actualizado es consonante con los hallazgos actuales

del aprendizaje autorregulado, que vinculan las características cognitivas, metacognitivas y motivacionales (Gutierrez de Blume, 2017).

Por otro lado, una investigación actualizada sobre los modelos de escritura para plantear un modelo integrador en el contexto de una segunda lengua (Abdel Latif, 2021) categorizó estos y otros modelos de escritura en modelos globales, cuando comprenden los aspectos generales de la escritura y modelos locales, cuando se enfocan en aspectos más específicos (por ejemplo, el modelo de Scardamalia y Bereiter se clasifica como un modelo específico). A través de un análisis de protocolos de pensamiento en voz alta, el autor propone una nueva dimensión: la búsqueda de contenido, y asigna funciones más especializadas al monitoreo y la revisión.

En cuanto a la medición de la competencia escrita, es importante señalar que la relación entre el desempeño y la habilidad de escritura es problemática. Se ha demostrado que la habilidad de escritura solo explica el 20% de la varianza en el desempeño, con otro 10% explicado por el género (van den Bergh et al., 2016). La competencia en la escritura, por tanto, se ha definido más recientemente como la flexibilidad y la adaptación a las circunstancias. En esta línea, un estudio de Allen et al. (2019) utilizó la evaluación automatizada de la escritura y un sistema de tutoría inteligente para investigar las relaciones entre la flexibilidad de escritura en un mismo texto (entre el primer y el segundo borrador)

y entre textos distintos (diferentes géneros), las habilidades de alfabetización de las y los escritores, y la provisión de retroalimentación superficial. El estudio demostró que las y los escritores demuestran flexibilidad en todas las tareas. Además, se encontraron similitudes y diferencias entre y dentro de los textos. Al revisar, las y los estudiantes modificaron todas las dimensiones excepto la sintaxis, centrándose tanto en las dimensiones textuales superficiales como profundas. Sin embargo, la mayor variación se encontró entre estudiantes, pero no como producto de la revisión. También se demostró que la capacidad de lectura solo predice algunas de las características lingüísticas de la escritura (narratividad, simplicidad sintáctica, cohesión referencial) y solo una parte de la varianza entre géneros.

Otro hallazgo importante de este estudio (Allen et al., 2019) fue que la flexibilidad lingüística se relaciona con la calidad del ensayo de diferentes maneras: La flexibilidad en la narratividad y la cohesión referencial se relacionan positivamente con las puntuaciones holísticas y algunas subescalas, mientras que la flexibilidad en la simplicidad sintáctica y la cohesión profunda se relacionan negativamente con la puntuación. Complementariamente a estos resultados, un estudio anterior sobre las puntuaciones holísticas y analíticas encontró una relación positiva entre el uso instruccional de las rúbricas que guían estas puntuaciones, así como también la disponibilidad de ellas durante el proceso de escritura y la calidad de los textos escritos (Sundeen, 2014). Tanto el uso

instruccional de rúbricas como la disponibilidad de ellas durante el proceso tienen un efecto positivo sobre la calidad.

Otro hallazgo importante del estudio de Allen et al. (2019) fue que la retroalimentación superficial no tuvo ningún efecto en los ensayos ni en la variabilidad entre géneros o revisiones, lo que demuestra que la retroalimentación de nivel inferior no tiene ningún efecto en la calidad de la escritura. Esto está en consonancia con otro estudio (Carter et al., 2018) que testeó la apreciación de los estudiantes de enfermería por el uso de ejemplares como una técnica de retroalimentación para mejorar la escritura académica. El estudio encontró que los ejemplares no afectaron significativamente el rendimiento.

En síntesis, si bien no existe un consenso sobre el modelamiento de los procesos cognitivos involucrados en la escritura a diferencia de la lectura, actualmente se puede concebir la producción escrita como un proceso donde interactúa el conocimiento —tanto retórico como de contenido— del escritor y sus representaciones, la representación de la tarea de escritura, y la generación de texto orientado a cumplir metas establecidas por el escritor en base a la situación. Asimismo, se ha instalado la noción de que existen distintas formas de producir textos en estadios iniciales y avanzados de experiencia o competencia que, según esta revisión de los modelos de escritura, se distinguen en cuanto a la

planificación y a la revisión del texto escrito, ambas condicionadas a la motivación del escritor. Finalmente, se ha demostrado que no existe una relación sencilla entre la competencia en escritura y el desempeño. Sobre el desempeño, las investigaciones más recientes han utilizado herramientas basadas en Procesamiento del Lenguaje Natural (PLN), área que será abordada a continuación.

**Procesamiento del Lenguaje Natural y Evaluación de la Escritura.** Los sistemas automatizados de evaluación de la escritura se han desarrollado, principalmente en inglés, para brindar al estudiantado más oportunidades de mejorar su escritura y descomprimir la carga laboral de docentes, considerando el acceso cada vez mayor a la educación superior (Allen et al., 2019). Estos sistemas pueden proveer retroalimentación sobre la escritura durante el proceso, ayudando a estudiantes a mejorar aspectos de sus textos. Los aspectos sobre los cuales se centran estos sistemas automáticos están basados en algoritmos desarrollados por PLN, y se centran principalmente a aspectos superficiales de la escritura como la complejidad, la cohesión, el vocabulario y la longitud (Correnti et al., 2022).

La puntuación automática de ensayos también se ha beneficiado del PLN. Una de las problemáticas planteadas en torno a la calificación de la escritura por medio de rúbricas es la subjetividad de su puntuación, excepto cuando se realiza

por equipos de expertos. En este sentido, se ha demostrado que entre evaluadores existen discrepancias que emergen de sus propias subjetividades. Las fuentes de estas discrepancias han sido clasificadas en sus diversas perspectivas sobre la institución donde trabajan, sobre sus roles como docentes, y sobre proceso de escritura, esto sumado a sus diversas características identitarias (Mumford & Atay, 2021). Para soslayar estas dificultades, los sistemas automáticos de puntuación de ensayos han sido empleados en pruebas estandarizadas de alto impacto, y han utilizado la PLN y el aprendizaje automático para modelar los puntajes de los evaluadores (Allen et al., 2019).

Por otra parte, se ha demostrado que la disponibilidad de estos sistemas durante el proceso formativo beneficia la autonomía del estudiantado (Geng & Razali, 2022). Sin embargo, estudios advierten que ellos tienen en cuenta principalmente aspectos retóricos, por lo que no consideran el contenido de la escritura, ni tampoco aspectos socioculturales del desarrollo de la escritura a nivel universitario (Correnti et al., 2022). Esto plantea un problema adicional cuando la tecnología se utiliza en la evaluación formativa, ya que propicia al estudiantado a modelar su escritura superficialmente para cumplir con la complejidad requerida por el sistema (S. Zhang, 2021).

La investigación más reciente en esta área se ha concentrado en el desarrollo de formas efectivas para ayudar al estudiantado a mejorar las

habilidades de lectoescritura. Investigaciones recientes han explorado los tipos de retroalimentación y los posibles usos de sistemas de corrección automatizada para mejorar la escritura académica de estudiantes en un segundo idioma (Calma et al., 2022), y el impacto de estas herramientas en la escritura académica en inglés como segundo idioma, donde se ha demostrado efectos en el dominio de la escritura, el compromiso conductual, cognitivo y emocional, la autoeficacia y las emociones positivas y negativas (Nazari et al., 2021). Es importante señalar que estos estudios se realizan en inglés como primera o segunda lengua, área donde existe un amplio desarrollo de tecnologías para evaluar la escritura, así como también sistemas de tutoría que atienden aspectos superficiales y estratégicos de los procesos de lectura y escritura (Allen et al., 2019; McNamara et al., 2004). En español, los desarrollos en tecnología hasta ahora han traducido tutores de comprensión lectora al inglés (McCarthy et al., 2018, 2020), y un estudio previo creó una herramienta de complejidad de textos para evaluar la adecuación de los textos proporcionados en la educación primaria y secundaria (Palma et al., 2019). La creación de herramientas de evaluación automática de la escritura en español aún no ha recibido suficiente atención por parte de la investigación y el desarrollo. Esto sería útil para definir la complejidad desde una perspectiva textual, lo que significaría beneficios para estudiantes y docentes en español como primera lengua, segunda lengua, lengua extranjera, o lengua de herencia.

## **Interacciones entre la Comprensión y la Expresión**

La lectura y la escritura son claramente procesos entrelazados; no existe duda sobre la relación entre ellos (Bazerman, 1980; Fitzgerald & Shanahan, 2000). Al respecto, un número especial del *Journal of Memory and Language* propuso tres hipótesis: (1) son dimensiones separadas, con sus propios procesos y representaciones independientes, sin influencia entre ellas; (2) son dimensiones separables, con representaciones y procesos compartidos que confluyen bajo ciertas circunstancias; y (3) son dimensiones inseparables, con procesos y representaciones compartidos e indivisibles (Meyer et al., 2016). Los estudios en este número demuestran que no existen actualmente datos empíricos que sustenten que estas dimensiones sean separadas; muy por el contrario, hay robusta evidencia que demuestra las confluencias entre comprensión y producción. Sobre estas relaciones, la evidencia hasta fines del siglo pasado identifica tres tipos que han sido investigadas: relaciones retóricas, conexiones procedimentales y conocimientos compartidos (Fitzgerald & Shanahan, 2000).

La investigación sobre relaciones retóricas concibe la lectura y la escritura como actividades comunicativas que se desarrollan participando en ellas como emisor y como receptor. Por su parte, la perspectiva de conexiones procedimentales incluye estudios en que se emplean la lectura y la escritura en

combinación para lograr metas específicas, como por ejemplo la lectura de un texto escrito o la escritura de resúmenes sobre un texto leído. Finalmente, el enfoque de conocimientos compartidos parte de la premisa que la lectura y la escritura son procesos cognitivos que dependen de representaciones mentales compartidas en distintos niveles, tanto de conocimiento lingüístico como de limitaciones contextuales. Esta última perspectiva fue elaborada por Fitzgerald y Shanahan (2000), quienes identificaron cuatro áreas de conocimientos compartidos:

1. Metaconocimiento: conocimiento sobre las funciones y propósitos de la lectura y escritura, conciencia sobre la interacción entre lectores y escritores, metacompreensión y estrategias de identificación o producción de palabras, monitoreo sobre el propio conocimiento.
2. Conocimiento de dominio sobre la sustancia y contenido: conocimiento previo y conocimiento agregado en la tarea de lectura y escritura, incluyendo el significado de palabras y el significado creado a través del contexto del texto.
3. Conocimiento sobre los atributos universales del texto: conocimiento grafofónico como la conciencia fonológica, conciencia grafémica y morfología; conocimiento de la sintaxis y puntuación; conocimiento del formato textual como la sintaxis oracional y la organización textual.

4. Conocimiento procedimental y habilidades para negociar la lectura y la escritura: saber cómo acceder, emplear y generar conocimiento en las áreas anteriores; integración de varios procesos. (Fitzgerald & Shanahan, 2000, p. 41)

Estos autores reportan distintos estudios realizados bajo esta perspectiva, identificando tres hallazgos principales: (1) la lectura y la escritura se influyen mutuamente, y esta influencia cambia según la edad, siendo la interacción entre reconocimiento de palabras y ortografía más preponderante en la lectura y escritura iniciales, y dando paso a interacciones en aspectos más estructurales en estadios posteriores; (2) las experiencias instruccionales afectan a la lectura y la escritura de forma distinta, teniendo más transferencia a la lectura que a la escritura; (3) el conocimiento sobre la lectura y la escritura se transfiere al aprendizaje de otros procesos.

Haciendo una revisión de la teoría que existe sobre la lectura y la escritura en distintos niveles educativos, los autores también postulan que los cuatro tipos de conocimientos compartidos tienen distinta importancia en distintas etapas del desarrollo (Fitzgerald & Shanahan, 2000). Para estudiantes de 18 años o superior, que es el enfoque del presente estudio, los autores postulan que los conocimientos compartidos más importantes son el conocimiento de dominio (conocimiento previo y conocimiento adquirido en la lectura y escritura), el

metaconocimiento (en particular, la conciencia sobre las formas de interactuar entre lectores y escritores) y el conocimiento procedimental y habilidades (saber cómo percibir cosas desde el punto de vista de otra persona, saber cómo analizar y criticar).

En este mismo enfoque de conocimientos compartidos también existe evidencia más actualizada. En primer lugar, un reciente metaanálisis de 151 estudios hemodinámicos demuestra la existencia de un sistema compartido de comprensión y producción sintáctica con dos regiones corticales de diferente activación en la comprensión y producción, así como también en el procesamiento de oraciones simples y complejas (Indefrey, 2018).

En relación con esto y sus aplicaciones en educación, la concepción de procesos y representaciones compartidos también ha propiciado la proliferación de estudios tendientes a demostrar el impacto de las intervenciones en escritura sobre la comprensión lectora, el impacto de las intervenciones en lectura sobre la producción escrita, y el impacto de las intervenciones combinadas sobre la comprensión lectora y la producción escrita. Estos diversos estudios fueron recogidos en tres meta análisis: (1) un primer meta análisis de 95 estudios experimentales o cuasiexperimentales en primer a doceavo grado que incluyeron al menos una medición en lectura para estimar el impacto de la escritura sobre la lectura comprobó que la escritura sobre material leído mejora la comprensión

lectora de textos expositivos y narrativos (Graham & Hebert, 2011); (2) un segundo meta análisis de 91 estudios experimentales o cuasiexperimentales sobre el impacto que las intervenciones en lectura tienen sobre la escritura demostró considerable efecto de la enseñanza de la lectura sobre el proceso de escritura (Graham, Liu, Bartlett, et al., 2018); (3) un meta análisis de 47 experimentos o cuasiexperimentos que combinaban proporcionalmente la intervención en lectura o escritura con hasta un 60% del tiempo destinado a una u otra, demostró un efecto importante de estas intervenciones sobre el desempeño en lectura y escritura en estudiantes de primer a doceavo grado (Graham, Liu, Aitken, et al., 2018). Además, este estudio demostró que las intervenciones con igual proporción en lectura y escritura tuvieron efectos mayores sobre la lectura, incluso que aquellos en los cuales la instrucción favorecía solamente a la lectura.

En síntesis, no existe duda que la lectura y escritura están relacionadas, ya que utilizan conocimientos compartidos en distintos niveles, incluyendo conocimiento previo, metac conocimiento, conocimiento lingüístico y conocimiento sobre las interacciones. Sin embargo, los datos empíricos sobre esta interacción solo muestran estas interacciones en algunos aspectos. Por ejemplo, Berninger et al. (2002), en un estudio de modelamiento de ecuaciones estructurales longitudinales en primer a sexto grado, comprobó tres tipos de relaciones: (1) el reconocimiento de palabras predice la escritura, y esta predicción es recíproca

solo en segundo grado; (2) la comprensión lectora y la ortografía tienen relaciones recíprocas y bidireccionales en todos los grados; y (3) la comprensión lectora predice la calidad de la composición, pero esta relación es recíproca solo en los grados 4 a 6.

### **Competencias Lectoescriturales en el Discurso Especializado**

Las competencias en las habilidades de lectura y escritura en el contexto de la educación superior tienen dos aristas principales: las características de los textos especializados y los niveles de competencias de los sujetos, las que serán abordadas a continuación.

### ***Características de los Textos Especializados***

La investigación psicolingüística sobre la lectura y los modelos de comprensión lectora hacen la distinción entre los textos narrativos y los textos expositivos (Weaver & Kintsch, 1996). En términos generales, los textos narrativos tienen como foco las relaciones sociales o interpersonales y resuelven problemas cotidianos, mientras que el foco de los textos expositivos son los conceptos e ideas y las relaciones lógicas entre ellos (Kim et al., 2021; Wu et al., 2020). Más precisamente, la diferencia entre un texto narrativo y un texto expositivo se relaciona al género, entendido como el propósito cultural de un texto

y su estructura esquemática (Eggins, 2004). En este sentido, el propósito cultural de un texto narrativo es contar una historia real o ficticia, por lo que usa principalmente el tiempo pasado y un orden cronológico de eventos, describe personajes, eventos, lugares, intenciones y emociones, y emplea correferencia para establecer conexión entre los elementos expuestos (Medina & Pilonieta, 2006). Por el contrario, el propósito del texto expositivo es informar, y se componen de aseveraciones generales y específicas sobre un tema (Cash & Shumm, 2006).

Sobre las diferencias más específicas entre estos textos, una narrativa está compuesta por un escenario y uno o más episodios, los que a su vez están compuestos por un inicio, una reacción compleja, un camino hacia el objetivo y un desenlace. Según Thorndyke, la narrativa está compuesta por un escenario, un tema, una trama y una resolución (Graesser et al., 1996). Los textos narrativos, al describir personajes, eventos y lugares concretos (sean ficticios o reales), se benefician de la visualización al poder anclar la información en la memoria y por lo mismo son más fáciles de recordar (Vieiro Iglesias & Gómez-Veiga, 2004). También son más fáciles de comprender por su cercanía con la oralidad y por su predictibilidad (Graesser et al., 1996).

Por el contrario, la estructura esquemática de un texto expositivo está compuesta por seis pasos: la presentación de un tema, la descripción de los

atributos de este tema, descripción de los eventos o procesos característicos del tema, una comparación de categorías, una síntesis final y un epílogo, siendo los tres primeros elementos obligatorios (Cash & Shumm, 2006). Por su parte, Kintsch propone tres estructuras posibles en los textos expositivos:

- Relaciones general-particular: identificación, definición, clasificación e ilustración.
- Relaciones objeto-objeto: comparaciones y contrastes, propósitos y evaluaciones.
- Relaciones objeto-parte: análisis general, análisis estructural, análisis funcional, análisis causal. (Weaver & Kintsch, 1996)

A diferencia de los textos narrativos, el texto expositivo no usa correferencia sino co-clasificación, que alude a la agrupación conceptual de entidades o eventos (Cash & Shumm, 2006). La relación de la comprensión de textos expositivos con la cognición anclada es, sin embargo, un tema no resuelto, ya que se desconoce cómo el cerebro implementa operaciones simbólicas. Sobre esto, se postula la existencia de símbolos amodales (predicación, combinación conceptual, recursión), mecanismos de simulación como la metáfora conceptual, o el anclaje de conceptos abstractos en simulaciones introspectivas de experiencias y situaciones, que incluyen la percepción interna de estados

motivacionales, afectivos, metas, creencias, operaciones cognitivas, y metacognición (Barsalou, 2010).

Un estudio más actualizado sobre las diferencias entre el procesamiento de textos narrativos y expositivos demostró que la teoría de la mente y la explicitación de estados mentales tienen una clara incidencia en la comprensión de textos narrativos (Kim et al., 2021). En particular la explicitación de estados mentales tiene una fuerte relación con la comprensión de discursos narrativos y una débil relación con la comprensión de textos expositivos, lo que se relaciona a las simulaciones introspectivas descritas por Barsalou (2010). En contraparte, un estudio longitudinal de los niveles 1 a 4 en primaria demostró que las funciones ejecutivas tienen un rol más fundamental en la comprensión de textos expositivos que en textos narrativos (Wu et al., 2020). Anteriormente, se ha comprobado que existe mayor complejidad de procesamiento y comprensión en la prosa más densamente informativa (Parodi, 2005).

Además de los tipos de textos, un área paralela es su grado de especialización. Si bien no pueden marcarse límites definidos entre un discurso más o menos especializado, la literatura postula características principales de los textos especializados:

1. usan medios lingüísticos de una forma y frecuencia específicos;

2. son un continuo de subtipos lingüísticos orientados a un tema específico;
3. son funcionales, en cuanto su propósito es facilitar la comunicación efectiva sobre una temática específica;
4. sufren variaciones desde una perspectiva diacrónica;
5. no son solamente una variación estilística o terminológica; y
6. emplean características no verbales en su concreción (Schröder, 1991).

Sobre los costos de procesamiento, un reciente estudio comprobó que los lectores de textos más especializados deben emplear una combinación de estrategias de comprensión superficiales y profundas cuando se enfrentan a textos sobre los que no manejan abundante conocimiento previo (Latini et al., 2021). Estos autores postulan que un amplio repertorio de estrategias superficiales y profundas y la evitación de pensamientos irrelevantes son los factores más correlacionados con la comprensión efectiva de textos especializados.

En resumen, el texto expositivo especializado tiene costos de procesamiento distintos a los textos menos especializados y se desconoce si su fijación en la memoria puede hacerse valer del anclaje, a diferencia de los textos narrativos, que pueden hacer uso de la visualización por su contenido concreto.

La lectura y escritura expositiva y especializada naturalmente requieren un lector y escritor que emplea procesos superficiales y profundos de forma estratégica, tema que será abordado más adelante.

Sobre la especialización desde el punto de vista de la escritura, la complejidad ha sido ampliamente estudiada en educación, distinguiendo claramente las trayectorias de desarrollo en la complejidad sintáctica (Hillocks, 1986). Su desarrollo se ha medido utilizando la noción de unidades mínimas terminales o unidades T, que conceptualiza una cláusula y todos sus elementos subordinados. Las medidas comunes de complejidad sintáctica han sido la longitud promedio de la cláusula, el número de cláusulas por unidad T y la longitud promedio de la oración, siendo todas ellas estadísticamente significativas al comparar distintos niveles de escolaridad, especialmente la longitud promedio de la unidad T y la longitud promedio de la cláusula. En sus estudios, Hunt (1965) también encontró diferencias significativas en el uso de cláusulas adjetivas y nominales, así como también entre nominales y número de modificadores para sustantivos, lo que es consonante con los hallazgos de O'Donnell et al. (1967). Es importante tener en cuenta que la longitud de la unidad T, la subordinación y el uso de adverbiales son más frecuentes en el texto expositivo que en el narrativo, y cuando se relacionan con la calidad de la escritura, la calidad del texto expositivo está directamente relacionada con la

complejidad sintáctica, mientras que no existe tal relación en los textos narrativos, aunque la evidencia sobre este asunto no es concluyente (Hillocks, 1986).

La riqueza léxica, otro aspecto de la complejidad, se ha estudiado con respecto a la calidad de la escritura y el dominio general en estudiantes de inglés de segunda lengua y lengua extranjera, y sigue siendo un aspecto importante de la investigación relacionada con la escritura en general (Li & Zhang, 2021). La investigación ha demostrado un aumento en las diferentes medidas de riqueza léxica en función del nivel de grado o competencia (H. Zhang et al., 2021). En este contexto, se ha encontrado que la variación léxica se correlaciona positivamente con la calidad de la escritura, aunque se ha reconocido que el uso real del vocabulario también depende del tema, la habilidad y el propósito comunicativo (Laufer & Nation, 1995). Los estudios sobre la riqueza léxica asumen que un uso sofisticado, rico y preciso del vocabulario proporciona evidencia sobre el dominio de la escritura y mejora la calidad (H. Zhang et al., 2021). Las medidas más comunes para la riqueza léxica son la originalidad léxica, que es el porcentaje del número total de tokens que son únicos; densidad léxica, que es el porcentaje de tokens léxicos en el número total de tokens; sofisticación léxica, que es el porcentaje de tokens avanzados en el número total de tokens léxicos, siendo “avanzados” una definición relativa al nivel del investigador y de los estudiantes; y la variación léxica, que es la relación tipo/token (*type-token ratio* [TTR]) entre las diferentes palabras del texto y el número total de palabras. Es

importante tener en cuenta que estas medidas muestran qué tan bien se expresan los estudiantes con el vocabulario que conocen; no el número de palabras real que conocen (Laufer & Nation, 1995).

Algunas dimensiones de la complejidad de la escritura, sin embargo, van más allá de aspectos superficiales como la riqueza léxica y la complejidad sintáctica. Desde una perspectiva sistémico-funcional (Eggins, 2004; Halliday, 2004; Halliday & Hasan, 1976, 1989), el lenguaje está constreñido por circunstancias sociales, situacionales y culturales que impregnan los textos y discursos. El aspecto cultural más amplio se denomina género, que puede identificarse por medio de tres características textuales (Eggins, 2004): (1) su configuración de registro, entendida como el medio y la forma de interacción (modo), las diferencias de roles y autoridad entre los participantes (tenor) y el área de especialidad en la que se propone el texto (campo); (2) su estructura esquemática, entendida como los pasos, movimientos o fases en el discurso, cada uno con su propósito; y (3) patrones de realización del texto, o la forma en que las metafunciones se manifiestan en el texto. En el contexto de los textos académicos, donde el modo es escrito y no interactivo, se puede especular que los escritores más experimentados mostrarían menos uso de verbos modales, lo que manifiesta certeza y autoridad, y más palabras académicas en términos de campo, que generalmente son más largas y abstractas, y tienen más sílabas.

Respecto a estas características, un análisis de corpus de 90 textos auténticos orales y escritos especializados y no especializados encontró 11 características comunes entre 65 características del español analizadas (Parodi, 2006). Los resultados mostraron que los textos especializados tienen más presencia de verbos modales de obligación, modo subjuntivo, nominalizaciones, participios en función adjetiva y frases preposicionales del complemento del nombre. Por el contrario, tienen menor presencia de desinencias de la tercera persona singular, pretérito indefinido, forma estativa activa “estar”, verbos privados, pronombres de negación y verbos modales de volición, todos los cuales proporcionan evidencia de aspectos generales de la complejidad que no están relacionados ni con la riqueza léxica ni con la complejidad sintáctica. Es importante señalar, sin embargo, que estos corpus incluyeron entrevistas, por lo que algunas de estas características pueden ser específicas del modo.

### ***Niveles de Competencia de los Sujetos y Metacognición***

A lo largo de la historia del modelamiento de los procesos de comprensión y producción de la lengua ha resultado interesante identificar aspectos de la comprensión y producción que distinguen dos tipos de lectores o escritores. Por una parte, existen lectores o escritores con mayor destreza, con más experiencia, o cuyos procesos de lectura o escritura son más estratégicos y eficientes en el uso de recursos cognitivos. Por otra parte, también existen lectores o escritores

con menor destreza, menos experiencia, o cuyos procesos de lectura o escritura son menos estratégicos y eficientes.

La investigación ha demostrado diferencias entre estos lectores y escritores con habilidades superiores e inferiores. Los lectores competentes automatizan la decodificación superficial, pudiendo destinar recursos cognitivos a manejar múltiples posibles hipótesis sobre el significado textual en paralelo, lo que posibilita inferencias más acertadas (van Dijk & Kintsch, 1983). Además, son capaces de generar esquemas de recuperación de información más efectivos (Ericsson & Kintsch, 1995), y tienen más éxito en un proceso de comprensión que implica la generación de inferencias gatilladas por vacíos semánticos en el texto (Gutierrez De Blume et al., 2021; McNamara, 2004; Soto et al., 2020; Soto, Gutiérrez de Blume, Jacovina, et al., 2019; van Dijk & Kintsch, 1983). Respecto a la producción escrita, este proceso se describe en escritores novatos como un proceso lineal y espontáneo, muy distinto a la producción por parte de escritores expertos, que está orientada a la resolución de problemas retóricos y de contenido, e interacción entre estos espacios de resolución de problemas (Scardamalia & Bereiter, 1992). Asimismo, se ha identificado una progresión en la composición escrita que va desde un proceso no directivo y espontáneo, más parecido a una corriente del pensamiento, hacia un proceso que organiza las ideas, las jerarquiza y requiere de planificación y revisión (Hayes, 2012).

Sobre estas diferencias, la teoría de la comprensión y producción de la lengua en contextos académicos propone un sistema de monitoreo y regulación cognitivo, lo que alude a la existencia de procesos metacognitivos durante la lectura y la escritura, entendiendo la metacognición como la conciencia de los procesos cognitivos, su monitoreo y regulación. En términos generales, la metacognición implica la consideración deliberada de los procesos cognitivos (Flavell, 1978, 1979). En sus inicios, conceptualizada como la conciencia de los procesos cognitivos propios, los procesos cognitivos de otras personas, las demandas que implican las tareas cognitivas, y las formas de abordarlas, evolucionó en dos componentes clave: el conocimiento sobre la cognición o conciencia metacognitiva, compuesta por el conocimiento declarativo, procedimental y condicional de los procesos cognitivos; y la regulación metacognitiva, que implica la planificación, estrategias de manejo de información, monitoreo de la comprensión, estrategias de depuración, y evaluación (Schraw & Dennison, 1994).

Una distinción importante es que la conciencia metacognitiva está compuesta por conocimiento estable, falible y de desarrollo tardío, mientras que la regulación metacognitiva se compone de saberes inestables, no necesariamente conscientes, y relativamente independientes de la edad (Brown, 1987, citado en Martí, 1995). También es importante destacar que la regulación metacognitiva y las estrategias que propicia no son automáticas y requieren

esfuerzo consciente, no siendo muchas veces observables en tareas de orden superficial (McNamara et al., 2007). Se ha demostrado que tanto la conciencia como la regulación metacognitiva son importantes para mejorar el desempeño en tareas cognitivas y transferir lo aprendido a situaciones nuevas, lo que es especialmente relevante para personas con discapacidad intelectual, ya que compensarían los déficits en habilidades cognitivas, mejorando su desempeño en lectura, que es una habilidad central para el desenvolvimiento en la sociedad actual (Soto, Rodríguez Poblete, et al., 2018).

Es importante considerar que las dificultades para abordar una tarea cognitiva pueden derivarse de problemas al ejecutar procesos cognitivos —dificultad cognitiva— o de problemas para determinar el proceso cognitivo apropiado —dificultad metacognitiva— (Martí, 1995). En el contexto de la investigación sobre lectura, no está claro el aporte de la cognición y metacognición, por cuatro razones principales: (1) Es complejo determinar el peso relativo de una u otra, (2) se requiere cierta conciencia de los sujetos para medir estos procesos, (3) la naturaleza de los métodos supone cierta separación o aislamiento de un aspecto específico de la metacognición, y (4) existe variedad de métodos para investigar la relación entre comprensión lectora y metacognición (Soto, Gutiérrez De Blume, et al., 2018).

Existe amplio consenso en la forma de recolectar información sobre la conciencia metacognitiva de los sujetos a través de cuestionarios y escalas. Sin embargo, la regulación metacognitiva, por su naturaleza sincrónica, inestable y contextual, tiene asociados distintos métodos y conceptualizaciones. Respecto a la conceptualización de la regulación metacognitiva, una de sus definiciones es como un proceso que incluye la evaluación y regulación (Hacker, 1998), lo que hace aún más compleja su medición. Una segunda postura teórica es definir la regulación como un proceso que involucra dos subprocesos: la evaluación, o detección de un problema en la comprensión, y la regulación, o reparo de este problema (Schraw & Dennison, 1994). Respecto a los procedimientos para medir la regulación, se ha empleado métodos relacionados al paradigma del error —detección de inconsistencias (evaluación) y reparo de ellas (regulación)—. En este paradigma, se han utilizado protocolos de pensamiento en voz alta, detección de inconsistencias y juicios de comprensión. Además, se han empleado métodos en línea como las regresiones oculares y los tiempos de lectura que evidencian la detección de estas inconsistencias (Otero, 1998).

Por otra parte, se ha comprobado que la efectiva regulación depende de la precisión en evaluación (Otero & Campanario, 1990) y depende de la definición de estándares de coherencia sobre los cuales se juzga la comprensión. Esta precisión recibe el nombre de calibración o precisión de la metacompreensión en el contexto de la lectura y se ha operacionalizado a través de la diferencia entre

los juicios de comprensión y el desempeño real en ella. En este contexto, se distingue la precisión de la metacomprensión absoluta global como la diferencia entre el juicio global de metacomprensión y la comprensión global lograda, y la precisión de la metacomprensión absoluta local o relativa como la diferencia entre el juicio de comprensión para cada ítem de una prueba de comprensión y el resultado en cada uno de ellos (Soto, Gutiérrez de Blume, Jacovina, et al., 2019). La literatura también señala que la precisión de la metacomprensión es, por lo general, baja (Soto et al., 2020). Sin embargo, no existe evidencia sobre la calibración del monitoreo en la expresión escrita.

En síntesis, los lectores y escritores con habilidades superiores automatizan los procesos superficiales y destinan recursos cognitivos a procesos estratégicos para abordar la lectura y la escritura. Sin embargo, la búsqueda y empleo de procesos estratégicos depende de la precisión con la que se juzga la comprensión. Si este juicio es preciso, se activarán procesos regulatorios que remedien la comprensión ineficiente cuando sea necesario. Por el contrario, si este juicio no es preciso, se activarán procesos regulatorios cuando no son necesarios, y no se activan cuando sí lo son. La precisión metacognitiva ha sido estudiada solamente en el contexto de la comprensión lectora, pero puede inferirse de esta revisión que los mismos procesos existen en la producción escrita, lo que se abordará con más detalle a continuación.

## **Procesos Mentales en la Comprensión y Producción Especializada**

Como se ha detallado anteriormente, existe amplio consenso en que la comprensión lectora tiene como resultado la creación de un modelo de situación, el que, por el contrario, sería el punto de partida para la expresión escrita. Se asume que los modelos de situación son representaciones mentales coherentes. Sin embargo, poco se sabe sobre su naturaleza más allá de que están compuestos por entidades mentales y sus relaciones (Tapiero, 2007). Quizá la noción más cercana a la naturaleza de un modelo de situación es la idea de esquemas, que son concebidos como configuraciones cognitivas que “modifican las impresiones producidas por los impulsos entrantes de tal manera que las sensaciones finales de la posición o localidad aparecen en la conciencia cargados con una relación a algo que ha sucedido anteriormente” (Bartlett, 1932, p. 200). En otras palabras, los esquemas son filtros a la experiencia externa y por lo tanto modulan la representación mental de ellas.

Los esquemas, siendo representaciones estables, interactúan con el texto en la comprensión para generar instancias de esquemas particulares, llamados modelos de situación. En efecto, ante la imposibilidad de los esquemas para adaptarse a situaciones nuevas o contextos determinados es como Kintsch y Van Dijk acuñan el concepto de “modelo de situación” (Tijero Neyra, 2009). Se ha postulado que la naturaleza de los modelos de situación es diversa, pudiendo por

ejemplo ser perceptual (espacial o visual) o ser similar a un mapa conceptual, dependiendo de la naturaleza del texto. Sin embargo, se ha postulado que un esquema en base a proposiciones puede representar todos estos tipos de modelos de situación (Tapiero, 2007).

En la investigación más reciente sobre las distinciones entre lectores y escritores de competencias superiores e inferiores, y sus representaciones mentales, las nociones de monitoreo metacognitivo (Gutierrez de Blume & Schraw, 2015; Soto, Gutiérrez De Blume, et al., 2018) y estándares de coherencia (Gutierrez De Blume et al., 2021; Otero, 2002; van den Broek et al., 2011) han cobrado mayor importancia. En el contexto de la comprensión lectora, se postula que los lectores evalúan el nivel de su comprensión o coherencia del modelo de situación que se va creando (Goodness [G]) en relación con estándares mínimos definidos en base a la tarea cognitiva, el texto, y la situación de lectura (Minimum Acceptability Level [MAL]). Por otra parte, respecto al monitoreo y regulación de la escritura, un estudio reciente (Escorcía & Gimenes, 2020) identificó tres variables importantes en la planificación: (1) Conocimiento condicional metacognitivo, referido a la conciencia de cuándo y para qué utilizar las estrategias cognitivas; (2) Autorregulación encubierta, referida al ajuste de las estrategias cognitivas o afectivas durante la tarea; y (3) Autorregulación ambiental, definida como el ajuste contextual de la tarea cognitiva.

El interés por explicar lo que diferencia a lectores-escriutores de habilidades superiores e inferiores en sus representaciones mentales y su monitoreo es un área de desarrollo paralela a la investigación sobre la relación entre los procesos de lectura y escritura. Si bien la investigación sobre lectura y sobre escritura se han desarrollado sin conexiones permanentes entre sí, se ha demostrado el paralelismo en estrategias metacognitivas empleadas en tareas de lectura y escritura por las mismas personas (por ejemplo, planificación de la escritura-previsualización de la lectura) (Tierney & Shanahan, 1996) y se ha propuesto la existencia de interacciones entre lectura, escritura y regulación de representaciones mentales que potencian el aprendizaje (Gutierrez De Blume et al., 2021). Sin embargo, estas investigaciones se han realizado distinguiendo perfiles de interacción entre lectura y escritura diferenciados según la competencia de lectores, pero a la fecha no se ha investigado sobre el rol de la metacognición como potenciadora de esta interacción.

Como puede observarse en esta revisión, un aspecto central que diferencia los procesos mentales de lectores y escriutores de habilidades superiores e inferiores es cómo definen y supervisan los estándares de coherencia sobre los cuales se juzga la calidad de la comprensión lectora y la expresión escrita. Es importante recordar que los estándares de coherencia se fijan en base a la tarea cognitiva, el texto, y la situación, que corresponden a representaciones mentales filtradas por los conocimientos previos del lector y

escritor, por lo que la precisión metacognitiva para establecer y supervisar estos estándares cobra mayor relevancia.

## **El sistema de Educación Superior en Chile**

En Chile, evaluar la alfabetización del estudiantado al ingresar a la educación superior requiere un esfuerzo importante. Aunque el sistema de educación superior generalmente distingue a las universidades privadas y públicas, así como a las tradicionales y no tradicionales, pueden clasificarse además como de investigación, masivas, orientadas a la acreditación, elitistas y no elitistas (Muñoz & Blanco, 2013). El proceso de selección para matricularse en universidades en Chile considera un cierto puntaje en una prueba de selección, generalmente en los componentes de lenguaje y matemática, pero también es importante considerar que, en este país, la calidad de la escolaridad depende en gran medida de variables socioeconómicas que dificultan el acceso de jóvenes de familias más pobres a la educación superior (Programa de las Naciones Unidas para el Desarrollo [PNUD], 2017). Este escenario también significa más desafíos para las universidades masivas y menos selectivas, ya que quienes ingresan a ellas tienen puntajes más bajos en lenguaje y matemática, lo que implica esfuerzos adicionales en estas instituciones para ayudar al estudiantado a desarrollar las habilidades necesarias para un aprendizaje efectivo en sus áreas de especialidad, considerando que la lectura y

la escritura no son solo habilidades que deben alcanzarse, sino también habilidades que median el desarrollo de competencias profesionales y académicas en un escenario de educación superior.

### **Síntesis y Propósito Investigativo**

La comprensión lectora está compuesta por tres niveles: la estructura superficial, donde el lector accede y obtiene información literal; la base textual, donde el lector integra e interpreta información local y global del texto; y el modelo de situación, donde el lector reflexiona y valora la información del texto en base a sus conocimientos previos. La producción escrita, por su parte, es un proceso donde interactúa el conocimiento del escritor, la representación de la tarea de escritura, y la generación de texto orientado a cumplir metas establecidas por el escritor en base a la situación. No existe duda que ambas están relacionadas, pero, no se cuenta actualmente con datos empíricos sobre cómo esta interacción varía en distintos estadios de desarrollo de la lectura y escritura, o de los distintos tipos de texto, que implican una aproximación evidentemente distinta a estas tareas.

Sobre los tipos de texto y lector, el texto expositivo especializado tiene costos de procesamiento distintos a los textos menos especializados y se desconoce si su fijación en la memoria puede hacerse valer del anclaje, a

diferencia de los textos narrativos que pueden hacer uso de la visualización por su contenido concreto. La lectura y escritura expositiva y especializada naturalmente requieren un lector y escritor estratégico, que automatiza los procesos superficiales y destina recursos mentales a procesos metacognitivos, cuya efectividad, a su vez, depende de la precisión metacognitiva con la que se definen y supervisan los estándares de coherencia sobre los cuales se juzga la comprensión y la expresión.

Considerando estos antecedentes, la investigación busca comprobar el potencial mediador de la precisión metacognitiva en interacción entre las habilidades de lectura y escritura en contextos académicos para estudiantes de primer año de una universidad privada no selectiva.

## **Objetivos**

### ***Objetivos Generales***

- Determinar dimensiones de complejidad escrita con sustento teórico a partir de los índices de TRUNAJOD para las tareas de escritura académica de estudiantes de primer año de una universidad privada no selectiva.

- Determinar el rol mediador de la precisión metacognitiva en el potencial de la escritura académica como variable predictora de la lectura académica para estudiantes de primer año de una universidad privada no selectiva.

### ***Objetivos Específicos***

- Determinar la complejidad en escritura académica a partir de los índices de lecturabilidad de TRUNAJOD.
- Determinar la validez de constructo de los modelos factoriales que expliquen una proporción significativa de la varianza en los índices de TRUNAJOD para textos académicos de estudiantes de primer año de una universidad privada no selectiva.
- Proponer un modelo factorial con sustento teórico que explique una proporción significativa de la varianza en los índices de TRUNAJOD para textos académicos de estudiantes de primer año de una universidad privada no selectiva.
- Determinar la existencia de diferencias significativas en las dimensiones de la complejidad escrita medida a través del modelo factorial de índices TRUNAJOD para textos de distintos niveles de complejidad.
- Determinar las competencias de lectura académica y precisión metacognitiva en estudiantes de primer año de una universidad privada no selectiva.

- Identificar las variables de la escritura que predicen significativamente las competencias de lectura académica.
- Determinar el aporte de la precisión metacognitiva como variable mediadora en la predicción de la lectura académica.

## **Hipótesis**

### ***Hipótesis de Investigación***

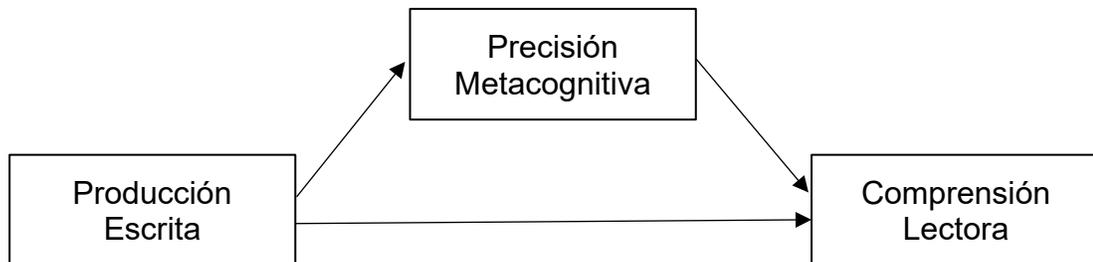
A continuación, se presentan las hipótesis de trabajo para la investigación, así como también el modelo hipotético de ruta que se busca comprobar en la segunda hipótesis de trabajo, y las hipótesis nulas que se busca refutar.

Hi<sub>1</sub>: Una proporción significativa de la varianza en los indicadores de complejidad de TRUNAJOD puede ser explicada por una estructura de dimensiones de complejidad más simple y con sustento teórico.

Hi<sub>2</sub>: La producción escrita, con la mediación de la precisión metacognitiva, predice significativamente el desempeño en lectura académica.

## Ilustración 1

### *Modelo Hipotético de Ruta*



### ***Hipótesis Nula***

Ho<sub>1</sub>: Existen modelos que explican una proporción significativa de la varianza en los indicadores de TRUNAJOD, pero estos no tienen sustento teórico, o bien, aquellos modelos que tienen sustento teórico no predicen una proporción significativa de la varianza en los indicadores de TRUNAJOD.

Ho<sub>2</sub>: Aunque existen relaciones entre la comprensión lectora, la producción escrita y la precisión metacognitiva, esta última no aporta significativamente a la predicción de la comprensión lectora desde la producción escrita.

## Capítulo II: Metodología

### Características de la Investigación

Con tal de cumplir con los objetivos y poner a prueba las hipótesis planteadas, se realizaron dos estudios desde un paradigma de investigación postpositivista, con un enfoque cuantitativo, y un diseño transeccional, de medición única (*single shot*), donde se realizaron mediciones de todas las variables, para luego efectuar los análisis estadísticos correspondientes a las hipótesis.

Con el ideal de la validez ecológica, se planteó un énfasis de la investigación sobre la lectura y escritura naturales. Si bien no existe consenso o definición clara de lo que es la lectura natural, una oración o texto natural, o cómo investigar al respecto (Haberlandt, 1994), la presente investigación buscó indagar sobre los procesos de lectura y escritura y su mediación a partir de la precisión metacognitiva en un escenario de lectoescritura que forma parte de los procesos de aprendizaje en la universidad, reduciendo la diversidad de situaciones y propósitos de los lectores, al mismo tiempo evitando la manipulación de los textos a leer y escribir. Esto implica una amenaza no menor a la validez de la investigación, que se pretende controlar a través de métodos complementarios, según se explicará a continuación.

## **Población y Muestra**

El universo corresponde a estudiantes universitarios en Chile que asisten a una universidad privada no selectiva (N = 148.966; Consejo Nacional de Educación [CNED], 2020). Dentro de este universo, se realizó un muestreo por conveniencia, escogiendo a los estudiantes de la Universidad Santo Tomás, que tiene presencia a nivel nacional con 13 sedes de Arica a Puerto Montt (N = 5.198; CNED, 2020). De estos estudiantes, la población estuvo conformada por estudiantes de primer año que al momento de la recogida de datos estuviesen inscritos en la asignatura “Taller de Competencias Comunicativas”, de la unidad de Formación General (N = 4.804). Se escogió esta asignatura, ya que es una cátedra que aborda el desarrollo de la lectoescritura, y las pruebas a aplicar corresponden a habilidades trabajadas en ella. Para cumplir con el tamaño muestral que permitiera generalizar los resultados a la población con un intervalo de confianza del 5% se calculó que se necesitaban al menos 356 estudiantes para generalizar con un nivel de confianza del 95%, si se considera un cálculo de muestra conservador (proporción 50%). El conteo final de estudiantes que participaron en la investigación puede resumirse de la siguiente manera:

- Estudiantes inscritos en la asignatura, que contestaron la prueba de lectura: 4095.

- Estudiantes que aceptaron participar en la investigación y contestaron la escala de monitoreo de la lectura: 481.
- Estudiantes de quienes se cuenta con puntuación de la escritura a través de rúbricas, además de sus puntajes en la prueba de lectura y la escala de monitoreo de la lectura: 152.
- Estudiantes de quienes se cuenta con todas las medidas anteriores y además sus textos escritos: 107.

Al determinar el intervalo de confianza con un cálculo conservador, y considerando la población y la muestra luego de la mortalidad muestral y la eliminación de casos atípicos en todas las medidas (N = 102), se obtuvo un intervalo de confianza de 9,41, lo que implica un margen de error de 9,4% para la investigación. Si bien es un intervalo mayor al aceptado en ciencias sociales, esta muestra incluye datos de lectura y escritura, además de los textos de estudiantes participantes.

## **Estudio 1**

El primer estudio<sup>1</sup> buscó comprobar la primera hipótesis de trabajo: “una proporción significativa de la varianza en los indicadores de complejidad de

---

<sup>1</sup> Este estudio y sus resultados han sido publicados en Pino Castillo, P. A., Soto, C., Asún, R. A., & Gutiérrez, F. (2023). Profiling support in literacy development: Use of natural language processing to identify learning needs in higher education. *Assessing Writing*, 58, 100787. <https://doi.org/10.1016/j.asw.2023.100787>

TRUNAJOD puede ser explicada por una estructura de dimensiones de complejidad más simple y con sustento teórico”. Para estos efectos, se empleó la herramienta TRUNAJOD con tal de obtener una evaluación automática y estandarizada de textos. TRUNAJOD es una herramienta ideada mediante un método basado en datos (data-driven) que analiza automáticamente textos escritos en español y extrae información sobre su complejidad. La herramienta proporciona un conjunto de índices superficiales y profundos y se ha utilizado para desarrollar análisis de lecturabilidad. En un estudio previo (Palma et al., 2019), los índices se agruparon en 5 dimensiones (similitud léxica, conectividad, coherencia referencial, concreción y narratividad) que se calibraron utilizando un corpus de textos escolares en español, lo que permitió la clasificación y descripción de un puntaje global de lecturabilidad. Este puntaje se empleó para medir la selección de materiales didácticos que se utilizan en la escuela. Un estudio posterior (Gutierrez De Blume et al., 2021) empleó índices de TRUNAJOD en textos narrativos e informativos producidos por estudiantes con tal de predecir la producción escrita a partir de la comprensión lectora. Sin embargo, la herramienta no se había empleado para analizar textos académicos producidos en el nivel de educación superior, lo que requirió un análisis diferente, considerando que los géneros tipifican situaciones, roles, acciones, audiencias potenciales, posturas, identidades y expectativas textuales del lector (Bazerman, 2016), y que el presente estudio emplea un enfoque de complejidad de escritura

desde la perspectiva del escritor en lugar de una de lecturabilidad desde la perspectiva del lector.

### ***Corpus***

Para el primer estudio, los datos consistieron en un conjunto de 107 bibliografías comentadas escritas por estudiantes, que formaban parte de la asignatura “Taller de Competencias Comunicativas”, curso de alfabetización académica al comienzo de todas las carreras de la universidad. Los estudiantes completaron un formulario de consentimiento informado para participar voluntariamente en el estudio. La participación consistió en el envío de su bibliografía comentada de mitad de semestre, las que escribieron siguiendo las instrucciones de seleccionar un tema de investigación para un ensayo de fin de semestre, leer un artículo de investigación relacionado con este tema y escribir la bibliografía que consistió en un resumen y evaluación del artículo leído. 107 bibliografías comentadas fueron procesadas por TRUNAJOD, que proporcionó 49 índices de complejidad de escritura superficial y profunda.

### ***Procedimiento de Análisis***

Tomando en cuenta la naturaleza del procesamiento por índices de TRUNAJOD, los textos fueron corregidos primero en su estructura superficial.

Esto incluyó la eliminación de encabezados en el texto y la corrección de la ortografía puntual, acentual y literal. Esto fue llevado a cabo por el mismo editor para garantizar la coherencia en todos los textos. Después de esto, los archivos de texto sin formato fueron procesados por TRUNAJOD que produce 49 índices superficiales y profundos. Se realizaron tres análisis factoriales exploratorios (AFE) para identificar posibles estructuras latentes. Las tres versiones fueron presentadas a jueces expertos, que fueron seleccionados porque formaban parte del equipo responsable de desarrollar el modelo teórico detrás de TRUNAJOD (Palma et al., 2019). Se les pidió que evaluaran la robustez teórica de cada modelo.

Una vez establecido el modelo, se calcularon los resultados factoriales para cada estudiante considerando las cargas factoriales del AFE. Luego, los textos se agruparon en (1) menor complejidad y (2) mayor complejidad utilizando el promedio de los factores, ponderando cada índice con su carga factorial, y dividiendo la muestra en estos dos grupos utilizando la media como punto de corte. Se utilizó una prueba t de muestras independientes para determinar si las diferencias entre estos grupos eran estadísticamente significativas en los diferentes factores y en la complejidad de escritura promedio.

## **Estudio 2**

El segundo estudio buscó comprobar la segunda hipótesis de trabajo: “la producción escrita, con la mediación de la precisión metacognitiva, predice significativamente el desempeño en lectura académica”. Con este propósito, se emplearon las medidas de comprensión lectora, precisión metacognitiva, evaluación de escritura a través de rúbricas, y dimensiones de TRUNAJOD generadas en el estudio 1. Las dimensiones ya identificadas fueron seleccionadas considerando investigaciones anteriores en el área, y por esta misma razón se adicionó un índice nuevo de TRUNAJOD que consideraba los errores ortográficos.

### ***Instrumentos y Materiales***

Para evaluar la comprensión lectora y la producción escrita, se emplearon instrumentos desarrollados por la Coordinación Nacional de Lenguaje de la universidad. Estos instrumentos fueron creados en base a los lineamientos de la prueba PISA de lectura. La prueba de comprensión lectora consiste en 30 preguntas, distinguiendo preguntas sobre la comprensión textual, preguntas sobre la comprensión inferencial, y preguntas sobre la comprensión crítica. Por su parte, la tarea de escritura considerada en esta investigación consiste en la producción de una bibliografía comentada en base a un artículo leído por cada

estudiante a mediados del semestre, y su evaluación a través de rúbricas analíticas de desempeño.

**Pruebas de Comprensión Lectora.** Con tal de potenciar la validez ecológica de la investigación y al mismo tiempo supervisar la validez de estos instrumentos, las pruebas se sometieron a análisis a través de juicios de expertos, donde se determinó el nivel de comprensión de cada pregunta en las 6 versiones de la prueba de lectura por facultad: nivel 1 (acceder y obtener), nivel 2 (integrar e interpretar) o nivel 3 (reflexionar y valorar). En promedio, la magnitud de acuerdo, medida a través del índice de Kappa de Fleiss (Landis & Koch, 1977), fue buena ( $\kappa = 0,72$ ) para todas las pruebas. Hubo una magnitud de acuerdo muy buena en el nivel 1 de las pruebas ( $\kappa = 0,84$ ) y buena en los niveles 2 ( $\kappa = 0,66$ ) y 3 ( $\kappa = 0,64$ ). El pormenorizado de estos análisis se resume a continuación. Los resultados de todos estos análisis se encuentran en los anexos.

**Tabla 2**

*Kappa de Fleiss para la Validación por Juicios de Expertos*

	General	Acceder y obtener	Integrar e interpretar	Reflexionar y valorar
Ciencias Sociales	0.676	0.765	0.586	0.694
Derecho	0.628	0.804	0.583	0.448
Economía y Negocios	0.717	0.885	0.653	0.538
Ingeniería	0.672	0.788	0.585	0.646
Recursos Naturales	0.797	0.837	0.753	0.810
Salud	0.834	0.972	0.775	0.715

Promedio	0.720	0.841	0.655	0.641
----------	-------	-------	-------	-------

**Evaluación de la Producción Escrita.** Para operacionalizar el desempeño en producción escrita en línea con los propósitos de la presente investigación, se consideraron, en primer lugar, los puntajes de escritura otorgados por los docentes de la asignatura. A pesar de ser una variable subjetiva con distintas miradas sin consenso comprobable, se consideró en el modelo predictor, dada su indudable importancia al reflejar el alineamiento con los propósitos de escritura del estudiantado, y considerando el rol de la disponibilidad de rúbricas para la calidad de los textos escritos (Sundeen, 2014).

Las dimensiones evaluadas en la producción escrita corresponden a componentes de contenido y componentes retóricos. En la bibliografía comentada se evalúa la presentación del autor, resumen de la lectura y comentario (contenido), así como también los aspectos ortográficos, adecuación, registro, léxico y formato (retórico). Tanto las lecturas como los trabajos de escritura son diversificados, adaptándose a la facultad a la que pertenecen los estudiantes. Estas seis versiones distintas de lecturas y trabajos escritos están alineadas en torno a las competencias a evaluar y los indicadores de desempeño.

**Complejidad de la Escritura.** Por otra parte, para los textos escritos se tomó en consideración el análisis de complejidad de textos a través de

TRUNAJOD. Como se explica más adelante, en los resultados del estudio 1, se identificaron 4 dimensiones de complejidad a través de un análisis factorial exploratorio, de las cuales se escogió la que, según la teoría existente, presentaba más potencial para conformar el modelo de ruta. Como medida adicional a los índices de TRUNAJOD agrupados en las dimensiones expuestas, se empleó un índice llamado “errores ortográficos promedio”. Con tal de alinear este índice a la complejidad, se transformó matemáticamente en una medida positiva, y se le denominó “ausencia de errores ortográficos” para la presente investigación.

De estas dimensiones se incluyó solamente la complejidad sintáctica y la ausencia de errores ortográficos. La ortografía, por una parte, teniendo en cuenta que se ha comprobado sus interacciones con la comprensión lectora (Berninger et al., 2002; Fitzgerald & Shanahan, 2000). Por otra parte, se incluyó la complejidad sintáctica dada la evidencia de su relación con la calidad del texto expositivo (Hillocks, 1986), además de la evidencia de un sistema compartido de comprensión y producción sintáctica en regiones corticales del cerebro (Indefrey, 2018), y el hecho de que estudios anteriores confirman la relación entre la lectura y la sintaxis (Allen et al., 2019).

**Precisión en el Monitoreo Metacognitivo.** La precisión en el monitoreo metacognitivo ha sido operacionalizada en las investigaciones del área a través

de dos indicadores: la precisión metacognitiva absoluta local, que se calcula a través de la diferencia absoluta entre las respuestas correctas en una prueba, y el juicio de los sujetos de haber respondido correctamente; y la precisión metacognitiva absoluta global, que corresponde a la diferencia absoluta entre el juicio de éxito global en una prueba de comprensión lectora luego de su aplicación y el desempeño global en ella. Si bien se calcularon tanto la precisión absoluta local como la global, se decidió solo emplear la precisión metacognitiva local, considerando la alta multicolinealidad entre ambas medidas, y el hecho de que la precisión local es una medida más detallada y comprehensiva de la precisión en la metacomprensión.

### ***Procedimiento de Análisis***

Antes de los análisis, se realizó la eliminación de casos atípicos identificados por sus respuestas al cuestionario de juicios de comprensión lectora, dejando fuera casos que respondieron 0% o 100% en todas las preguntas (N = 5). Luego de esto, se realizaron los análisis de normalidad de los datos a través de la prueba de Shapiro-Wilk, además otros supuestos estadísticos necesarios para el análisis de datos, como la normalidad, linealidad, multicolinealidad y homocedasticidad.

Inicialmente, se había planeado realizar un Modelo de Ecuaciones Estructurales (MES) para comprobar las relaciones entre todas las variables y los índices de TRUNAJOD. Sin embargo, el MES necesita variables manifiestas y variables latentes para combinar análisis factoriales y regresión, y comprobar el modelo a través de estos análisis. Considerando la alta multicolinealidad de los índices de TRUNAJOD que imposibilitaría el análisis MES, se procedió a realizar un análisis de ruta solo con variables manifiestas, y comprobar la existencia de variables mediadoras a través de una prueba de Sobel.

La prueba de Sobel es una prueba que permite determinar la significancia estadística de una variable mediadora en la relación entre una variable dependiente (criterio o resultado) y una independiente (predictor). Para comprobar el efecto mediador en una regresión, es necesario comprobar que la inclusión del mediador reduce el efecto de la variable independiente y el efecto del mediador sigue siendo estadísticamente significativo. Se habla de mediación completa cuando la introducción del mediador provoca que la variable independiente deje de ser estadísticamente significativa, y se habla de mediación parcial cuando el efecto de la variable independiente se reduce, pero sigue siendo significativo. Para el modelo hipotético a comprobar a través de la prueba de Sobel, se consideraron las siguientes variables:

- Comprensión lectora: pruebas de comprensión lectora de 30 ítems con aspectos superficiales, inferenciales y evaluativos de la lectura. Esta variable se hipotetizó como la variable de criterio o resultado.
- Precisión absoluta de la metacomprensión: se optó por utilizar la precisión absoluta local, calculada a partir de la diferencia absoluta entre de los juicios de comprensión y el desempeño por cada ítem de la prueba de lectura. Se desestimó el uso de la precisión global por la alta multicolinealidad entre ambas medidas, y considerando que la precisión local es una medida más comprehensiva de la metacomprensión. Esta variable se hipotetizó como mediadora.
- Desempeño escrito medido a través de rúbricas: medida de competencia en la escritura académica, evaluada por profesores de la asignatura según lineamientos de la Coordinación Nacional de Lenguaje. Considerando que las rúbricas son presentadas y revisadas antes de la evaluación escrita, se hipotetizó que su rol es mediador entre la producción escrita y la comprensión lectora, ya que consisten en una evaluación externa sobre el alineamiento de la producción escrita con estándares explicitados, en línea con la precisión metacognitiva. Anteriormente, se ha comprobado el impacto positivo de las rúbricas en la producción escrita cuando los estudiantes tienen acceso a ellas antes de la tarea (Sundeen, 2014).
- Complejidad sintáctica: factor compuesto por índices de TRUNAJOD que apuntan a la presencia de estructuras sintácticas más elaboradas, tales

como la longitud de oración, densidad de cláusula, palabras antes de la raíz, y marcadores discursivos, entre otros. Se consideró solo esta dimensión de la complejidad, ya que dentro de las 4 dimensiones que mide TRUNAJOD, es la que tiene más sustento teórico como posible predictor de la lectura (Allen et al., 2019; Hillocks, 1986; Indefrey, 2018).

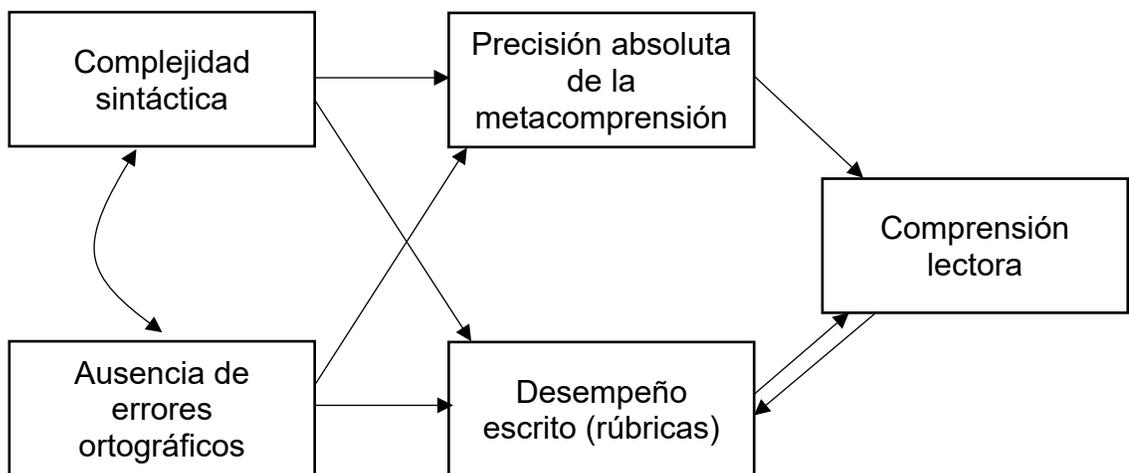
- Ausencia de errores ortográficos: índice en base a indicador automatizado de promedio de errores ortográficos en el texto. Para los propósitos de este experimento, se invirtió el índice con tal de medir la ausencia de errores. Este índice se agregó en el modelo predictor dado su sustento teórico (Berninger et al., 2002; Fitzgerald & Shanahan, 2000).

El análisis de ruta se empleó para evaluar la mediación de predictores y comprobar la hipótesis de trabajo. Para estos efectos, se examinó el efecto directo de un predictor sobre el resultado o criterio. A continuación, se investigó el efecto directo del predictor sobre el mediador. Luego, se evaluó el efecto directo del mediador sobre el resultado o criterio. Finalmente, se examinó un modelo de mediación completo con todas las variables incluidas. Existe mediación parcial si el coeficiente de regresión estandarizado entre el predictor y el resultado disminuye en función del mediador, pero sigue siendo estadísticamente significativo. La mediación completa, por otro lado, está presente cuando el coeficiente de regresión estandarizado entre el predictor y el resultado disminuye y además se vuelve estadísticamente no significativo debido

a la introducción del mediador. El modelo de análisis de ruta hipotético se ilustra a continuación.

## Ilustración 2

### *Modelo de Análisis de Ruta Hipotético*





## **Capítulo III: Resultados**

### **Resultados del Estudio 1**

Los resultados de los diferentes análisis de este estudio se presentan a continuación de la siguiente manera: (1) resultados de los análisis factoriales exploratorios y la opinión de expertos sobre el modelo final; y (2) estadísticas descriptivas para los grupos de escritores novatos y experimentados y diferencias estadísticas entre los grupos.

#### ***Resultados de Análisis Factoriales***

Como se mencionó anteriormente, se realizaron tres análisis factoriales exploratorios para su evaluación por jueces expertos. Como regla general, todas las cargas factoriales por debajo de ,3 fueron eliminadas de los modelos a evaluar, teniendo como resultado tres versiones estadísticamente robustas que consisten en cuatro, cinco y seis factores, considerando 46, 48 y 38 índices, respectivamente.

Después de la revisión por jueces expertos, la versión más sólida teóricamente fue la de cuatro factores, que explicó el 52,3% de la varianza. Esta decisión se tomó considerando que, dada la naturaleza de los índices de

coherencia local que se calculan por repetición de palabras, se consideró más apropiado que estuvieran inversamente relacionados con los índices de variabilidad de palabras, ambos pertenecientes a la dimensión de riqueza léxica, lo que no era el caso en el modelo de 6 factores. En esta decisión, también se tuvo en cuenta que se evita la repetición de palabras en la escritura académica. Otro aspecto que se consideró es que índices como la concreción y la imaginabilidad, como características del tipo de palabras que se utilizaban pertenecientes a discursos más especializados, se interpretaban mejor junto con índices como la longitud de las palabras y la densidad léxica, como aspectos de la coherencia registral, lo que descartaba el modelo de 5 factores. La lista de factores y los índices que los componen se detallan en la tabla a continuación, junto con sus cargas factoriales.

**Tabla 3**

*Modelo Exploratorio de 4 Factores para la Complejidad Escrita*

Índices TRUNAJOD	Factor 1	Factor 2	Factor 3	Factor 4
TTR lemma	.954			
TTR noun	.856			
TTR function	.784			
TTR content	.840			
TTR prp	.510			
TTR argument	.876			
SM semantic_giveness	-.818			
EG local_coherence_PU	-.825			
EG local_coherence_PW	-.937			
EG local_coherence_PACC	-.891			
EG local_coherence_PU_dist	-.845			

EG local_coherence_PW_dist	-.665	
EG local_coherence_PACC_dist	-.556	
SM word2vec_sent_sim	.770	
SM coherence_fasttext	-.677	
SP average sentence length	.942	
SP words before root	.366	
SP clause density	.738	
DM polysemic discourse markers	.626	
DM cause discourse markers	.620	
DM vague meaning (closed_class_words)	.926	
DM all types of discourse markers	.951	
TTR verb		.449
TTR adv		.396
SM verb_syn_overlap		.403
SP syntactic similarity		-.892
SP density 1 and 2 person		-.444
SP negation density		-.518
SP NOUN PROPN density		.868
SP NOUN density		.688
SP VERB AUX density		-.852
SP average clause length		.720
GV Pronoun Noun ratio		-.471
TTR adj		-.342
TTR lexical diversity MTL		.324
SM noun_syn_overlap		.418
SM lexical_overlap		.522
LSM concreteness		-.424
LSM imageability		-.569
LSM avg_context_availability		-.419
SP lexical density		.538
SP Verbs / Nouns		.386
SP average word length in letters		.605
SP average word length in syllables		.494
SP ADJ density		.579
GV Pronoun density		-.388

---

La interpretación teórica de estos cuatro factores dio como resultado las dimensiones que explican la complejidad de la escritura, las que se presentan a continuación junto con una descripción relacionada con los índices que contienen.

**Riqueza Léxica.** Esta dimensión comprende el rango de vocabulario empleado por el escritor, medido en diferentes índices de relación tipo-token (*type-token ratio* [TTR]) que indican la proporción de palabras (tokens) en un tipo específico de palabras. Un TTR más alto muestra que las palabras no se repiten. Esta dimensión también está inversamente relacionada con indicadores basados en repetición de palabras, como los de coherencia, y el índice llamado *semantic givenness*, que indica la similitud en el significado en oraciones contiguas.

**Complejidad Sintáctica.** Esta dimensión comprende índices relacionados con estructuras gramaticales y textuales más elaboradas, como la longitud de las oraciones, el número de cláusulas por oración, el uso de marcadores de discurso y las palabras de conexión. También se relaciona positivamente con el uso de construcciones sintácticas familiares, que muestran precisión gramatical, e inversamente relacionado con la coherencia utilizando *FastText*, que utiliza n-gramas de caracteres en lugar de tokens, por lo que está inversamente relacionado con el uso de pronombres de referencia.

**Estructura Informativa.** Considerando el género de las bibliografías comentadas, que es un texto expositivo más que narrativo, esta dimensión agrupa índices relacionados con el uso de la tercera persona, por lo que se relaciona negativamente con indicadores de textos narrativos. Estos incluyen, por ejemplo, el uso de primera y segunda persona, el uso de nombres comunes y propios, y uso de verbos auxiliares. Por otro lado, se encuentra más diversidad de verbos y adverbios, lo que significa que las oraciones expresan diferentes acciones o no acciones y sus modificadores, a diferencia de un texto narrativo en el que las acciones tienden a ser más similares en todas las oraciones. Finalmente, esta dimensión se relaciona con menor uso de la negación, que denota un discurso asertivo, y estructuras sintácticas menos similares en las oraciones, debido a los diferentes verbos y sus argumentos asociados.

**Discurso Especializado.** Considerando que la lectura de artículos y la redacción de bibliografías comentadas está circunscrita en una disciplina particular, esta dimensión muestra cuán especializado es el discurso en términos de la longitud de las palabras en letras o sílabas, lo que demuestra un lenguaje más disciplinar. Esta dimensión también se caracteriza por la menor presencia de palabras concretas e imaginables, lo que es esperable en textos académicos. Asimismo, existe una menor disponibilidad de contexto, lo que significa que no es probable que el significado de las palabras se deduzca del contexto, el sinónimo o el contenido de las palabras. Esto muestra que las entidades a las

que se hace referencia se repiten a través de las mismas palabras o sus sinónimos. Finalmente, también se encuentra una alta proporción de adjetivos, pero baja diversidad, lo que va de la mano con la repetición de entidades a las que se hace referencia, ya sea por los mismos sustantivos o sus sinónimos.

### ***Diferencias entre Textos de Mayor y Menor Complejidad***

Una vez establecidas las dimensiones de complejidad de escritura y los jueces expertos coincidieron en que había suficiente apoyo teórico para la agrupación de índices en estas categorías, se utilizó el promedio ponderado basado en las cargas factoriales AFE y los índices TRUNAJOD. Los resultados de toda la muestra se muestran en la tabla a continuación.

**Tabla 4**

#### *Suma de Promedios Ponderados para la Complejidad Escrita*

	N	M	SD
Riqueza Léxica	107	-3,75	20,40
Complejidad Sintáctica	107	88,60	28,68
Estructura Informativa	107	12,02	3,93
Discurso Especializado	107	18,84	4,22

La suma de los promedios ponderados se utilizó para dividir la muestra en (1) textos de menor complejidad ( $N = 58$ ) y (2) textos de mayor complejidad ( $N = 49$ ). La estadística descriptiva para estos dos grupos se presenta en la tabla 5 y

una prueba t de muestras independientes se presenta en la tabla 6, junto con la d de Cohen para el tamaño del efecto. Todos los datos se distribuyeron normalmente.

**Tabla 5**

*Estadística Descriptiva para Textos de Menor (1) y Mayor (2) Complejidad*

Dimensión	Grupo	N	M	SD
Riqueza Léxica	1	58	-9,55	17,33
	2	49	3,12	21,76
Complejidad Sintáctica	1	58	72,62	21,87
	2	49	107,52	23,98
Estructura Informativa	1	58	11,61	3,23
	2	49	12,51	4,61
Discurso Especializado	1	58	18,55	3,62
	2	49	19,19	4,86
Complejidad Escrita Promedio	1	58	93,23	17,13
	2	49	142,33	19,21

**Tabla 6**

*Prueba t para Muestras Independientes*

Dimensión	t	df	p	d
Riqueza Léxica	-3,353	105	< ,001	-0,651
Complejidad Sintáctica	-7,867	105	< ,001	-1,526
Estructura Informativa	-1,187	105	0,119	-0,230
Discurso Especializado	-0,772	105	0,221	-0,150
Complejidad Escrita Promedio	-13,973	105	< ,001	-2,711

*Nota.*  $H \mu_1 < \mu_2$

Las tablas 5 y 6 muestran la estadística descriptiva e inferencial para los puntajes totales agrupados en los cuatro factores. Como puede verse, los textos de menor complejidad (1) obtuvieron puntuaciones de complejidad de la escritura y todas sus dimensiones más bajas que los textos de mayor complejidad (2). Sin embargo, estas diferencias solo fueron estadísticamente significativas para la complejidad de escritura en su conjunto, y para las dimensiones de riqueza léxica y complejidad sintáctica. Además, la complejidad escrita promedio y la complejidad sintáctica mostraron un gran tamaño de efecto, mientras que la riqueza léxica solo mostró uno moderado.

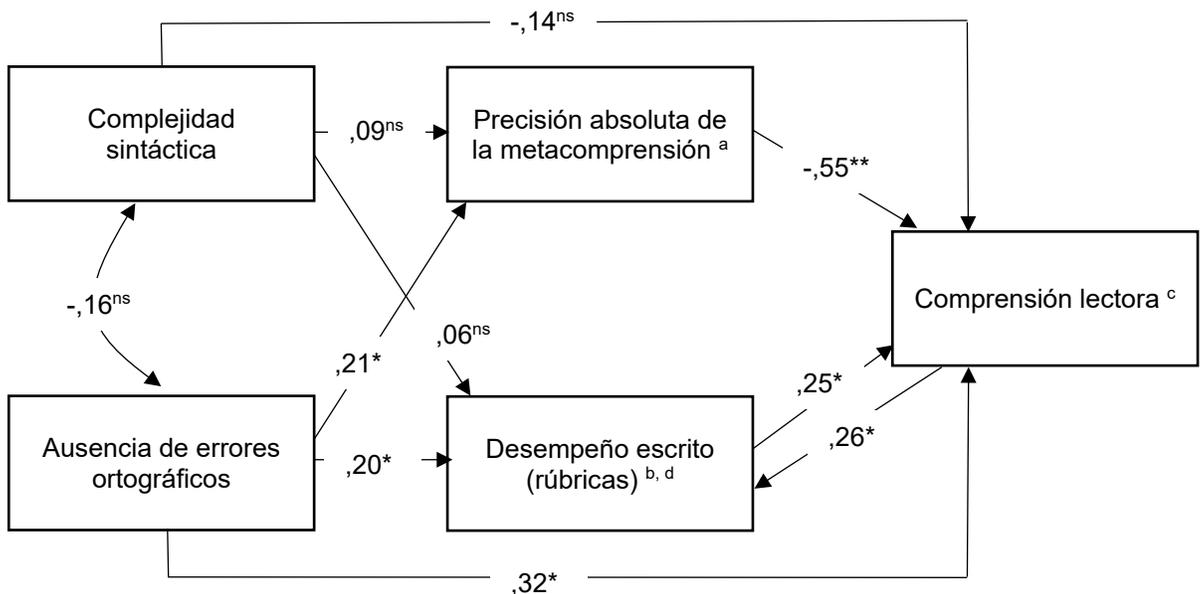
## **Resultados del Estudio 2**

Los resultados del análisis de ruta revelaron que la complejidad sintáctica no fue un predictor significativo de ninguno de los mediadores (precisión local absoluta de la metacompreensión o desempeño escrito) o de la comprensión lectora. La ausencia de errores ortográficos, por otro lado, predijo significativamente tanto los mediadores ( $\beta = ,21$  para precisión absoluta de la metacompreensión;  $\beta = ,20$  para desempeño escrito) como la comprensión lectora ( $\beta = ,32$  para comprensión lectora). Además, ambos mediadores predijeron significativamente la comprensión lectora (precisión absoluta de la metacompreensión;  $\beta = -,55$ ; desempeño escrito,  $\beta = ,25$ ), mostrando la comprensión lectora una relación recíproca con el desempeño escrito ( $\beta = ,26$ ).

Sin embargo, las pruebas de mediación de Sobel para la significancia de los efectos mediadores indirectos indicaron que solo la precisión local absoluta de la metacompreensión ( $\beta = ,20$ ,  $z = ,023$ ) mediaba de forma significativa y parcial la relación entre la ausencia de errores ortográficos y la comprensión lectora. Las estimaciones de los parámetros del modelo final y los tamaños del efecto se presentan en la ilustración 3 a continuación.

### Ilustración 3

#### *Análisis de Ruta Final con Parámetros del Modelo*



*Nota.* Tamaño del efecto,  $R^2$ : <sup>a</sup> ,125; <sup>b</sup> ,099; <sup>c</sup> ,414; <sup>d</sup> ,062 (comprensión lectora prediciendo desempeño escrito). \*  $p < ,05$  \*\*  $p < ,01$  <sup>ns</sup> Estadísticamente no significativo.

## Conclusiones

Los resultados obtenidos proveen evidencia para rechazar probabilísticamente ambas hipótesis nulas establecidas para la presente investigación:

Ho<sub>1</sub> Existen modelos que explican una proporción significativa de la varianza en los indicadores de TRUNAJOD, pero estos no tienen sustento teórico, o bien, aquellos modelos que tienen sustento teórico no predicen una proporción significativa de la varianza en los indicadores de TRUNAJOD.

Ho<sub>2</sub> Aunque existen relaciones entre la comprensión lectora, la producción escrita y la precisión metacognitiva, esta última no aporta significativamente a la predicción de la comprensión lectora desde la producción escrita.

Esto implica aprobar probabilísticamente ambas hipótesis de trabajo, lo que se traduce, en primer lugar, en que (Hi<sub>1</sub>) una proporción significativa de la varianza en los indicadores de complejidad de TRUNAJOD puede ser explicada por una estructura de dimensiones de complejidad más simple y con sustento teórico. Más específicamente, 52,3% de la varianza observada puede explicarse por una estructura de 4 dimensiones con sustento teórico, a saber, (1) riqueza

léxica, (2) complejidad sintáctica, (3) estructura informativa y (4) discurso especializado. Adicionalmente, se encontraron diferencias significativas entre textos de mayor y menor complejidad en el promedio total de complejidad escrita, así como también en las dimensiones superficiales de riqueza léxica y complejidad sintáctica.

En segundo lugar, esto también se traduce en que (H<sub>2</sub>) la producción escrita, con la mediación de la precisión metacognitiva, predice significativamente el desempeño en lectura académica. En términos más específicos, la ausencia de errores ortográficos, con la mediación parcial de la precisión absoluta de la metacomprensión, es capaz de predecir la comprensión lectora de forma significativa, explicando un 20% de la varianza. Por el contrario, la complejidad sintáctica no fue capaz de predecir la comprensión lectora, incluso con la mediación de la precisión metacognitiva. Finalmente, el desempeño en escritura evaluado a través de rúbricas fue capaz de aportar predicción significativa de la comprensión lectora, y esta relación existe también a la inversa, donde la comprensión lectora aporta a la predicción de la producción escrita evaluada a través de rúbricas.

En otras palabras, la ortografía es capaz de predecir la comprensión lectora, y esta predicción es potenciada por la precisión de la metacomprensión. Esto quiere decir que para los estudiantes que son capaces de juzgar más

precisamente su comprensión lectora, el efecto predictor de la ortografía sobre su comprensión lectora es aún mayor.

En términos de los procesos mentales que subyacen a estas mediciones, podemos argumentar que la precisión metacognitiva potencia las interacciones entre el sistema de producción escrita y comprensión lectora. En particular, aspectos superficiales de la escritura como la ortografía son capaces de predecir la competencia lectora, y esta interacción es más potente en personas que tienen una conciencia más precisa sobre sus propias competencias lectoras.

## **Discusión General**

En cuanto a los resultados sobre la complejidad de textos y las diferencias significativas en riqueza léxica y complejidad sintáctica, los resultados confirman la teoría existente al respecto de la complejidad en una primera lengua y una lengua extranjera. En primer lugar, se ha demostrado que la riqueza léxica mejora gradualmente a medida que los estudiantes avanzan en su educación, y la densidad léxica muestra un incremento más lento en comparación con la diversidad y la sofisticación (Zou, 2022). Al respecto, se ha argumentado que la exposición lingüística es un elemento crítico en el desarrollo de la riqueza léxica (Li & Zhang, 2021; H. Zhang et al., 2021), lo que también puede aplicarse al discurso especializado, con la diferencia de que último depende de la exposición

a un léxico específico perteneciente a la disciplina. En resumen, el desarrollo de la riqueza léxica y el discurso especializado parecen depender de la lectura, lo que respalda la noción de que la lectura y la escritura son actividades que se aprenden mejor juntas (Graham, Liu, Aitken, et al., 2018). En cuanto a la complejidad sintáctica, se ha demostrado que la subordinación y el uso de adverbiales son características más frecuentes en los textos expositivos que en los narrativos, y la calidad del texto expositivo está directamente relacionada con la complejidad sintáctica (Hillocks, 1986). Además, estudios previos han demostrado que la instrucción directa y centrada en aspectos de forma tiene un efecto sobre la complejidad sintáctica, así como la diversidad léxica en inglés con fines académicos (Mazgutova & Kormos, 2015), y sobre la complejidad sintáctica en chino como segundo idioma (Zhou & Lü, 2022). Tomando estos hallazgos en su conjunto, es de esperar que estudiantes de primer año de una universidad privada no selectiva presenten diferencias en estos índices, considerando las desigualdades del sistema escolar chileno (PNUD, 2017).

Por otra parte, la dimensión de estructura informativa está relacionada con aspectos generales del texto que coinciden con la coherencia registral en Lingüística Sistémico Funcional (Eggins, 2004; Halliday, 2004; Halliday & Hasan, 1976, 1989). En otras palabras, la estructura del texto informativo es la capacidad del escritor para cumplir con el nivel de impersonalidad y autoridad que requiere un texto informativo (coherencia en modo y tenor). Considerando que se han

encontrado grandes diferencias en las prácticas de escritura y lectura entre la escuela y la educación superior (Baker, 2018), es esperable que en esta dimensión no existan diferencias significativas en complejidad de textos, ya que estas nuevas prácticas afectan a todo el estudiantado en una forma similar.

En cuanto a los hallazgos sobre la predicción de la lectura y el rol potenciador de la precisión metacognitiva, estos resultados demuestran que aspectos superficiales de la producción escrita, como lo es la ortografía, son predictores significativos para la competencia de lectura en contextos académicos gracias a la mediación de la precisión metacognitiva. Esto es consonante con hallazgos anteriores sobre las relaciones entre comprensión lectora y producción escrita, que comprueban estas interacciones preponderantemente en un nivel superficial (Berninger et al., 2002). Más específicamente, los estudios longitudinales sobre las conexiones entre lectura y escritura (Abbott et al., 2010), han solamente comprobado una relación entre la ortografía y el reconocimiento de palabras, lo que coincide con los hallazgos de Fitzgerald y Shanahan en estadios iniciales (2000). La importancia del presente estudio sobre esta área de conocimiento es la confirmación de que esta predicción es potenciada por la precisión metacognitiva. Esto indica que un aspecto central en la interacción entre comprensión y expresión es la precisión con la que se juzga la comprensión, lo que implica que potenciar la precisión de la comprensión no solo beneficiaría la competencia en lectura, sino que también

el traspaso de competencias de escritura hacia competencias de lectura. Este nuevo hallazgo complementa la perspectiva de que la relación entre lectura y escritura se vuelve más potente cuando los individuos tienen una mejor representación mental del texto (Gutierrez De Blume et al., 2021), e identifica que la precisión metacognitiva es un aspecto específico que potencia esta relación.

Las principales limitaciones de esta investigación fueron su potencial generalizador y la validez instrumental de las medidas de comprensión lectora y producción escrita. Como se explicó anteriormente, estos resultados pueden generalizarse a la población con un margen de error de 9,4%, lo que implica que la ausencia de errores ortográficos potenciada por la precisión metacognitiva predice entre un 10,6% y un 29,4% de la comprensión lectora. Por otro lado, si bien las pruebas de comprensión lectora se sometieron a validación de juicios de expertos con una buena magnitud de acuerdo entre jueces ( $\kappa = 0,72$ ), los puntajes de producción escrita responden a una evaluación a través de rúbricas evaluativas sin cálculo de validez inter-observador.

### **Implicancias para la Teoría**

Considerando las dimensiones de complejidad escritas encontradas en este estudio y las diferencias entre escritura más y menos compleja, los hallazgos confirman trayectorias de aprendizaje en complejidad sintáctica, teniendo en

cuenta también que la calidad del texto expositivo está directamente relacionada con la complejidad sintáctica (Hillocks, 1986). Curiosamente, también se observa que la riqueza léxica era un factor importante al separar textos de mayor y menor complejidad, e investigaciones anteriores han demostrado una relación directa entre la riqueza léxica y la calidad de la escritura (Li & Zhang, 2021; H. Zhang et al., 2021), aunque es importante tener en cuenta que el uso real del vocabulario depende del tema y el contexto, así como del conocimiento del vocabulario. Por lo tanto, el rendimiento de la escritura no proporciona una imagen completa del desarrollo léxico y la sofisticación (Laufer & Nation, 1995).

Adicionalmente, aunque la coherencia registral en modo, tenor y campo se identificaron como dos dimensiones separadas, a saber, la estructura informativa y el discurso especializado, es importante señalar que estos no mostraron diferencias significativas entre textos de mayor y menor complejidad, lo que puede apuntar al hecho de que la complejidad de la escritura es más o menos independiente de estos aspectos, al menos en el contexto de primer año de educación superior en una universidad privada no selectiva. Se necesita investigación adicional para relacionar estos índices de complejidad de la escritura con la calidad de la escritura evaluada por especialistas en lenguaje y en las disciplinas.

A la luz de la perspectiva sociocultural del desarrollo de la alfabetización en la educación superior, estos resultados apuntan a la primacía de un enfoque de habilidades de estudio en lugar de la socialización o alfabetización académica (Lea & Street, 1998, 2006). Sin embargo, los estudios futuros deben comparar la complejidad de la escritura obtenida por la aplicación de esta herramienta y la evaluación de la producción escrita por parte de evaluadores, para determinar si la conciencia crítica de la identidad y las relaciones de poder, y la forma en que estos aspectos se cristalizan en configuraciones lingüísticas particulares, es un aspecto relevante a desarrollar en estudiantes universitarios que están en proceso de aprendizaje en sus disciplinas. Es importante mencionar que las herramientas automáticas de evaluación de escritura no pueden reemplazar la evaluación del profesorado y la detección de las necesidades de aprendizaje, especialmente en aspectos globales como la autoridad y la aculturación. Sin embargo, este desarrollo puede facilitar la detección de necesidades relativas a aspectos superficiales que pueden escapar a la evaluación humana.

Por otro lado, en cuanto a las interacciones entre escritura, lectura y precisión metacognitiva, los hallazgos de esta investigación confirman no solo la importancia de la metacognición en procesos lectoescriturales más estratégicos (Gutierrez De Blume et al., 2021; McCarthy et al., 2018, 2020; Soto, Gutiérrez De Blume, et al., 2018; Soto, Gutiérrez de Blume, Jacovina, et al., 2019; Soto, Rodríguez Poblete, et al., 2018), sino que además de ser una variable que

potencia la lectura y la escritura, potencia además la interacción entre ambas habilidades, reforzando la predicción de la comprensión lectora.

Adicionalmente, estos resultados aportan a la consideración de un sistema mediador de las interacciones entre comprensión y producción, que ya ha sido propuesto en estudios anteriores (Gutierrez De Blume et al., 2021). Más allá de considerar la existencia de sistemas de interacción, como ha sido propuesto anteriormente (Fitzgerald & Shanahan, 2000), o paralelismos entre aspectos de la comprensión y la producción como la planificación y la previsualización (Tierney & Shanahan, 1996), estos hallazgos demuestran que la precisión metacognitiva puede potenciar su interacción, más allá de las habilidades y sus componentes similares.

En el contexto de la lectura, existe amplio conocimiento sobre la precisión de la metacompreensión que es, por lo general, baja (Soto et al., 2020), y está en la base de la evaluación de la comprensión en relación con estándares mínimos de aceptabilidad (Otero, 2002) y las acciones regulatorias para remediar la comprensión deficiente (Otero & Campanario, 1990). Sin embargo, el presente estudio, además, comprueba que esta precisión está en la base de la interacción entre la producción escrita y la comprensión lectora. En este sentido, los hallazgos de la presente investigación invitan a reconsiderar la investigación

sobre lectura, escritura y estándares de coherencia de forma conjunta para evaluar la relación entre estos tres constructos.

En síntesis, estos hallazgos proponen el impacto de la precisión metacognitiva para potenciar la lectura y la escritura en contextos académicos. Al demostrar que la precisión con la cual los lectores juzgan su comprensión potencia la transferencia de competencias de escritura, al menos en sus aspectos superficiales, hacia la competencia de comprensión lectora, estos resultados invitan a replantear esta interacción en distintos niveles, en base a la conciencia metacognitiva de los sujetos más allá de sus niveles de competencia.

Adicionalmente, los resultados de la presente investigación invitan a considerar la existencia de diversos estándares en torno al proceso de escritura: estándar de coherencia, de aspectos formales, de contenido. A partir de estos resultados, puede inferirse que el modo en que estos van cambiando dependerá del período de desarrollo en lectoescritura en que se encuentre el sujeto y del tipo de tareas propuestas, como se ha demostrado anteriormente (Gutierrez De Blume et al., 2021). La literacidad va evolucionando a través de la experiencia, y por tanto se produce una variación en la forma en que ciertos componentes de esta van tomando mayor relevancia. Como se ha demostrado en investigaciones anteriores, los lectores con mayor experiencia tienden a automatizar aspectos superficiales del proceso, por lo que pueden destinar recursos cognitivos a

aspectos más estratégicos (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983), lo que también se ha comprobado en la escritura (Hayes, 2012; Scardamalia & Bereiter, 1992). Si a esto se añade la consideración sobre el tipo de tarea, se delinea un panorama variable, en que cierto tipo de relación entre componentes adquiere un mayor protagonismo, perfilando un especial vínculo entre lectura y escritura, a partir de la interacción de variables particulares.

Por otro lado, y como ya sea visto en diferentes estudios (Gutierrez De Blume et al., 2021; Soto et al., 2022), las relaciones entre lectura y escritura están mediadas por componentes metacognitivos que afectan esta relación, especialmente en tareas que requieren un intenso monitoreo en torno al proceso de ejecución para el logro de un determinado desempeño. En este sentido, los hallazgos del presente estudio permiten inferir aspectos de los procesos cognitivos pertinentes a la muestra, compuesta por estudiantes que ingresan a la educación superior, y que en general presentan un bajo desempeño en lectoescritura. En este contexto, los resultados apuntan a que, al aproximarse inicialmente a un discurso especializado propio de las diferentes disciplinas, los estudiantes con bajo desempeño en lectoescritura pierden el control sobre los aspectos más profundos en el procesamiento de la información, lo que los lleva a apoyarse en aspectos más superficiales de los textos (Scardamalia & Bereiter, 1992), sobre los cuales el monitoreo metacognitivo cobra mayor importancia al potenciar la predicción de la lectura. Es justamente en ese plano en el que el

presente estudio permite evidenciar que se producen las conexiones más importantes. Dadas las tareas de alta complejidad que se les propone al estudiante, es en la dimensión más superficial (vinculado a la ortografía y aspectos formales) donde juega un rol preponderante la precisión metacognitiva para potenciar la transferencia a la competencia lectora.

### **Implicancias para la Enseñanza y el Aprendizaje**

Los hallazgos sobre la complejidad escrita se suman a las aplicaciones continuas de la PLN en el desarrollo de la alfabetización. A la fecha, los estudios han utilizado estas tecnologías para evaluar la complejidad de la escritura como apoyo para la instrucción de lectura (Palma et al., 2019), el impacto de las correcciones automatizadas en la competencia, el compromiso y los factores afectivos (Nazari et al., 2021) y los tipos de retroalimentación y posibles aplicaciones de estos (Calma et al., 2022). A diferencia de estudios anteriores, este trabajo utilizó el PLN para perfilar la complejidad de la escritura y detectar de una manera más objetiva los aspectos descendidos en la escritura de estudiantes de educación superior, lo que se ha hecho anteriormente en inglés (Allen et al., 2019). Estos resultados son similares a los de Allen et al. (2019), en que los textos de mayor complejidad tienen palabras menos concretas y una estructura de texto menos narrativa. Sin embargo, es interesante notar que, en el estudio mencionado, la complejidad sintáctica fue menor en textos más

complejos para ensayos argumentativos, mientras que los hallazgos de la presente investigación muestran lo contrario, con estructuras sintácticas más complejas utilizadas en textos más complejos. Estos resultados pueden contribuir al desarrollo de una herramienta automatizada de evaluación de la escritura en español, particularmente en los aspectos superficiales, considerando las discrepancias entre tutores al evaluar la escritura (Mumford & Atay, 2021). Una mayor investigación sobre esta área emergente también puede identificar si estas dimensiones son exclusivas de la escritura académica, o si son comunes a las prácticas de escritura en diferentes contextos y géneros en español.

En cuanto al uso y desarrollo de herramientas basadas en PLN, es importante tener en cuenta que estudios previos han encontrado ventajas y advertencias en su implementación. Las principales ventajas han sido el impacto en el dominio de la escritura de los estudiantes, ya que se centran en aspectos superficiales de la escritura (Geng & Razali, 2022; S. Zhang, 2021), y en su autonomía cuando estas herramientas se utilizan como retroalimentación formativa, que es una diferencia crucial entre las prácticas de alfabetización en la escuela y en el nivel de educación superior (Baker, 2018). Sin embargo, como no evalúan el contenido de la escritura, estas herramientas no pueden reemplazar una evaluación integral. Además, existe el peligro adicional de que el estudiantado aprenda por el uso excesivo a centrarse exclusivamente en los aspectos de su escritura que se evalúan automáticamente, dejando de lado las

características de la escritura que van más allá de lo superficial, como la adecuación social y cultural del texto escrito (Correnti et al., 2022). Todo esto indica que el uso más seguro de estas herramientas es como evaluación diagnóstica, ya que pueden ayudar al profesorado a dirigir sus procesos de enseñanza y atender las necesidades de los estudiantes de manera más efectiva (Geng & Razali, 2022).

La identificación de dimensiones de complejidad textual tiene implicancias importantes para una universidad privada no selectiva, en la que el dominio inicial de lectura y escritura de los estudiantes necesita más apoyo en consideración de sus diversos antecedentes que crean amplias desigualdades en su punto de partida en la universidad (PNUD, 2017), y los deficientes resultados de alfabetización en nuestro país (OCDE, 2016). A la fecha, esto se ha hecho desde una perspectiva general, considerando la evaluación de la escritura por parte de docentes y con cursos de apoyo a la alfabetización para todos los estudiantes. En este sentido, este desarrollo proporcionará información para diseñar e implementar estrategias de apoyo que puedan atender más específicamente las necesidades de los estudiantes en estas habilidades centrales que median el desarrollo de sus competencias académicas y profesionales.

Considerando los resultados que distinguen mayor y menor complejidad en esta muestra y considerando la investigación existente, pareciera ser que la

mejor opción para atender las deficiencias en la complejidad sintáctica al comienzo de un programa universitario es la instrucción centrada en la forma para estudiantes que obtienen un puntaje inferior al promedio en complejidad sintáctica. Por otro lado, aunque la evidencia sobre el impacto de la enseñanza sobre la conciencia del registro es limitada, estos resultados, en general, podrían ayudar a las instituciones a detectar las necesidades de estudiantes al ingresar a la educación superior y perfilar el apoyo que necesitan en su desarrollo de escritura académica. Esta es una contribución importante para el proceso de enseñanza-aprendizaje en el sentido de que revela que se debe dedicar más tiempo y atención a estos aspectos superficiales para reducir las brechas en los niveles iniciales de competencia escrita de los estudiantes. Además, la media en cada dimensión para textos de menor complejidad se puede utilizar como un método de detección temprana para los estudiantes que necesitan un apoyo más específico en estos aspectos de la escritura.

Por otro lado, los resultados de esta investigación aportan con nueva evidencia sobre las interacciones entre lectura y escritura ampliamente estudiadas, donde se ha comprobado que la enseñanza en lectura aporta a la escritura y viceversa (Graham, Liu, Aitken, et al., 2018; Graham, Liu, Bartlett, et al., 2018; Graham & Hebert, 2011). Si bien esta interacción está comprobada, la presente investigación aporta con el rol mediador de la precisión metacognitiva, que potencia el rol predictor de la lectura desde la escritura, en particular la

ortografía. Este hallazgo es de suma importancia para las intervenciones en la enseñanza y aprendizaje de la lectoescritura, ya que comprueba un tercer foco no estudiado anteriormente: la precisión de la metacompreensión. Tomando en consideración estos resultados, las intervenciones en lectoescritura debieran, además de centrarse en estas habilidades y su enseñanza combinada y balanceada (Graham, Liu, Aitken, et al., 2018), prestar atención a los juicios de comprensión de los estudiantes, que estarían en la base de esta interacción, potenciándola.

Por otro lado, considerando que el proceso de lectoescritura en la universidad es ampliamente distinto al que se exige en la enseñanza escolar, requiriendo más autonomía y descubrimiento (Baker, 2018) y teniendo como centro la conciencia de la propia identidad y las relaciones de poder en la comunidad académica (Lea & Street, 2006), estos hallazgos proveen más evidencia para la nivelación de estudiantes al ingresar a la educación superior, considerando que la precisión metacognitiva potencia el ciclo de aprendizaje entre lectura y escritura. Este hallazgo, que ya había sido delineado en estudios anteriores (Gutierrez De Blume et al., 2021), es también consonante con la existencia de un sistema de aprendizaje autorregulado (Gutierrez de Blume, 2021), compuesto por la cognición, la metacognición y la motivación, que se requiere potenciar para abordar las prácticas más autodirigidas de lectura y escritura en la universidad.

## **Futuras Direcciones**

Los hallazgos de la presente investigación invitan a continuar explorando las interacciones entre comprensión y expresión, añadiendo la variable de la precisión metacognitiva. Si bien existe consenso en el estudio de la precisión metacognitiva de la comprensión, un futuro desafío para la investigación en esta área es crear métodos y herramientas para sustanciar la precisión metacognitiva de la expresión. Asimismo, considerando que la psicolingüística ha privilegiado la comprensión por sobre la expresión, y la comprensión escrita por sobre la oral (Vieiro Iglesias & Gómez-Veiga, 2004), un camino futuro en esta agenda investigativa sería explorar la expresión y comprensión oral, buscando comprobar la existencia de sistemas similares en ellas.

La precisión de la meta-expresión bien podría constituir una variable que explica aspectos de la escritura que aún no han podido operacionalizarse o medirse. Por ejemplo, considerando que recientemente se ha concebido la competencia en escritura como la flexibilidad y adaptabilidad a diversas tareas (Allen et al., 2019), la meta-expresión podría ser la variable que explique esta flexibilidad, y podría constituir un nuevo constructo que dé cuenta más precisamente del desempeño en la escritura, sobre todo considerando que la

competencia en escritura solo explica el 20% del desempeño (van den Bergh et al., 2016).

## Bibliografía

- Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. <https://doi.org/10.1037/a0019318>
- Abdel Latif, M. M. M. (2021). Remodeling writers' composing processes: Implications for writing assessment. *Assessing Writing, 50*. <https://doi.org/10.1016/j.asw.2021.100547>
- Allen, L. K., Likens, A. D., & McNamara, D. S. (2019). Writing flexibility in argumentative essays: a multidimensional analysis. *Reading and Writing, 32*(6), 1607–1634. <https://doi.org/10.1007/s11145-018-9921-y>
- Baker, S. (2018). Shifts in the treatment of knowledge in academic reading and writing: Adding complexity to students' transitions between A-levels and university in the UK. *Arts & Humanities in Higher Education, 14*(4), 388–409. <https://doi.org/10.1177/1474022217722433>
- Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science, 2*(4), 716–724. <https://doi.org/10.1111/j.1756-8765.2010.01115.x>
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

- Bazerman, C. (1980). *A Relationship between Reading and Writing: The Conversational Model* (Vol. 41, Issue 6).  
<http://www.jstor.org>URL:<http://www.jstor.org/stable/375913>
- Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In A. C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). The Guilford Press.
- Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities*, 35(1), 39–56.  
<https://doi.org/10.1177/002221940203500104>
- Calma, A., Cotronei-Baird, V., & Chia, A. (2022). Grammarly: An instructional intervention for writing enhancement in management education. *International Journal of Management Education*, 20(3), 100704.  
<https://doi.org/10.1016/j.ijme.2022.100704>
- Carlino, P. (2003). Alfabetización académica: Un cambio necesario, algunas alternativas posibles. *Educere*, 6(20), 409–420.
- Carlino, P. (2005). *Escribir, leer, y aprender en la universidad. Una introducción a la alfabetización académica*. Fondo de Cultura Económica.
- Carlino, P. (2013). Alfabetización académica diez años después. *Revista Mexicana de Investigación Educativa RMIE*, 18(57), 355–381.  
<http://www.redalyc.org/pdf/140/14025774003.pdf>

- Carter, R., Salamonson, Y., Ramjan, L. M., & Halcomb, E. (2018). Students use of exemplars to support academic writing in higher education: An integrative review. *Nurse Education Today*, 65(February), 87–93. <https://doi.org/10.1016/j.nedt.2018.02.038>
- Cash, M. M., & Shumm, J. S. (2006). Making sense of knowledge: Comprehending expository text. In J. S. Schumm (Ed.), *Reading assessment and instruction for all learners*. The Guilford Press.
- Cassany, D., & López, C. (2010). De la universidad al mundo laboral: Continuidad y contraste entre las prácticas letradas académicas y profesionales. In G. Parodi (Ed.), *Alfabetización académica y profesional en el siglo XXI: Leer y escribir desde las disciplinas* (pp. 347–374). Ariel.
- Consejo Nacional de Educación. (2020). *Consejo Nacional de Educación*. Índices Educación Superior. <https://www.cned.cl/indices-educacion-superior>
- Correnti, R., Matsumura, L. C., Wang, E. L., Litman, D., & Zhang, H. (2022). Building a validity argument for an automated writing evaluation system (eRevise) as a formative assessment. *Computers and Education Open*, 3(November 2020), 100084. <https://doi.org/10.1016/j.caeo.2022.100084>
- De Vega, M., & Cuetos, F. (1999). Introducción: Los desafíos de la psicolingüística. In F. Cuetos & M. De Vega (Eds.), *Psicolingüística del español* (pp. 13–52). Editorial Trotta.
- Eggs, S. (2004). *An introduction to Systemic Functional Linguistics* (2nd ed.). Bloomsbury Academic.

- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211–245.
- Escorcía, D., & Gimenes, M. (2020). Metacognitive components of writing: Construction and validation of the Metacognitive Components of Planning Writing Self-inventory (MCPW-I). *Revue Européenne de Psychologie Appliquée*, 70(1), 100515. <https://doi.org/10.1016/j.erap.2019.100515>
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35(1), 39–50. <https://doi.org/10.1207/S15326985EP3501>
- Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural/process theories of complex human behavior* (pp. 34–78). Sijthoff & Noordhoff.
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring A New Area of Cognitive — Developmental Inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066x.34.10.906>
- Geng, J., & Razali, A. B. (2022). Effectiveness of the Automated Writing Evaluation program on improving undergraduates' writing performance. *English Language Teaching*, 15(7), 49. <https://doi.org/10.5539/elt.v15n7p49>
- Graesser, A., Golding, J. M., & Long, D. L. (1996). Narrative representation and comprehension. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 171–205). Lawrence Erlbaum Associates.

- Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018). Effectiveness of Literacy Programs Balancing Reading and Writing Instruction: A Meta-Analysis. *Reading Research Quarterly*, 53(3), 279–304. <https://doi.org/10.1002/rrq.194>
- Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for Writing: A Meta-Analysis of the Impact of Reading Interventions on Writing. *Review of Educational Research*, 88(2), 243–284. <https://doi.org/10.3102/0034654317746927>
- Gutierrez de Blume, A. P. (2017). The effects of strategy training and an extrinsic incentive on fourth- and fifth-grade students' performance, confidence, and calibration accuracy. *Cogent Education*, 4(1), 1–17. <https://doi.org/10.1080/2331186X.2017.1314652>
- Gutierrez de Blume, A. P. (2021). Autorregulación del aprendizaje: Desenredando la relación entre cognición, metacognición y motivación. *Voces y Silencios. Revista Latinoamericana de Educación*, 12(1), 81–108. <https://doi.org/10.18175/vys12.1.2021.4>
- Gutierrez de Blume, A. P., & Schraw, G. (2015). Effects of strategy training and incentives on students performance, confidence, and calibration. *Journal of*

*Experimental Education*, 83(3), 386–404.

<https://doi.org/10.1080/00220973.2014.907230>

- Gutierrez De Blume, A. P., Soto, C., Ramírez Carmona, C., Rodriguez, F., & Pino Castillo, P. A. (2021). Reading competence and its impact on writing: An approach towards mental representation in literacy tasks. *Journal of Research in Reading*, 44(3), 1–19. <https://doi.org/10.1111/1467-9817.12359>
- Haberlandt, K. (1994). Methods in reading research. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (1st ed., pp. 1–31). Academic Press, Inc.
- Hacker, D. J. (1998). Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 1–23). Lawrence Erlbaum Associates.
- Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed.). Hodder Arnold. <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/143984111X>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Halliday, M. A. K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social semiotic perspective* (2nd ed.). Oxford University Press.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. <https://doi.org/10.1177/0741088312451260>
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive Processes in*

*Writing* (pp. 3–30). Lawrence Erlbaum Associates. <https://doi.org/BF456W8C6761980>

Hillocks, G. (1986). *Research on written composition*. ERIC Clearinghouse of Reading and Communication Skills.

Hunt, K. W. (1965). Grammatical structures written at three grade levels. In *National Council of Teachers of English, Research Report No. 3*. <https://eric.ed.gov/?id=ED113735>

Indefrey, P. (2018). The relationship between syntactic production and comprehension. In S.-A. Rueschemeyer & G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics* (pp. 483–505). Oxford University Press.

Kim, Y. S. G., Dore, R., Cho, M., Golinkoff, R., & Amendum, S. (2021). Theory of mind, mental state talk, and discourse comprehension: Theory of mind process is more important for narrative comprehension than for informational text comprehension. *Journal of Experimental Child Psychology, 209*. <https://doi.org/10.1016/j.jecp.2021.105181>

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182. <https://doi.org/10.1037//0033-295x.95.2.163>

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394.  
<https://doi.org/10.1037/0033-295X.85.5.363>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.  
<https://doi.org/10.2307/2529310>
- Latini, N., Bråten, I., & Haverkamp, Y. E. (2021). Breadth and depth of strategic processing during text comprehension. *Learning and Individual Differences*, 91. <https://doi.org/10.1016/j.lindif.2021.102058>
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322.  
<https://doi.org/10.1093/applin/16.3.307>
- Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157–172.  
<https://doi.org/10.1080/03075079812331380364>
- Lea, M. R., & Street, B. V. (2006). The “academic literacies” model: Theory and applications. *Theory into Practice*, 45(4), 368–377.
- Levelt, W. J. M. (1938). *Speaking: From intention to articulation*. Halliday Lithograph.
- Li, X., & Zhang, H. (2021). Developmental features of lexical richness in English writings by Chinese L3 beginner learners. *Frontiers in Psychology*, 12(September), 1–11. <https://doi.org/10.3389/fpsyg.2021.752950>

- Lillis, T., & Scott, M. (2015). Defining academic literacies research: issues of epistemology, ideology and strategy. *Journal of Applied Linguistics and Professional Practice*, 4(1), 5–32. <https://doi.org/10.1558/japl.v4i1.5>
- MacArthur, A. C., & Graham, S. (2016). Writing research from a cognitive perspective. In A. C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 24–40). The Guilford Press.
- Martí, E. (1995). Metacognición: Entre la fascinación y el desencanto. *Infancia y Aprendizaje*, 72, 9–32. <https://doi.org/10.1174/02103709560561131>
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. <https://doi.org/10.1016/j.jslw.2015.06.004>
- McCarthy, K. S., Soto, C., Gutierrez De Blume, A. P., & Palma, D. (2020). Improving Reading Comprehension in Spanish using iSTART-E: A Pilot Study. *International Journal of Computer-Assisted Language Learning and Teaching*, April. <https://www.researchgate.net/publication/340952008>
- McCarthy, K. S., Soto, C., Malbrán, C., Fonseca, L., Simian, M., & McNamara, D. S. (2018). iSTART-E: Reading comprehension strategy training for Spanish speakers. *Artificial Intelligence in Education: 19th International Conference, AIED 2018, London, UK, June 27–30, 2018, Proceedings, Part II*, 215–219. <https://doi.org/10.1007/978-3-319-93846-2>

- McNamara, D. S. (2004). Aprender del texto: Efectos de la estructura textual y las estrategias del lector. *Revista Signos*, 37(55), 19–30. <https://doi.org/10.4067/S0718-09342004005500002>
- McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers*, 36(2), 222–233.
- McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 51, pp. 297–384). Elsevier Academic Press. [https://doi.org/10.1016/S0079-7421\(09\)51009-2](https://doi.org/10.1016/S0079-7421(09)51009-2)
- McNamara, D. S., Ozuru, Y., Best, R., & O'Reilly, T. (2007). The 4-pronged comprehension strategy framework. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions and technologies* (pp. 465–496). Lawrence Erlbaum Associates, Inc.
- Medina, A. L., & Pilonieta, P. (2006). Once upon a time: Comprehending narrative texts. In J. S. Schumm (Ed.), *Reading assessment and instruction for all learners* (pp. 222–261). The Guilford Press.
- Meyer, A. S., Huettig, F., & Levelt, W. J. M. (2016). Same, different, or closely related: What is the relationship between language production and comprehension? *Journal of Memory and Language*, 89, 1–7. <https://doi.org/10.1016/j.jml.2016.03.002>

- Mumford, S., & Atay, D. (2021). Teachers' perspectives on the causes of rater discrepancy in an English for Academic Purposes context. *Assessing Writing*, 48(March 2020), 100527. <https://doi.org/10.1016/j.asw.2021.100527>
- Muñoz, M., & Blanco, C. (2013). Una taxonomía de las universidades chilenas. *Calidad En La Educación*, 38, 181.213. <https://doi.org/10.31619/caledu.n38.109>
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, 7(5), 1–9. <https://doi.org/10.1016/j.heliyon.2021.e07014>
- O'Donnell, R. C., Griffin, W. J., & Norris, R. C. (1967). *Syntax of kindergarten and elementary school children: A transformational analysis*.
- Organisation for Economic Co-operation and Development. (2016). Skills Matter: Further results from the survey of adult skills. In OECD Skills Studies (Ed.), *Skills matter: Further results from the survey of adult skills*. OECD Publishing. <https://doi.org/10.1787/9789264258051-en>
- Organisation for Economic Co-operation and Development. (2018). *Writing in a changing world*. <https://doi.org/https://doi.org/10.1787/c70a7cb3-en>.
- Organización para la Cooperación y Desarrollo Económicos. (n.d.). *El programa PISA de la OCDE: Qué es y para qué sirve*. OCDE Publishing.

- Organización para la Cooperación y Desarrollo Económicos. (2017). *Marco de Evaluación y de Análisis de PISA para el Desarrollo : Lectura, matemáticas y ciencias*. OECD Publishing.
- Otero, J. C. (1998). Influence of knowledge activation and context on comprehension monitoring of science texts. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 145–164). Lawrence Erlbaum Associates Publishers.
- Otero, J. C. (2002). Noticing and fixing difficulties while understanding science texts. In J. C. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 281–308). Lawrence Erlbaum Associates.
- Otero, J. C., & Campanario, J. M. (1990). Comprehension evaluation and regulation in learning from science texts. *Journal of Research in Science Teaching*, 27(5), 447–460. <https://doi.org/10.1002/tea.3660270505>
- Palma, D., Soto, C., Véliz, M., Riffo, B., & Gutiérrez, A. (2019). A data-driven methodology to assess text complexity based on syntactic and semantic measurements. *Human Interaction and Emerging Technologies*, 509–515.
- Parodi, G. (2005). La comprensión del discurso especializado escrito en ámbitos técnico-profesionales: ¿Aprendiendo a partir del texto? *Revista Signos*, 38(58), 221–267. <https://doi.org/10.4067/s0718-09342005000200005>
- Parodi, G. (2006). Discurso especializado y lengua escrita: Foco y variación. *Estudios Filológicos*, 41, 165–204. <https://doi.org/10.4067/S0071-17132006004100012>

- Pino Castillo, P. A., Ormeño Hofer, C., Zúñiga Herrera, M., & Zanocco Soto, P. (2019). El significado del desarrollo anidado de las literacidades y numeracidades académicas en el currículum universitario. *Revista Dilemas Contemporáneos: Educación, Política y Valores*, 28(1), 1–24. <https://doi.org/10.46377/dilemas.v28i1.1607>
- Programa de las Naciones Unidas para el Desarrollo. (2017). *Desiguales: Orígenes, cambios y desafíos de la brecha social en Chile*. Programa de las Naciones Unidas para el Desarrollo.
- Rebolledo Luna, V., Gutiérrez De Blume, A. P., & Soto, C. (2021). Reading Comprehension in Students with Autism Spectrum Disorder: The Role of Metacognitive Skills, Motivation, and Use of Strategies. *Global Journal of Intellectual & Developmental Disabilities*, 9(2), 1–3. <https://doi.org/10.19080/gjidd.2021.09.555758>
- Scardamalia, M., & Bereiter, C. (1992). Dos modelos explicativos de composición escrita. *Infancia y Aprendizaje*, 58, 43–64.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Schröder, H. (1991). Linguistic and text-theoretical research on languages for special purposes: A thematic and bibliographical guide. In H. Schröder (Ed.), *Subject-oriented texts: Languages for special purposes and text theory* (Vol. 16, pp. 1–48). Walter de Gruyter.

- Scovel, T. (1998). *Psycholinguistics*. Oxford University Press.
- Soto, C., Gutiérrez De Blume, A. P., Asún, R., Jacovina, M., & Vásquez, C. (2018). A deeper understanding of metacomprehension in reading: Development of a new multidimensional tool. *Frontline Learning Research*, 6(1), 31–52. <https://doi.org/10.14786/flr.v6i1.328>
- Soto, C., Gutierrez de Blume, A. P., Carrasco Bernal, M. A., & Contreras Castro, M. A. (2020). The role of meta-cognitive cues on the comprehension of proficient and poor readers. *Journal of Research in Reading*, 43(3), 272–289. <https://doi.org/10.1111/1467-9817.12303>
- Soto, C., Gutiérrez de Blume, A. P., Jacovina, M., McNamara, D., Benson, N., & Riffo, B. (2019). Reading comprehension and metacognition: The importance of inferential skills. *Cogent Education*, 6(1). <https://doi.org/10.1080/2331186x.2019.1565067>
- Soto, C., Gutierrez de Blume, A. P., Rebolledo, V., Rodríguez, F., Palma, D., & Gutiérrez, F. (2022). Metacognitive monitoring skills of reading comprehension and writing between proficient and poor readers. *Metacognition and Learning*. <https://doi.org/10.1007/s11409-022-09317-8>
- Soto, C., Gutiérrez de Blume, A. P., Rodríguez, M. F., Asún, R., Figueroa, M., & Serrano, M. (2019). Impact of Bridging Strategy and Feeling of Knowing Judgments on Reading Comprehension Using COMPRENDE: an Educational Technology. *TechTrends*, 63(5), 570–582. <https://doi.org/10.1007/s11528-019-00383-5>

- Soto, C., Rodríguez Poblete, M. F., & Gutierrez de Blume, A. P. (2018). Exploring the meta-comprehension abilities of students with intellectual disabilities. *International Journal of Special Education*, 33(2), 233–247.
- Sundeen, T. H. (2014). Instructional rubrics: Effects of presentation options on writing quality. *Assessing Writing*, 21, 74–88. <https://doi.org/10.1016/j.asw.2014.03.003>
- Tapiero, I. (2007). *Situation models and levels of coherence: Toward a definition of comprehension*. Lawrence Erlbaum Associates, Inc.
- Tierney, R. J., & Shanahan, T. (1996). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *The handbook of reading research* (pp. 246–280). Longman.
- Tijero Neyra, T. (2009). Representaciones mentales: discusión crítica del modelo de situación de Kintsch. *Onomázein*, 19(1), 111–138.
- van den Bergh, H., Rijlaarsdam, G., & van Steendam, E. (2016). Writing process theory: A functional dynamic approach. In A. C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 57–71). The Guilford Press.
- van den Broek, P., Bohn-Gettler, C., Kendeou, P., & Carlson, S. (2011). When a Reader Meets a Text: The Role of Standards of Coherence in Reading Comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.),

- Text relevance and learning from text* (pp. 123–139). Information Age Publishing, Inc.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.
- Vieiro Iglesias, P., & Gómez-Veiga, I. (2004). *Psicología de la lectura: Procesos, teorías y aplicaciones instruccionales*. Pearson Prentice Hall.
- Weaver, C. A., & Kintsch, W. (1996). Expository text. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 230–245). Lawrence Erlbaum Associates.
- Wu, Y., Barquero, L. A., Pickren, S. E., Taboada Barber, A., & Cutting, L. E. (2020). The relationship between cognitive skills and reading comprehension of narrative and expository texts: A longitudinal study from Grade 1 to Grade 4. *Learning and Individual Differences, 80*.  
<https://doi.org/10.1016/j.lindif.2020.101848>
- Zhang, H., Chen, M., & Li, X. (2021). Developmental features of lexical richness in English writings by Chinese beginner learners. *Frontiers in Psychology, 12*(June), 1–15. <https://doi.org/10.3389/fpsyg.2021.665988>
- Zhang, S. (2021). Review of automated writing evaluation systems. *Journal of China Computer-Assisted Language Learning, 1*(1), 170–176.  
<https://doi.org/10.1515/jccall-2021-2007>

- Zhou, J., & Lü, C. (2022). Enhancing syntactic complexity in L2 Chinese writing: Effects of form-focused instruction on the Chinese topic chain. *Frontiers in Psychology, 13*(April), 1–12. <https://doi.org/10.3389/fpsyg.2022.843789>
- Zou, L. (2022). Lexical development of English major students. *International Journal of Education and Humanities, 2*(3), 29–31. <https://doi.org/10.54097/ijeh.v2i3.338>

## Anexos

### Análisis de Juicios de Expertos para las Pruebas de Lectura

#### Ciencias Sociales

##### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.676	.041	16.443	.000	.673	.678

a. Sample data contains 30 effective subjects and 5 raters.

##### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Asymptotic		Asymptotic 95% Confidence Interval		
			Standard Error	z	Lower Bound	Upper Bound	
1	.273	.765	.058	13.251	.000	.761	.769
2	.407	.586	.058	10.142	.000	.582	.589
3	.320	.694	.058	12.014	.000	.690	.697

a. Sample data contains 30 effective subjects and 5 raters.

## Derecho

### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.628	.043	14.642	.000	.625	.631

a. Sample data contains 30 effective subjects and 5 raters.

### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
				z	Sig.	Lower Bound	Upper Bound
1	.287	.804	.058	13.932	.000	.801	.808
2	.547	.583	.058	10.098	.000	.579	.587
3	.167	.448	.058	7.760	.000	.444	.452

a. Sample data contains 30 effective subjects and 5 raters.

## Economía y Negocios

### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.717	.043	16.481	.000	.714	.719

a. Sample data contains 30 effective subjects and 5 raters.

### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
				z	Sig.	Lower Bound	Upper Bound
1	.367	.885	.058	15.332	.000	.882	.889
2	.480	.653	.058	11.306	.000	.649	.656
3	.153	.538	.058	9.315	.000	.534	.541

a. Sample data contains 30 effective subjects and 5 raters.

### Ingeniería

### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.672	.042	15.957	.000	.669	.674

a. Sample data contains 30 effective subjects and 5 raters.

### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
				z	Sig.	Lower Bound	Upper Bound
1	.327	.788	.058	13.646	.000	.784	.791
2	.473	.585	.058	10.141	.000	.582	.589
3	.200	.646	.058	11.186	.000	.642	.649

a. Sample data contains 30 effective subjects and 5 raters.

## Recursos Naturales y Medicina Veterinaria

### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.797	.041	19.410	.000	.794	.800

a. Sample data contains 30 effective subjects and 5 raters.

### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
				z	Sig.	Lower Bound	Upper Bound
1	.287	.837	.058	14.497	.000	.833	.841
2	.413	.753	.058	13.035	.000	.749	.756
3	.300	.810	.058	14.021	.000	.806	.813

a. Sample data contains 30 effective subjects and 5 raters.

## Salud

### Overall Agreement<sup>a</sup>

	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
			z	Sig.	Lower Bound	Upper Bound
Overall Agreement	.834	.042	19.888	.000	.831	.837

a. Sample data contains 30 effective subjects and 5 raters.

### Agreement on Individual Categories<sup>a</sup>

Rating Category	Conditional Probability	Kappa	Standard Error	Asymptotic		Asymptotic 95% Confidence Interval	
				z	Sig.	Lower Bound	Upper Bound
1	.407	.972	.058	16.842	.000	.969	.976
2	.387	.775	.058	13.425	.000	.771	.779
3	.207	.715	.058	12.391	.000	.712	.719

a. Sample data contains 30 effective subjects and 5 raters.