



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

**EL EFECTO DE LA PROFICIENCIA EN L2 Y LA
FUNCIÓN EJECUTIVA DE INHIBICIÓN SOBRE EL
PROCESAMIENTO ARITMÉTICO EN ORACIONES CON
RELACIÓN CATEGÓRICA Y FUNCIONAL: EVIDENCIAS
EMPÍRICAS CON ESTUDIOS CONDUCTUALES**

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

POR: MARTÍN ANDRÉS FLORES QUIROZ

Profesor Guía: Mabel Urrutia Martínez

Diciembre 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 Chipo, mi pilar y mejor amiga, que me enseñó
el valor de la perseverancia.*

*A William y Felipe,
mis hermanos, que me apoyaron
y creyeron más que nadie.*

*A Carolina y Francoise,
por ser una luz cuando
yo estaba completamente perdido.*

*A Catalina, por recordarme
lo bonito e importante que
es esmerarse de manera sincera.*

Agradecimientos

Agradezco a mi profesora guía, Dra. Mabel Urrutia Martínez, por su paciencia y guía en este proceso; al Programa de Doctorado en Lingüística de la Universidad de Concepción por el apoyo de todos estos años, y al Programa de Pedagogía en Inglés de la Universidad Central de Chile, por la compañía a lo largo de este proceso.

Este trabajo fue financiado por la Agencia Nacional de Investigación y Desarrollo (ANID) / Beca Doctorado Nacional 2016-21160966, FONDECYT REGULAR 1231433, 1210653 y de los Fondos Basales para Centros de Excelencia PIA-CONICYT FB0003

Tabla de Contenido

Derecho de Autor	2
Agradecimientos	4
Tabla de Contenido	5
Índice de Tablas	9
Índice de Figuras	11
Resumen	12
Abstract	14
Introducción	16
Capítulo I: Problema de Investigación y Marco Teórico	19
Los Problemas Matemáticos Escritos	23
Desalineación Semántica y Problemas Matemáticos Escritos	28
Los Efectos del Bilingüismo en el Procesamiento de Problemas Matemáticos Escritos	38
Estudios Realizados en Estados Unidos	39
Estudios Realizados en Otros Países	47
El Rol de la Inhibición como Función Ejecutiva	55
La Relación entre la Inhibición y el Procesamiento del Lenguaje	59
Inhibición del Procesamiento Matemático Abstracto	66
El Rol de la Experticia Matemática	68
Modelos teóricos	69
La Interfaz Sintáctica-Semántica	70
La teoría de la Carga Cognitiva	73
El Modelo de Activación Interactiva Bilingüe	82
El Modelo de Activación Interactiva Bilingüe del Desarrollo	87
El Modelo de Congruencia Semántica (SECO)	91
La Eficiencia de las Adecuaciones Lingüísticas y Estrategias para Enfrentar los Problemas Matemáticos Escritos	98
Metaanálisis sobre la Resolución de Problemas Matemáticos Escritos y el Lenguaje de las Pruebas	101
Metaanálisis del Rol del Lenguaje en la Resolución de Problemas Matemáticos Escritos	103
Estrategia de Búsqueda	105
Criterios de Elegibilidad	105

Procedimiento de Análisis	107
Tamaño Medio del Efecto y Prueba de Heterogeneidad	109
Análisis de Moderadores	110
Sesgo de Publicación	111
Discusión	112
Objetivos	116
Objetivos Generales	116
Objetivos Específicos	117
Hipótesis	117
Hipótesis de Investigación	117
Capítulo II: Metodología y Resultados	119
Metodología General	119
Estudio de Valoración de Pares de Palabras en Relación Categórica y Funcional	123
Diseño de Estudio	123
Participantes	124
Procedimiento	125
Material	127
Resultados	129
Análisis	130
Estudio de Resolución de Problemas Matemáticos Escritos en Relación Categórica y Funcional	134
Diseño de Estudio	134
Participantes	135
Procedimiento	137
Material	139
Material Matemático	139
Material Lingüístico	140
Resultados	141
Resultados Porcentaje de Aciertos	141
Resultados de Tiempo de Reacción	144
Análisis	147
ANOVA Porcentaje de Aciertos	147
ANOVA de Tiempos de Reacción	151
Estudio Stroop	154
Diseño	154

Participantes	154
Procedimiento	156
Material	158
Resultados	159
Resultados Porcentaje de Acierto	159
Resultados Tiempos de Reacción	161
Análisis	164
ANOVA Porcentaje de Aciertos	164
ANOVA Tiempos de Reacción	167
Capítulo III: Discusión General y Conclusiones	174
Discusión General	174
Discusión Estudio de Evaluación de Pares de Palabras en Relación Categorica y Funcional	174
Discusión Estudio de Resolución de Problemas Matemáticos Escritos en Relación Categorica y Funcional	180
Discusión Estudio Stroop	187
Conclusiones	195
Contribuciones	199
Limitaciones metodológicas	200
Futuras Investigaciones	202
Bibliografía	204
Anexos	220
Modelo Consentimiento Informado	221
Escala de Likert Evaluación Funcional	225
Escala de Likert Evaluación Categorica	229

Índice de Tablas

Tabla 1 Gráfico de tallo y hoja de todos los tamaños de efecto (gs)	108
Tabla 2 Ejemplos de lista de evaluación categórica	126
Tabla 3 Ejemplos de lista de evaluación funcional	126
Tabla 4 Moda, media y mediana para los sets categóricos y funcionales en la lista de evaluación categórica.	127
Tabla 5 Moda, media y mediana para los sets categóricos y funcionales en la lista de evaluación funcional.	127
Tabla 6 Prueba t de la lista de evaluación categórica para dos muestras, suponiendo varianzas iguales	128
Tabla 7 Prueba t de la lista de evaluación funcional para dos muestras, suponiendo varianzas iguales	129
Tabla 8 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.	138
Tabla 9 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.	139
Tabla 10 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.	140
Tabla 11 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.	141
Tabla 12 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.	142

Tabla 13 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.	142
Tabla 14 Pruebas de normalidad para porcentaje de aciertos	144
Tabla 15 Pruebas de normalidad para tiempos de reacción	147
Tabla 16 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de bajo conocimiento en inglés.	155
Tabla 17 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en inglés.	156
Tabla 18 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en matemática.	156
Tabla 19 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de bajo conocimiento en inglés.	157
Tabla 20 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.	158
Tabla 21 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.	159
Tabla 22 Pruebas de normalidad para porcentaje de aciertos	160
Tabla 23 Pruebas de normalidad para tiempos de reacción	163

Índice de Figuras

Figura 1 Diagrama de las relaciones A) categóricas y B) funcionales (Guthormsen, 2007, p. 15) .	31
Figura 2 Interacciones meta-reflexivas en el espacio del problema (Campbell et al., 2007, p. 9).	79
Figura 3 Arquitectura del modelo de activación interactiva bilingüe (Grainger et al., 2010, p. 269).	83
Figura 4 El modelo de activación interactiva bilingüe del desarrollo para el aprendizaje de vocabulario en una L2.	87
Figura 5 El modelo SECO (Gros et al., 2020, p. 5).	91
Figura 6 Gráfico de embudo	109
Figura 7 Secuencia de presentación de estímulos en el experimento resolución de problemas matemáticos en relación categórica y funcional.	135
Figura 8 Secuencia de presentación de estímulos en el experimento Stroop.	153

Resumen

El lenguaje y la matemática han sido conceptualizados como elementos independientes que no se influyen mutuamente. Esto se ve de manera pronunciada en el caso de la matemática, que es conceptualizada de manera abstracta. A pesar de esa perspectiva, los resultados de un vasto corpus de estudios realizados en Estados Unidos sugieren una realidad distinta ya que los resultados en pruebas matemáticas de aprendientes de inglés como lengua extranjera o segunda lengua son pronunciadamente menores que los resultados de los hablantes nativos de inglés. Junto con eso, independiente del conocimiento matemático, distintos modelos teóricos plantean cómo los bilingües, dependiendo de la etapa de desarrollo en que se encuentran, podrían tener mayores dificultades en el procesamiento lingüístico de su L2, dependiendo de sus capacidades de inhibir información no relevante proveniente de su L1, abriendo una arista que no había sido considerada anteriormente. A través de tres estudios cuantitativos, se intenta determinar cómo se evalúan las relaciones sintácticas y semánticas de pares de palabras, para luego determinar cómo estas podrían afectar la resolución de problemas matemáticos escritos sencillos, y, finalmente, determinar los efectos cognitivos de inhibición en distintos grupos de estudiantes universitarios. Los resultados muestran diferencias significativas en la evaluación semántica y sintáctica de pares de palabras, estas relaciones sintácticas y semánticas tienen un efecto significativo en los porcentajes de

acierto en la resolución de problemas matemáticos escritos, y que el grupo con alto conocimiento en matemática tiene tiempos de reacción menores que el grupo de bajo conocimiento en inglés al responder a una tarea Stroop. Estos hallazgos sugieren que sí hay una relación entre el conocimiento lingüístico y el procesamiento matemático. Además, la capacidad de inhibición de los participantes sí puede estar más desarrollada debido a su experticia inhibiendo información irrelevante en sus determinados campos de estudio.

Palabras clave: desalineación semántica, problemas matemáticos escritos, relaciones sintáctico-semánticas, inhibición

Abstract

Language and mathematics have been conceptualized as independent elements that do not influence each other. This is particularly evident in mathematics, which is conceptualized in an abstract manner. Despite this perspective, results from a vast corpus of studies in the United States suggest a different reality since the scores in mathematical tests of English as a foreign or second language learners are significantly lower than those of native English speakers. Regardless of mathematical knowledge, several theoretical models suggest that bilinguals, depending on their stage of development, could have greater difficulties in the linguistic processing of their L2, depending on their ability to inhibit information from their L1, which suggests a possible explanation for this difference in results. By conducting three quantitative studies, this research intends to determine how syntactic and semantic relationships between word pairs were evaluated, how these might affect solving simple written mathematical problems, and finally, to determine the cognitive effects of inhibition in different groups of university students. The results indicate significant differences in the semantic and syntactic evaluation of word pairs, that these syntactic and semantic relationships have a significant effect on the accuracy rates when solving written mathematical problems, and that the group with high mathematical knowledge has shorter reaction times than the group with low English knowledge when completing a Stroop task. These findings suggest that there is a relationship between linguistic

knowledge and mathematical processing. Furthermore, the participants' inhibition capacity may be more developed due to their expertise in inhibiting irrelevant information in their respective fields of study.

Keywords: semantic misalignment, mathematical word problems, syntactic-semantic relations, inhibition

Introducción

La relación entre el lenguaje y la resolución de problemas matemáticos escritos ha sido ampliamente estudiada en el extranjero durante los últimos 30 años, debido a la creciente necesidad de asegurar que los estudiantes cumplan con los objetivos curriculares correspondientes a su nivel. La aparición de pruebas estandarizadas que se aplican a nivel nacional en muchos países (SATs en Estados Unidos (2023), SIMCE en Chile (2022), PISA para los países miembros de la OCDE, (2023)) se utilizan como una manera de medir el avance de los estudiantes y de tomar decisiones estratégicas de cambios curriculares, paradigmas de enseñanza, e incluso, financiamiento. Con estas pruebas estandarizadas se ha creado un amplio corpus de investigación que indica que, a pesar de haber un buen desempeño en los problemas matemáticos de computación, los problemas matemáticos escritos suelen producir resultados bajo lo esperado, con brechas de rendimiento que solo se acrecientan a medida que los estudiantes avanzan a años superiores (Abedi & Lord, 2001; Kieffer et al., 2009; Kieffer et al., 2012; Pennock-Roman & Rivera, 2011; Rios et al., 2020).

Estos preocupantes resultados se ven exacerbados cuando se incluye a estudiantes que no tienen como lengua materna el idioma de enseñanza y evaluación. Como demuestra el gran corpus de estudios realizados principalmente en Estados Unidos (Abedi & Lord, 2001; Banks et al., 2016; Lager,

2006; Martin & Fuchs, 2019; Ockey, 2007; Powell et al., 2020; entre otros), los logros matemáticos de los hablantes no nativos suelen ser considerablemente inferiores a los de sus pares nativos. Más aún, las adecuaciones lingüísticas realizadas para asegurar que las evaluaciones se enfocan en el constructo matemático y no en el lingüístico no han sido fructíferas como han demostrado los metaanálisis realizados en las últimas dos décadas (Kieffer et al., 2012; Pennock-Roman & Rivera, 2011; Ríos et al., 2020).

La mayoría de estos estudios se han centrado en estudiantes en edad escolar, habiendo una cantidad muy reducida de investigaciones que se han enfocado en estudiantes universitarios aprendientes de una lengua extranjera. Junto con eso, hasta donde sabemos, no hay estudios en Chile que hayan investigado este tema ni con estudiantes aprendientes de español ni con estudiantes universitarios aprendientes de una lengua extranjera. Considerando que muchas universidades extranjeras tienen exámenes de ingreso que requieren demostrar no solo dominio de la lengua, sino que también de un manejo apropiado de las habilidades matemáticas, se hace necesario estudiar este fenómeno desde una perspectiva psicolingüística para determinar cómo el procesamiento del lenguaje y las relaciones semánticas de las palabras utilizadas pueden afectar los logros matemáticos.

A pesar de que el lenguaje en que se mide el conocimiento matemático tiene una influencia sobre los logros académicos, no se puede descartar el rol de las funciones ejecutivas, como la inhibición, en el procesamiento del lenguaje. Particularmente, cuando se considera que al procesar información en una lengua distinta a la lengua materna se genera una carga cognitiva superior, puesto que se debe inhibir información que no es relevante o congruente como ocurre al resolver problemas matemáticos (Federmeier et al., 2020; Müller & Kerns, 2015).

Esta tesis se divide en tres capítulos. Primero, el problema de investigación y marco teórico, donde se plantea el problema de investigación, su justificación y una revisión de la literatura de los conceptos clave. Segundo, la metodología y resultados, que describe las características principales de esta investigación, como el diseño, los participantes, instrumentos de recolección de datos y procedimientos de análisis; y los resultados de la investigación. Tercero, se presentan las conclusiones, que dan cuenta del logro de los objetivos, los hallazgos principales, las limitaciones metodológicas de esta investigación, y futuras investigaciones en esta área.

Capítulo I: Problema de Investigación y Marco Teórico

En las últimas décadas ha cobrado importancia el estudio de la relación entre el procesamiento lingüístico y el procesamiento matemático. Como Guthormsen et al. (2016) indican, muchos autores consideraban que el procesamiento matemático, al ser de corte abstracto, era independiente del procesamiento lingüístico. Por lo tanto, la habilidad matemática de las personas se medía sin tomar en cuenta su conocimiento de mundo.

Contraria a esta postura, la evidencia de las últimas décadas indica que hay una relación estrecha entre ambos tipos de procesamiento, donde los problemas escritos son considerados una veta de estudio. En este tipo de problemas, las personas deben integrar la información matemática, presentada en el problema con las restricciones conceptuales provenientes del lenguaje. De este modo, se vive un proceso de actualización, de acuerdo a Daroczy et al. (2015), mediado por el lenguaje, como en el caso de: Donald Trump ha estado casado 3 veces ¿Cuántas esposas tiene? En este caso, no se puede considerar solamente la información numérica, ya que las restricciones culturales, mediadas por el lenguaje, dan a entender que Donald Trump tiene una esposa y no tres en este momento.

El procesamiento de problemas escritos ha sido investigado mayoritariamente en estudiantes en edad escolar, hablantes monolingües y bilingües. La evidencia recopilada ha indicado que las dificultades con los problemas escritos pueden ser provocadas por el tipo de modelamiento matemático que se utiliza (Bernardo & Calleja, 2005; Guthormsen et al., 2016), por la forma en que el problema es presentado lingüísticamente (Bernardo & Calleja, 2005; Daroczy et al., 2015) o por la proficiencia en el idioma de los participantes (Bernardo, 2002; Bernardo & Calleja, 2005; Kempert et al., 2011). Dentro del grupo de hablantes bilingües, se ha estudiado el procesamiento matemático de manera conductual en Filipinas, donde el filipino y el inglés tienen carácter oficial; y en Alemania, con hablantes de turco y alemán, donde solo este último tiene carácter oficial.

Para el caso particular de este proyecto, es relevante estudiar cómo el conocimiento matemático se puede ver afectado por el idioma en que son presentados los problemas matemáticos. En el caso de Chile, hay tres grandes grupos bilingües: chilenos pertenecientes a pueblos originarios, chilenos hablantes de inglés y la creciente población migrante en Chile como son los haitianos. Por un motivo de acceso a los participantes, el foco del presente estudio será en chilenos hablantes de inglés, lo que no descarta continuar este estudio con los otros dos grupos bilingües anteriormente mencionados a futuro. En esta línea, Chung et al. (2019) indicaron que la transferencia de habilidades

de una lengua a otra es un proceso complejo, puesto que depende de muchos factores (cercanía entre lenguas, edad de aprendizaje, complejidad de la lengua, proficiencia, entre otros).

El procesamiento matemático no solo tendría una relación con el procesamiento lingüístico. Concretamente, Wang et al. (2023), al estudiar la adición y división con sustantivos semánticamente incongruentes con estas operaciones, concluyeron que la inhibición cumple un rol fundamental para discernir la información relevante y hacer caso omiso de la información contradictoria.

A través de tres estudios, se propone estudiar 1) cómo se evalúan las relaciones semánticas y sintácticas entre sustantivos en español; 2) cómo el procesamiento sintáctico-semántico puede afectar la resolución de problemas matemáticos escritos en inglés y en español; y 3) la capacidad de inhibición de tres grupos de estudiantes universitarios.

El amplio corpus de estudios realizados en Estados Unidos en las últimas décadas haría creer que ya habría soluciones prácticas para las dificultades que los hablantes no nativos presentan al momento de resolver problemas matemáticos en una lengua que aún no dominan. Desafortunadamente, la evidencia, que será discutida en detalle en el marco teórico, ha indicado que las

adecuaciones lingüísticas realizadas en Estados Unidos tienen un efecto muy pequeño o inexistente. Además, los estudios se han enfocado en su mayoría en describir y acompañar a estudiantes en edad escolar, sin considerar el creciente número de estudiantes universitarios que prosigue sus estudios en una lengua extranjera, con la expectativa impuesta de que su rendimiento sea el mismo que en su lengua materna.

En Chile, considerando la revisión bibliográfica realizada, no existen estudios que evalúen las relaciones sintáctico-semánticas entre sustantivos en español. Tampoco existen estudios que consideren a estudiantes universitarios al momento de resolver problemas matemáticos escritos en distintos idiomas. Finalmente, el realizar este tipo de estudios será útil para entender la realidad específica de lenguas distintas al inglés, particularmente en el contexto sudamericano. Por lo tanto, existe una necesidad de iniciar esta línea de estudios. Finalmente, es necesario estudiar cómo las funciones ejecutivas, particularmente la inhibición, pueden incidir en el procesamiento lingüístico y matemático de los estudiantes universitarios.

Por lo tanto, estudiar la realidad de estudiantes universitarios al enfrentarse a problemas matemáticos en español y en inglés, junto con manipular las relaciones semánticas de los sustantivos que son parte de los problemas matemáticos, podría permitirnos establecer cómo el procesamiento del lenguaje

podría incidir en la ejecución de las habilidades matemáticas desde una perspectiva psicolingüística.

A continuación, se presentará una breve discusión empírico-teórica que da cuenta del estado del arte sobre los problemas matemáticos escritos, la desalineación semántica, los efectos de ser bilingüe en la resolución de problemas matemáticos escritos, la inhibición en relación a los problemas escritos, y los modelos teóricos que se pueden relacionar con esta problemática. Luego, se presentan los resultados de un metaanálisis realizado para estudiar el tamaño del efecto de las diferencias reportadas entre hablantes nativos y no nativos al resolver problemas matemáticos escritos. Para finalizar este capítulo, se presentan los objetivos de esta investigación y las hipótesis para los distintos estudios.

Los Problemas Matemáticos Escritos

Los problemas matemáticos escritos han sido utilizados para enseñar y practicar matemática desde hace milenios, como la evidencia encontrada en papiros egipcios demuestra, junto a la existencia de antiguos libros de aritmética en China e India desde los primeros días de la imprenta (Verschaffel et al., 2020). Teniendo esta gran tradición como parte de su historia, los problemas matemáticos escritos siguen siendo utilizados hoy en día con un propósito

principal: una forma práctica de aplicar el conocimiento matemático aprendido a situaciones de la vida diaria.

Según Driver y Powell (2017, p.41),

A word problem is a mathematics calculation embedded within sentences (...) To solve word problems, students use text, typically presented in English, to identify missing information, make a plan to solve the problem, and perform one or more calculations to get the solution.

Verschaffel et al. (2020) complementaron esta definición al indicar que este tipo de problema matemático es radicalmente distinto a las computaciones sencillas donde se presenta la información numérica y la operación a realizar sin que haya una necesidad de hacer referencia a posibles restricciones del mundo real que podrían tener un efecto en la forma de resolver el problema.

A pesar de su aplicación práctica a situaciones cotidianas, los problemas matemáticos escritos han sido considerados sólo como un acercamiento o una simulación de problemas del mundo real auténticos que podrían ser resueltos al utilizar distintas operaciones matemáticas. Los problemas auténticos suelen considerar información que va mucho más allá de la que se encuentra descrita en un problema matemático escrito, siendo estos últimos considerados como de

naturaleza vaga o ambigua, donde es posible tener más de una respuesta correcta (Verschaffel et al., 2020). A modo de ejemplo, si uno está considerando tomar un préstamo, no es suficiente con saber qué monto se necesita y cuáles son las prestaciones y beneficios de solo un banco en particular. En el mundo real, uno debe considerar las tasas de interés de los distintos bancos para los distintos tipos de préstamo, las distintas opciones de cuotas, cuál es el sueldo que uno recibe en ese momento, entre otros. Es así como estas tareas de modelamiento matemático, donde se debe planear una manera de resolver el problema de acuerdo a la información entregada, se pueden poner en un continuo de autenticidad en el que los problemas escritos se encontrarían en el polo negativo y los problemas del mundo real en el polo positivo.

Aunque este tipo de problemas ha sido utilizado por una gran cantidad de tiempo, los educadores han dado cuenta de las distintas complicaciones que este tipo de práctica genera, abogando por un cambio a problemas que sean mejores simulaciones a los problemas prácticos del mundo real que los estudiantes podrían enfrentar y donde podrían aplicar las matemáticas. Es así como décadas de investigación desde distintas perspectivas se han concentrado en estudiar los efectos de las características computacionales, lingüísticas y de presentación y las características de los sujetos en la resolución exitosa de estos problemas.

Desde el punto de vista de la comprensión, Verschaffel et al. (2020, p.3) indicaron que:

word problems that can be solved by the same arithmetic operation (i.e., a direct addition or subtraction with the two given numbers in the problem) but that belong to different semantic problem types, yield different degrees of difficulty, different ways of representing and solving these problems, and different types of errors.

Por lo tanto, para resolver exitosamente un problema matemático, no es suficiente con identificar la operación matemática a realizar ni los números necesarios, ya que, dependiendo de las palabras utilizadas, la dificultad y la forma de resolverlo pueden variar, generando distintos tipos de errores.

También la capacidad de la memoria de trabajo, el sistema cognitivo capaz de manipular información de manera breve, y la inhibición, la capacidad de dejar de prestar atención a estímulos irrelevantes, pueden afectar la resolución exitosa de estos problemas. En cuanto a la memoria de trabajo, hay diferencias notorias entre los participantes que resuelven exitosamente los problemas escritos y los que no (Verschaffel et al.,2020). Desafortunadamente, estos estudios han sido de naturaleza correlacional, sin haber una gran cantidad de estudios en su rol causal, de acuerdo a la revisión realizada por Verschaffel et al. (2020). Ellos sí destacaron que en los estudios donde la memoria de trabajo ha sido limitada de

manera experimental, hay un aumento considerable de respuestas erróneas. En cuanto a la inhibición, su rol para ignorar la información saliente pero no necesaria, ha sido destacado de manera constante para las habilidades matemáticas generales. A modo de ejemplo, es necesario inhibir la relación entre el adverbio más y la adición en algunos problemas donde es necesario restar. En este punto, Verschaffel et al. (2020) indicaron que no tienen conocimiento de experimentos que hayan estudiado de manera específica la relación causal entre inhibición y problemas matemáticos escritos, quedando esta como una tarea a futuro.

La relación entre matemática y lenguaje puede ser conceptualizada desde la perspectiva del procesamiento simbólico, puesto que ambas requieren la interpretación y manipulación adecuada de símbolos para transmitir significado. En ambos casos, se requieren reglas dependientes de la operación matemática a realizar y de reglas gramaticales y restricciones semánticas, respectivamente.

Esta relación entre matemática y lenguaje no es sólo simbólica en naturaleza, ya que el procesamiento semántico es crucial en ambos ámbitos. En matemática, el procesamiento semántico es esencial para entender conceptos, relaciones y propiedades matemáticas, tales como los símbolos matemáticos, + ó *, a modo de ejemplo. En cuanto al lenguaje, el procesamiento semántico es clave para interpretar desde palabras hasta oraciones y entender estructuras

sintácticas de manera simplificada. En ambos casos, es necesaria la activación de estructuras de conocimiento relevante y la integración de información (Guthormsen et al., 2016).

Por otra parte, para resolver problemas matemáticos escritos es necesario activar recursos de la memoria de trabajo, junto con procesos ejecutivos que se enfoquen en inhibir los prejuicios semánticos o información superficial que estén incluidos en la presentación del problema (Duque de Blas et al., 2021).

Desalineación Semántica y Problemas Matemáticos Escritos

A pesar de que en décadas pasadas se ha planteado que el conocimiento abstracto, como el matemático, no se ve afectado por el conocimiento de mundo (Guthormsen et al., 2016), en las últimas décadas se ha dado cuenta del rol de la experiencia y entrenamiento para resolver de manera exitosa los problemas matemáticos escritos.

En esta línea, Bassok et al. (1998) plantearon que la experiencia juega un rol clave al momento de sacar provecho de las similitudes estructurales entre estímulos que son distintos en cuanto a contenido, contexto y redacción. Es así como el entrenamiento que fomenta la construcción de representaciones

abstractas a partir de un estímulo base aumenta la probabilidad que la transferencia de conocimiento a otro estímulo ocurra.

Para que esta transferencia se dé, el contenido debe tener un efecto clave. Por ende, se sugirió que las personas mantienen contenido clave del estímulo base y del estímulo objetivo y comparan correspondencias e incompatibilidades en contenido y estructura para ofrecer soluciones análogas a distintos problemas (Bassok, 2001).

Esta relación entre contenido y estructura es tan estrecha que puede afectar la forma en que se interpreta la estructura de la base y el estímulo objetivo, afectando así los contextos donde es aceptable una solución análoga en desmedro de otra. Es así como Bassok y colegas (1995) mostraron que en una tarea experimental los estudiantes universitarios crean de manera espontánea sumas entre palabras con una relación semántica simétrica (X doctores trabajando con Y doctores), mientras que crean divisiones para palabras con una relación asimétrica (X estudiantes reciben Y premios). Esta interpretación de la relación entre contenido y problema matemático es lo que Bassok et al. (1998) denominaron alineación semántica.

Este concepto se refiere al proceso que permite establecer inferencias matemáticas basadas en la información presentada en el problema matemático

escrito. Ellos plantearon que este tipo de procesamiento análogo permite mapear dos tipos de conocimiento que son distintos en naturaleza: el conocimiento semántico concreto y el conocimiento matemático abstracto. Para poner a prueba esta inferencia matemática, Bassok et al. (1998) diseñaron tres experimentos para evaluar cómo estudiantes universitarios utilizaban el conocimiento semántico para crear y evaluar problemas matemáticos escritos.

Para los tres experimentos, ellos crearon sets de palabras que tenían dos posibles tipos de relación entre sí: categórica y funcional. En primer lugar, la relación categórica es una relación semántica simétrica que se da entre palabras que pertenecen al mismo grupo, o cohipónimos, como el caso de naranjas y manzanas que son cohipónimos del superordinado fruta. En segundo lugar, la relación funcional es una relación semántica asimétrica que se da entre palabras, en que una cumple una función en respecto a otra. Es así como las palabras tomates y cajas se encuentran en una relación funcional, ya que las cajas contienen a los tomates (Bassok et al., 1998).

La hipótesis detrás de la creación de estos sets fue que los estudiantes crearían problemas matemáticos distintos dependiendo de los sets que recibiesen. Este fue el caso para los experimentos 1 y 2 donde los estudiantes debían crear problemas matemáticos. Para el caso de los cohipónimos, los estudiantes en su mayoría (82%) crearon problemas de adición, debido al

carácter conmutativo de esta operación matemática. En el caso de la relación funcional, los estudiantes (79%) crearon problemas de división.

El último experimento que ellos reportaron se centró en la evaluación de los participantes de la posibilidad que estos subsets utilizados en los experimentos 1 y 2 apareciesen en problemas de adición o división. La evaluación se realizó a través de una escala de Likert, donde 1 era 'muy improbable' y 7, 'muy probable'. En general, la adición (4.75) fue considerada más probable que la división (3.81) para los distintos subsets. Se destaca que los subsets que contenían cohipónimos estaban mayormente asociados con la adición, mientras que los pares de palabras en relación funcional estaban mayormente asociados con la división.

Guthormsen (2007, p.15) describen estas relaciones en la siguiente figura.

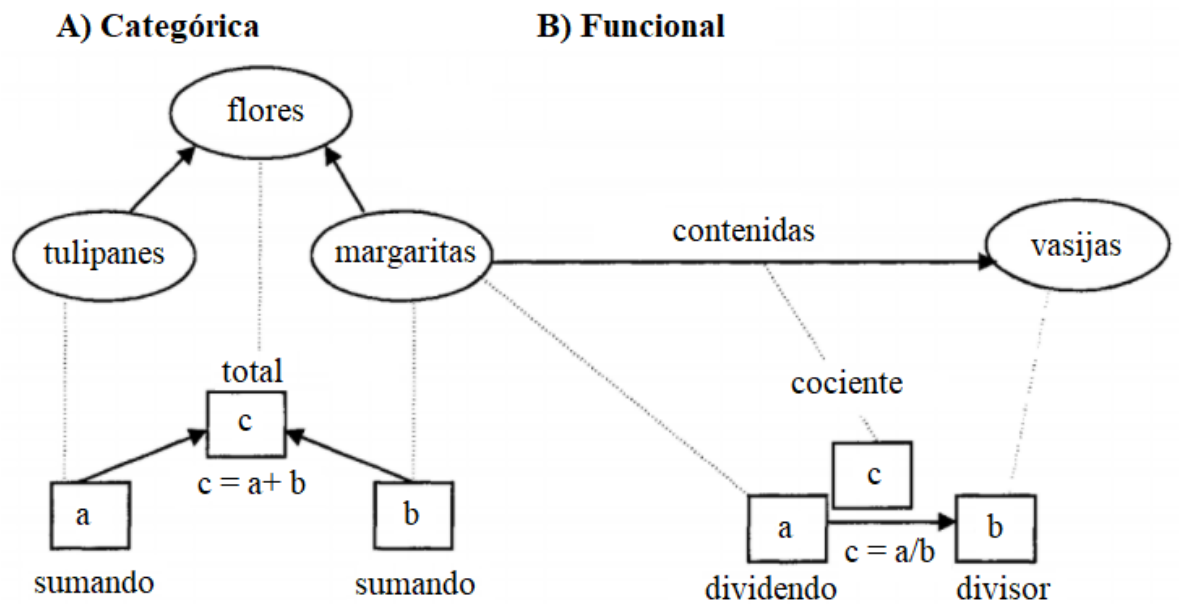


Figura 1 Diagrama de las relaciones A) categóricas y B) funcionales (Guthormsen, 2007, p. 15) ¹.

En esta figura se destaca la relación que hay entre los distintos componentes de las operaciones matemáticas, más allá de los componentes semánticos ya indicados. En la relación categórica, los sumandos también tienen una relación simétrica, ya que su orden es conmutable sin afectar el resultado final y ambos cumplen el mismo rol ($4 + 3 = 3 + 4$). Mientras que, en la relación funcional, el dividendo y el divisor se encuentran en una relación asimétrica al no ser conmutables ($12 \div 3 = 4$, $3 \div 12 \neq 4$) y no tener el mismo rol dentro del problema.

¹ Traducción al español realizada por el autor de esta tesis

Esta preferencia semántica de ciertas operaciones matemáticas ha sido explorada en posteriores oportunidades por Bassok y colegas. En 2008, Bassok et al. investigaron si el efecto de la suma estudiado por LeFevre y colegas (1988) podía ser modulado por la presencia de primes semánticamente alineados y desalineados. El efecto de la suma es la activación no intencionada de la adición, es decir, si se ve $5 + 3$, se activará de manera inmediata el número 8. Esta activación no intencionada genera que, en tareas de coincidencia de números, donde luego de presentar la suma se debe decidir si un estímulo objetivo correspondía a uno de los números presentados en la operación matemática, los participantes se demoren más en rechazar estímulos objetivos que equivalen a la suma de los dos números presentados previamente. Para ambos experimentos reportados, los primes semánticamente alineados activaron la adición, generando mayores tiempos de respuesta.

Para finalizar, en la enseñanza de matemática, Bassok ya en 2001 señalaba que la corta extensión de los problemas matemáticos escritos podría explicar los problemas que generan en niños y adultos. Su corta extensión hace necesario omitir información que en una situación del mundo real sí es relevante, lo que crea problemas para muchos estudiantes. En segundo lugar, estos problemas suelen traer pistas que permiten a los estudiantes procesar los problemas de manera abstracta sin necesidad de poner atención a todo el contenido lingüístico presente. Esto, en ocasiones, genera que los estudiantes

con una comprensión lectora baja no estén en condiciones de corregir sus resultados cuando aparecen anomalías como en problemas de adición, que tengan la palabra menos. Por ejemplo, Luis tiene 52 bolitas menos que Pedro. Si Luis tiene 28 bolitas ¿Cuántas tiene Pedro? La presencia de la palabra menos haría que los estudiantes con baja comprensión lectora sustraigan, utilizando el número de la primera oración (52) como minuendo y el número de la segunda oración (28) como sustraendo, obteniendo el resultado erróneo 24.

Guthormsen et al. (2016) reportaron tres experimentos en que evaluaron cómo el conocimiento de las relaciones semánticas entre palabras puede afectar el procesamiento matemático. En primer lugar, evaluaron cómo la adición de palabras categóricamente relacionadas o cohipónimos, versus la adición de palabras funcionalmente relacionadas afectan las respuestas conductuales y de electrofisiología mediante la técnica Event Related Potential (ERP o potenciales evocados en español) de un grupo de participantes. En este experimento, los participantes debían dar respuestas escritas a problemas como:

1a) 4 rosas + 3 tulipanes = _____

1b) 4 rosas + 3 vasijas = _____

La respuesta ERP de los participantes sugiere que la adición de palabras funcionalmente relacionadas (oración 1b) genera un componente N400, lo que muestra que los participantes consideran la adición como una operación

matemática que se realiza comúnmente entre palabras categóricamente relacionadas. Es decir, la adición de manera abstracta se puede realizar, pero se puede observar una desalineación por la falta de alienación semántica. Además, la respuesta conductual de los participantes es claramente inferior en la adición de palabras funcionalmente relacionadas.

En el segundo experimento reportado, se volvieron a utilizar los sets de palabras categórica o funcionalmente relacionados en el caso de la división. Al igual que en el anterior experimento, los participantes debían escribir una respuesta para problemas como:

2a) 12 rosas / 3 vasijas = _____

2b) 12 rosas / 3 tulipanes = _____

En este caso, la respuesta ERP sugiere que hay desalineación semántica (presencia de N400) al tratar de dividir objetos categóricamente relacionados (oración 2b), pero los resultados no son significativos al tener una muestra muy reducida de participantes.

El experimento final reportado se hace cargo del bajo número de participantes y de cómo el tener que escribir una respuesta podría afectar la medición del EEG. Por lo tanto, la tarea experimental solicitada a los participantes

era la de evaluar si la respuesta a un problema escrito de adición o división era aceptable o no, como:

3a) Doce rosas más tres tulipanes es igual a quince.

3b) Doce rosas más tres vasijas es igual a quince.

3c) Doce rosas divididas en tres vasijas es igual a cuatro.

3d) Doce rosas divididas en tres tulipanes es igual a cuatro.

Las respuestas conductuales y ERP mostraron que los participantes podían agruparse en dos tipos de estrategia para resolver los problemas: modelamiento matemático y no-modelamiento matemático. El primer grupo mostraba la presencia de P600 cuando el segundo objeto de los problemas de adición estaba funcionalmente relacionado (oración 3b) y cuando el segundo objeto de los problemas de división estaba categóricamente relacionado (oración 3d), además de mostrar la presencia de N400 cuando aparecía el resultado final de este tipo de oraciones. Esto sugiere que los participantes evaluaron la desalineación semántica como un problema estructural de la operación matemática, luego del segundo objeto y como una violación a la integración conceptual al momento de evaluar la aceptabilidad del resultado final.

El segundo grupo, por el contrario, no tenía presencia de P600 y tendía a evaluar como aceptables problemas que tenían desalineación semántica. Esto

sugiere que ellos se concentraban en los aspectos abstractos del problema, sin tomar en cuenta cómo estos se anclaban a la realidad mediada por el lenguaje.

Bassok (2001) destacó el gran efecto que tiene la alineación semántica sobre la manera en que las personas razonan sobre los problemas escritos, independiente si interactúan con estructuras matemáticas conocidas o nuevas. Esto, a la vez, sugiere que las personas esperan correspondencias sistemáticas entre los roles que los objetos tienen en la vida real y los roles que asumen en los problemas escritos. Por lo tanto, en un intento por alinear relaciones semánticas con ciertas estructuras matemáticas, se podrían cometer errores.

Recientemente, Wang et al. (2023) destacaron el interés que ha existido en la relación entre lenguaje y matemática, particularmente en el área de la cognición matemática. En su revisión de literatura, ellos indicaron que autores previos han argumentado que el desempeño matemático está estrechamente relacionado con la habilidad de entender el idioma, afectando la capacidad del procesamiento computacional de problemas matemáticos.

Para evaluar esto, realizaron un experimento de adiciones y divisiones semánticamente alineadas y desalineadas en participantes de edad escolar y edad adulta. Se consideraron a estos participantes, ya que el desempeño aritmético y el control inhibitorio mejoran con la edad. En general, sus resultados

apuntaron a que las condiciones semánticas congruentes facilitaban la solución exitosa de sumas, pero no de divisiones para ambos grupos de participantes. En particular, los adultos obtuvieron tiempos de reacción menores y tasas de acierto mayores que los participantes en edad escolar en ambas operaciones matemáticas. Para la adición, Wang et al. sugirieron que la inhibición juega un rol esencial, ya que es necesario suprimir la información incongruente que podría influir en cómo se resuelve un problema matemático. En este caso, los participantes debían suprimir la tendencia a resolver una adición como división, debido a la presencia de palabras funcionalmente relacionadas.

Para explicar por qué la solución de divisiones era más exitosa y rápida en la condición incongruente, Wang et al. (2023) plantearon que las palabras funcionalmente relacionadas tienen menor similitud semántica que las palabras categóricamente relacionadas, lo que dificultaría la velocidad y precisión al resolver estos problemas. Junto con eso, se ha demostrado que resolver divisiones genera una carga cognitiva mayor que las adiciones. Por lo tanto, no se contaría con recursos cognitivos para procesar la inconsistencia semántica y se enfocarían solo en los aspectos numéricos del problema.

Los Efectos del Bilingüismo en el Procesamiento de Problemas Matemáticos Escritos

El efecto del dominio del idioma y su relación con los problemas escritos se ha estudiado principalmente en los Estados Unidos, donde, en la actualidad, 1 de cada 10 estudiantes habla inglés como segundo idioma (Powell et al., 2020). A la vez, no es casualidad que la mayoría de los estudios se hayan llevado a cabo allí porque hay varias políticas que las escuelas estatales deben seguir para asegurarse de que los estudiantes de inglés están aprendiendo el contenido incluso cuando su lengua materna difiere del inglés (Kieffer et al., 2009). En esta línea, Kieffer et al. (2009, p. 1170), después de realizar un metaanálisis sobre la eficiencia de las adaptaciones del idioma, concluyen que "there is indeed a substantial link between students' English language proficiency and their performance on tests of math, science, and social studies.". A saber, los aprendientes de inglés logran puntuaciones significativamente más bajas que sus pares que son hablantes nativos. Además, Kieffer y colegas (2009, p. 1170) señalaron que "several correlational studies have found that assessments and individual test items that have more linguistic complexity yield larger performance gaps between ELLs and non-ELLs." Por lo tanto, debido al dominio del idioma de los estudiantes, se podría estar evaluando injustamente el conocimiento matemático que los aprendientes de inglés pueden poseer, debido a la varianza irrelevante del constructo (Banks et al., 2016).

Para entender de mejor manera este fenómeno, es necesario tener en cuenta las conclusiones a las que llegaron los investigadores en Estados Unidos y en otros países, como se verá en las siguientes subsecciones.

Estudios Realizados en Estados Unidos

Desde 1996, en los Estados Unidos se ha informado que los aprendientes de inglés obtienen puntuaciones más bajas que sus pares, con una brecha que crece a medida que los años de instrucción pasan. Por lo tanto, ha existido una creciente necesidad de garantizar que los hablantes no nativos de inglés alcancen el mismo nivel de conocimiento que sus pares nativos, especialmente teniendo en cuenta que en los Estados Unidos uno de cada diez estudiantes era un aprendiente de inglés en los años 2014-2015 (Martin & Fuchs, 2019).

Abedi y Lord (2001), después de destacar el papel del dominio del lenguaje en las pruebas matemáticas (como se informó en investigaciones anteriores), intentaron nivelar el campo de juego, simplificando los problemas escritos para determinar si los puntajes mejoran para los aprendientes de inglés, como una manera de medir la capacidad matemática y no la comprensión lectora. Después de crear el material de evaluación simplificado, estudiaron su efecto mediante una entrevista y una prueba posterior.

A un grupo de estudiantes se les mostraron preguntas de prueba modificadas y no modificadas para recopilar sus impresiones. La mayoría de los estudiantes declararon que consideraron que las versiones modificadas eran más fáciles de entender y, por lo tanto, podían ahorrar tiempo a la hora de leer y solucionar los problemas. Más aún, destacaron que eran más directas y con menor cantidad de palabras complicadas que las versiones no modificadas, algo que podría influir en el rendimiento de los hablantes no nativos de manera positiva.

Después de lograr estas apreciaciones positivas, Abedi y Lord (2001) probaron las preguntas de prueba modificadas y no modificadas en una gran muestra (n=1174) de estudiantes de 8vo básico, de diferentes clases de matemáticas, orígenes sociales y lenguas maternas. A modo de resumen, la comparación de los resultados produjo los siguientes resultados: 1) los aprendientes de inglés obtuvieron puntuaciones más bajas que los hablantes nativos, 2) los estudiantes de clases avanzadas de matemáticas tuvieron un mejor desempeño que los estudiantes de clases no avanzadas, 3) los estudiantes de un nivel socioeconómico más bajo tuvieron peores resultados que sus compañeros, y 4) no hubo diferencias en puntaje relacionadas con el género. A pesar de estos resultados, la versión modificada de los problemas escritos sí benefició de manera significativa a los hablantes no nativos, los estudiantes de clases menos avanzadas y los estudiantes de clases sociales bajas. Por lo tanto,

estos resultados sugieren que los grupos marginados se benefician de problemas de palabras simplificados que no disminuyeron la dificultad matemática, lo que les permitió mostrar sus conocimientos matemáticos de una manera más justa.

Lager (2006), mediante la aplicación de una prueba y una entrevista posterior, intentó proporcionar información sobre los problemas que tienden a ser respondidos incorrectamente y las razones lingüísticas por las que son difíciles para un grupo de estudiantes nativos del inglés y aprendientes de inglés de diferentes escuelas en sexto y octavo básico. Al replicar los hallazgos de investigaciones anteriores, los estudiantes nativos lograron resultados estadísticamente significativos más altos que los hablantes no nativos, con un tamaño de efecto medio ($d = .617$). Sin embargo, la puntuación promedio para los hablantes nativos fue de 3,32 de 9 puntos totales, lo que indica claramente que también tuvieron problemas con la prueba.

Por lo tanto, después de que los estudiantes tomaron la prueba y destacaron las palabras que no podían entender, se entrevistó a una muestra de los participantes. De las entrevistas surgieron algunas categorías que podrían proporcionar información sobre las diferentes dimensiones que deben tenerse en cuenta al enseñar problemas escritos. En primer lugar, algunas de las palabras más resaltadas, como patrón, figura, entre otras, tienen características semánticas que pueden conducir a malentendidos. Particularmente, patrón

puede ser parte de un nuevo vocabulario técnico, mientras que figura puede tener varios significados (polisemia). En segundo lugar, la notación matemática, como los paréntesis, también tiene diferentes significados en las matemáticas, lo que además puede conducir a confusión. Por último, los alumnos que no proporcionaron la respuesta correcta tenían la impresión de que entendían las instrucciones correctamente. Este dilema, como afirma Lager (2006), podría producir más dificultades para construir conocimiento aritmético porque "they will not be motivated to ask the appropriate questions or to seek outside help to correct them." (p. 191).

Teniendo en cuenta lo anterior, Lager (2006) sugirió la necesidad de ampliar el léxico de los estudiantes. Este léxico más amplio también debe considerar el nivel de abstracción que tienen las palabras porque los conceptos concretos son más fáciles de entender para los estudiantes de educación básica. En cuanto a la enseñanza, Lager sugirió que los profesores de matemáticas también deben ser profesores de idiomas en el sentido de que necesitan ser conscientes del lenguaje utilizado en sus clases para anticipar las necesidades de cada estudiante y así identificar posibles malentendidos. Además, deben pedir a los alumnos que expliquen el fundamento de sus respuestas para asegurar la comprensión. Por último, los profesores deben aprovechar las fortalezas lingüísticas de los estudiantes no nativos que ya saben que un concepto puede

tener diferentes significados en diferentes idiomas, algo similar a lo que ocurre en las matemáticas.

Ockey (2007) se centró en analizar los diferentes ítems en una prueba estandarizada con el fin de identificar los posibles sesgos contra un grupo dado de estudiantes. Su primer objetivo fue establecer si había alguna diferencia en los puntajes entre hablantes nativos y no nativos en una prueba estandarizada. No es de extrañar que la diferencia de puntuaciones se haya establecido una vez más. En particular, los hablantes nativos tenían una puntuación media tres puntos más alta que los aprendientes de inglés, con un tamaño de efecto medio ($d = .573$). Con esta evidencia, se analizaron los distintos ítems de la prueba para identificar los posibles sesgos contra los hablantes no nativos. Al realizar dos análisis diferentes, sólo se identificó un ítem de 20 que tenía un funcionamiento diferencial que perjudicaba a los aprendientes de inglés. Este resultado puede deberse a la casualidad o a las limitaciones de los análisis utilizados o a los distintos niveles de dominio del idioma de los hablantes no nativos. A pesar de esto, Ockey (2007) concluye que parece ser el caso de que las pruebas estandarizadas son un medio justo para evaluar los problemas escritos para estudiantes nativos y no nativos del inglés, incluso cuando estos no han sido simplificados como sugería Abedi y Lord (2001) y sin haber habido intervenciones orientadas a la resolución de problemas escritos en el aula como sugería Lager (2006).

En relación con estudios más recientes, se pueden destacar las conclusiones de Banks et al. (2016), Martin y Fuchs (2019) y Powell et al. (2020). En primer lugar, Banks y colegas afirmaron que hay una brecha de puntuación persistente entre los estudiantes de segundo idioma, en este caso el inglés, y los estudiantes nativos. Por lo tanto, evaluaron el posible efecto que ciertos componentes lingüísticos pueden tener en las puntuaciones entre aprendientes de inglés y hablantes nativos. Curiosamente, hubo pocos elementos que mostraron una interacción, como los ítems en voz pasiva o con cláusula condicional. aunque las puntuaciones fueron considerablemente más altas en el grupo de hablantes nativos.

Martin y Fuchs (2019) informaron sobre los resultados de un grupo de estudiantes de primero básico en varias pruebas tomadas en primavera y otoño para establecer una comparación. Estos estudiantes fueron clasificados en función de su dominio del idioma y su estado como en riesgo de una discapacidad matemática. En esta comparación, en el grupo de riesgo, los hablantes nativos y los aprendientes de inglés tuvieron un desempeño similar en problemas computacionales, mientras que, en los problemas escritos, los hablantes nativos tuvieron un desempeño considerablemente mejor que los hablantes no nativos ($d=0,64$). En el grupo de no riesgo, estas diferencias fueron aún más pronunciadas, ya que los hablantes nativos lograron puntuaciones más altas en habilidades computacionales y problemas escritos en ambas pruebas,

acrecentando la brecha en resultados durante todo el año. Estos resultados sugieren la necesidad de centrarse en el vocabulario académico utilizado en matemáticas para reducir la brecha entre los hablantes nativos y aprendientes de inglés.

Powell et al. (2020) describieron claramente los diferentes desafíos que los aprendientes de inglés y los estudiantes con dificultades matemáticas tienen que enfrentar cuando se trata de aprender el lenguaje de las matemáticas. En primer lugar, hay una amplia gama de vocabulario matemático académico que los estudiantes necesitan saber para entender con éxito los problemas. Este vocabulario puede solaparse con el lenguaje cotidiano, creando conflictos de significado, similitud fonológica e incluso falsos cognados en otros idiomas. Como se mencionó en investigaciones anteriores (Abedi & Lord, 2001; Lager, 2006; Ockey, 2007), los hablantes no nativos tienen dificultades con los problemas matemáticos escritos desde temprana edad, aumentando estas dificultades y ampliando la brecha con sus compañeros a medida que pasan a cursos superiores. Al mismo tiempo, los estudiantes con dificultades matemáticas tienen un menor rendimiento en el conteo de números, comparación, solución de problemas escritos, comprensión racional de números y álgebra. Un tercio de ellos también puede tener dificultades en la lectura, logrando puntuaciones aún más bajas (Abedi & Lord, 2001; Lager, 2006; Ockey, 2007).

Después de aplicar tres pruebas de problemas matemáticos escritos de un solo dígito, ecuaciones abiertas y vocabulario matemático a estudiantes de tercero básico, Powell et al. (2020) concluyeron que los hablantes nativos obtuvieron mejores resultados en todas ellas, en diferentes cursos, en comparación con los aprendientes de inglés. Sin embargo, los estudiantes sin dificultades matemáticas, sin importar si eran nativos o no, obtuvieron una puntuación inferior al 50% en las pruebas de vocabulario, lo que indica que los estudiantes no conocen todo el vocabulario matemático necesario para ser exitoso en los problemas matemáticos, lo que podría acrecentar las diferencias en resultados para los hablantes no nativos.

Como se puede notar, los distintos estudios a partir de 2001 en adelante han dado cuenta de la robusta conclusión que los hablantes nativos tienen mejores resultados en los problemas matemáticos escritos al ser comparados con aprendientes de inglés. Junto con esta conclusión, también se puede destacar que en la mayoría de los estudios no hubo diferencias en problemas computacionales y que no ha sido posible establecer sesgos que afecten a los aprendientes de inglés. En la siguiente subsección, se discutirán los estudios de este fenómeno en otros países.

Estudios Realizados en Otros Países

Bernardo (2002), en su estudio de hablantes de filipino e inglés, indicó que si el idioma en que se presenta el problema matemático es el mismo que la L1 del estudiante (inglés o filipino), los estudiantes obtienen resultados que son comparativamente superiores a los obtenidos cuando los estudiantes son evaluados en su L2. En este caso particular, los estudiantes debían dar soluciones para distintos problemas matemáticos del nivel NB2 (3° y 4° básico) y se evaluó la comprensión del problema y la solución dada, dependiendo del idioma utilizado. En este mismo estudio, Bernardo sugirió que el nivel de proficiencia en el idioma podía tener un efecto significativo al momento de solucionar problemas escritos, algo que no había sido considerado con anterioridad.

En la República de Irlanda, Ní Ríordáin y O'Donoghue (2009) señalaron la necesidad de tener resultados específicos para cada país para evitar un sesgo de generalización. Por lo tanto, ellos se centraron en las posibles variaciones en las puntuaciones matemáticas que pueden estar presentes cuando los estudiantes que provienen de escuelas donde el idioma de la enseñanza es el irlandés, se trasladan a escuelas donde el idioma de instrucción es el inglés. En la primera transición, de sexto básico a séptimo básico, el dominio del idioma inglés se correlacionó moderadamente con las puntuaciones que los estudiantes

obtuvieron tanto en estudiantes monolingües como bilingües. En la segunda transición, de octavo básico a primero medio, las diferencias en el dominio del idioma inglés fueron más pronunciadas. El dominio del inglés de los estudiantes monolingües se encontró muy fuertemente correlacionado con su desempeño en problemas escritos. Mientras que el dominio del inglés de los estudiantes bilingües se correlacionó moderadamente con su rendimiento en problemas escritos. Los resultados de la segunda transición están en concordancia con investigaciones previas que afirman que las diferencias en las puntuaciones de problemas matemáticos escritos son más marcadas en la educación media.

Al medir la correlación entre el dominio irlandés y las puntuaciones en los problemas escritos en inglés, hay una correlación moderada para la primera transición. Mientras tanto, en la segunda transición, las correlaciones fueron débiles y no significativas. Esto que sugiere que los bilingües no se benefician de saber irlandés durante la segunda transición al resolver problemas escritos en inglés.

Además, Ní Ríordáin y O'Donoghue (2009) después de clasificar a los bilingües en diferentes grupos en función de su dominio en ambos idiomas, evaluaron si las diferencias en las puntuaciones promedio eran significativas. Este análisis mostró que los bilingües que tienen un alto dominio en ambos idiomas, obtienen una puntuación superior a la de los bilingües que tienen una

lengua dominante, un bajo dominio de ambos idiomas e incluso los monolingües en la primera transición. Esto sugiere que los bilingües altamente competentes tienen ventajas a la hora de resolver problemas escritos.

En la segunda transición, los bilingües de alto nivel tienen puntuaciones significativamente más altas que todos los grupos, excepto el grupo de los bilingües dominante en inglés. Esto replicaba los resultados de la primera transición, con dos diferencias importantes. En primer lugar, las pruebas estadísticas indicaron que no hay ninguna diferencia significativa entre las puntuaciones promedio de los bilingües altamente competentes y los bilingües que son dominantes en inglés. En segundo lugar, los bilingües que son dominantes en irlandés no exhibieron grandes beneficios en la segunda transición como lo hicieron en la primera.

En Filipinas, al aplicar la misma prueba en filipino e inglés por dos medios diferentes, escrito y leído en voz alta, Bautista et al. (2009) intentaron probar las estrategias encontradas en la literatura sobre las dificultades con los problemas escritos. Su análisis dio dos resultados importantes. En primer lugar, el lenguaje de la prueba tuvo un efecto significativo. Los estudiantes filipinos tenían mejores calificaciones cuando la prueba estaba en filipino, independientemente de los medios para presentar el problema. Esto significa que el uso de su lengua materna llevó a respuestas más correctas y menos errores de operación,

beneficiando en menor proporción a los estudiantes que tenían menores rendimientos. En segundo lugar, los estudiantes que tenían un mayor dominio del idioma inglés tenían mejores puntuaciones que los de menor dominio, lo que puede implicar que su mayor dominio del inglés fue un elemento clave para lograr puntuaciones más altas.

Al notar que el dominio del idioma es un factor clave para resolver problemas escritos, Kempert et al. (2011) en Alemania estudiaron cómo este efecto puede ilustrarse en un grupo de estudiantes turco-alemanes en comparación con un grupo de monolingües alemanes. Al mismo tiempo, compararon las puntuaciones de los bilingües en función de su competencia lingüística para identificar posibles diferencias. Los problemas escritos utilizados estaban en alemán para ambos grupos. Para los bilingües, también se presentaron problemas escritos en turco para ilustrar aún más las posibles diferencias. En concordancia con la literatura, los monolingües obtuvieron mejores resultados en los problemas escritos que los bilingües, aunque ambos grupos tenían un mejor desempeño en los problemas sin distractores. Particularmente para los bilingües, los que tenían un mayor dominio del alemán lo hicieron mucho mejor que sus compañeros menos competentes. Esto también se produce en los problemas escritos en turco, aunque en menor medida. Los resultados conductuales sugieren que el nivel de proficiencia en el idioma del problema escrito mostraba una correlación directa, significativa y fuerte.

Pavón y Cabezuelo (2019) se centraron en la relación entre dominio del idioma y la resolución de problemas escritos en los programas de inmersión en Andalucía, España, en participantes de 2º medio. Mediante una prueba de comprensión de lectura en inglés, una prueba de problemas escritos tipo PISA con problemas seleccionados en inglés y español y un cuestionario, los autores intentaron proporcionar más información sobre la posible relación del idioma de las pruebas, el dominio del idioma y la dificultad de los problemas escritos.

En relación con el idioma de las pruebas, los participantes lograron puntuaciones más altas cuando los problemas estaban en español (su L1) que cuando estaban en inglés. Esta diferencia fue significativa con un tamaño del efecto pequeño ($d = .498$). Al analizar los resultados basados en el nivel de dominio del inglés de los participantes, los resultados no fueron significativos, lo que significa que la comprensión de la lectura podría no afectar la resolución de problemas escritos. Este resultado no está en concordancia con la percepción de los estudiantes sobre el papel de la comprensión de la lectura para resolver con éxito los problemas escritos. Vale la pena señalar que la proficiencia en inglés de los participantes osciló desde el nivel A2 hasta el nivel B2. Particularmente, solo cinco participantes se encontraban en el nivel A2. Esto podría explicar por qué la comprensión de lectura no sería un factor a tener en cuenta para las puntuaciones obtenidas en los problemas escritos en este estudio puesto que la

mayoría de los participantes tendría las herramientas lingüísticas para resolver esta tarea.

En relación con la complejidad del problema escrito, hay una diferencia significativa en las puntuaciones en función de la dificultad y el lenguaje de la prueba, pero ambos factores no interactuaron. Además, el idioma sólo representa el 7,9% de las diferencias, lo que indica que las diferencias en las puntuaciones se explican principalmente por la dificultad. Las diferencias pronunciadas en dificultad fueron significativas, lo que concuerda con las puntuaciones más bajas logradas en inglés. Cuando se trataba de problemas menos difíciles, no hubo diferencias significativas en las puntuaciones logradas en inglés y español. Los estudiantes también informaron que resolver problemas en inglés tomó más tiempo que en español. Los hallazgos relacionados con el lenguaje de las pruebas son consistentes con investigaciones anteriores (Abedi & Lord, 2001; Banks et al., 2016; Lager, 2006; Martin & Fuchs, 2019; Ockey, 2007; Powell et al., 2020). Además, la dificultad del problema es otro factor que debe tenerse en cuenta en futuras investigaciones.

Trakulphadetkrai et al. (2020) intentaron identificar el papel que el lenguaje (como capacidad general del lenguaje), la comprensión lectora y la memoria de trabajo tienen en el conocimiento matemático. Para lograr este objetivo, estudiaron varias escuelas en el Reino Unido, donde evaluaron particularmente

a estudiantes no nativos y nativos en la comprensión de la lectura (prueba YARC), la capacidad general del lenguaje (C-test) y la memoria funcional (prueba AMWA), mientras que también midieron el conocimiento matemático, pidiendo a los estudiantes que respondiesen cinco problemas sin palabras y 15 problemas escritos.

Al comparar las puntuaciones de toda la prueba de matemáticas, no se encontraron resultados significativos. Al analizar sólo las puntuaciones de los problemas escritos, el rendimiento de los estudiantes no nativos fue significativamente diferente con los hablantes nativos, logrando puntuaciones más altas. Estos resultados también tuvieron un tamaño de efecto medio ($d = 0,677$). Al analizar las diferencias en el rendimiento en la comprensión lectora, la capacidad general del lenguaje y la memoria de trabajo, los hablantes nativos tuvieron un mayor desempeño en la comprensión lectora y la capacidad general del lenguaje, sin diferencias significativas en la memoria de trabajo.

Por último, el análisis de la contribución de la comprensión lectora, la capacidad general del lenguaje y la memoria de trabajo mostró que hay una correlación significativa entre la comprensión lectora y el desempeño en problemas matemáticos escritos, lo que es consonante con investigaciones previas. Ambos grupos tenían una correlación significativa entre la capacidad de lectura de una sola palabra, la capacidad general del lenguaje y la memoria de

trabajo, que apoya aún más el papel de la competencia lingüística para resolver problemas escritos.

Es importante destacar que el análisis de regresión identificó la capacidad general del lenguaje como la variable que explica la mayor variación en los puntajes obtenidos en los problemas escritos. Curiosamente, mientras que los puntajes en habilidades lingüísticas generales predicen el desempeño de los hablantes nativos, los puntajes en comprensión lectora predijeron el desempeño de los aprendientes de inglés. Esto sugiere que la relación entre habilidad lingüística y el desempeño en problemas matemáticos escritos puede que no siga la misma trayectoria para hablantes nativos y aprendientes de inglés. Esta diferencia en la manera en que el lenguaje afecta a hablantes nativos y no nativos podría explicar por qué hay diferencias tan pronunciadas en los puntajes obtenidos, incluso cuando las pruebas estandarizadas no tienen un prejuicio en contra de los aprendientes de inglés, como sugirió Ockey (2007) en su análisis.

Al igual que los estudios realizados en Estados Unidos, estas investigaciones conductuales muestran evidencia que sugiere que el idioma de los problemas escritos tiene un efecto en su comprensión y resolución positiva. Más aún, también muestran la necesidad de no incurrir en un sesgo de generalización y de estudiar realidades distintas a la estadounidense.

Junto con el rol que tiene el lenguaje al momento de resolver problemas matemáticos escritos, es necesario considerar la manera en que las funciones ejecutivas, particularmente la inhibición, podría afectar nuestra capacidad de resolver problemas matemáticos con información conflictiva.

El Rol de la Inhibición como Función Ejecutiva

Müller y Kerns (2015) consideraron como función ejecutiva a los procesos cognitivos necesarios para el control descendente de una acción, pensamiento o emoción que están relacionados con el córtex prefrontal, aunque actualmente se plantea que hay una activación neural distribuida en todo el cerebro (Chikara & Ko, 2020; Kang et al., 2022). Ellos también destacaron que esta definición no está exenta de controversia, ya que no ha sido posible llegar a un acuerdo para dar una definición más precisa. En esta misma línea, destacan que ha habido al menos 33 conceptos relacionados con la función ejecutiva; que algunos autores plantean que varios de estos conceptos son parte de otros procesos cognitivos, e.g. la inhibición es parte de la memoria de trabajo; y que hay una gran cantidad de teorías y enfoques que han tratado de explicar, con distintos niveles de éxito, el desarrollo de la función ejecutiva a medida que los seres humanos crecen y se obtienen resultados experimentales.

A pesar del confuso panorama presentado anteriormente, en esta sección nos enfocaremos en la inhibición como una función ejecutiva para luego relacionarla con el procesamiento matemático y el lenguaje.

La inhibición es definida por MacLeod (2007, p. 5) como “the stopping or overriding of a mental process, in whole or in part, with or without intention.” En este sentido, la inhibición es una función ejecutiva que puede influir sobre la atención selectiva, la memoria de trabajo o de largo plazo y procesos cognitivos como el procesamiento matemático abstracto. Müller y Kerns (2015) indicaron algunas críticas que se hacen a esta función ejecutiva y que futuras investigación tendrían que resolver. En primer lugar, muchas tareas que involucran mecanismos de inhibición también podrían ser explicados por mecanismos no inhibitorios. A modo de ejemplo, ellos indican que en la tarea Stroop, niños entre 7 y 9 años mostraban una supresión mayor de la lectura de la palabra que sus pares de 9 a 11 años. Este resultado es contradictorio con la idea de que, a menor edad, se tendría menor control para inhibir estímulos no deseados. En segundo lugar, la inhibición es un mecanismo negativo que evita la ejecución de una acción no deseada pero no permite explicar cómo se llega a la acción deseada. Finalmente, también indicaron que algunos investigadores han sugerido que la inhibición no es un proceso independiente, sino que es un efecto secundario de la activación de la memoria de trabajo, como fue mencionado anteriormente.

De todos modos, como Van Dooren e Inglis (2015) indicaron, la inhibición está presente en procesos atencionales en la memoria. En cuanto a los procesos atencionales, se puede mencionar el caso del priming negativo donde la inhibición de un estímulo en favor de otro estímulo objetivo hace que el procesamiento del estímulo inhibido se vea deteriorado (Van Dooren & Inglis, 2015). Particularmente, si la tarea experimental solicita enfocarse en las palabras que parten con una letra en particular dentro de una lista, esa letra se convierte en el centro de atención, inhibiendo al resto. Luego, si se les pide a los participantes enfocarse en otra letra, su rendimiento será más lento debido a la inhibición realizada anteriormente. En cuanto a la memoria, se puede mencionar cómo la práctica constante de una operación matemática con ciertos números objetivo dificulta la resolución exitosa de otras operaciones matemáticas que utilizan esos mismos números. A modo de ejemplo, si se les pide a los participantes practicar multiplicaciones con algunos números objetivo (4×2 , 6×7) para luego pedirles que realicen adiciones que podrían incluir a los números objetivo ($4 + 2$, $6 + 7$) o números de control, los participantes tendrán tiempos de respuesta mayores y una precisión menor en la adición si aparecen los números objetivo, debido a un olvido momentáneo producido por la inhibición de esta operación matemática en la sección de práctica (Van Dooren & Inglis, 2015).

En cuanto a los tipos de inhibición, Müller y Kerns (2015) indicaron que hay una gran cantidad de taxonomías que describen estos procesos, pero

destacan dos en particular: *la inhibición de una respuesta y el control de interferencia*. En la primera, es necesario que una respuesta automática, dominante o predominante se detenga, se atrase o se omita. A modo de ejemplo, este proceso inhibitorio está presente cuando un infante es capaz de seguir las instrucciones de su cuidador para detener una acción, como bajar la voz. En el segundo proceso, es necesario suprimir una respuesta dominante, mientras se realiza una respuesta no dominante que compite con la anterior. Un ejemplo claro de este proceso es la tarea Stroop donde se debe omitir la lectura de la palabra que está escrita mientras se dice el color que esta tiene. Ambos procesos mejoran a medida que los niños crecen hasta llegar a la adultez.

A pesar de que la inhibición puede ocurrir de manera no intencionada, como se mencionó anteriormente, Kunda y Spencer (2003) sugirieron que inhibir un proceso cognitivo es demandante y que esta función cognitiva puede fallar a medida que el uso de recursos cognitivos aumenta. No solo los recursos cognitivos a utilizar pueden afectar la inhibición, ya que también el sufrir de ansiedad (Wood et al., 2001) o de lesiones prefrontales (Alvarez & Emory, 2006) puede afectar negativamente la capacidad de los individuos para inhibir procesos cognitivos en tareas tales como evaluar la pertinencia semántica de dos palabras mientras se practica un set de números o al resolver un problema que requiere como estrategia hacer movimientos contraintuitivos.

La Relación entre la Inhibición y el Procesamiento del Lenguaje

Müller y Kerns (2015) recogen las ideas de Vygotsky y Luria para establecer la noción que el lenguaje permite a los infantes internalizar y regular su propio comportamiento y lo que se espera de ellos al tener el lenguaje una base y función social. Para ilustrar esta idea, discutieron los resultados de una tarea go/no go, donde se debe realizar una acción si aparece un estímulo determinado e inhibir una acción si aparece un estímulo distinto, realizada por Luria donde se estudió el efecto de etiquetar las acciones que los participantes iban a realizar o inhibir. Los participantes de 3 años tuvieron dificultades al momento de realizar la versión básica de la tarea donde sólo debían apretar un botón cuando una luz roja aparecía, mientras que los participantes de mayor edad no tuvieron problemas para sortear esta tarea. En el caso en que se les pedía a los participantes de 3 años acompañar su acción motora con una verbalización sencilla de la acción que realizaban, mejoraron sus resultados considerablemente. En el caso en que se les pedía acompañar sus acciones no motoras (no presionar el botón) con una verbalización, los participantes obtenían peores resultados que en la versión básica del experimento. Los autores utilizaron estos resultados para proponer la idea que los infantes de 3 años utilizaban el lenguaje para regular su comportamiento, aunque sin ser capaces de poder abstraerlo a situaciones que no requerían acciones motoras. Este no es el caso para los niños preescolares mayores.

Obermeyer et al. (2020), de acuerdo a resultados experimentales anteriores, indicaron que la inhibición verbal deficiente podría estar relacionada con tareas semánticas del lenguaje, resultando en una habilidad reducida para inhibir representaciones no relevantes, en un incremento en la competencia de ítems en la selección léxica y en una habilidad reducida para resolver la ambigüedad a nivel léxico y sintáctico. Para estudiar este fenómeno, se enfocaron en el desempeño de pacientes afásicos en distintas tareas lingüísticas: nombrar, repetición de palabras, repetición de no palabras, categorización y rima; y en tres tareas de función ejecutiva: la tarea Simon, la tarea Flanker y el cambio de número-letra. Las tareas lingüísticas variaban en carga cognitiva al poner distintos tiempos entre el estímulo y el momento de responder (1 a 5 segundos) y una tarea de repetición de números al azar por 5 segundos entre el estímulo y el momento de responder. Agregar esta tarea aumenta la carga cognitiva. Las dos primeras tareas de función ejecutiva se relacionan con inhibición y la última con cambio de atención. Luego de analizar los datos, se determinó que sólo la inhibición tiene un efecto en las tareas de categorización y de repetición de palabras siempre que haya una carga cognitiva mayor, mientras que el cambio de atención no tuvo un efecto en ninguna de las tareas. Con estos resultados, Obermeyer y colegas plantearon que la inhibición es una habilidad ejecutiva necesaria para tareas que requieren procesamiento léxico-semántico, especialmente si hay una carga cognitiva superior.

En esta línea, Federmeier y colegas (2020) recogen los resultados de una variedad de estudios que indican el importante rol de las habilidades cognitivas generales sobre las tareas lingüísticas. En las tareas de producción, se da cuenta de la acción inhibitoria a través de la actividad Alpha en una tarea de nombrar imágenes (Federmeier et al., 2020). Al momento de nombrar una imagen, se ve mayor activación de Alpha en las áreas motoras del cerebro, lo que da cuenta de una acción inhibitoria de áreas cerebrales que no son relevantes para completar exitosamente la tarea. En los participantes que tenían menor activación Alpha en áreas irrelevantes, tenían tiempos de respuesta mayores que los participantes que tenían mayor activación (Federmeier et al., 2020).

No solo esta actividad cerebral participa en la inhibición, ya que en tareas de nombrar un estímulo donde hay más de una opción posible, también se ve mayor activación de N2, que está relacionada con la inhibición de actividad motora en tareas de go/no-go (Federmeier et al., 2020). En tareas de comprensión, la desambiguación de significados se relaciona con un efecto ERP negativo frontal mayor y también con el tamaño de la sacada visual para medir la efectividad de la selección de significados (Federmeier et al., 2020). En estos casos, el suprimir el significado de una palabra contextualmente inadecuada, como en el caso de *duck*, que puede significar pato o agacharse, requiere que se suprima el movimiento de los ojos, produciendo fijaciones más lentas.

Kroll y Bialystok (2013) indicaron que hay evidencia para sugerir una relación bidireccional entre las lenguas que un bilingüe maneja, incluso en contextos donde solo una lengua está siendo utilizada. Ellos plantearon que para seleccionar adecuadamente cuál de las dos lenguas se va a utilizar es necesario un mecanismo en el sistema de control ejecutivo. Para plantear esto, ellos se basaron en la evidencia que se ha reportado en estudios con Resonancia Magnética funcional (fMRI) que indica que los sistemas ejecutivos frontales utilizados para cambiar de lengua son los mismos utilizados en la atención selectiva en tareas de función ejecutiva no verbales. A la vez, estas redes son utilizadas de manera más eficiente por los hablantes bilingües. Estas ventajas en función ejecutiva por parte de los bilingües se han establecido en estudios conductuales y de neuroimagen. De acuerdo a estos resultados, Kroll y Bialystok (2013) plantearon que el uso de dos sistemas lingüísticos generaría una reorganización de recursos tanto lingüísticos como no lingüísticos, particularmente los relacionados con la función ejecutiva.

Junto con lo planteado anteriormente, Kroll y Bialystok (2013) también indicaron que la activación paralela de ambas lenguas está presente incluso cuando se está utilizando solamente una lengua, teniendo resultados experimentales para el reconocimiento auditivo de palabras, reconocimiento visual de palabras y producción oral. Aunque ambas lenguas estén activadas en estas tareas, difieren en la manera en la que se manifiesta la coactivación de

lenguas. A modo de ejemplo, en el reconocimiento visual de palabras, la evidencia de neuroimagen indica activación de ambas lenguas incluso si la información disponible ya indica claramente que solo se trata de una lengua, lo que intuitivamente sugeriría que solo debería haber acceso léxico a una lengua. Por el contrario, en algunas condiciones experimentales, los bilingües sí son capaces de acceder selectivamente a una sola lengua en tareas de reconocimiento auditivo de palabras.

En esta línea, Kroll y Bialystok (2013) destacaron los resultados de una tarea de reconocimiento visual de palabras, donde los participantes debían decidir si dos palabras se relacionaban semánticamente. Algunos pares contenían homógrafos interlingües, palabras que comparten forma, pero no significado en dos lenguas como pie-toe (donde pie hace referencia a una masa dulce en inglés y no significa lo mismo que pie, la parte del cuerpo en español), siendo seguidos por un par (foot-hand) que contenía la interpretación inicial del homógrafo en español traducido al inglés, foot en este caso. Los resultados experimentales indican que los bilingües tenían tiempos de respuesta mayores si el par contenía un homógrafo interlingual, siendo estos tiempos de respuestas mayores tras pasados colateralmente al par de palabras que contenía el homógrafo traducido. Estos resultados apoyan la idea de un control inhibitorio por parte de los participantes para evaluar correctamente la relación semántica de las palabras sin dejarse influenciar por las lenguas que dominaban.

Esta idea de control inhibitorio sobre la L1 también se da en bilingües altamente proficientes, como Declerck et al. (2015) plantearon. En una tarea de nombrar números en alemán, turco e inglés, los participantes tuvieron mayores tiempos de respuesta y porcentajes de error menores en sus lenguas dominantes, alemán y turco, que, en su lengua menos dominante, inglés. Estos resultados sugieren que las lenguas que tienen una mayor activación, ya sea por uso o proficiencia, requieren de una inhibición mayor para poder sortear de manera exitosa la tarea experimental. A la vez, esta inhibición es persistente y, por lo tanto, se traspa a la siguiente prueba experimental, generando mayores costos cognitivos que se ven expresados en los resultados obtenidos.

Tomando estos resultados como punto de inicio, Declerck y Phillipp (2018) se propusieron estudiar no solamente la producción, sino que también la comprensión en hablantes bilingües. Para este propósito, implementaron una tarea de nombrar imágenes, otra de nombrar las palabras que aparecían en la pantalla y otra de categorización de palabras escritas. En esta oportunidad, se enfocaron en hablantes nativos de alemán que también dominaban el inglés y el francés, en menor medida. En las tareas de producción, se replicaron los resultados anteriores, donde hubo costos al momento de cambiar de lengua debido a la inhibición de las lenguas dominantes. En la tarea de comprensión, por el contrario, solo se obtuvieron resultados que sugieren una inhibición parcial. Es decir, el control inhibitorio se hizo presente en costos cognitivos al momento

de categorizar las palabras en francés, la lengua menos dominante de las tres. Para explicar estos resultados, los autores recurren al modelo BIA-d de Grainger y colegas (2010). En este modelo, los procesos de activación en cada tarea son distintos, generando inhibición en cantidades distintas. En primer lugar, las tareas de producción tienen un origen endógeno, donde el propósito de hablar en una lengua en particular activará los nodos lingüísticos, inhibiendo a las otras lenguas. Por el contrario, las tareas de comprensión tienen un origen exógeno, donde la lengua del estímulo presentado activa los nodos lingüísticos, inhibiendo a las otras lenguas.

Fricke y colegas (2019) apoyaron lo planteado anteriormente al indicar que en tareas de producción los bilingües proficientes inhiben su L1 cuando planean lo que van a decir en su L2. Junto con eso, también indicaron que estos mecanismos de regulación varían dependiendo del contexto en el cual se utiliza una lengua. Un ejemplo de esto es la activación entre lenguas que los aprendientes de una L2 pueden experimentar si se encuentran inmersos en un contexto de L1 o de L2. En el primero, la L1 afectaría la producción fonética de la L2, mientras que, en el segundo caso, no habría este efecto. Por lo tanto, en el contexto de inmersión en una L2, los aprendientes inhiben de mayor manera su L1.

Inhibición del Procesamiento Matemático Abstracto

La relación entre los problemas escritos y el bilingüismo, anteriormente descrita, y el tipo de modelamiento matemático utilizado por los participantes, ha sido discutida como una explicación posible para los resultados obtenidos por los estudiantes bilingües. Bernardo y Calleja (2005) presentaron problemas escritos a estudiantes filipinos bilingües en inglés y filipino que necesitaban tener en cuenta las restricciones del mundo real para ser resueltos satisfactoriamente. A modo de ejemplo, consideremos la siguiente oración traducida al español (Bernardo & Calleja, 2005, p. 119):

4) Steve compró cuatro tablonos de madera de 2.5 metros cada uno.
¿Cuántas planchas de 1 metro puede cortar?

La mayoría de los estudiantes responden 10 planchas, ya que $(4 \times 2.5) / 1 = 10$, sin considerar que por cada plancha de 2.5 metros solo se puede obtener 2 planchas de un metro. Por lo tanto, la respuesta correcta es 8 planchas de 1 metro en total. Los autores indicaron que muy pocos estudiantes son capaces de aplicar las restricciones del mundo real para este tipo de problemas y que las respuestas incorrectas son mayores cuando la proficiencia en el idioma del problema escrito es baja. Una explicación complementaria es el tipo de modelamiento matemático utilizado por los estudiantes, en donde no son

capaces de inhibir la información matemática para poder integrar la información lingüística.

Van Dooren e Inglis (2015) destacaron que, luego de haber sido ampliamente investigado durante la última década, el desempeño matemático se encuentra influenciado por las habilidades de la función ejecutiva, entre ellas la inhibición. Esta relación la ilustraron con los resultados de una tarea Stroop en niños de tercero básico, donde existe una relación significativa entre los puntajes obtenidos en tareas matemáticas y el control inhibitorio, siendo este hallazgo replicado en una cantidad considerable de estudios.

Star y Pollack (2015) plantearon que la inhibición puede ser útil en el procesamiento matemático, ya sea para inhibir las intuiciones o estrategias automáticas de los estudiantes o para apoyar la resolución de problemas matemáticos de mayor complejidad. En el primer caso, los estudiantes deben inhibir información perceptual o atencional como enfocarse en el área cuando la tarea solicita que se enfoquen en el perímetro de figuras geométricas. En el segundo caso, los estudiantes deben inhibir información que proviene de procesos analíticos como utilizar el conocimiento de los números naturales en situaciones no adecuadas.

Sobre este último punto, Gómez et al. (2015) indicaron el importante rol del control inhibitorio para poder resolver satisfactoriamente la tarea de decidir cuál fracción es mayor, como en $5/6$ versus $8/19$. En este ejemplo, el divisor de la primera fracción es menor que el divisor de la segunda, pero la primera fracción en su totalidad es mayor que la segunda fracción. En este estudio conductual, los estudiantes que eran capaces de inhibir el conocimiento de las magnitudes de los divisores para considerar las fracciones en su totalidad obtenían mejores resultados en la tarea, lo que también se correlacionaba con sus resultados académicos.

El Rol de la Experticia Matemática

En relación a la experticia matemática como un factor que puede modular la respuesta ERP a problemas matemáticos incongruentes, donde la solución no corresponde con la operación ($8 * 4 = 26$), Bechtold et al. (2019) sugirieron que la experiencia individual de los participantes es clave. En un estudio de ERP donde los participantes tenían que completar una tarea de decisión léxica entre palabras abstractas y pseudopalabras, los participantes expertos en matemática tuvieron un mayor porcentaje de acierto que los no expertos al enfrentarse a conceptos matemáticos abstractos.

En cuanto a las mediciones ERP, los expertos en matemática presentaron una reducción en la amplitud de N400, lo que indica una facilitación para acceder

a las características semánticas de los ítems léxicos. Asimismo, los expertos en matemática tuvieron una amplitud mayor de un componente positivo tardío parietal que indica que la experticia matemática motivó un acceso fácil a los conceptos matemáticos. Estos resultados sugieren que los expertos en matemática se vieron enfrentados a un menor esfuerzo para procesar la información semántica que los no expertos.

Modelos teóricos

En esta sección, se describirán distintos modelos teóricos que pueden dar cuenta de cómo se procesa la información matemática y lingüística para resolver problemas matemáticos escritos.

La Interfaz Sintáctica-Semántica

La interfaz sintáctica-semántica da cuenta de la estrecha relación que hay entre la semántica y la sintaxis. En términos sencillos, esta interfaz representa la interacción que hay entre la organización estructural del lenguaje, la sintaxis, y el significado que expresan las expresiones lingüísticas, la semántica. Particularmente, Hackl (2013) planteó que la meta empírica de esta interfaz es explicar la manera en que el significado de una expresión lingüística depende de sus propiedades sintácticas y la manera en que las propiedades sintácticas son

una función de su significado. Los mecanismos subyacentes de esta interfaz permiten explorar la manera en que la sintaxis y la semántica se informan y limitan mutuamente para generar lenguaje coherente.

Estudios conductuales (Brunellière & Frauenfelder, 2014; Gutiérrez, 2011; Wang, 2021) han dado cuenta de la estrecha relación que hay entre la semántica y la sintaxis, particularmente en el área de comprensión y producción del lenguaje. A modo de ejemplo, en experimentos que incluyeron oraciones con incongruencias semánticas y sintácticas, estos elicitaban tiempos de reacción mayores que las oraciones congruentes (Brunellière & Frauenfelder, 2014; Gutiérrez, 2011; Wang, 2021). Esto sugiere que la integración de información semántica y sintáctica tiene un efecto sobre el procesamiento del lenguaje. De todas maneras, es necesario considerar que estos resultados también pueden deberse a restricciones de la memoria de trabajo y la carga cognitiva en este tipo de procesamiento.

A pesar de esta estrecha relación, estudios de neuroimagen han demostrado que hay patrones de activación cerebral distintivos para procesar los componentes sintácticos y semánticos del lenguaje. Por ejemplo, el procesamiento sintáctico ha sido asociado con activación en el área de Broca. En cambio, el procesamiento semántico requiere de una red distribuida que incluye los lóbulos temporales, particularmente, el giro temporal medio izquierdo

para la extracción de información semántica, y corteza cingulada posterior para los procesos de integración semántica (Carter et al., 2019; Hagoort, 2003). Esto también se ve refrendado en estudios de ERP, donde N400 es un componente sensible a las anomalías semánticas y P600 como un componente sensible a las anomalías sintácticas (Hammer et al., 2008).

Es importante considerar esta interfaz en relación al fenómeno de desalineación semántica postulado por Bassok y colegas (1998). Ellos plantearon que los sustantivos que aparecen en problemas matemáticos tienen una tendencia a alinearse con algunas operaciones matemáticas, como la adición y la división. Particularmente, los sustantivos que son cohipónimos tienen una preferencia por la adición y son clasificados como categóricos. Mientras que los sustantivos que presentan un rol de función en relación a otros sustantivos tienen una preferencia por la división y son clasificados como funcionales.

Estas relaciones han sido conceptualizadas como semánticas en naturaleza por Bassok y colegas en su amplio corpus de investigación en el tema (1998, 2001, 2008, 2016), pero desde la perspectiva de la interfaz sintáctica-semántica se puede reinterpretar ambas relaciones. La relación categórica, donde dos sustantivos son parte de un mismo set y, por lo tanto, cohipónimos, puede ser visto como una relación semántica donde ambos sustantivos comparten una gran cantidad de características semánticas para pertenecer a un

superordinado. Este sería el caso de perros y gatos, que pertenecen al superordinado mascota. Por otra parte, la relación funcional, donde un sustantivo cumple un rol o una función en relación a otro, puede ser vista como una relación sintáctica. Este sería el caso de rosas y vasijas donde la palabra vasija cumple una función o un rol en relación a la palabra rosa, ya que las flores son contenidas en recipientes. Esto es similar a las funciones que se establecen en una oración dependiendo del verbo que se utilice y cuáles son sus argumentos válidos.

La teoría de la Carga Cognitiva

Como se ha establecido en las secciones anteriores, los problemas matemáticos escritos son complicados no sólo para los hablantes no nativos de una lengua. Estos problemas no son complicados necesariamente por las operaciones matemáticas que se deben realizar, sino que por el lenguaje necesario para poder entender los distintos componentes y sus interacciones. Estas dificultades lingüísticas pueden deberse a palabras polisémicas, con significados técnicos fuera del alcance de los hablantes, falsos cognados e incluso las relaciones semánticas de las distintas palabras. Al mismo tiempo, estas dificultades no solo son parte de las evaluaciones, ya que también son parte del proceso de enseñanza, como se ha podido establecer al no haber resultados exitosos en las adecuaciones lingüísticas estudiadas hasta ahora.

La teoría que se presentará a continuación es de corriente simbólica, como se detallará a continuación. Esta forma de cognitivismo postula que la actividad mental es simbólica, puesto que se requiere la manipulación de símbolos, patrones físicos que pueden ser combinados para formar estructuras simbólicas complejas, de la misma manera que nosotros manipulamos fonemas para pronunciar palabras (Newell & Simon, 1976). Este proceso de manipulación/combinación sigue reglas estrictas que dictan qué combinaciones son aceptables, del mismo modo que sólo ciertas combinaciones sintácticas son gramaticales. Junto con esto, las reglas son de carácter recursivo, permitiendo combinaciones más complejas aceptables. Desde esta perspectiva, el comportamiento inteligente puede ser descrito como la manipulación exitosa de estructuras simbólicas de acuerdo a un set de reglas. Finalmente, los procesos que permiten la generación y transformación de estructuras simbólicas complejas también pueden ser representadas como símbolos y estructuras simbólicas dentro del sistema (Bermúdez, 2014).

La teoría de la carga cognitiva (Cognitive Load Theory, CLT), postulada por Sweller en los 80, se presenta como una opción para explicar por qué existen dificultades para aprender y enseñar en contextos educacionales basados en la capacidad de la memoria de trabajo. De acuerdo a Paas et al. (2003), la información que debe ser procesada para aprender y aplicar ese conocimiento de manera adecuada, depende del nivel de interacción que los distintos

componentes tengan. Esta teoría propone que la información va en un continuo desde baja a alta interacción. Los elementos de baja interacción pueden ser aprendidos y comprendidos de manera individual sin la necesidad de tomar en cuenta los otros elementos. En el caso de los elementos de alta interacción, es necesario comprender todos los elementos y sus interacciones antes de poder decir que todos los elementos de alta interacción han sido comprendidos. Por lo tanto, el material que contenga estos elementos será más difícil de entender.

Paas et al. (2003) argumentaron que el procesamiento de estos tipos de interacción requiere el trabajo en conjunto de la memoria de trabajo y la de largo plazo. La memoria de trabajo solo puede procesar un número limitado de elementos. Kirschner et al. (2006) plantearon que el número máximo de elementos que pueden ser almacenados, por un tiempo limitado de 30 segundos sin que estos sean ensayados, es siete, aunque si la tarea cognitiva es exigente, solo se podrían procesar de dos a tres elementos, un número considerablemente inferior al de la mayoría de las actividades intelectuales humanas. Para poder ampliar la capacidad de procesamiento es necesario recurrir a la memoria de largo plazo. Este tipo de memoria almacena una gran cantidad de esquemas, definidos por los autores como constructos cognitivos que incorporan elementos múltiples de información para cumplir una función específica. De acuerdo a Kirschner et al. (2006), la memoria de largo plazo de los expertos en un área contiene una gran cantidad de información que nos permite identificar la situación

en la que estamos y nos indica, generalmente de manera inconsciente, qué hacer y cuándo hacerlo. Sin esta información, seríamos incapaces de cruzar la calle sin ponernos en peligro o resolver problemas matemáticos.

Para aumentar la capacidad de procesamiento de la memoria de trabajo, los esquemas pueden ser trasladados de la memoria de largo plazo a la de trabajo. Es así como la memoria de trabajo puede procesar un elemento que está compuesto de un número mayor de elementos que interactúan entre sí. Esto reduce considerablemente la carga cognitiva de la memoria de trabajo, más aún si la activación del esquema se hace de manera automática.

Para explicar cómo estas interacciones afectan a la memoria de trabajo, Sweller (1988) postuló las nociones de carga cognitiva intrínseca, carga cognitiva extrínseca o ineficaz y carga cognitiva relevante o eficaz. La carga cognitiva intrínseca se relaciona con el tipo de exigencias impuestas a la memoria de trabajo, ya que la interacción de los elementos es intrínseca al material utilizado. La manera en la que los elementos interactúan puede variar, afectando la carga cognitiva impuesta. Más aún, esta interacción no puede ser alterada por manipulaciones al momento de ser enseñadas ya que dependen completamente de las instrucciones inherentes del problema matemático a resolver, como al realizar una simple adición o resolver una ecuación diferencial.

A la vez, es necesario considerar la manera en la que la información es presentada en las actividades de enseñanza, ya que también podría crear una carga cognitiva. Cuando esta carga extra es innecesaria, esta interfiere con la automatización y aprendizaje de esquemas. Por lo tanto, se le cataloga como carga extrínseca o ineficaz. Paas et al. (2003) destacaron que muchos métodos de enseñanza convencional generan una carga extrínseca puesto que, al ser desarrollados, no se toma en cuenta la estructura de la información y/o la arquitectura cognitiva. Este tipo de carga depende completamente de cómo los profesores diseñan los materiales de enseñanza.

El último tipo de carga cognitiva también se relaciona con la manera en que se presenta la información. En este caso y de manera contraria a la carga extrínseca, la carga relevante o eficaz, potencia el aprendizaje al dedicar los recursos de la memoria de trabajo al aprendizaje, al segmentar la información relevante y automatizar esquemas para liberar recursos cognitivos.

Estos tres tipos de carga cognitiva tienen un carácter aditivo y no pueden superar los recursos de la memoria de trabajo para poder lograr el aprendizaje de nueva información. A pesar de tener este carácter aditivo, hay una relación asimétrica entre los tipos de carga cognitiva. En primer lugar, la carga intrínseca crea una carga base que no puede ser reducida a no ser que se creen esquemas adicionales que reduzcan el número de elementos que están interactuando. Si

luego de procesar esta carga base quedan recursos de la memoria de trabajo disponibles, estos pueden ser dedicados para procesar los dos tipos de carga cognitiva restantes. Si el diseño de los materiales reduce la carga extrínseca, quedan mayores recursos disponibles para procesar la carga relevante y, en consecuencia, hacer el proceso de aprendizaje más expedito. A medida que se van aprendiendo y automatizando más esquemas, la carga intrínseca es reducida, liberando más capacidad de la memoria de trabajo aún. Finalmente, estos nuevos recursos permiten aprender esquemas más avanzados.

En el caso de los experimentos utilizados en este estudio, la carga intrínseca estaría expresada en la interacción entre números, sustantivos y operaciones matemáticas. Al incluir divisiones sencillas que dan como resultado números enteros, se espera que la carga intrínseca impuesta por la operación matemática sea baja. Como el foco está en las divisiones, se espera que haya un aumento de la carga extrínseca en los casos en que los sustantivos se encuentren en una relación categórica, ya que el esquema de las relaciones semánticas categóricas se asocia con la adición y no con la división, siendo expresada esta carga en tiempos de reacción y/o tasas de error mayores. Junto con esto, el inhibir este tipo de información incongruente generaría una carga extrínseca mayor. En los casos en que haya una relación funcional entre los sustantivos, se espera que haya una carga cognitiva menor, expresada en tiempos de reacción menores y/o tasas de acierto mayores. Esto ocurriría ya que

se automatizan esquemas debido a la congruencia semántica de los sustantivos utilizados, creando una carga cognitiva eficaz que facilitaría la resolución del problema matemático escrito.

La sobrecarga cognitiva ocurre cuando los elementos que se deben mantener activados en la memoria de trabajo sobrepasan la capacidad de esta. En el caso particular de los problemas escritos, los participantes deben mantener en mente la estrategia que utilizará para resolver el problema y cómo se relacionan los distintos elementos presentados, mientras se resuelve el problema en cuestión (Campbell et al., 2007). Para estudiar este fenómeno, estos autores estudiaron cómo se resuelven problemas escritos de pruebas estandarizadas para proponer un esquema que refleje el tipo de conocimiento que se debe activar y cómo estos interactúan para resolver los problemas escritos.

Para ilustrar esto, consideremos el siguiente problema²:

La familia de Sandy lava su ropa en una lavandería cuyos equipos funcionan a base de monedas. Cada carga en la lavadora cuesta 1,25 US\$, mientras que cada carga en la secadora cuesta 25 centavos por 10 minutos. La familia de Sandy tiene que hacer 5 cargas de ropa en la lavadora y cada carga debe estar 30 minutos en la secadora. ¿Cuál expresión matemática le dará a la familia de Sandy el costo total de hacer estas cargas de lavado?

² Traducción al español realizada por el autor de esta tesis

- A. $(1,25 \text{ US\$} + 0,25 \text{ US\$}) \times 3 \times 5$
- B. $[1,25 \text{ US\$} + (3 \times 0,25 \text{ US\$})] \times 5$
- C. $[(3 \times 1,25 \text{ US\$}) + 0,25 \text{ US\$}] \times 5$
- D. $3 \times (1,25 \text{ US\$} + 0,25 \text{ US\$}) \times 5$

A pesar de que las operaciones matemáticas involucradas son sencillas (adición y multiplicación), solo la mitad de los 70.000 estudiantes que rindió la prueba de la cual este problema era parte fueron capaces de escoger la respuesta correcta (opción B).

Para resolver exitosamente este problema, es necesario que experiencias de vida, el lenguaje, procesos cognitivos y el conocimiento y la habilidad para aplicar el contenido matemático interactúen. Para ilustrar esta relación, ellos proponen el siguiente esquema (Campbell et al., 2007, p. 9):

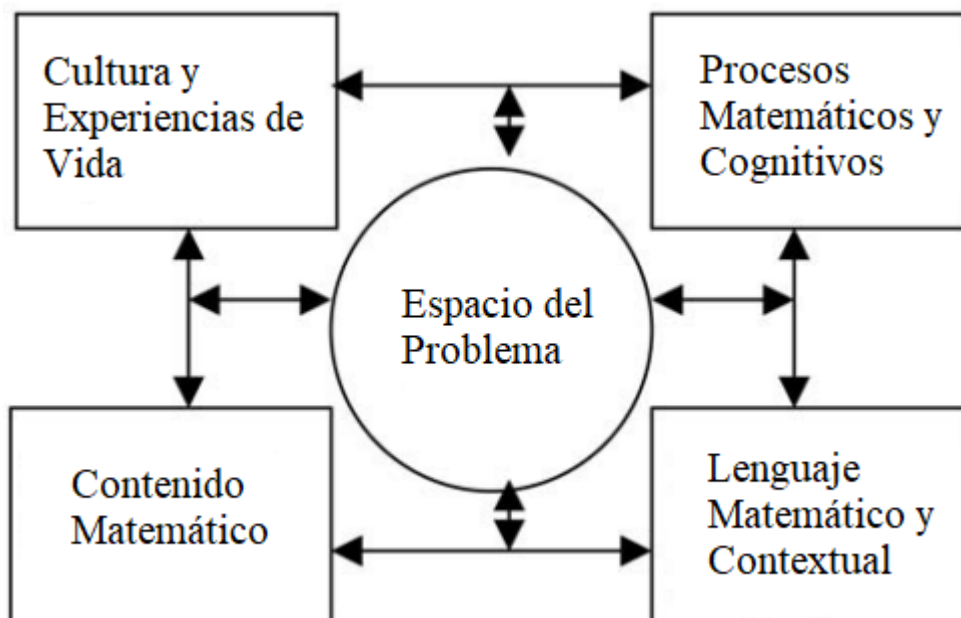


Figura 2 Interacciones meta-reflexivas en el espacio del problema (Campbell et al., 2007, p. 9)³.

Al resolver un problema, se parte en la esquina superior derecha, al identificar la complejidad de las demandas cognitivas asociadas al problema escrito la primera vez que se lee. Luego de eso, las experiencias previas con situaciones de evaluación interactúan con la conciencia de los procesos cognitivos y el lenguaje de las estructuras matemáticas. Esta interacción influye en la decisión de manejar la carga cognitiva. Finalmente, el contenido matemático es procesado de manera automática, ya que las experiencias de vida y la comprensión del lenguaje se utilizan para entender qué es lo que se debe hacer con esta información.

³ Traducción al español realizada por el autor de esta tesis

Campbell y colegas (2007) plantearon que los problemas matemáticos escritos pueden generar mayores demandas cognitivas principalmente enfocadas en el lenguaje utilizado. Por ejemplo, la palabra load (carga) en inglés se aprende inicialmente como verbo y su uso como sustantivo podría activar otro tipo de significados. Junto con eso, el confundir qué clase de palabra es load también podría generar problemas de índole sintáctico, ya que se generaría la expectativa de que a continuación vendría una frase nominal con la función de objeto directo (Campbell et al., 2007).

Esto también podría ocurrir para expression (expresión matemática) que podría ser interpretado como expresión facial. Estos dos ejemplos se relacionan con los efectos de desalineación semántica postulados por Bassok et al. (1998), ya que el conocimiento de mundo de los participantes podría generar dificultades al solucionar problemas matemáticos desde la perspectiva del lenguaje. Por lo tanto, los estudiantes, nativos y no nativos del inglés, podrían entender los conceptos matemáticos y cómo resolver el problema, pero el uso de lenguaje poco familiar podría interferir en la resolución exitosa y en la manera en la que los estudiantes son evaluados.

El Modelo de Activación Interactiva Bilingüe

El modelo que se discutirá a continuación es de corriente conexionista. La primera característica de un modelo conexionista es que la información se encuentra distribuida a través de los distintos pesos de las conexiones entre neuronas, es decir, un valor numérico positivo o negativo que representan la fuerza excitatoria o inhibitoria de la conexión (Schmidhuber, 2015).

La segunda característica es que no existen reglas o símbolos distintivos como en el caso de la corriente simbólica. Es decir, no existen reglas que permitan la manipulación de símbolos, a excepción de la función de activación y el algoritmo de aprendizaje (McClelland & Clereemans, 2009). En este caso, la función de activación corresponde al valor que se debe obtener para que una neurona artificial dispare. Mientras que el algoritmo de aprendizaje corresponde a las adecuaciones numéricas que los pesos deben realizar para cumplir con su función excitatoria o inhibitoria a través de la exposición repetitiva de estímulos (McClelland & Clereemans, 2009).

La tercera característica de este modelo es su capacidad de modificar el peso de sus conexiones de acuerdo a los inputs recibidos. Este tipo de aprendizaje puede ser supervisado, puesto que la red neuronal recibe retroalimentación sobre las diferencias entre el output producido y el output

esperado, pero también puede ser no supervisado, en este caso se utiliza un algoritmo de clasificación donde una unidad de output solo dispara ante ciertos inputs (Schmidhuber, 2015). En este tipo de red neuronal, existen conexiones inhibitorias entre las unidades de output, lo que establece una especie de competencia entre unidades de output para que solo una gane ante un input específico (Schmidhuber, 2015). Este último tipo de red neuronal comparte características con el modelo que se presentará a continuación.

Grainger et al. (2010) plantean que cuando se trata de explicar la comprensión bilingüe hay dos grandes propuestas. La primera propuesta, la hipótesis selectiva del lenguaje, establece que no existe una interferencia entre lenguas si el contexto que se está tratando de comprender es completamente monolingüe (Macnamara, 1967). Es así como la información proveniente del estímulo se asociaría directamente a las representaciones lingüísticas de esa lengua en particular, sin haber activación de la otra lengua. Por el contrario, la segunda propuesta, la hipótesis de acceso no selectivo, establece que la información proveniente del estímulo puede activar representaciones lingüísticas de ambos idiomas, debido a posibles solapes ortográficos o fonológicos (Grainger & Dijkstra, 1992). Esta propuesta en particular es la que sirve como cimiento para el modelo de activación interactiva bilingüe, propuesto por Grainger y colegas (2010).

En este modelo, la selección de una lengua opera de manera descendente, al amplificar el procesamiento de las representaciones lingüística de una lengua y/o al inhibir las representaciones lingüísticas de la otra lengua. La arquitectura de este modelo es ilustrada en la siguiente figura (Grainger et al., 2010, p. 269).

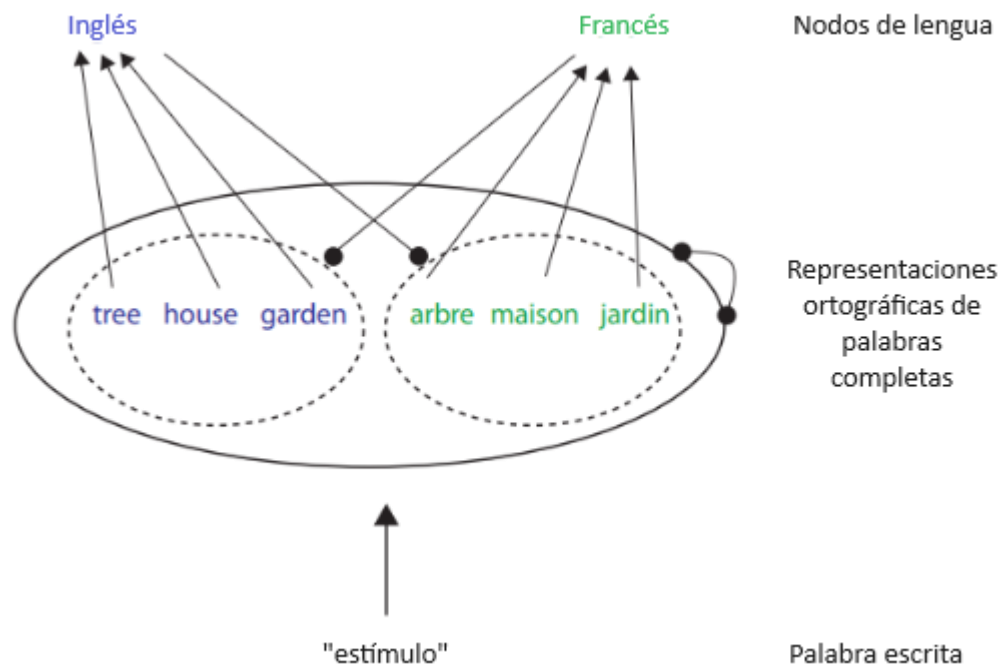


Figura 3 Arquitectura del modelo de activación interactiva bilingüe (Grainger et al., 2010, p. 269)⁴.

En esta figura se puede apreciar cómo los nodos de cada lengua integran información del estímulo visual y contextual para determinar la probabilidad de

⁴ Traducción al español realizada por el autor de esta tesis

que una palabra pertenezca a una lengua en particular. La activación de los nodos es lo que permite determinar cuál idioma se está procesando. Por lo tanto, al momento de procesar la palabra y activar el nodo de lenguaje con mayor activación, sólo las representaciones lingüísticas de una lengua se mantendrán activadas. Es así como este modelo propone una selección tardía a través de un control descendente desde los nodos de cada lengua hasta las representaciones al nivel de la palabra impresa.

Debido a las características de esta propuesta, hay dos consecuencias importantes para el procesamiento léxico bilingüe. En primer lugar, es imposible apagar por completo la lengua no seleccionada, puesto que la alimentación de información activa nodos de lengua que son independientes entre sí. Por lo tanto, siempre ambas lenguas estarán activadas, aunque en distintos grados. En segundo lugar, las representaciones coactivadas por la lengua no-objetivo son parte de la interferencia generada por las representaciones lingüísticas de las palabras no-objetivo. El daño de esta interferencia entre lenguas es limitado gracias al control inhibitorio descendente de los nodos de cada lengua. A continuación, se presentará evidencia para fundamentar lo planteado anteriormente.

Grainger et al. (2010) plantearon que existe evidencia conductual y de electrofisiología que sugiere la existencia de interferencia entre lenguas. Desde

los experimentos conductuales, el uso de palabras de la lengua no meta, ortográficamente similares a las palabras objetivo, tuvo efectos significativos en los tiempos de reacción y tasas de acierto. Para estos resultados, los autores proponen que este efecto de vecindad entre lenguas sugiere que, al procesar una palabra, se activan palabras ortográficamente similares no solo dentro de la lengua objetivo, sino que también en la otra lengua. Desde los experimentos de electrofisiología, los vecinos entre lenguas generaron ondas ERP más negativas en la ventana del componente N400 que las palabras que no tenían vecinos entre lenguas. Esta evidencia sugiere que el procesamiento bilingüe no es de acceso selectivo, aunque sí existe una convergencia rápida de la representación léxica apropiada en la lengua objetivo, debido a la operación conjunta de mecanismos de inhibición descendente y lateral. La inhibición lateral se ve representada en los efectos de interferencia provocados por los vecinos lingüísticos, mencionados anteriormente.

El caso de la inhibición descendente se ve reflejado en los efectos de cambiar idiomas. Por ejemplo, en una tarea de decisión léxica donde los participantes debían decidir si el estímulo presentado era una palabra, independiente de la lengua en que fuese presentada. De esta manera, se podría creer que la pertenencia de una palabra a una L1 o L2 podría ser ignorada. A pesar de esto, los resultados indican que los tiempos de decisión eran más lentos en una lengua si la palabra anterior pertenecía a una lengua distinta, siendo estos

costos lingüísticos mayores para las palabras objetivo en la L2. Grainger y colegas (2010) indicaron que estos resultados han sido replicados posteriormente en otros estudios, lo que indica que existe una activación ascendente al ser expuesto al estímulo escrito. Dichos estímulos activan los nodos de cada lengua, que luego inhiben de manera descendente las representaciones léxicas.

Lo descrito anteriormente, ilustra los principios básicos del modelo de activación interactiva bilingüe de Grainger y colegas (2010), que está fundamentado en evidencia empírica. Aun así, los autores destacaron la naturaleza estática de este modelo, por lo que no sería aplicable en aprendientes de una L2 que aún no alcanzan un alto nivel de proficiencia. Por ese motivo, ellos proponen el modelo de activación interactiva bilingüe del desarrollo, particularmente para aprendientes tardíos de una L2.

El Modelo de Activación Interactiva Bilingüe del Desarrollo

Este modelo en particular se enfoca en el aprendizaje de vocabulario por parte de aprendientes tardíos de una L2. Junto con este énfasis, el modelo se enfoca en representar de manera simplificada las formas visuales/ortográficas de las palabras y sus significados. La siguiente figura ilustra el proceso de aprendizaje de una palabra en la L2 (Grainger et al., 2010, p.277).

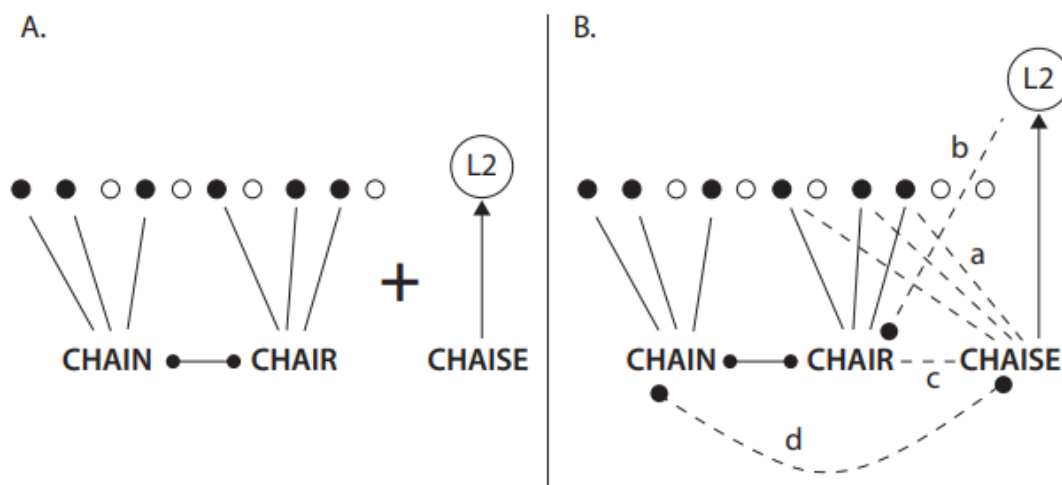


Figura 4 El modelo de activación interactiva bilingüe del desarrollo para el aprendizaje de vocabulario en una L2⁵.

En una primera etapa, el significado de la palabra nueva en la L2 (chaise) es anclado al significado de una palabra en la L1 (chair), al ser una traducción. Esto genera una coactivación de la palabra en la L2 y su traducción en la L1, junto con la representación semántica correspondiente. Desde una perspectiva conexionista, estas conexiones entre la palabra en la L2, su traducción a la L1 y su nodo de lengua se verían reforzadas por una actualización Hebbiana. Es así como la etapa inicial del aprendizaje de vocabulario en una L2 involucraría la presentación de la palabra en la L2, su significado anclado a su correspondiente traducción en la L1 y la información que indicaría que esta nueva palabra

⁵ Los círculos negros corresponden a conexiones inhibitorias, mientras que los círculos blancos corresponden a conexiones excitatorias.

pertenece a una L2. Las coactivaciones que son compatibles con esta información se verían reforzadas por el aprendizaje Hebbiano.

En una segunda etapa, cuando una palabra de la L2 (chaise) llega a un punto crítico de activación sin anclarse a la traducción en la L1 (chair), el proceso de anclaje deja de utilizarse. Luego de esto, las conexiones entre la palabra en la L2, las características semánticas y el nodo de la L2 siguen siendo reforzados a través del aprendizaje Hebbiano. Este cambio de foco hacia la autonomía en la L2 se ve reforzado por la inhibición descendente del nodo de la L2 hacia la traducción en la L1. Junto con esto, la conexión excitatoria entre las representaciones lingüísticas entre la L1 y la L2 decrecen. Una vez que todos estos pasos han ocurrido, las conexiones excitatorias entre las representaciones léxicas de la L2 y la semántica se incrementan, las conexiones inhibitorias entre el nodo de la L2 y las palabras de la L1 crecen y luego decrecen una vez que el proceso de anclaje deja de ser utilizado y se desarrollan las conexiones inhibitorias de la palabra en la L2 hacia otras palabras ortográficamente similares en la L1 y L2. Como último punto, los autores destacan que la integración gradual de las palabras en la L2 a un lexicón integrado se ve acompañado de un aumento en la interferencia entre lenguas.

Finalmente, este modelo está basado en la arquitectura de otros modelos de redes neuronales de comprensión y producción de palabras, donde las

palabras están conectadas a representaciones semánticas distribuidas. Cada palabra se encuentra unida por conexiones excitatorias mutuas a ciertas características semánticas. Mientras que las palabras que son coactivadas por el mismo estímulo y que no comparten características semánticas tienen conexiones inhibitorias mutuas, siguiendo los principios de las redes neuronales winner-take-all, donde la palabra con mayor activación es la que termina siendo activada y seleccionada en desmedro de sus competidores.

Este modelo se puede aplicar de manera particular a la desalineación semántica. En primer lugar, los sustantivos que son parte del problema matemático pueden activar expectativas sobre la operación matemática que se debe resolver. Guthormsen et al. (2016) ya han demostrado cómo la incongruencia entre expectativas generadas por los sustantivos en el problema matemático genera un componente N400, particularmente en los casos en que la presencia de cohipónimos activa la adición, pero se debe resolver una división. O cuando la relación asimétrica entre sustantivos sugiere que uno cumple una función en relación a otro, activando la división, pero se debe resolver una adición. En ambos casos, las activaciones ascendentes generadas por los sustantivos de los problemas matemáticos activan a su vez el tipo de operación matemática a realizar. En algunos casos, como en la desalineación semántica, la activación de la operación matemática debe ser inhibida para resolver el problema matemático de manera satisfactoria.

El Modelo de Congruencia Semántica (SECO)

Como ya se discutió con anterioridad en este marco teórico, la resolución de problemas matemáticos escritos requiere un procesamiento que va más allá de lo meramente procedimental (Gros et al., 2020). Por ejemplo, determinar la diferencia de número en el siguiente problema matemático, hay 5 aves y 3 gusanos, ¿Cuántos pájaros hay más que gusanos? Es mucho más difícil para niños menores de cinco años que responder a la pregunta ¿Cuántas aves no comerán un gusano? Esta dificultad no se puede explicar desde lo procedimental, ya que la resolución de este problema matemático requiere utilizar la misma operación matemática (5 menos 3 es igual a 2) (Gros et al., 2020). Junto con eso, ellos mencionan que la elección entre dos algoritmos para solucionar el mismo problema matemático depende de los costos cognitivos asociados a cada estrategia.

Por lo tanto, Gros y colegas (2020) plantearon el modelo de congruencia semántica (SECO), que es un modelo teórico que representa cómo el conocimiento de mundo (representada en la semántica de una lengua) y la semántica matemática interactúan para codificar, recodificar y resolver problemas matemáticos escritos. Este modelo cuenta con seis componentes que pueden interactuar entre sí en distintas etapas para resolver un problema matemático de manera exitosa.

La siguiente figura representa visualmente estos componentes.

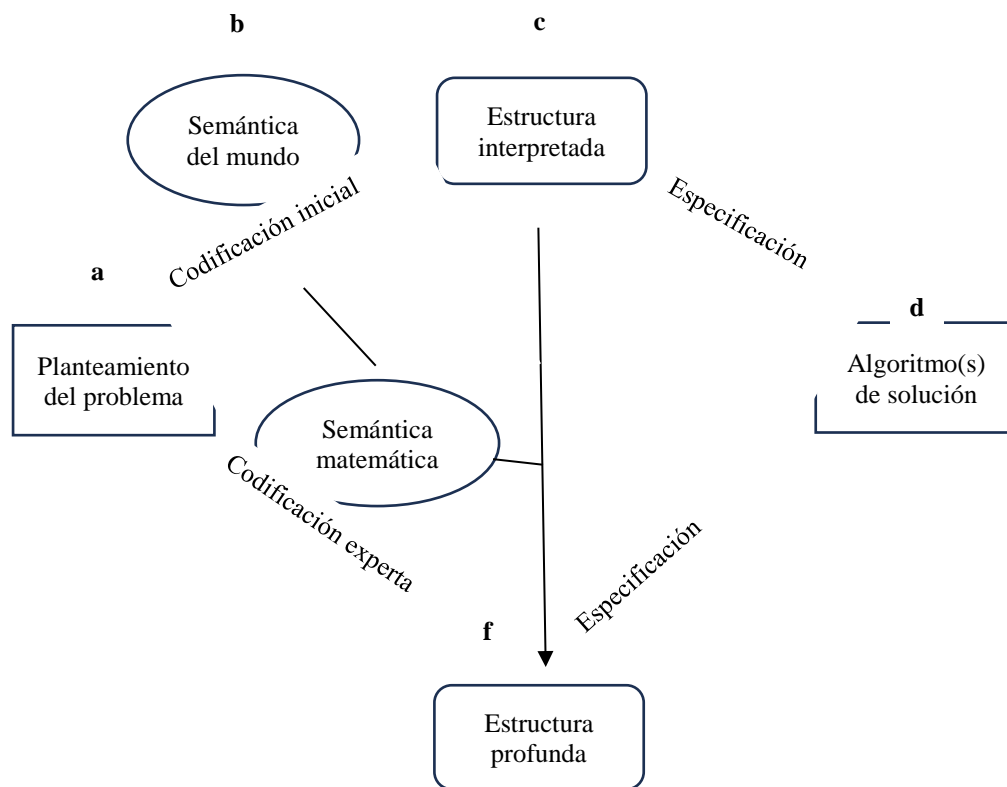


Figura 5 El modelo SECO (Gros et al., 2020, p. 5)⁶.

En primer lugar, el componente del planteamiento del problema es un texto que describe los elementos del problema, la situación dónde y cómo interactúan y sus valores asociados. En segundo lugar, el componente de la semántica del mundo representa el conocimiento de la vida cotidiana que el individuo tiene sobre los elementos presentes en el planteamiento del problema. Esto incluye la

⁶ Traducción al español realizada por el autor de esta tesis

manera en que estos distintos elementos se relacionan entre sí en el día a día del individuo. En tercer lugar, el componente de la semántica matemática hace referencia al conocimiento matemático del individuo que se puede aplicar al problema matemático. En cuarto lugar, el componente de la estructura interpretada se abstrae del planteamiento del problema, donde se integra la información presentada en el texto junto con las propiedades, asociaciones y restricciones de los elementos en relación a la semántica del mundo. En quinto lugar, el componente de algoritmo(s) de solución representa los pasos que se deben seguir para resolver de manera correcta el problema matemático. Finalmente, el componente de la estructura profunda se refiere a la integración de los elementos relevantes del problema matemático y sus relaciones enfocándose exclusivamente en la semántica de la matemática.

Estos seis componentes pueden interactuar en cuatro procesos para resolver el problema matemático. Estos procesos se describen en el orden que aparecen en la imagen, pero, como Gros et al. (2020) plantearon, estos procesos pueden ocurrir en paralelo. El primer proceso es la *codificación inicial*. Este proceso describe cómo el planteamiento del problema se abstrae a una estructura interpretada que depende de la semántica del mundo y matemática evocadas por la redacción del planteamiento del problema. En este proceso, la semántica del mundo limita la representación de la situación descrita al destacar o minimizar las relaciones de los elementos presentes en el problema. A la vez,

la semántica matemática da forma a las relaciones matemáticas que evoca el planteamiento del problema.

El segundo proceso es la *especificación*. Este proceso describe cómo la estructura interpretada puede materializarse en un algoritmo para resolver el problema. Si las relaciones y los valores representados tienen un significado matemático claro, esto se puede traducir en operaciones matemáticas relevantes. Gros y colegas (2020) destacaron que no toda estructura interpretada lleva a un algoritmo adecuado, ya que parte de la información que se codifica podría ser irrelevante. Por el contrario, la estructura profunda sí especifica los algoritmos relevantes para resolver el problema matemático, dado que está libre de la influencia de la semántica del mundo.

El tercer proceso es la *codificación experta*. Este proceso describe lo que hipotéticamente ocurre cuando un experto codifica un problema matemático dentro de su dominio de experticia. Particularmente, un experto podría usar su conocimiento de la semántica matemática para hacer caso omiso de la semántica del mundo y abstraer una estructura profunda sin considerar una estructura interpretada.

Finalmente, el cuarto proceso es la *recodificación*. En los casos en que la estructura interpretada no permite llegar a un algoritmo relevante y no ocurre la

codificación experta, se puede abstraer una nueva representación del problema al enfocarse en la semántica matemática para crear una nueva estructura que sea similar a la estructura profunda y que permita llegar al algoritmo necesario. Los autores plantean que este proceso es cognitivamente costoso y que no ocurre de manera sistemática.

Estos componentes y procesos se integran de la siguiente manera. Al leer un planteamiento del problema (a), los individuos en un principio codificarán una estructura interpretada (c) considerando la semántica del mundo (b) y la semántica matemática (e) evocada por la redacción del problema matemático. Esta estructura interpretada es idiosincrática, ya que dependerá de la semántica del mundo y matemática del individuo. Esta estructura puede derivar en un algoritmo (d). En los casos en que no se llega a un algoritmo exitoso, ocurre un proceso de recodificación que se enfoca en la semántica matemática (e) para poder llegar a la estructura profunda del problema (f) y a un nuevo algoritmo de solución. Por último, el modelo considera la posibilidad de que un experto acceda directamente a la estructura profunda (f) a partir del planteamiento del problema (a) sin extraer una estructura interpretada (c) a partir de la semántica del mundo (b).

Algunos componentes y procesos de este modelo se alinean con el modelo de Construcción-Integración de Kintsch (1998) y particularmente con lo

que él planteó para los problemas matemáticos escritos. Por ejemplo, a partir de los distintos elementos del problema matemático, se crea una representación lingüística que depende de la semántica del mundo para generar una estructura interpretada que podría llevar a un algoritmo de solución. Estos componentes del modelo SECO son similares a los planteados por Kintsch para el proceso de construcción e integración al momento de comprender un texto, ya que, desde las palabras, se genera una representación lingüística que se convierte en una red proposicional que puede ser influenciada por la experiencia del lector para finalizar en un proceso de integración que genera una representación textual.

Una diferencia es que las simulaciones que discutió Kintsch (1998) para resolver problemas matemáticos indican que un fallo en la representación lingüística generaría graves problemas para llegar a la representación textual necesaria para resolver el problema en cuestión, mientras que Gros et al. (2020) plantearon que los expertos en matemáticas podían evitar la información lingüística y enfocarse solo en la semántica de la matemática para resolver el problema.

Para ejemplificar el modelo SECO, consideremos el siguiente problema matemático. Para una entrega de premios, un florista prepara para cada uno de los 14 candidatos cinco rosas y siete tulipanes ¿Cuántas flores en total utiliza el florista? (Gros et al., 2020). Este problema en particular derivaría en dos posibles algoritmos (expansión o factorización) dependiendo de la semántica del mundo y

la semántica de la matemática de la persona que resuelve este problema. En relación al primer algoritmo posible, debido a la redacción de este problema (a), la semántica del mundo (b) codifica una representación donde rosas y tulipanes son sets separados de elementos. Esto lleva a una estructura interpretada (c) que produce un algoritmo de expansión (d), donde para resolver este problema, primero se multiplica cada elemento por separado y los resultados de estas multiplicaciones se suman (i.e. $(14 \times 5) + (14 \times 7)$). En relación al segundo algoritmo posible, el individuo debería inferir que rosas y tulipanes pueden agruparse bajo la entidad flores, lo que implicaría un uso de la semántica del mundo (b) distinta al primer algoritmo. En este segundo algoritmo, se puede recodificar la estructura interpretada (c) solamente considerando la semántica matemática (e) para llegar a una estructura profunda (f) que permite agrupar estos elementos, donde se suman primero el número de rosas y tulipanes para luego multiplicar ese valor por los 14 candidatos (i.e. $14 \times (5 + 7)$).

En relación a la desalineación semántica, los individuos interpretarían la estructura de los problemas de acuerdo a su semántica del mundo. Es decir, los cohipónimos, como rosas y tulipanes, tienen una tendencia mayor a ser parte de adiciones, mientras que las rosas y los ramos tienen una tendencia mayor a ser parte de divisiones. Por lo tanto, para resolver exitosamente un problema matemático donde ocurre desalineación semántica sería necesario recurrir a la semántica matemática para identificar la operación matemática a realizar, sin

importar los sustantivos que aparezcan, para poder llegar al algoritmo de solución apropiado.

El modelo de Gros y colegas (2020) es capaz de dar cuenta de los resultados obtenidos en investigaciones previas para demostrar la influencia que tiene el conocimiento de mundo sobre las representaciones y estrategias de solución empleadas por los aprendientes, junto con las dificultades que aparecen entre la desalineación semántica y los algoritmos y representaciones matemáticas (Gros et al., 2020).

La Eficiencia de las Adecuaciones Lingüísticas y Estrategias para Enfrentar los Problemas Matemáticos Escritos

Como se ilustró anteriormente, hay un vasto conjunto de estudios conductuales que indican una relación robusta entre el nivel de proficiencia en el idioma de los problemas escritos y los puntajes obtenidos por los participantes. Junto con estas investigaciones, el foco ha estado en identificar las posibles adecuaciones lingüísticas que se pueden realizar a las pruebas para mejorar el desempeño de los hablantes no nativos y las estrategias que podrían ayudar a enfrentar de mejor manera los problemas escritos. Por lo tanto, en esta subsección se describirán los resultados principales que se han obtenido desde 2009 en adelante.

Bautista et al. (2009) evaluaron si el modo de presentación de los problemas matemáticos, escrito o leído en voz alta, podría mejorar los resultados de los estudiantes que tenían una menor proficiencia del idioma del inglés en Filipinas. Aunque los estudiantes obtuvieron un puntaje más alto en el modo escrito de la prueba, esta diferencia no fue significativa, lo que descarta el rol del modo de evaluación y lo centra en el dominio del idioma.

Driver y Powell (2017) decidieron centrar su atención en estrategias que tuvieron en cuenta los grupos cultural y lingüísticamente diversos de aprendientes de inglés mientras que también se centraron en los estudiantes con posibles dificultades matemáticas. Ellas observan que pocos estudios se han centrado en la instrucción, prestando principalmente atención a las puntuaciones y posibles procesos relacionados con la resolución de este tipo de problemas. Por lo tanto, implementaron una intervención que pone como énfasis principal la comprensión de la estructura del problema, conocida como instrucción basada en esquemas. Al brindarles tutorías a 9 estudiantes no nativos con dificultades matemáticas, durante 10 semanas, tres veces a la semana durante 20 a 25 minutos cada sesión, fueron capaces de mejorar las puntuaciones de los participantes después de analizar las pruebas previas y posteriores. Esta mejora fue significativa y tuvo un gran efecto (η^2 parcial = ,79). También aplicaron un cuestionario para evaluar la tutoría, la cual recibió comentarios positivos que expresaron satisfacción por parte de los estudiantes. Desde la

preprueba a la post prueba, los participantes pasaron del grupo de riesgo de dificultades matemáticas a lograr resultados similares a los de los estudiantes nativos sin dificultades matemáticas, en comparación con una muestra normativa.

Orosco y Abdulrahim (2018) señalaron la necesidad de llevar a cabo intervenciones para ayudar a los aprendientes de inglés cuando se enfrentan a problemas escritos, para evitar generarles una dificultad de aprendizaje matemático. En este sentido, también expresaron la necesidad de centrarse en el desarrollo profesional de los docentes en formación y proporcionar las herramientas adecuadas para hacer frente a este problema. En particular, propusieron el uso de la enseñanza de estrategias de comprensión (comprehension strategy instruction, CSI) como un medio para mejorar la resolución de problemas escritos en este grupo particular de estudiantes. CSI se presenta como una manera de generar andamios instructivos para mejorar la comprensión y los puntajes de los estudiantes en problemas matemáticos escritos. Durante diez semanas, los estudiantes que fueron identificados como aprendientes de inglés con dificultades matemáticas, recibieron tutorías dos veces por semana. Las puntuaciones en las pre y post pruebas indican que el grupo experimental aumentó sus puntuaciones de manera significativa con un tamaño de efecto pequeño ($g=.37$).

Estos resultados sugieren que la formación del profesorado podría ser la clave para reducir la brecha entre hablantes nativos y aprendientes de inglés, al centrarse particularmente en el lenguaje relevante que se utilizará y se enseñará durante las lecciones y las estrategias de comprensión para los estudiantes.

Metaanálisis sobre la Resolución de Problemas Matemáticos Escritos y el Lenguaje de las Pruebas

Dado que la relación entre la resolución de problemas escritos y el lenguaje ha mostrado resultados sólidos en las últimas décadas, la investigación se ha centrado en las diferentes adaptaciones que se pueden hacer a las pruebas para que sean justas para los aprendientes de inglés, sin afectar el constructo que están midiendo. Hasta la fecha, dos reseñas narrativas (Abedi et al., 2004; Sireci et al., 2003), dos revisiones sistemáticas (Acosta et al., 2008; Baker et al., 2016) y varios metaanálisis (Kieffer et al., 2009; Kieffer et al., 2012; Pennock-Roman & Rivera, 2011; Ríos et al., 2020) se han centrado en resumir y analizar los efectos de las adaptaciones de idioma en las pruebas estandarizadas, particularmente en ciencias, matemáticas, artes e inglés. Estas revisiones y análisis han considerado como un hecho la relación entre el dominio del idioma y la resolución de problemas escritos en los aprendientes de inglés, y se han centrado en formas de reducir esta brecha. Ese objetivo en sí mismo es importante, teniendo en cuenta que para 2030, se proyecta que los aprendientes

de inglés en Estados Unidos representarán el 40% de todos los estudiantes en edad escolar (Powell et al., 2020).

Kieffer y colegas (2009) comentaron los resultados de un metaanálisis con respecto a la diferencia en las puntuaciones entre los aprendientes de inglés y los hablantes nativos de inglés en los Estados Unidos. En relación con las matemáticas, señalan que el tamaño medio del efecto en las evaluaciones no nacionales es de $g = ,604$, mientras que en las evaluaciones nacionales el tamaño medio era $d = ,831$ para los estudiantes de cuarto básico y $d = 1,006$ para los estudiantes de octavo básico. Ellos proponen la noción de que otras variables podrían explicar esta diferencia, como la clase social y la calidad de las escuelas. Sin embargo, informan de un tamaño de efecto intermedio para la diferencia de puntuación en favor de los hablantes nativos en condiciones experimentales.

En relación con los metaanálisis sobre adaptaciones del lenguaje, sorprendentemente, los investigadores han encontrado una eficacia limitada para mejorar las puntuaciones en los aprendientes de inglés (Kieffer et al., 2009) o ninguna eficacia "statistically different from zero" (Ríos et al., 2020, p. 72), lo que indica claramente que las medidas adoptadas para nivelar el campo para los aprendientes de inglés no han funcionado. Más aún, Ockey (2007) ha sugerido que las pruebas matemáticas estandarizadas son justas al evaluar las habilidades aritméticas porque los reducidos sesgos encontrados pueden haber

sido debido a la casualidad o defectos metodológicos con la muestra. Además, investigaciones recientes han encontrado resultados contradictorios relacionados con el papel de la comprensión lectora al resolver con éxito problemas de palabras (Pavón & Cabezuelo, 2019; Trakulphadetkrai et al., 2020) con asignaturas que reciben instrucción en inglés en España y el Reino Unido, respectivamente. Estas conclusiones sugieren que los hallazgos fundamentales de estas investigaciones podrían haber pasado por alto otros elementos clave.

Por último, Ríos et al. (2020, p. 73) han señalado recientemente que centrarse en poblaciones particulares puede mejorar el conocimiento que tenemos con respecto a la eficacia de las adaptaciones del lenguaje para los aprendientes de inglés porque es necesario "to recognize EL heterogeneity and begin to study specific EL subpopulations (...) by collecting large sample sizes that account for idiosyncrasies". Esto pone de relieve la necesidad de tener un conocimiento más contextualizado sobre los efectos que el lenguaje de instrucción puede tener en los estudiantes que tienen una lengua materna diferente.

Metaanálisis del Rol del Lenguaje en la Resolución de Problemas Matemáticos Escritos

Aunque la relación entre el dominio del lenguaje y la resolución de problemas escritos se ha estudiado extensamente, las primeras investigaciones sobre este tema se centraron principalmente en dos dimensiones del problema que deben revisarse. En primer lugar, la atención se centró en informar diferencias significativas entre los aprendientes de inglés y las puntuaciones de los hablantes nativos y no en el tamaño del efecto. Este foco se debe a que se podría llevar a conclusiones engañosas al considerar solamente el valor p , ya que, con un número suficientemente alto de participantes, se puede obtener significancia incluso cuando las diferencias entre grupos experimentales son menores (Field & Gillet, 2010). En segundo lugar, los estudios se llevaron a cabo principalmente en Estados Unidos, con el inglés como el idioma de instrucción y evaluación, lo que deja de lado la investigación realizada en el resto del mundo con otros idiomas de instrucción, lo que podría crear una imagen incompleta del problema. Esta situación ya ha sido indicada como un posible sesgo por Banks et al. (2016).

Por lo tanto, para finalizar la discusión teórica de esta tesis doctoral, se realizó un metaanálisis, con el propósito de determinar el tamaño de efecto promedio de los estudios que se han centrado en documentar la diferencia en las puntuaciones en la resolución de problemas escritos entre hablantes nativos y

hablantes no nativos, tratando de identificar variables moderadoras que deben ser consideradas en intervenciones y adaptaciones. En consecuencia, también se incluyeron estudios que han utilizado un idioma diferente al inglés para probar las diferencias.

A continuación, se presentarán las estrategias de búsqueda de artículos, los criterios de elegibilidad y el procedimiento de análisis de los artículos para obtener el tamaño del efecto y posibles variables moderadoras. Luego, se presentarán los resultados principales y la discusión de estos.

Estrategia de Búsqueda

Se siguieron dos estrategias principales de búsqueda para identificar estudios apropiados para este metaanálisis. En primer lugar, se buscaron dos bases de datos: Google Scholar y ERIC. Los términos de búsqueda utilizados fueron (math OR mathematical OR mathematics OR arithmetic) AND "word problem" AND (bilingualism OR bilingual OR language). Los resultados se limitaron a estudios entre 2001 y 2020. Tanto los términos de búsqueda como el intervalo de fechas son amplios para incluir tantos estudios como sea posible, particularmente fuera de los Estados Unidos. La búsqueda de Google Scholar obtuvo 2630 resultados, mientras que la búsqueda ERIC dio 218 resultados. La

segunda estrategia consistió en examinar las referencias de los estudios que cumplían los criterios de admisibilidad (detallados a continuación).

Criterios de Elegibilidad

Se incluyeron estudios si, en primer lugar, 1) midieron cuantitativamente el desempeño de los estudiantes de segundo idioma en contraste con los hablantes nativos o 2) si comparaban el rendimiento de los estudiantes de segundo idioma en su L1 y L2. En segundo lugar, era necesario que los estudios no hubiesen sido realizados por el gobierno, ya que el tamaño de la muestra superaría inmediatamente el tamaño del efecto de otros estudios. Si los estudios se llevaron a cabo en los Estados Unidos, tuvieron que ser publicados después de 2006 para incluir estudios que no fueron considerados por Kieffer et al. (2009). Además de considerar todos estos requisitos, los estudios incluidos eran de naturaleza cuasiexperimental.

Los estudios fueron excluidos si: 1) no proporcionaron puntuaciones separadas cuando tenían muestras mixtas (evaluando problemas de palabras para estudiantes de segundo idioma y estudiantes con necesidades educativas especiales, por ejemplo); no proporcionaron número de participantes, puntuaciones medias y desviaciones estándar para calcular el tamaño del efecto; o 3) se encontraron metodológicamente defectuosos.

De los 2848 estudios encontrados, y después de eliminar los duplicados y estudios que no fueron revisados por pares o eran parte de una tesis, 35 fueron leídos completamente. Después de comprobar los criterios de elegibilidad, 10 estudios compusieron la muestra final. Estos estudios fueron escritos por 10 autores principales, publicados entre 2001 y 2020 en revistas revisadas por pares. La muestra final se compuso de 26 tamaños de efectos para 2.528 hablantes nativos y 1.786 estudiantes de segundo idioma. La mayoría de los estudios informaron que se centraban en evaluar un idioma diferente del inglés o evaluar a los estudiantes en su lengua materna e inglés (69,2% de los tamaños de efecto). Además, el 69,03% de los participantes formaban parte de estudios donde el inglés fue el idioma de instrucción y evaluación. Los idiomas distintos del inglés incluyeron al filipino (3 estudios), alemán y turco (1 estudio) y español (1 estudio). El estudio que se centró en el alemán y el turco fue el único estudio que no midió el inglés también.

Procedimiento de Análisis

Con el fin de calcular el tamaño del efecto para la población, se utilizó el paquete MetaXL de EpiGear para Excel (Barendregt & Doi, 2016). Dado que las muestras de los estudios individuales eran pequeñas, se calculó la *g* de Hedges (Kieffer et al., 2009).

Los procedimientos seguidos para llevar a cabo el metaanálisis fueron los propuestos por Field y Gillet (2010) y se describen a continuación. Después de tabular las muestras y los tamaños de los efectos para cada estudio, se ejecutó un método de efectos aleatorios en MetaXL. Este método fue elegido porque permite "generalize beyond the studies included in the meta-analysis" (Field & Gillet, 2010, p. 673). Este nivel de generalización es relevante al tratar de nivelar el campo para los hablantes no nativos. Después de calcular el tamaño medio del efecto para todos los estudios, se llevó a cabo un análisis de variables moderadoras para identificar la influencia de las variables en los tamaños del efecto. Esto se hizo usando Stata SE versión 13 (StataCorp, 2013). En este caso, se consideraron dos variables moderadoras: la edad y el idioma de las pruebas. La edad se clasificó en dos categorías: de 6 a 12 años y de 13 a 18 años. El idioma utilizado para las pruebas se clasificó entre inglés y otros. Si un estudio evaluó a los participantes en inglés y su idioma nativo, esto también se consideró parte de la otra categoría de idioma. El nivel de proficiencia en un idioma no se consideró como un posible moderador, ya que casi ningún estudio proporcionó esta información. Por último, el sesgo de publicación se abordó mediante una gráfica de embudo y una prueba de heterogeneidad. En primer lugar, el propósito del gráfico de embudo era identificar si los tamaños de los efectos se distribuían uniformemente a través de la mediana, lo que indicaría que no se encontró ningún sesgo de publicación. Al mismo tiempo, se llevó a cabo una prueba de heterogeneidad para identificar la proporción de la variación en las estimaciones

del tamaño del efecto, debido a la heterogeneidad en lugar del azar (Ríos et al., 2020).

La siguiente gráfica de tallo y hoja resume las frecuencias de los tamaños de efecto de los diferentes estudios incluidos después de ser calculados.

Tabla 1 Gráfico de tallo y hoja de todos los tamaños de efecto (gs)

Tallo	Hoja
-,4	9
-,3	1
-,1	4
,0	7
,1	3, 5, 6
,2	4
,3	0, 3, 4, 6
,4	1
,5	3, 7
,6	0, 1, 3, 8
,7	5, 6, 7
,8	4
,9	
1,	01, 5
2,	7

Como se puede ver, el tamaño más pequeño del efecto fue $g = ,07$ y el más grande, $g = 2,7$. La mayoría de los tamaños de efectos estaban en el rango de ,3 a ,7.

Tamaño Medio del Efecto y Prueba de Heterogeneidad

El tamaño medio calculado del efecto fue de $g = ,360$ [IC del 95%: $,176; ,543$]. Esto implica que el índice I^2 fue 85,02%, lo que sugiere una heterogeneidad considerable, consolidando la necesidad de un análisis moderador para identificar si otras variables pudiesen explicar los tamaños de los efectos.

Análisis de Moderadores

El análisis de moderadores indica que la edad moderó significativamente el 3,5% de los tamaños de los efectos ($p = ,002$). En relación con el lenguaje de instrucción y pruebas, el análisis del moderador indica que el lenguaje de instrucción y pruebas moderó significativamente el 4,7% de los tamaños de efecto ($p < ,001$).

En síntesis, cuando los participantes están en el rango de edad de 6 a 12 años y el idioma de instrucción y prueba es el inglés, el 5,7% de los tamaños de los efectos fueron moderados por estas variables ($p = ,001$). Si el idioma de instrucción y prueba es otro, solo el 2,7% de los tamaños de los efectos fueron moderados significativamente ($p = ,05$).

Cuando los participantes están en el rango de edad de 13 a 18 años, el idioma de instrucción y las pruebas modera un porcentaje inferior de tamaños del efecto (0,5% para inglés y aproximadamente 0% para otros), sin que el valor alcance significancia ($p=,66$ y $p=,991$, respectivamente).

Sesgo de Publicación

El siguiente gráfico de embudo ilustra la distribución de los tamaños de efecto a lo largo de la mediana.

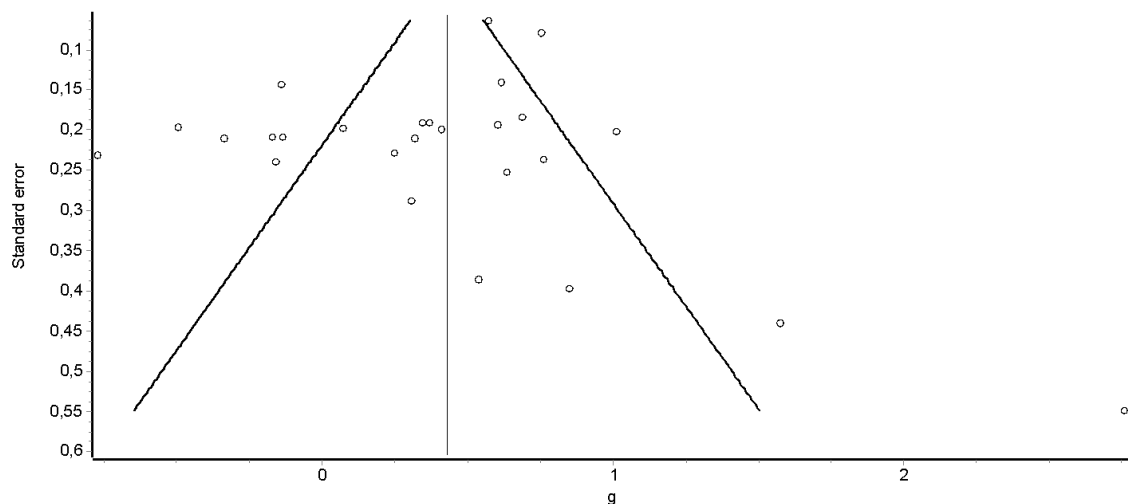


Figura 6 Gráfico de embudo

Este gráfico de embudo ilustra que más de la mitad de los tamaños de efecto se distribuyeron dentro de la región triangular. Los 12 tamaños de efecto

fuera de la región triangular pueden estar relacionados con la considerable heterogeneidad reportada anteriormente.

Discusión

El tamaño del efecto poblacional reportado en este metaanálisis fue de $g = ,360$, un tamaño de efecto intermedio. Según Coe (2002), este tamaño del efecto significaría que el 62% de los aprendientes de un idioma estarían por debajo de la puntuación media de los hablantes nativos. Este resultado contrasta notablemente en comparación con el tamaño del efecto de población reportado por Kieffer et al. (2009) de $g = ,604$, que significa que el 73% de los aprendientes de un idioma estarían por debajo de la puntuación promedio de los hablantes nativos. Esta diferencia en el tamaño de los efectos en la población podría explicarse por los estudios que se incluyeron. Principalmente, el estudio de Kieffer et al. no incluyó idiomas de pruebas que no fueran el inglés y se centró en estudios realizados en los Estados Unidos. Sin embargo, cuando el lenguaje de la instrucción y la lengua materna coinciden, las brechas serían menores, aunque no tan pronunciadas como se informó anteriormente.

El metaanálisis realizado muestra que la mayoría de los estudios incluidos han medido la relación entre la resolución de problemas matemáticos escritos y el lenguaje en una lengua de instrucción y evaluación distintas al inglés (17 de

26 tamaños de efecto). Sin embargo, sólo un estudio reportó tamaños de efectos donde el inglés no era el idioma de instrucción o de evaluación. Esto muestra claramente cómo la mayoría de las investigaciones se han centrado en el inglés.

Esto podría sugerir que el lenguaje de instrucción y evaluación podría ser un moderador importante del tamaño del efecto poblacional, lo que sería congruente con las conclusiones alcanzadas por investigaciones anteriores. Sin embargo, el análisis de variables moderadoras mostró que el lenguaje de instrucción y evaluación modera una pequeña proporción del tamaño del efecto poblacional. Además, lo mismo puede decirse sobre el rango de edad de los participantes. Ambos resultados pueden explicarse por las características de la muestra. En primer lugar, el inglés, como único medio de instrucción y evaluación, era parte de aproximadamente un tercio de los tamaños de efectos reportados. Además, el tamaño de la muestra para el rango de edad de 13 a 18 años era un cuarto en comparación con todos los participantes, centrándose únicamente en dos estudios. Estas diferencias podrían haber influido en los resultados de este metaanálisis, posiblemente infrarrepresentando el papel del lenguaje de la instrucción y evaluación, y el rango de edad. Esto debe ser considerado en investigaciones futuras como un efecto moderador.

En relación con el sesgo y la heterogeneidad de la publicación, hubo una heterogeneidad considerable en este metaanálisis ($I^2 = 85,02\%$), que se expresa

también en el número de estudios que están fuera de los rangos esperados en el gráfico de embudo (casi 50%). Como se mencionó en el párrafo anterior, el rango de edad y el idioma de instrucción y evaluación sólo representan un pequeño porcentaje de los tamaños de efecto (5,7%). La heterogeneidad puede explicarse por los diferentes diseños de investigación de los estudios (Higgins & Thompson, 2002) y por la diferente información reportada por los estudios que no permitieron proporcionar más variables moderadoras (Ríos et al., 2020).

Los resultados de este metaanálisis corroboran los hallazgos desagregados de investigaciones anteriores: los participantes logran puntuaciones más altas cuando su lenguaje de instrucción, evaluación y lengua materna son los mismos. Cuando esto se complementa con la baja eficiencia de las estrategias de adaptación para mejorar las puntuaciones de los hablantes no nativos (Kieffer et al., 2009; Ríos et al., 2020). Dichos resultados señalan la necesidad de seguir investigando no sólo adecuaciones para nivelar las condiciones de juego, sino también las puntuaciones de los hablantes no nativos en problemas escritos en idiomas diferentes del inglés para describir el alcance de las diferencias en las puntuaciones.

Obviamente, debido a la naturaleza de los criterios de inclusión, este metaanálisis no podía abarcar una descripción equilibrada del rango de edad y el lenguaje de instrucción y evaluación. Tener un tamaño de muestra más

equilibrado podría proporcionar una mejor imagen de la interacción entre el lenguaje y los problemas escritos. Al mismo tiempo, este metaanálisis ilustra que la edad y el lenguaje de instrucción y las pruebas sólo representan menos del 6% del tamaño del efecto poblacional. Por lo tanto, es importante explorar otras variables que podrían explicar por qué los hablantes nativos logran puntuaciones más altas en matemáticas. Kieffer et al. (2009) postularon la noción de que los factores sociodemográficos podían explicar la diferencia en los tamaños de los efectos, particularmente en los estudios controlados y las pruebas nacionales. Como tal, las investigaciones futuras también deben informar estas variables que se incluirían en futuros metaanálisis.

Desafortunadamente, no se pudieron incluir estudios prometedores en el área porque faltaba información clave sobre muestras, promedios y desviaciones estándar. Se debe prestar atención a estos datos, incluso cuando los resultados no son significativos, para ampliar las perspectivas relacionadas con este fenómeno.

Por último, las investigaciones futuras deben prestar atención al dominio del idioma como una variable que podría explicar no sólo la diferencia en las puntuaciones, sino también por qué las adaptaciones no son tan eficientes como es necesario. Por lo tanto, los investigadores deben medir el dominio del idioma de forma independiente mediante una prueba estandarizada para identificar

claramente el dominio L2 de los hablantes no nativos. Además, los estudiantes adultos deben ser incluidos, teniendo en cuenta que algunas universidades podrían requerir evaluar los conocimientos matemáticos en un idioma extranjero.

Como se ha establecido en este metaanálisis, el foco ha estado mayoritariamente en evaluar a participantes en edad escolar, sin considerar su nivel de proficiencia y sin prestar atención al procesamiento del lenguaje como tal. Esto cimienta la necesidad de realizar investigaciones en esta área incluyendo participantes en un rango etario distinto y poniendo el foco en el procesamiento del lenguaje y en variables lingüísticas más allá del idioma de instrucción y evaluación.

Sobre la base de este marco teórico y metaanálisis realizados se proponen los siguientes objetivos e hipótesis para esta investigación.

Objetivos

Objetivos Generales

- Determinar la relación lingüística categórica y funcional de la semántica matemática y su relación con los efectos cognitivos de inhibición en estudiantes con distinta proficiencia, tanto lingüística como matemática.

Objetivos Específicos

- Comparar las posibles diferencias significativas en los puntajes atribuidos a pares de palabras que se encuentran en una relación matemática categórica, en comparación a los puntajes atribuidos a pares de palabras en relación matemática funcional en español para estudiantes universitarios.
- Comparar la relación categórica y funcional de frases lingüísticas en términos de porcentajes de acierto en estudiantes y tiempos de reacción con distinta proficiencia lingüística y matemática.
- Analizar los efectos cognitivos de inhibición en estudiantes con diferentes proficiencias lingüísticas y matemáticas.

Hipótesis

Hipótesis de Investigación

Para la tarea de evaluación de pares de palabras en relación categórica y funcional, se espera que los pares de palabras que sean cohipónimos obtengan puntajes mayores en la escala de Likert de relación categórica. Junto con esto, se espera que los pares de palabras que sean cohipónimos obtengan puntajes inferiores en la escala de Likert de relación funcional.

Para el experimento de resolución de problemas matemáticos en relación categórica y funcional, se espera que los expertos en matemática obtengan porcentajes de acierto mayores y tiempos de reacción menores que los hablantes bilingües, sin importar su proficiencia. Junto con eso, se espera que los pares de palabras en relación funcional faciliten la resolución de los problemas matemáticos, debido a la relación semántico-matemática en la que se encuentran. Por lo tanto, todos los participantes obtendrán mayores porcentajes de acierto y menores tiempos de reacción.

Para el experimento Stroop, se espera que los expertos en matemática y los hablantes de alta proficiencia en inglés obtengan porcentajes de acierto superiores y tiempos de reacción menores que los hablantes de baja proficiencia en inglés, debido a un mayor control ejecutivo. A la vez, se espera que todos los participantes obtengan mayores porcentajes de acierto y menores tiempos de reacción en la condición congruente, debido a un proceso de facilitación.

Capítulo II: Metodología y Resultados

Metodología General

Para los propósitos de esta tesis doctoral, se realizaron tres estudios: primero, el estudio de valoración de pares de palabras en relación categórica y funcional; segundo, el estudio de resolución de problemas matemáticos escritos en relación categórica y funcional; y tercero, el estudio Stroop. Estos estudios difieren en características entre sí en cuanto al método utilizado.

La primera diferencia es el grado en que el estudio refleja los procesos mentales y/o neuronales que se estudian (Mertins, 2016). Estos grados pueden clasificarse en métodos offline, métodos online y métodos online verdaderos. Los métodos offline se centran en medir la competencia lingüística de los participantes, ya que no hay un acceso directo a los procesos mentales y reflejan la toma de decisiones conscientes. Un ejemplo de este tipo de método es la aplicación de cuestionarios, donde los participantes pueden tomarse algo de tiempo para contestar. Los métodos en línea se centran en un acceso mediado a los procesos mentales subyacentes, principalmente porque estos procesos son automatizados e inconscientes. Ejemplos de este método son tareas experimentales que requieren una respuesta con un breve tiempo de latencia, como en tareas de seguimiento ocular o experimentos conductuales que miden

tiempos de reacción. Finalmente, los métodos online directos tienen un acceso inmediato a los procesos mentales/neuronales relevantes, como en el caso de técnicas de resonancia magnética funcional y electroencefalografía (Mertins, 2016).

En esta tesis doctoral, el estudio de valoración de pares de palabras en relación categórica y funcional corresponde a un método offline. Mientras que los estudios de resolución de problemas escritos en relación categórica y funcional y Stroop son métodos online.

En esta tesis doctoral se utilizaron dos diseños de investigación. Se utilizó el diseño descriptivo de encuesta transversal para el estudio de valoración de pares de palabras en relación categórica y funcional. En este tipo de diseño no experimental, se recolecta información una sola vez con el propósito de medir actitudes, creencias, opiniones o prácticas actuales (Creswell & Creswell, 2017). Para los estudios de resolución de problemas escritos en relación categórica y funcional y Stroop, se utilizó el diseño cuasiexperimental. De acuerdo a Creswell y Creswell (2017), este tipo de diseño carece de algunas características habituales en los experimentos como la distribución al azar de los participantes en distintos grupos, lo que podría aumentar los riesgos a la validez interna del estudio. A nombrar, en nuestros estudios, los participantes no fueron distribuidos al azar. De todas maneras, este tipo de diseño es capaz de determinar los efectos

de causalidad entre variables independientes y dependientes, a través de la manipulación de variables (Creswell & Creswell, 2017).

Junto con estos diseños de investigación, también se utilizó el diseño factorial mixto. Hernández y colegas (2014) describen el diseño factorial como un tipo de diseño experimental donde se manipulan al menos dos variables independientes para analizar qué efecto tienen estas variables independientes sobre las variables dependientes. A la vez, cada variable independiente debe incluir dos o más niveles. Finalmente, este diseño es de naturaleza mixta al incluir comparaciones intersujetos e intrasujetos (Hernández et al., 2014). El detalle de cada variable independiente y sus factores es detallado en las secciones que corresponden a cada uno de los estudios.

Para los estudios cuasiexperimentales de esta tesis, se midieron tiempos de reacción y porcentajes de acierto. Los tiempos de reacción son una forma de medir una respuesta conductual, generalmente en milisegundos, desde que se presenta una tarea experimental hasta que esta se completa (Baayen & Milin, 2010). Luce (1991) plantea que los tiempos de reacción dan cuenta del tiempo que le toma a un participante procesar la información que se presenta en el estímulo. Los psicólogos experimentales concuerdan que existen tres tipos básicos de tiempos de reacción: los simples, los de reconocimiento y los de elección. Los tiempos de reacción simple se obtienen cuando un participante

responde a un estímulo perceptual como la luz o el sonido; los tiempos de reacción de reconocimiento se obtienen en tareas en que se presentan dos estímulos y el participante debe reaccionar a uno e ignorar el otro; y los tiempos de reacción de elección donde los participantes deben escoger una respuesta entre varias opciones disponibles (Baayen & Milin, 2010). En esta tesis doctoral, el estudio de resolución de problemas matemáticos escritos medirá tiempos de reacción de elección, mientras que el estudio Stroop medirá tiempos de reacción de reconocimiento.

Los porcentajes de acierto expresan la precisión con la que los participantes completaron la tarea experimental (Castelain & Marín, 2014). Smith y Kosslyn (2008) plantean que medir la exactitud de una tarea experimental permite evaluar de manera objetiva cuán eficaz fue el proceso cognitivo que se está midiendo. Desde su perspectiva, esta medida conductual se utiliza para evaluar procesos cognitivos que abarcan desde la discriminación de información hasta los procesos cognitivos que requieren recordar (Smith & Kosslyn, 2008). Ellos también plantean que existe una relación entre los tiempos de reacción y los porcentajes de acierto, donde los participantes le darían más importancia a uno de estos componentes en desmedro del otro. Es decir, los participantes podrían enfocarse en reaccionar rápidamente afectando negativamente el porcentaje de acierto o viceversa (Smith & Kosslyn, 2008). En esta tesis doctoral, ambos estudios conductuales medirán porcentajes de acierto.

En las secciones siguientes se presentará el diseño, procedimiento, participantes, el material experimental, los resultados y una discusión particular para cada uno de los estudios realizados. Estos estudios responden uno a uno a los objetivos específicos anteriormente planteados. Igualmente, se indicará el diseño que se utilizó en cada estudio en la siguiente sección. En cuanto al procedimiento, el primer estudio que recolectó información fue el de valoración de pares de palabras en relación categórica y funcional, seguido por el de resolución de problemas matemáticos escritos y finalizando con el estudio Stroop. Cada estudio fue tomado en una sesión diferente con cada grupo de participantes.

Estudio de Valoración de Pares de Palabras en Relación Categórica y Funcional

Diseño de Estudio

En este estudio se implementó un diseño descriptivo de encuesta transversal, factorial mixto intersujetos (Hernández et al., 2014), 2 tipos de relación semántica a evaluar (categórica vs. funcional) e intrasujetos, 2 tipos de relación semántica de los sustantivos (categórica vs. funcional).

Con respecto a las variables se identifican dos variables independientes: 1) tipo de relación semántica a evaluar y 2) el tipo de relación semántica de los sustantivos. La variable dependiente que se medirá será el puntaje asignado a cada par de palabras para determinar su frecuencia.

Participantes

Para el estudio de valoración de pares de palabras en relación categórica y funcional, la técnica de muestreo utilizada fue no probabilística. En este tipo de muestreo, como Creswell (2012) planteó, la selección de la muestra se realiza en base a sus características, disponibilidad y conveniencia. Por lo tanto, los criterios de inclusión en este estudio fueron 1) ser estudiante universitario y 2) ser estudiante de pedagogía en inglés. Se determinó estas características porque los estímulos serán utilizados a futuro en experimentos con estudiantes de pedagogía en inglés mayores de 18 años. El número total de participantes fueron 80 estudiantes (53 mujeres) de una universidad privada de la ciudad de Santiago, Chile, de un rango etario entre 19 a 25 años. La mitad de los participantes respondió una escala de Likert donde debían evaluar si los pares de palabras presentados correspondían a una relación categórica, mientras que la mitad restante debía evaluar si los pares de palabras presentados correspondían a una relación funcional.

Es importante destacar que estos participantes corresponden a una muestra distinta a los participantes de los estudios de resolución de problemas matemáticos escritos y Stroop.

Procedimiento

Antes de participar en este estudio, los participantes leyeron y firmaron un protocolo y consentimiento informado, aprobado por el comité de ética de la Universidad. Las escalas de Likert fueron respondidas por los participantes de manera voluntaria en el mismo día y horario luego de que terminaron sus clases lectivas de manera presencial. Para realizar esta evaluación, los participantes indicaron en una escala de Likert si la relación semántica entre los pares de palabras era representativa, donde 1 indicaba que no era representativa y 7 que sí era representativa. Antes de realizar esta valoración, se le presentaron a los participantes ejemplos de la relación semántica a evaluar que era muy poco representativos (1), medianamente representativos (4) y altamente representativos (7).

Las instrucciones de las escalas de Likert se ejemplifican a continuación. Para simplificar la comprensión de la tarea, en vez de hablar de relación funcional o relación categórica, se usaron palabras más sencillas de entender.

“En este cuestionario, tu tarea consiste en evaluar hasta qué punto existe una relación de dependencia entre los siguientes pares de palabras. Para realizar la evaluación, debes calificar de 1 a 7 los pares de palabras, donde 1 equivale a una relación categórica extremadamente baja y 7 a una relación categórica extremadamente alta. Por favor, marca con una X la casilla correspondiente y trata de utilizar todo el rango de calificaciones disponible.

A continuación, te daremos algunos ejemplos antes de comenzar con la evaluación:

Relación de dependencia alta: camisa y colgador. El colgador tiene una relación de dependencia alta con la camisa ya que su única función es colgar ropa. Por lo tanto, un posible puntaje podría ser 7.

Relación de dependencia intermedia: camisa y sillón tienen una relación de dependencia intermedia ya que en ciertos contextos el sillón podría ser utilizado para dejar ropa. Por lo tanto, un posible puntaje podría ser 4.

Relación de dependencia baja: fútbol y tenis tienen una relación de dependencia baja ya que se pueden realizar de manera independiente. Por lo tanto, un posible puntaje podría ser 1.

Palabra 1	Palabra 2	Calificación						
		menos dependencia			mayor dependencia			
chalecos	ganchos	1	2	3	4	5	6	7
químicos	físicos	1	2	3	4	5	6	7
galletas	brownies	1	2	3	4	5	6	7
hamsters	jerbos	1	2	3	4	5	6	7

“

Después de tabular los datos, se realizó una prueba t para determinar si había diferencias significativas en las evaluaciones que los participantes hicieron para los sets de palabras en cada lista. Finalmente, se descartaron las palabras funcionales en la lista de evaluación categórica que obtuvieron un promedio igual o superior a 4 y se descartaron las palabras categóricas que obtuvieron un valor inferior a 4. Para la lista de evaluación funcional, se descartaron las palabras

categorías que obtuvieron un promedio igual o superior a 4 y se descartaron las palabras funcionales que obtuvieron un valor inferior a 4.

Material

Se aplicaron dos escalas de Likert, una para determinar si los pares de palabras presentados se encontraban en una relación categórica y otra para determinar si los pares de palabras presentados se encontraban en una relación funcional, de acuerdo a las categorías de Bassok y colegas (1998; 2001). Para seleccionar las palabras que los participantes evaluaron, se tomó como referencia la lista utilizada por Guthormsen et al. (2016) para sus experimentos en inglés. Esta lista se encuentra compuesta de 64 tríos de palabras, expresadas en una palabra base, una palabra que se relaciona intuitivamente de manera categórica con la palabra base y una palabra que se relaciona intuitivamente de manera funcional con la palabra base. Luego de traducir las palabras, se adaptaron algunos ítems al uso lingüístico de nuestra cultura, posteriormente se evaluó que no hubiese diferencias significativas entre frecuencia de uso y largo de palabra, tomando como referencia para frecuencia de uso los datos de buscapalabras (Davis & Perea, 2005) y LIFCACH (Lista de Frecuencias del Castellano de Chile) (Sadowsky & Martínez-Gamboa, 2012). Para evaluar que no hubiese diferencias significativas, se realizaron pruebas t de Student entre las palabras base, categórica y funcional. No hubo diferencias significativas para

frecuencia de uso ni largo de palabras en ninguna de las comparaciones realizadas.

Para que los participantes evaluaran las relaciones entre estos tríos de palabras, se utilizó el método propuesto por Gurthormsen et al. (2016). Primero, se realizaron dos listas donde se incluía la columna de palabra base, la mitad de las palabras en relación categórica y la mitad de las palabras en relación funcional. Esto se ejemplifica en las siguientes tablas.

Tabla 2 Ejemplos de lista de evaluación categórica

Palabra base		Tipo de relación
Abejas	Avispas	Categórica
Naranjas	Manzanas	Categórica
Presos	Celdas	Funcional
Poodles	Caniles	Funcional
Ranas	Sapos	Categórica

Tabla 3 Ejemplos de lista de evaluación funcional

Palabra base		Tipo de relación
Abejas	Colmenas	Funcional
Naranjas	Canastos	Funcional
Presos	Gendarmes	Categórica
Poodles	Terriers	Categórica
Ranas	Pozas	Funcional

Para la lista de evaluación categórica, un grupo de participantes evaluó si los pares de palabras se encontraban en una relación categórica. Mientras que,

para la lista de evaluación funcional, el otro grupo evaluó si los pares de palabras se encontraban en una relación funcional.

Resultados

La tabla 4 ilustra las medidas de estadística descriptiva de moda, media y mediana en la lista de evaluación categórica.

Tabla 4 Moda, media y mediana para los sets categóricos y funcionales en la lista de evaluación categórica.

	Sets categóricos	Sets funcionales
Moda	5.325	4.275
Media	5.60546875	4.18515625
Mediana	5.675	4.3875

Basados en la moda, media y mediana, se pueden identificar diferencias entre los puntajes otorgados a los sets categóricos y funcionales en la lista de evaluación categórica, con una clara tendencia a favorecer a los sets categóricos con puntuaciones más altas.

La tabla 5 ilustra las medidas de estadística descriptiva de moda, media y mediana en la lista de evaluación funcional.

Tabla 5 Moda, media y mediana para los sets categóricos y funcionales en la lista de evaluación funcional.

	Sets categóricos	Sets funcionales
Moda	3.275	4.95
Media	3.960336538	5.02265625
Mediana	3.525	5.0125

Basados en la moda, media y mediana, se pueden identificar diferencias entre los puntajes otorgados a los sets categóricos y funcionales en la lista de evaluación funcional, con una clara tendencia a favorecer a los sets funcionales con puntuaciones más altas.

Análisis

El análisis propuesto fue la t de Student, de acuerdo a lo realizado por Guthormsen et al. (2016) al momento de evaluar sus palabras base, categóricas y funcionales en inglés. Según Hernández et al. (2014), esta prueba estadística se utiliza para determinar si existe una variación significativa entre dos variables o grupos. En este caso en particular, era necesario evaluar si existían diferencias significativas entre las evaluaciones realizadas para cada lista.

La tabla 6 muestra los resultados de la t de Student en la lista de evaluación categórica.

Tabla 6 Prueba t de la lista de evaluación categórica para una muestra pareada

	Sets categóricos	Sets funcionales
Media	5.60546875	4.18515625
Desviación estándar	0.7443804085	1.068441346
Observaciones	32	32
Grados de libertad	31	
Estadístico t	6.170045192	
P(T<=t) una cola	0.00000003	
Valor crítico de t (una cola)	1.669804163	
P(T<=t) dos colas	0.0000001	
Valor crítico de t (dos colas)	1.998971517	

Como se puede notar, sí hubo diferencias significativas entre los puntajes otorgados a los sets de palabras en relación categórica (M=5,6, SD=0,74) en comparación a los puntajes otorgados a los sets de palabras en relación funcional (M=4,1, SD=1,06) en la lista que evaluaba la relación categórica, $t(31) = 6,17$, $p=0.00000003$. Por lo tanto, se puede declarar que los participantes evaluaron las relaciones semánticas entre sets de palabras como efectivamente distintas para la lista categórica, otorgando puntuaciones mayores a los sets categóricamente relacionados.

La tabla 7 muestra los resultados de la t de Student en la lista de evaluación funcional.

Tabla 7 Prueba t de la lista de evaluación funcional para para una muestra pareada

	Sets categóricos	Sets funcionales
Media	3.960336538	5.02265625
Desviación estándar	1.195869245	0.7643614079
Observaciones	32	32
Grados de libertad	31	
Estadístico t	-4.234114408	
P(T<=t) una cola	0.000038644	
Valor crítico de t (una cola)	1.669804163	
P(T<=t) dos colas	0.000077289	
Valor crítico de t (dos colas)	1.998971517	

Como se puede notar, sí hubo diferencias significativas entre los puntajes otorgados a los sets de palabras en relación categórica (M=3,9, SD=1,19) en comparación a los puntajes otorgados a los sets de palabras en relación funcional (M=5,0, SD=0,76) en la lista que evaluaba la relación funcional, $t(31) = -4,23$, $p=0.000038644$. Por lo tanto, se puede declarar que los participantes evaluaron las relaciones semánticas entre sets de palabras como efectivamente distintas para la lista funcional, otorgando puntuaciones más altas a los sets relacionados funcionalmente.

Las valoraciones realizadas por los participantes de este estudio van en línea con las reportadas por Guthormsen et al. (2016) en estudiantes universitarios, para listas de pares de palabras en inglés. Particularmente, para la lista que evaluaba la relación categórica, las autoras reportaron que los pares

de palabras categóricas obtuvieron valoraciones significativamente superiores ($M = 5,83$, $SD = 0,89$) que las valoraciones de los pares de palabras funcionales ($M = 3,14$, $SD = 1,28$), $t(35) = 9,00$, $p < .001$. En el caso de nuestro estudio, los puntajes obtenidos por los pares de palabras en relación categórica ($M=5,6$, $SD=0,74$) también son significativamente mayores que los puntajes obtenidos por los pares de palabras en relación funcional ($M=4,1$, $SD=1,06$), $t(31) = 6,17$, $p=0.00000003$. De todos, es necesario considerar el tamaño del efecto de cada estudio para realizar una comparación adecuada.

En el caso del estudio de Guthormsen et al. (2016), el tamaño del efecto de esta comparación fue de $d = 2,31$. Esto significa que 98% (Coe, 2002) de las valoraciones de los pares de palabras funcionales estarían por debajo de la media de las valoraciones de los pares de palabras en relación categórica. En nuestro estudio, el tamaño del efecto fue de $d = 1,54$. Esto significa que 92% (Coe, 2002) de las valoraciones de los pares de palabras funcionales estarían por debajo de la media de las valoraciones de los pares de palabras en relación categórica. Esto indica que las diferencias en valoración en el estudio de Guthormsen et al. (2016) fueron más pronunciadas que en nuestro estudio.

En cuanto a la evaluación realizada para la lista de relación funcional, Guthormsen et al. (2016) reportaron que las valoraciones de los pares de palabras en relación funcional ($M = 5,57$, $SD = 1,10$) fueron significativamente

superiores que las valoraciones de los pares de palabras en relación categórica ($M = 2,31$, $SD = 1,26$), $t(40) = 9,41$, $p < .001$. Así también, en el caso de nuestro estudio, las valoraciones de los pares de palabras en relación funcional ($M=5,0$, $SD=0,76$) fueron significativamente mayores que las valoraciones de los pares de palabra en relación categórica ($M=3,9$, $SD=1,19$), $t(31) = -4,23$, $p=0.000038644$. De todos, es necesario considerar el tamaño del efecto de cada estudio para realizar una comparación adecuada.

En el caso del estudio de Guthormsen et al. (2016), el tamaño del efecto de esta comparación fue de $d = 2,61$. Esto significa que 99% (Coe, 2002) de las valoraciones de los pares de palabras categóricos estarían por debajo de la media de las valoraciones de los pares de palabras en relación funcional. En nuestro estudio, el tamaño del efecto fue de $d = 1,03$. Esto significa que 84% (Coe, 2002) de las valoraciones de los pares de palabras categóricos estarían por debajo de la media de las valoraciones de los pares de palabras en relación funcional. Esto indica que las diferencias en valoración en el estudio de Guthormsen et al. (2016) fueron más pronunciadas que en nuestro estudio.

A modo de síntesis, estos resultados indican que los participantes de este estudio efectivamente discriminan entre las relaciones semánticas funcionales y categóricas de manera significativa.

Estudio de Resolución de Problemas Matemáticos Escritos en Relación Categórica y Funcional

Diseño de Estudio

En este estudio se implementó un diseño cuasiexperimental, factorial mixto intersujetos (Hernández et al., 2014), con 3 grupos de participantes (español, inglés, experto en matemáticas) e intrasujetos con 2 tipos de relación semántica (categórica vs. funcional).

Con respecto a las variables, se identifican tres variables independientes: 1) idioma del problema escrito, 2) nivel de conocimiento en matemáticas y 3) tipo de relación semántica. Las variables dependientes que se medirán serán tiempo de respuesta en milisegundos y porcentaje de aciertos.

Participantes

Para este experimento, los participantes fueron estudiantes universitarios de tres universidades de Chile. Se calculó el tamaño de la muestra, de acuerdo al programa g power (Faul et al., 2007), que arrojó un total de 66 participantes para la muestra total con un tamaño de efecto d de 0.25, probabilidad de error α

de 0.005 y el poder $1-\beta$ de 0.95; sin embargo, se agregó más cantidad de participantes, con un tamaño final de 80 participantes.

Los participantes que se describen a continuación son los mismos participantes del estudio Stroop. En el caso de los bilingües, el grupo de bajo conocimiento en inglés constó de 24 participantes (13 mujeres) y el grupo de alto conocimiento en inglés constó de 27 (19 mujeres) participantes. Finalmente, el grupo de alto conocimiento en matemática consistió en 29 participantes (12 mujeres). Al igual que en el estudio de valoración de pares de palabras, el muestro fue no probabilístico (Creswell, 2012).

Particularmente, se invitó a participar a estudiantes con las siguientes características. Para el grupo de bilingües, la característica principal era estar estudiando pedagogía en inglés o traducción. Luego, para pertenecer al grupo de bajo conocimiento en inglés era necesario estar en el nivel A1 o A2 de acuerdo al Marco Común Europeo de Referencia (Council of Europe, 2020). Mientras que para pertenecer al grupo de alto conocimiento en inglés era necesario estar en el nivel B2 o C1 de acuerdo al Marco Común Europeo de Referencia (Council of Europe, 2020). Para pertenecer al grupo de alto conocimiento en matemática, los estudiantes universitarios de carreras científico-matemáticas rindieron una prueba de fracciones (Gómez et al., 2014).

El nivel de inglés de los estudiantes fue establecido a través de la prueba English Unlimited Placement test de Cambridge University Press (Clementson et al., 2010). La sección que se utilizó fue la de Written test, donde los estudiantes deben escoger la opción correcta para llenar el vacío en una oración. Esta prueba tiene una duración máxima de 40 minutos, siendo las instrucciones principales 1) escoger la respuesta correcta y 2) parar cuando las preguntas sean muy difíciles.

La habilidad en matemática fue establecida a través de una prueba de fracciones (Gómez et al., 2014), donde los participantes debían escoger cuál de dos fracciones presentadas era la mayor sin utilizar calculadora. Esta prueba constó de 25 preguntas con una duración máxima de 10 minutos.

Finalmente, una de las participantes inicialmente consideradas tuvo que ser descartada ya que sus respuestas no fueron grabadas por el sitio web. De todos modos, la muestra final fue de 80 participantes.

Procedimiento

Antes de participar en este estudio, los participantes leyeron y firmaron un protocolo y consentimiento informado, aprobado por el comité de ética de la Universidad.

a) La tarea experimental fue realizada a través de la página web de pavlovia.org, debido a las dificultades generadas por la pandemia para acceder a los participantes de manera presencial. A la vez, la tarea experimental fue programada en PsychoPy2 (Peirce et al., 2019). Al momento de realizar esta tarea, los participantes compartieron sus pantallas para monitorear que no hubiese problemas.

b) Los participantes se sentaron frente a un monitor de 18 pulgadas a una distancia de 90 centímetros. Las instrucciones aparecieron en la pantalla, indicando que debían leer las oraciones cuidadosamente para luego indicar cuál respuesta de las opciones es la correcta. Los participantes se autoadministraron las oraciones al pulsar la barra espaciadora.

Luego de que los participantes estuvieran listos, cada ensayo siguió el siguiente orden, utilizando el paradigma de ventana móvil, que se muestra en la siguiente figura.

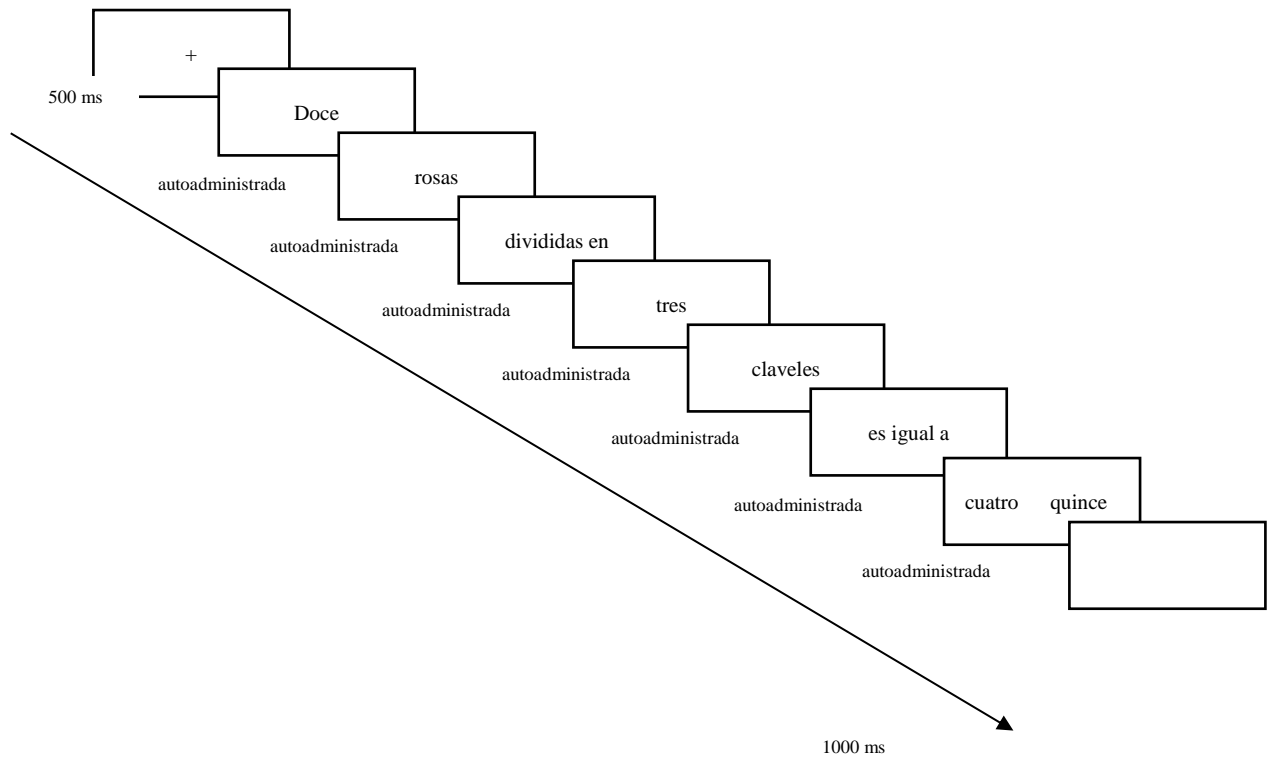


Figura 7 Secuencia de presentación de estímulos en el experimento resolución de problemas matemáticos en relación categórica y funcional.

c) Cada ensayo partió con un punto de fijación presentado por 500 ms.

d) Una vez que se autoadministraron la oración estímulo, aparecieron las dos respuestas posibles, una al extremo inferior izquierdo de la pantalla y otra al extremo inferior derecho de la pantalla.

e) Para responder, los participantes presionaron la tecla q para la respuesta que aparezca en el extremo izquierdo de la pantalla y la tecla p para la respuesta que aparezca en el extremo derecho de la pantalla. La instrucción que se les dio es que ellos debían decidir cuál es la respuesta correcta para el problema escrito. Los participantes tuvieron 4000 milisegundos para escoger su respuesta.

f) Luego de haber respondido, hubo un intervalo de 1000 ms antes de iniciar el siguiente ensayo.

Material

Material Matemático

Los números utilizados como material experimental oscilaron entre dos y veinticinco para evitar la carga cognitiva extra generada por números altos (Ashcraft, 1992). Adicionalmente, cada problema matemático debía dar como resultado un número entero positivo y en el caso de las divisiones, se excluyeron los números que al ser divididos se obtiene su raíz cuadrada, como en el caso de $25 / 5 = 5$. Estas decisiones fueron tomadas de acuerdo a lo planteado por Guthormsen et al. (2016).

Material Lingüístico

Los sustantivos utilizados después de cada número correspondían a una de tres categorías: *relación categórica*, *relación funcional* o *control*. Como se indicó en el marco teórico, la relación categórica es una relación de hiponimia, ya que ambos sustantivos corresponden a elementos de una misma categoría. En el caso de la relación funcional, ambos sustantivos se encuentran en una relación semántica en que uno cumple una función en relación a otro. En el caso de los sustantivos del contexto control, estos correspondieron a sustantivos que, al ser evaluados, no se encontraban en ninguna de las relaciones semánticas anteriormente mencionadas.

La relación semántica entre sustantivos fue determinada en el estudio de valoración de pares de palabras en relación categórica y funcional.

Al combinar ambos tipos de materiales, los participantes tuvieron que resolver problemas como los siguientes:

Doce	rosas	divididas en	tres	vasijas	es igual a	quince	cuatro

Donde cada celda representa el material lingüístico que se vería en la pantalla y que los participantes se autoadministrarían, hasta llegar a la última celda donde debían escoger la respuesta correcta.

Resultados

Resultados Porcentaje de Aciertos

La tabla 8 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.

Tabla 8 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.

	Contexto Categórico	Contexto Funcional	Control
Moda	100	100	100
Media	96,06	96,06	94,9
Mediana	100	100	94,44

Basados en la moda, media y mediana, se puede identificar que el grupo de bajo conocimiento en inglés, a pesar de tener una moda, media y mediana iguales para los distintos contextos experimentales, tiene un porcentaje de acierto menor para las oraciones de control utilizadas en media y mediana.

La tabla 9 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.

Tabla 9 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.

	Contexto categórico	Contexto Funcional	Control
Moda	100	100	100
Media	96,5	97,53	92,79
Mediana	100	100	94.44

Basados en la moda, media y mediana, se puede identificar que el grupo de alto conocimiento en inglés tuvo un desempeño distinto, dependiendo del contexto experimental, con resultados superiores para el contexto funcional, seguido por el contexto categórico y finalmente seguido por las oraciones de control.

La tabla 10 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.

Tabla 10 Moda, media y mediana para el porcentaje de acierto del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.

	Contexto Categórico	Contexto Funcional	Control
Moda	100	100	100
Media	98,46	98,65	96,36
Mediana	100	100	94.44

Basados en la moda, media y mediana, se puede identificar que el grupo de alto conocimiento en matemática tuvo un desempeño distinto, dependiendo del contexto experimental, con resultados superiores para el contexto funcional, seguido por el contexto categórico y finalmente seguido por las oraciones de control.

A modo de resumen, para la estadística descriptiva de los porcentajes de acierto de los distintos grupos experimentales, el grupo de expertos en

matemática tuvo un desempeño mayor en todos los contextos experimentales, incluidas las oraciones de control. En cuanto a los contextos experimentales, el contexto funcional fue la que tuvo mayor porcentaje de acierto para todos los grupos experimentales.

Resultados de Tiempo de Reacción

La tabla 11 ilustra la media y la mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.

Tabla 11 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de bajo conocimiento en inglés.

	Contexto Categórico	Contexto Funcional	Control
Media	836,21	844,22	840,77
Mediana	823,93	840,08	824,42

Basados en la media y mediana, se puede identificar que el grupo de bajo conocimiento en inglés tuvo tiempos de reacción mayores en el contexto funcional.

La tabla 12 ilustra la media y la mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.

Tabla 12 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en inglés.

	Contexto Categórico	Contexto Funcional	Control
Media	861,052	870,67	849,81
Mediana	853,08	882,72	846,1

Basados en la media y mediana, se puede identificar que el grupo de alto conocimiento en inglés tuvo tiempos de reacción mayores en el contexto funcional.

La tabla 13 ilustra la media y la mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.

Tabla 13 Media y mediana para los tiempos de reacción del experimento de resolución de problemas matemáticos en relación categórica y funcional para el grupo de alto conocimiento en matemática.

	Contexto Categórico	Contexto Funcional	Control
Media	778,24	784,88	799,52
Mediana	757,366	767,44	760,4

Basados en la media y mediana, se puede identificar que el grupo de alto conocimiento en matemática tuvo tiempos de reacción mayores para las oraciones de control.

A modo de resumen, para la estadística descriptiva de los tiempos de reacción de los distintos grupos experimentales, el grupo de expertos en matemática tuvo tiempos de reacción menores en todos los contextos experimentales, incluidas las oraciones de control. En cuanto a los contextos experimentales, el contexto funcional fue la que tuvo tiempos de reacción más altos en todas las medianas y en casi todos los promedios en todos los grupos experimentales.

Análisis

ANOVA Porcentaje de Aciertos

Antes de realizar una ANOVA de medidas repetidas, se realizó una prueba de normalidad para determinar si los resultados se encuentran distribuidos normalmente. Esta prueba se presenta en la tabla 14.

Tabla 14 Pruebas de normalidad para porcentaje de aciertos

Pruebas de normalidad

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Contexto categórico	,364	80	,000	,535	80	,000
Contexto funcional	,418	80	,000	,410	80	,000
Control	,260	80	,000	,716	80	,000

Puesto que la significancia de ambas pruebas es menor que .05, no es posible indicar que nuestros resultados están distribuidos normalmente. A pesar de eso, ANOVAs han demostrado no ser sensibles a desviaciones moderadas de la distribución normal y que la tasa de falsos positivos no se ve afectada por la violación del supuesto de normalidad (Glass et al., 1972; Lix et al., 1996).

Por lo tanto, se realizó una ANOVA de medidas repetidas para analizar si había diferencias significativas en los porcentajes de acierto y cómo interactúan

las distintas variables del experimento de resolución de problemas matemáticos en relación categórica y funcional en relación al porcentaje de aciertos.

En relación a la interacción entre grupos experimentales y contexto, los resultados muestran que no hay interacción significativa ($F(4,80) = 1,047$, $MSE=3356,232$, $p = 0,371$). En cuanto a los efectos principales, sí hay diferencias entre los distintos contextos ($F(2, 80) = 7,912$, $MSE=3356,232$, $p = 0,002$).

A pesar de la robustez de la prueba ANOVA ante distribuciones no normales, se realizó la prueba no paramétrica Kruskal-Wallis para verificar si había diferencias significativas en los porcentajes de aciertos de cada contexto experimental en los distintos grupos del estudio. En el caso del contexto funcional, no hubo diferencias significativas en los distintos grupos ($p = 0,525$). Para el contexto categórico, tampoco hubo diferencias significativas en los distintos grupos ($p = 0,408$). Finalmente, en el contexto control tampoco hubo diferencias significativas en los distintos grupos ($p = 0,326$).

Para ahondar en las diferencias significativas entre contextos identificadas en la prueba ANOVA, se realizó una prueba t de muestras emparejadas. En primer lugar, se realizó una comparación entre los porcentajes de acierto del contexto categórico y los porcentajes de acierto del contexto funcional. No hubo diferencias significativas entre los porcentajes de acierto de ambos contextos

($t(80) = 0,948$, $p = 0,346$). En segundo lugar, se realizó una comparación entre los porcentajes de acierto del contexto funcional y los porcentajes de acierto del contexto control. En este caso, se encontraron diferencias significativas entre los porcentajes de acierto de estos contextos ($t(80) = 3,085$, $p = 0,003$). Finalmente, se realizó una comparación entre los porcentajes de acierto del contexto categórico y los porcentajes de acierto del contexto control. Sí hubo diferencias significativas entre los porcentajes de acierto de estos contextos ($t(80) = 2,968$, $p = 0,004$). En ambas comparaciones, los participantes tuvieron mayores porcentajes de aciertos para el contexto categórico y funcional que para el contexto de control.

Guthormsen et al. (2016) indican que sus participantes tuvieron un mayor porcentaje de acierto cuando las operaciones matemáticas estaban alineadas con las relaciones semánticas presentes en el problema matemático. Es decir, los participantes tuvieron mayores porcentajes de acierto para problemas de adición si los sustantivos se encontraban en relación categórica y mayores porcentajes de acierto para problemas de división si los sustantivos se encontraban en relación funcional. Este resultado es congruente con nuestros resultados, puesto que los distintos grupos tuvieron mejores resultados para el contexto funcional, cuya relación está alineada con la división.

Estos resultados se pueden explicar desde la perspectiva del modelo SECO (Gros et al., 2020), donde el componente de la semántica del mundo facilita a los participantes la resolución de problemas, puesto que la operación matemática a resolver se encuentra alineada con las relaciones semánticas de los sustantivos presentes en el planteamiento del problema. Es así como los participantes son capaces de generar un algoritmo de solución que les permite llegar a una solución correcta en los casos donde la operación matemática y las relaciones de los sustantivos se encontraban alineadas, sin necesidad de recodificar la información del problema matemático. Esto se ve reflejado particularmente en el caso de la relación funcional, donde hubo porcentajes de acierto mayores que en el contexto control. Junto con este resultado, también hubo porcentajes de acierto mayores para la relación categórica al ser comparada con el contexto control. Esta diferencia se puede explicar al considerar que la relación categórica de los sustantivos, a pesar de crear una incongruencia con la división, no generaban una gran carga cognitiva y se podía acceder directamente a la semántica de la matemática y llegar al algoritmo necesario para resolver estos problemas matemáticos.

ANOVA de Tiempos de Reacción

Antes de realizar una ANOVA de medidas repetidas, se realizó una prueba de normalidad para determinar si los resultados se encuentran distribuidos normalmente. Esta prueba se presenta en la tabla 15.

Tabla 15 Pruebas de normalidad para tiempos de reacción

Pruebas de normalidad

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Contexto categórico	,082	77	,200	,965	77	,030
Contexto funcional	,067	77	,200	,975	77	,136
Control	,090	77	,194	,965	77	,031

Los tiempos de reacción del contexto funcional se encuentran distribuidos normalmente, ya que su significancia es mayor a .05. Por el contrario, los tiempos de reacción de los contextos categórico y control no se encuentran distribuidos normalmente. Aun así, como ya se mencionó anteriormente, es posible realizar una ANOVA ya que estas no son sensibles a desviaciones moderadas de la normalidad.

Por lo tanto, se realizó una ANOVA de medidas repetidas para analizar si había diferencias significativas en los porcentajes de acierto y cómo interactúan

las distintas variables del experimento de resolución de problemas matemáticos en relación categórica y funcional en relación a los tiempos de reacción.

En relación a la interacción entre grupos experimentales y contexto, los resultados muestran que no hay interacción significativa ($F(4,80) = 0,362$ $MSE=516578,728$, $p = 0,826$). En cuanto a los efectos principales, no hay diferencias entre los distintos contextos ($F(2, 80) = 0,965$, $MSE=516578,728$, $p = 0,38$).

Junto con la prueba ANOVA, se realizó la prueba no paramétrica Kruskal-Wallis para verificar si había diferencias significativas en los tiempos de reacción de cada contexto experimental en los distintos grupos del estudio. En el caso del contexto funcional, no hubo diferencias significativas en los distintos grupos ($p = 0,319$). Para el contexto categórico, tampoco hubo diferencias significativas en los distintos grupos ($p = 0,338$). Finalmente, en el contexto control no se encontraron diferencias significativas en los distintos grupos ($p = 0.558$).

Los participantes de nuestro estudio tuvieron tiempos de reacción superiores en el contexto funcional en todos los grupos, a pesar de que era posible esperar tiempos de reacción menores para el contexto funcional al estar alineada con la operación de división. Los resultados de Ronasi et al. (2018) también dan cuenta de tiempos de reacción mayores en su estudio para una

condición semántica que se encontraba alineada con la sustracción. Guthormsen et al. (2016) también sugieren que la relación funcional entre los sustantivos puede que esté menos restringida que la relación categórica. Esto quiere decir que es menos probable que los participantes anticipen los posibles sustantivos que puedan completar el rol de divisor, debilitando el efecto semántico. En su estudio, esto se traducía en una ausencia del componente N400 en problemas de división semánticamente desalineados. En nuestro estudio, esto se puede traducir en una tendencia a tener tiempos de reacción superiores en el contexto funcional aún cuando estos no sean significativos, en relación a los otros contextos experimentales. Junto con esta explicación, Guthormsen et al. (2016) plantean que en las divisiones es posible llegar a una respuesta correcta sin necesidad de modelar una estrategia de solución. Esto se puede relacionar con el componente de semántica matemática del modelo SECO (Gros et al., 2020), donde la experiencia de los participantes con las divisiones les permite saltarse la información lingüística y solo enfocarse en la información matemática, sin que se genere un efecto de facilitación por la presencia de sustantivos en contexto funcional. De todos modos, no se puede descartar que, debido al diseño y lógica de la tarea, los participantes luego de responder una cierta cantidad de ensayos deciden enfocarse solamente en la información matemática.

Estudio Stroop

Diseño

En este estudio se implementó un diseño cuasiexperimental, factorial mixto intersujetos (Hernández et al., 2014), 3 grupos de participantes (español, inglés y matemáticas) e intrasujetos 2 tipos de congruencia (congruente vs. incongruente).

Con respecto a las variables, se identifican dos variables independientes: 1) grupo experimental e 2) inhibición. Las variables dependientes que se medirán serán tiempo de respuesta y porcentaje de aciertos.

Participantes

Para este experimento, los participantes fueron estudiantes universitarios de tres universidades de Chile. Para calcular el tamaño de la muestra, se utilizó programa g power (Faul et al., 2007), que indicó un total de 66 participantes para la muestra total con un tamaño de efecto d de 0.25, probabilidad de error α de 0.005 y el poder $1-\beta$ de 0.95; sin embargo, se agregó más cantidad de participantes, con un tamaño final de 80 participantes.

Los participantes que se describen a continuación son los mismos participantes del estudio de resolución de problemas matemáticos escritos. En el caso de los bilingües, el grupo de bajo conocimiento en inglés constó de 24 participantes (13 mujeres) y el grupo de alto conocimiento en inglés constó de 27 (19 mujeres) participantes. Finalmente, el grupo de alto conocimiento en matemática consistió en 29 participantes (12 mujeres).

Particularmente, se invitó a participar a estudiantes con las siguientes características. Para el grupo de bilingües, la característica principal era estar estudiando pedagogía en inglés o traducción. Luego, para pertenecer al grupo de bajo conocimiento en inglés era necesario estar en el nivel A1 o A2 de acuerdo al Marco Común Europeo de Referencia (Council of Europe, 2020). Mientras que para pertenecer al grupo de alto conocimiento en inglés era necesario estar en el nivel B2 o C1 de acuerdo al Marco Común Europeo de Referencia (Council of Europe, 2020). Para pertenecer al grupo de alto conocimiento en matemática, los estudiantes universitarios de carreras científico-matemáticas rindieron una prueba de fracciones (Gómez et al., 2014).

La habilidad en matemática fue establecida a través de una prueba de fracciones, donde los participantes debían escoger cuál de dos fracciones presentadas era la mayor sin utilizar calculadora. Esta prueba constó de 25 preguntas con una duración máxima de 10 minutos.

El nivel de conocimiento en inglés de los estudiantes será establecido a través de la prueba English Unlimited Placement test de Cambridge University Press (Clementson et al., 2010). La sección que se utilizó fue la de Written test, donde los estudiantes deben escoger la opción correcta para llenar el vacío en una oración. Esta prueba tiene una duración máxima de 40 minutos, siendo las instrucciones principales 1) escoger la respuesta correcta y 2) parar cuando las preguntas sean muy difíciles.

Finalmente, una de las participantes inicialmente consideradas tuvo que ser descartada ya que sus respuestas no fueron grabadas por el sitio web. De todos modos, la muestra final fue de 80 participantes.

Procedimiento

Antes de participar en este estudio, los participantes leyeron y firmaron un protocolo y consentimiento informado, aprobado por el comité de ética de la Universidad. La tarea Stroop fue contestada en una sesión posterior a la del estudio de resolución de problemas matemáticos escritos.

a) La tarea experimental fue realizada a través de la página web de pavlovia.org debido a las dificultades generadas por la pandemia para acceder a los participantes de manera presencial. A la vez, la tarea experimental fue

programada en PsychoPy2 (Peirce et al., 2019). Al momento de realizar esta tarea, los participantes compartieron sus pantallas para monitorear que no hubiese problemas.

b) Los participantes se sentaron frente a un monitor de 18 pulgadas a una distancia de 90 centímetros. Las instrucciones aparecieron en la pantalla, indicando que debían leer la palabra cuidadosamente para luego identificar cuál es el color de la tinta que tiñe la palabra escrita. Los participantes se autoadministraron las palabras al pulsar la barra espaciadora.

Luego de que los participantes estuvieran listos, cada ensayo siguió el siguiente orden, utilizando el paradigma de ventana móvil, presentado visualmente en la siguiente figura. El fondo de la pantalla de la tarea experimental era negro, como en el estudio de Gajewski et al. (2020).

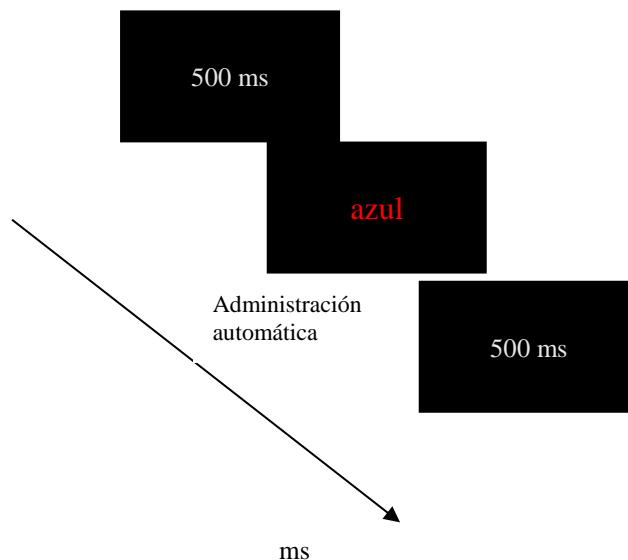


Figura 8 Secuencia de presentación de estímulos en el experimento Stroop.

c) Una vez que apareció la palabra, los participantes debían identificar el color de la tinta que tiñe la palabra, sin prestar atención al contenido semántico de la palabra.

d) Para realizar la identificación correcta, los participantes presionaron la tecla v para la tinta de color rojo, la tecla b para la tinta de color azul, la tecla n para la tinta de color verde y la tecla m para la tinta de color amarillo.

e) Luego de haber respondido, comenzó el siguiente ensayo de manera automática. Los participantes no tenían límite de tiempo para seleccionar su respuesta.

Material

La prueba Stroop consistió en la identificación de la tinta con la cual una palabra estaba teñida. En este experimento, se seleccionó cuatro colores a identificar: rojo, azul, verde y amarillo. Junto con eso, las palabras que aparecieron también eran rojo, azul verde y amarillo. Estas decisiones fueron tomadas de acuerdo al experimento realizado por Gajewski et al. (2020).

Resultados

Resultados Porcentaje de Acierto

La tabla 16 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de bajo conocimiento en inglés.

Tabla 16 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de bajo conocimiento en inglés.

	Contexto Congruente	Contexto Incongruente
Moda	100	100
Media	97,39	95,74
Mediana	95,83	97,91

Basados en la moda, media y mediana, se puede identificar que el grupo de bajo conocimiento en inglés tuvo un promedio de porcentaje de acierto superior en el contexto congruente, aunque la mediana del contexto congruente fue menor que la del contexto incongruente.

La tabla 17 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en inglés.

Tabla 17 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en inglés.

	Contexto Congruente	Contexto Incongruente
Moda	100	97,91666667
Media	97,53	95,13
Mediana	100	95,83

Basados en la moda, media y mediana, se puede identificar que el grupo de alto conocimiento en inglés tuvo un desempeño superior en el contexto congruente para todas las medidas de estadística descriptiva.

La tabla 18 ilustra las medidas de estadística descriptiva de moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en matemática.

Tabla 18 Moda, media y mediana para el porcentaje de acierto del experimento Stroop para el grupo de alto conocimiento en matemática.

	Contexto Congruente	Contexto Incongruente
Moda	100	97,91666667
Media	97,12	97,55
Mediana	100	95,83

Basados en la moda, media y mediana, se puede identificar que el grupo de alto conocimiento en matemática tuvo un promedio superior en el contexto

incongruente, mientras que la moda y mediana fueron inferiores a el contexto congruente.

A modo de resumen, para la estadística descriptiva de los distintos grupos experimentales, el grupo de expertos en matemática tuvo un desempeño mayor en el contexto incongruente, mientras que los otros grupos experimentales tuvieron un desempeño levemente superior en el contexto congruente.

Resultados Tiempos de Reacción

La tabla 19 ilustra la media y la mediana para los tiempos de reacción del experimento Stroop para el grupo de bajo conocimiento en inglés.

Tabla 19 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de bajo conocimiento en inglés.

	Contexto Congruente	Contexto Incongruente
Media	939,96	1089,06
Mediana	911,69	1010,51

Basados en la media y mediana, se puede identificar que el grupo de bajo conocimiento en inglés tuvo tiempos de reacción mayores para el contexto incongruente.

La tabla 20 ilustra la media y la mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.

Tabla 20 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.

	Contexto Congruente	Contexto Incongruente
Media	910,22	977,1
Mediana	855,11	899,64

Basados en la media y mediana, se puede identificar que el grupo de alto conocimiento en inglés tuvo tiempos de reacción mayores para el contexto incongruente.

La tabla 21 ilustra la media y la mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.

Tabla 21 Media y mediana para los tiempos de reacción del experimento Stroop para el grupo de alto conocimiento en matemática.

	Contexto Congruente	Contexto Incongruente
Media	841,37	863,39
Mediana	822,58	821,71

Basados en la media y mediana, se puede identificar que el grupo de alto conocimiento en matemática tuvo un promedio de tiempos de reacción mayores para el contexto incongruente. En cuanto a la mediana, esta fue ligeramente menor para el contexto incongruente.

A modo de resumen, para la estadística descriptiva de los tiempos de reacción de los distintos grupos experimentales, el grupo de expertos en matemática tuvo tiempos de reacción menores en todos los contextos experimentales. En cuanto a los contextos experimentales, el contexto incongruente fue la que tuvo tiempos de reacción más altos en todos los promedios y en casi todas las medianas en todos los grupos experimentales.

Análisis

ANOVA Porcentaje de Aciertos

Antes de realizar una ANOVA de medidas repetidas, se realizó una prueba de normalidad para determinar si los resultados se encuentran distribuidos normalmente. Esta prueba se presenta en la tabla 22.

Tabla 22 Pruebas de normalidad para porcentaje de aciertos

Pruebas de normalidad

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Contexto congruente	,306	80	,000	,499	80	,000
Contexto incongruente	,302	80	,000	,556	80	,000

Puesto que la significancia de ambas pruebas es menor que .05, no es posible indicar que nuestros resultados están distribuidos normalmente. De todas maneras, cómo se indicó anteriormente, es posible realizar una ANOVA de medidas repetidas por no ser sensibles a desviaciones de la normalidad.

Por lo tanto, se realizó una ANOVA de medidas repetidas para analizar si había diferencias significativas en los porcentajes de acierto y cómo interactúan las distintas variables del experimento Stroop en relación al porcentaje de aciertos.

En relación a la interacción entre grupos experimentales y contexto, los resultados muestran que no hay interacción significativa ($F(2,80) = 0,812$, $MSE=2719,946$, $p = 0,448$). En cuanto a los efectos principales, no hay diferencias entre los distintos contextos ($F(1, 80) = 1,541$, $MSE=2719,95$, $p = 0,218$).

Considerando que los resultados no se encuentran distribuidos normalmente, se realizó la prueba no paramétrica Kruskal-Wallis para verificar si había diferencias significativas en los porcentajes de acierto de cada contexto experimental en los distintos grupos del estudio. En el caso del contexto congruente, no hubo diferencias significativas en los distintos grupos ($p = 0,300$). Para el contexto incongruente, tampoco hubo diferencias significativas en los distintos grupos ($p = 0,598$).

Lubin et al. (2016) midieron si existían diferencias en las tasas de error entre un grupo de expertos en matemática y no expertos al completar una tarea Stroop para dar cuenta de sus habilidades de control inhibitorio. Sus resultados indican que no hay un efecto principal de grupo experimental, a pesar de que todos los participantes sí cometieron más errores en la condición incongruente. Estos resultados contrastan con nuestro experimento por dos motivos. En primer lugar, los expertos en matemática tuvieron un desempeño levemente superior en el contexto incongruente, en comparación a los no expertos que tuvieron mejor desempeño en el contexto congruente. En segundo lugar, en nuestro estudio no hubo diferencias significativas entre grupos.

Estos resultados se pueden deber a que la muestra de cada grupo de nuestro estudio no fue lo suficientemente alta, lo que queda demostrado por la distribución no normal de medianas. Junto con eso, se puede considerar el nivel

de habilidad matemática y lingüística de los participantes. Como se discute en el modelo de activación interactiva bilingüe cognitiva (Grainger et al., 2010) y el modelo SECO (Gros et al., 2020), la habilidad de los participantes puede permitir inhibir información no relevante para la tarea en cuestión. En este caso, nuestros participantes se encontraban en distinto nivel de habilidad lingüística y matemática, pero a pesar de eso, es posible que su experiencia inhibiendo información no relevante para sus respectivas áreas del conocimiento les diese una ventaja que se refleja en porcentajes de acierto similares en ambos contextos experimentales y sin diferencias significativas. Junto con eso, no se puede descartar el efecto techo que puede haber aparecido en esta tarea experimental ya que la mayoría de los participantes tuvo puntajes muy altos, donde la moda fue de 100 en todos los grupos experimentales, lo que dificulta la distribución normal de los datos y sus análisis (Salkind, 2010).

ANOVA Tiempos de Reacción

Antes de realizar una ANOVA de medidas repetidas, se realizó una prueba de normalidad para determinar si los resultados se encuentran distribuidos normalmente. Esta prueba se presenta en la tabla 23.

Tabla 23 Pruebas de normalidad para tiempos de reacción

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Contexto congruente	,129	74	,004	,926	74	,000
Contexto incongruente	,136	74	,002	,921	74	,000

Puesto que la significancia de ambas pruebas es menor que .05, no es posible indicar que nuestros resultados están distribuidos normalmente. De todas maneras, como se indicó anteriormente, es posible realizar una ANOVA de medidas repetidas por no ser sensibles a desviaciones de la normalidad.

Se realizó una ANOVA de medidas repetidas para analizar si había diferencias significativas en los tiempos de reacción y cómo interactúan las distintas variables del experimento Stroop en relación a los tiempos de reacción.

En relación a la interacción entre grupos experimentales y contexto, los resultados muestran que sí hay interacción significativa ($F(2,80) = 12,268$, $p = 0,000$). En cuanto a los efectos principales, sí hay diferencias entre los distintos contextos ($F(1,80) = 53,995$, $MSE=318471,011$, $p = 0,000$).

Junto con la prueba ANOVA, se realizó la prueba no paramétrica Kruskal-Wallis para verificar si también había diferencias significativas en los tiempos de reacción de cada contexto experimental en los distintos grupos del estudio. En

el caso del contexto congruente, no hubo diferencias significativas en los distintos grupos ($p = 0,505$). Para el contexto incongruente, sí hubo diferencias significativas en los distintos grupos ($p = 0,018$).

Resumen de contrastes de hipótesis

	Hipótesis nula	Prueba	Sig.	Decisión
1	La mediana de las diferencias entre T_FAC y T_INHI es igual a 0.	Prueba de Wilcoxon de los rangos con signo para muestras relacionadas	,000	Rechace la hipótesis nula.

Se muestran significaciones asintóticas. El nivel de significancia es ,05.

La prueba de Wilcoxon también indica que hay diferencias significativas en los tiempos de reacción del contexto congruente y el contexto incongruente. Particularmente, los grupos de bajo nivel de inglés y alto nivel de inglés tuvieron tiempos de reacción menores en el contexto congruente. Mientras que el grupo de alto conocimiento en matemática obtuvo tiempos de reacción menores en el contexto incongruente. Junto con estos resultados, el grupo de alto conocimiento en matemática obtuvo tiempos de reacción considerablemente menores en el contexto incongruente al comparar con los otros dos grupos.

En el estudio mencionado anteriormente de Lubin et al. (2016) se ilustra que sí existe una diferencia significativa en los tiempos de reacción de los participantes, donde requirieron más tiempo para la condición incongruente. A

pesar de estos resultados, ellos indican que no hubo un efecto principal de grupo ni interacción entre estos factores. En el caso de nuestro estudio, sí hubo una interacción significativa entre tiempos de reacción y grupos experimentales.

Para ahondar en las posibles diferencias entre los grupos, se hizo una comparación entre cada grupo mediante la comparación t de Student de medias para muestras independientes. Primero, se realizó una t de Student para analizar si había diferencias significativas en los tiempos de reacción entre los grupos de bajo conocimiento y alto conocimiento en inglés y considerando los contextos congruente e incongruente. No hubo diferencias significativas en los tiempos de reacción entre los grupos de bajo conocimiento y alto conocimiento en inglés al considerar el contexto congruente ($t(52) = 0,461$, $p = 0,647$) y tampoco al considerar el contexto incongruente ($t(52) = 1,423$, $p = 0,162$).

Segundo, se realizó una t de Student para analizar si había diferencias significativas en los tiempos de reacción entre los grupos de bajo conocimiento en inglés y alto conocimiento en matemática y considerando los contextos congruente e incongruente. No hubo diferencias significativas en los tiempos de reacción entre los grupos bajo conocimiento en inglés y alto conocimiento en matemática al considerar el contexto congruente ($t(53) = 1,548$, $p = 0,128$). Por el contrario, sí hubo diferencias significativas al considerar el contexto incongruente ($t(53) = 3,078$, $p = 0,003$). A nombrar, el grupo de alto conocimiento

en matemática tuvo tiempos de reacción considerablemente menores en el contexto incongruente.

Tercero, se realizó t de Student para analizar si había diferencias significativas en los tiempos de reacción entre los grupos alto conocimiento en inglés y alto conocimiento en matemática y considerando los contextos congruente e incongruente. No hubo diferencias significativas en los tiempos de reacción entre los grupos alto conocimiento en inglés y alto conocimiento en matemática, al considerar el contexto congruente ($t(56) = 0,904$, $p = 0,371$) y tampoco al considerar el contexto incongruente ($t(56) = 1,55$, $p = 0,126$).

Como se indicó con anterioridad, Lubin et al. (2016) descubrieron que hay diferencias significativas entre las condiciones congruente e incongruente. Particularmente, los tiempos de reacción fueron superiores para la condición incongruente. Estos resultados son similares con los de nuestro experimento, puesto que hubo diferencias significativas entre el grupo de bajo conocimiento en inglés y el grupo de alto conocimiento en matemática, para el contexto incongruente. En el caso del estudio de Lubin et al., estas diferencias fueron significativas para los grupos de expertos y no expertos en matemáticas. Mientras que, en nuestro estudio, solamente hubo diferencias significativas en la comparación realizada entre el grupo de bajo conocimiento en inglés con el grupo de alto conocimiento en matemática en el caso del contexto incongruente. En

términos concretos, los participantes del grupo de alto conocimiento en matemática se vieron menos afectados por el esfuerzo cognitivo de inhibir información irrelevante para completar la tarea experimental en comparación al grupo de bajo conocimiento en inglés que tuvo tiempos de reacción considerablemente mayores. Las otras comparaciones realizadas no generaron diferencias significativas.

A diferencia del caso de los porcentajes de acierto, al comparar los tiempos de reacción entre los grupos de estudio sí hubo diferencias significativas. Esto se puede deber a que los distintos niveles de habilidad en sus respectivas áreas de conocimiento tengan un efecto en la velocidad para completar la tarea. Por lo tanto, el grupo de bajo conocimiento en inglés debería obtener tiempos de reacción mayores que los grupos de alto conocimiento en inglés y alto conocimiento en matemática. En este caso particular, estos tiempos de reacción fueron significativamente mayores para la porción de la tarea experimental que requería inhibir una respuesta ante un estímulo incongruente en el caso del grupo de bajo conocimiento en inglés al compararlo con el grupo de alto conocimiento en matemática. Es así como los participantes del grupo de bajo conocimiento en inglés tuvieron mayores dificultades para inhibir la información irrelevante para esta tarea experimental y se demoraron más en completarla de manera satisfactoria.

Los modelos teóricos de activación interactiva bilingüe cognitiva (Grainger et al., 2010) y SECO (Gros et al., 2020) pueden explicar por qué se dieron estos resultados. En ambos modelos, la habilidad juega un rol central para explicar cómo los participantes son capaces de inhibir información que es irrelevante para la actividad que están realizando, ya sea activar los elementos lingüísticos necesarios para comunicarse en una segunda lengua o identificar el algoritmo adecuado para un problema matemático específico. Estas consideraciones teóricas explican por qué no hubo diferencias significativas entre el grupo de alto conocimiento en inglés y el grupo de alto conocimiento en matemática ya que ambos grupos, en sus respectivas áreas de estudio, tienen una vasta experiencia inhibiendo información. A la vez, estas consideraciones teóricas pueden explicar por qué sí hubo diferencias significativas entre el grupo de bajo conocimiento en inglés y alto conocimiento en matemática. Particularmente, el grupo de bajo conocimiento en inglés aún no tiene suficiente experiencia inhibiendo información en su área de estudio.

Capítulo III: Discusión General y Conclusiones

Discusión General

En esta sección, se discutirá el cumplimiento de las distintas hipótesis para cada experimento, se abordarán los resultados obtenidos desde los distintos modelos teóricos presentados en el marco teórico, para finalmente abordar los componentes lingüísticos y de función ejecutiva estudiados.

Discusión Estudio de Evaluación de Pares de Palabras en Relación Categórica y Funcional

En cuanto a la tarea de evaluación de pares de palabras en relación categórica y funcional, la primera hipótesis planteada fue que los pares de palabras que son cohipónimos obtendrían puntajes mayores en la escala de Likert de relación categórica. A la vez, se esperaba que los pares de palabras que son cohipónimos obtengan puntajes inferiores en la escala de Likert de relación funcional como segunda hipótesis.

Luego de recolectar la información, se realizó una t de Student para determinar si existían diferencias significativas entre las medias de las evaluaciones para cada lista. En primer lugar, se determinó si había diferencias

significativas en la lista que evaluaba si los pares de palabras se encontraban en relación categórica. En efecto, el análisis indica que sí existen diferencias significativas, ya que la valoración de los sets de palabras en relación categórica fue significativamente distinta a la valoración de los sets de palabras en relación funcional. En términos concretos, los participantes valoraron con puntajes mayores a los sets de palabras en relación categórica. Por lo tanto, se puede aceptar la primera hipótesis del estudio de valoración de pares de palabras en relación categórica y funcional, debido a que los cohipónimos obtuvieron valoraciones mayores en la lista que evaluaba la relación categórica de los pares de palabras.

En segundo lugar, se evaluó si había diferencias significativas en la lista que evaluaba si los pares de palabras se encontraban en relación funcional. En efecto, el análisis indica que sí existen diferencias significativas, puesto que la valoración de los sets de palabras en relación categórica fue significativamente distinta a la valoración de los sets de palabras en relación funcional. De manera concreta, los participantes valoraron con puntajes mayores a los sets de palabras en relación funcional. Por lo tanto, se puede aceptar la segunda hipótesis del estudio de valoración de pares de palabras en relación categórica y funcional, puesto que los cohipónimos obtuvieron valoraciones menores en la lista que evaluaba la relación funcional de los pares de palabras.

A pesar de que el fenómeno de desalineación semántica en problemas matemáticos escritos se ha estudiado ampliamente en Estados Unidos (Bassok et al., 1995; Bassok et al., 1998; Bassok, 2001) y también se ha estudiado en distintos países, como Corea del Sur (Lee et al., 2016), Rusia (Tyumeneva et al., 2018) y China (Qiu & Wang, 2021; Wang et al., 2023), solamente existe un estudio del que se tenga conocimiento a la fecha., el de Guthormsen et al. (2016) en Estados Unidos. A modo de ejemplo, Wang et al. (2023) mencionaron haber evaluado las relaciones categóricas y funcionales de los ítems que utilizaron en su estudio, pero no reportaron los puntajes otorgados ni los análisis. En el estudio de electrofisiología de Guthormsen et al. (2016), se realizó un estudio para validar el material lingüístico que utilizaron en sus experimentos. De hecho, el procedimiento que ellas utilizaron fue el mismo que se utilizó en este estudio. En una primera tarea, un grupo de participantes evaluó si los pares de palabras se encontraban en relación categórica. En esta tarea, Guthormsen et al. (2016) determinaron que sí existió una diferencia significativa en las evaluaciones realizadas para los pares categóricos y los pares funcionales, donde los pares categóricos obtuvieron puntajes mayores. En una segunda tarea, otro grupo de participantes evaluó si los pares de palabras se encontraban en relación funcional. En esta tarea, Guthormsen et al. (2016) determinaron que también hubo una diferencia significativa entre los pares categóricos y los pares funcionales, donde los pares funcionales obtuvieron puntajes mayores.

Por lo tanto, los resultados del estudio de valoración de pares de palabras en relación categórica y funcional en español concuerdan con los resultados obtenidos por Guthormsen et al. (2016). De todas maneras, es necesario indicar que los resultados obtenidos por ellas tuvieron un tamaño del efecto considerablemente mayor al de nuestro estudio, lo que sugiere diferencias más pronunciadas en las valoraciones realizadas por sus participantes.

En cuanto a los modelos teóricos discutidos en el marco teórico, los resultados de las escalas de Likert aplicadas en este estudio permiten ahondar de mejor manera en los supuestos teóricos de la interfaz sintáctica-semántica y el modelo de congruencia semántica (SECO) (Gros et al., 2020). En cuanto a la interfaz sintáctica-semántica, los resultados de este primer estudio dan cuenta de que efectivamente los hablantes de español perciben de manera distinta la relación categórica y la relación funcional, recordando que la relación categórica es una relación semántica en naturaleza y que la relación funcional es una relación sintáctica en naturaleza.

Como se mencionó anteriormente, un grupo de participantes del primer estudio tuvo que evaluar si los pares de palabras eran representativos de la relación categórica, donde los pares categóricos obtuvieron puntajes significativamente mayores que los pares funcionales. A la vez, otro grupo de participantes tuvo que evaluar si los pares de palabras eran representativos de

la relación funcional, donde los pares funcionales obtuvieron puntajes significativamente mayores que los pares categóricos. Estos resultados pueden relacionarse con la evidencia de estudios de neuroimagen que dan cuenta de patrones de activación cerebral distintos para procesar los elementos semánticos y sintácticos del lenguaje (Carter et al., 2019; Hagoort, 2003). Es decir, al momento de valorar los sets de palabras de una lista en particular, los participantes deberían haber activado distintas zonas de su cerebro, dependiendo de si la relación entre los pares de palabras era categórica (semántica) o funcional (sintáctica). Y, en consecuencia, tal activación cerebral podría dar cuenta de una manera distinta cómo fueron valorados los distintos pares de palabras. Naturalmente, esta interpretación debe ser refrendada por estudios de neuroimagen donde se estudien los patrones de activación al momento de valorar estas relaciones lingüísticas.

Junto con estos patrones de activación cerebral distintos, hay evidencia de estudios de electrofisiología que dan cuenta de que las anomalías semánticas y sintácticas generan componentes ERP distintos. Concretamente, el componente N400 ha sido descrito como un componente sensible a las anomalías semánticas, mientras que el componente P600 ha sido descrito como un componente sensible a las anomalías sintácticas (Hammer et al., 2008). Por lo tanto, al momento de valorar cuán representativo de una relación categórica es un par de palabras que tienen una relación funcional, se debería haber

generado un componente N400 ya que se generaría una anomalía semántica. Para el componente P600, se esperaría una situación similar al momento de valorar cuán representativo de una relación funcional es un par de palabras que tienen una relación categórica, dado que esto generaría una anomalía sintáctica. Del mismo modo que la interpretación sobre patrones de activación distintos, esta interpretación también debe ser refrendada por estudios de electrofisiología, donde se estudien los componentes que aparecen al momento de valorar estas relaciones lingüísticas.

En cuanto al modelo de congruencia semántica (SECO) (Gros et al., 2020), el componente de la semántica del mundo está estrechamente relacionado con la valoración realizada por los participantes para los distintos pares de palabras en las listas categórica y funcional. Particularmente, el componente de la semántica del mundo en este modelo teórico equivaldría a las experiencias de la vida cotidiana que los participantes tienen con los pares de palabras que tuvieron que evaluar. De esta manera, los participantes frecuentemente categorizan un par de palabras como *camelia* y *rosa* como representativo de una relación semántica donde ambos pares pertenecen a la misma categoría: flor. A la vez, un par de palabras como *rosa* y *vasija* sería categorizado como representativo de una relación sintáctica, ya que uno de los elementos dicta el rol que el otro tiene, como en el caso de *vasija*, que cumple el rol de contener a las flores. Esta semántica del mundo permite también a los

participantes el conceptualizar estas relaciones lingüísticas con problemas matemáticos específicos y así facilitar la resolución exitosa de los problemas matemáticos.

En resumen, ambas hipótesis fueron aceptadas, lo que permitió crear la batería lingüística definitiva que fue utilizada en el estudio de resolución de problemas matemáticos en relación categórica y funcional.

Discusión Estudio de Resolución de Problemas Matemáticos Escritos en Relación Categórica y Funcional

En cuanto a la tarea de resolución de problemas matemáticos escritos en relación categórica y funcional, la primera hipótesis plantea que los expertos en matemática deberían obtener porcentajes de acierto mayores y tiempos de reacción menores que los hablantes bilingües. A la vez, la segunda hipótesis plantea que los pares de palabras en relación funcional facilitarían la resolución de los problemas matemáticos, debido a que la relación funcional se alinea de mejor manera con la división. Por lo tanto, se espera que los participantes obtengan mayores porcentajes de acierto y menores tiempos de reacción cuando los pares de palabras presentados estén en relación funcional.

Luego de recolectar la información, se realizó una ANOVA de medidas repetidas para analizar la interacción entre grupos experimentales y contextos y

sus efectos principales. En primer lugar, se descartó que hubiese una interacción significativa entre grupos experimentales y contextos. En segundo lugar, se determinó que sí hubo diferencias significativas al considerar los distintos contextos de este estudio.

Después de establecer que sí existen diferencias significativas para los porcentajes de acierto en los distintos contextos, se realizaron t de Student para comparar las medias de los distintos contextos. En cuanto a la comparación de los porcentajes de acierto entre los contextos categórico y funcional, no hubo diferencias significativas. Al comparar los porcentajes de acierto del contexto categórico con el contexto control, sí hubo diferencias significativas. Finalmente, al comparar los porcentajes de acierto del contexto funcional con el contexto control, también hubo diferencias significativas. Concretamente, los porcentajes de acierto de los contextos categórico y funcional fueron mayores que los del contexto control.

Después de analizar los tiempos de reacción en los distintos contextos experimentales para los distintos grupos, se realizó una ANOVA de medidas repetidas para analizar si hubo diferencias significativas al considerar los tiempos de reacción. Este análisis dio cuenta que no hubo interacción significativa y que tampoco hubo un efecto principal en los distintos contextos experimentales.

En relación a las hipótesis para este experimento, la primera hipótesis fue rechazada, ya que no hubo diferencias significativas al comparar los porcentajes de acierto y tiempos de reacción de los distintos grupos experimentales. Es decir, el grupo de expertos en matemática no tuvieron porcentajes de acierto mayores y tiempos de reacción menores que los grupos de hablantes bilingües.

En cuanto a la segunda hipótesis, que indicaba que se facilitaría la resolución de los problemas matemáticos para los pares de palabras en relación funcional, esta fue parcialmente aceptada, puesto que hubo porcentajes de acierto significativamente mayores en el contexto funcional al compararlo con el contexto control. Esto también ocurrió al comparar el contexto categórico con el contexto control. Lamentablemente, no hubo diferencias significativas al comparar los porcentajes de acierto de los contextos funcional y categórico. Finalmente, tampoco hubo diferencias significativas en cuanto a los tiempos de reacción, por lo que se rechaza la idea de que el contexto funcional tendría menores tiempos de reacción en comparación con los otros contextos experimentales.

En el caso del estudio de Guthormsen et al. (2016), los participantes tuvieron mayor porcentaje de acierto que se alineaba con la operación a resolver. Es decir, hubo mayores porcentajes de acierto para la condición categórica si se resolvía una suma y hubo mayores porcentajes de acierto para la condición

funcional si se resolvía una división. En nuestra tarea experimental, los participantes tuvieron que resolver divisiones. Particularmente para nuestro estudio, el contexto funcional sí tuvo porcentajes de acierto significativamente mayores que el contexto control. Y se hubiese esperado, como indicaba la segunda hipótesis, que también el contexto funcional hubiese tenido porcentajes de acierto significativamente mayores que el contexto categórico. Esto no ocurrió en esta comparación. Por lo tanto, nuestros resultados concuerdan parcialmente con los resultados de Guthormsen et al. (2016).

Considerando los tiempos de reacción, y mencionando nuevamente que no hubo diferencias significativas en los distintos análisis que realizamos, el contexto funcional tuvo tiempos de reacción mayores que los otros contextos experimentales, a pesar de encontrarse alineada con la operación de división. Este resultado contraintuitivo ha sido reportado con anterioridad. Por ejemplo, Ronasi et al. (2018) indicaron haber tenido tiempos de reacción mayores en la condición semántica que en su estudio se alineaba con la operación de sustracción.

Estos resultados se pueden explicar si se considera lo planteado por Guthormsen et al. (2016) sobre la relación funcional. Al explicar sus resultados, ellos plantearon la idea de que la relación funcional entre sustantivos es menos restringida ya que hay una mayor libertad para que un sustantivo complete el rol

de divisor al ser comparada con la relación categórica. En el caso de la relación categórica, el universo de sustantivos que puede ser cohipónimo es considerablemente más reducido que el universo de sustantivos que pueden cumplir el rol de divisor. A modo de ejemplo, el sustantivo *pera* puede activar palabras *manzana* o *naranja* como miembros del superordinado *frutas*. Pero los sustantivos que pueden cumplir una función en relación a *pera* no son parte de un mismo superordinado. Por ejemplo, un *cuchillo* y un *plato* se pueden relacionar con *pera* como utensilios para comer, pero también una *pera* puede relacionarse con *caja*, *guantes* y *tijeras de podar* como implementos relacionados con su cosecha.

Por lo tanto, los posibles sustantivos que tienen alguna función en relación a *pera* son mucho más difíciles de clasificar en una categoría única. Esta explicación tiene un fundamento en los resultados ERP obtenidos por Guthormsen y colegas (2016). Particularmente, hubo una ausencia de N400 en los problemas de divisiones que estaban semánticamente desalineados. Finalmente, ellas también plantean que en la división se puede llegar a un resultado sin necesidad de modelar una estrategia de solución que considere todos los elementos del problema matemático.

Estos resultados se pueden relacionar principalmente con el modelo teórico de congruencia semántica (SECO) (Gros et al., 2020). Particularmente,

con dos componentes de este modelo: la *semántica del mundo* y la *semántica de la matemática*. En primer lugar, el porcentaje de aciertos significativamente superior en el contexto categórica en comparación al contexto control de los participantes se puede explicar por el proceso de codificación experta, dependiente de la semántica de la matemática. En este proceso, los individuos utilizan su experticia en matemática para crear un algoritmo de solución adecuado sin necesidad de tomar en cuenta la semántica del mundo.

En este caso en particular, al ser los problemas matemáticos a resolver sencillos, los participantes podían inhibir la tendencia a resolver las divisiones como adiciones cuando los sustantivos se encontraban en relación categórica, enfocándose netamente en los componentes matemáticos (valores de los números y operación matemática). Esta estrategia de enfocarse solo en los elementos matemáticos también podría explicar por qué se obtuvieron porcentajes de acierto mayores que en el contexto control, debido a que habría un proceso de liberación de recursos cognitivos al no prestar atención a la información que no era relevante para resolver el problema matemático.

Una explicación similar se puede plantear para entender por qué el porcentaje de aciertos significativamente mayor en el contexto funcional en comparación al contexto control. En este caso, además de la capacidad para obviar los sustantivos de una división como explicaban Guthormsen et al. (2016)

y como teorizaban Gros y colegas (2020) para la codificación experta, la semántica del mundo puede crear un proceso de facilitación, ya que los participantes tuvieron que resolver divisiones, que se encuentran naturalmente alineadas con la relación funcional. Por lo tanto, la codificación inicial y experta del modelo teórico SECO (Gros et al., 2020) trabajarían en conjunto para acceder a la respuesta correcta. Finalmente, estas interacciones pueden explicar por qué no hubo diferencias significativas entre los porcentajes de acierto de los contextos categórica y funcional, a causa de que ambas se vieron beneficiadas por distintos procesos de codificación para llegar al algoritmo de solución adecuado.

Interesantemente, estos beneficios solo se vieron significativamente reflejados en los porcentajes de acierto de los contextos categórico y funcional, pero no se vieron reflejados en los tiempos de reacción. Esto podría explicarse al considerar que el tamaño del efecto esperado al momento de calcular la muestra de este estudio fue demasiado optimista, lo que puede afectar directamente la posibilidad de alcanzar los niveles de significancia necesarios.

A modo de resumen, a diferencia del estudio de valoración, se rechazó completamente la primera hipótesis, debido a que no hubo interacción entre grupos experimentales y porcentajes de acierto y tiempos de reacción. Por lo tanto, no se podía explorar las diferencias entre grupos de manera independiente.

Por el contrario, se aceptó parcialmente la segunda hipótesis. Ahora bien, para ahondar en la función ejecutiva de inhibición, los participantes del estudio de resolución de problemas matemáticos en relación categórica y funcional también realizaron la tarea experimental Stroop.

Discusión Estudio Stroop

En cuanto a la tarea experimental Stroop, se plantearon dos hipótesis. La primera hipótesis plantea que, debido a un mayor control ejecutivo, los expertos en matemática y los hablantes de alta proficiencia en inglés deberían obtener porcentajes de acierto superiores y tiempos de reacción menores que los hablantes de baja proficiencia en inglés. La segunda hipótesis plantea que, debido a un proceso de facilitación, todos los participantes obtendrían mayores porcentajes de acierto y menores tiempos de reacción en el contexto congruente.

En cuanto a los porcentajes de acierto de este estudio, luego de recolectar la información, se realizó una ANOVA de medidas repetidas para analizar si hubo una interacción entre grupos experimentales y contextos y si hubo efectos principales. Primero, se descartó que hubiese una interacción significativa entre grupos experimentales y contextos. Segundo, también se descartó que hubiese un efecto principal significativo al considerar los contextos experimentales congruente e incongruente.

En relación a los tiempos de reacción, también se realizó una ANOVA de medidas repetidas para analizar si hubo una interacción entre grupos experimentales y contextos y si hubo efectos principales. Este análisis indica que sí se encontraron diferencias significativas en los tiempos de reacción al considerar los contextos congruente e incongruente. En primer lugar, el análisis indica que sí hubo una interacción entre grupos experimentales y contextos. En segundo lugar, también hubo un efecto principal de contextos experimentales.

Luego de haber establecido que sí existe una interacción entre grupos experimentales y contextos, además de un efecto principal de contexto experimental, se realizaron tres pruebas t de Student para comparar cada grupo experimental y determinar si había diferencias en las medias de los grupos al considerar el contexto congruente y el contexto incongruente. Primero, se comparó los grupos de bajo y alto conocimiento en inglés. Esta comparación determinó que no hubo diferencias significativas entre estos grupos al considerar los tiempos de reacción en los contextos experimentales. Segundo, se comparó el grupo de bajo conocimiento en inglés con el grupo de alto conocimiento en matemática. Esta comparación determinó que no hubo diferencias entre los grupos al considerar el contexto congruente, pero que sí hubo diferencias significativas en el contexto incongruente. Particularmente, el grupo de alto conocimiento en matemática tuvo tiempos de reacción significativamente

menores que el grupo de bajo conocimiento en inglés. Finalmente, se comparó los grupos de alto conocimiento en matemática y alto conocimiento en inglés. Esta comparación también determinó que no hubo diferencias significativas en los tiempos de reacción en los dos contextos experimentales.

Considerando las hipótesis de este estudio, la primera hipótesis fue parcialmente aceptada, ya que el grupo de alto conocimiento en matemática tuvo tiempos de reacción menores que los participantes del grupo de bajo conocimiento en inglés en el contexto incongruente. La segunda hipótesis fue rechazada, en vista de que no hubo diferencias significativas en los porcentajes de acierto y tiempos de reacción que favoreciesen al contexto congruente.

En cuanto a estudios que comparen expertos en matemática con no expertos, estudios como el de Lubin et al. (2016) indican que no existen diferencias significativas en los porcentajes de acierto de los grupos experimentales, siendo este un resultado similar al nuestro. A pesar de esto, ellos sí reportaron diferencias significativas en los porcentajes de acierto de los contextos experimentales, donde los porcentajes de acierto del contexto congruente fueron significativamente superiores al contexto incongruente. En nuestro estudio, tampoco hubo diferencias significativas en relación a los porcentajes de acierto y los contextos experimentales.

Al tomar en cuenta los tiempos de reacción, Lubin et al. (2016) reportaron diferencias significativas entre las condiciones congruente e incongruente. Concretamente, los tiempos de reacción en la condición congruente fueron significativamente menores a los tiempos de reacción de la condición incongruente. Los resultados de nuestro estudio concuerdan parcialmente con estos resultados, ya que el grupo de bajo conocimiento en inglés tuvo tiempos de reacción significativamente mayores que el grupo de alto conocimiento en matemática. Junto con este resultado, Lubin et al. (2016) reportaron que las diferencias en tiempos de reacción también fueron significativas para los grupos experimentales, donde el grupo de expertos matemáticos tuvo menores tiempos de reacción en comparación al grupo de no expertos en matemática. En nuestro caso, nuestros resultados también concuerdan parcialmente, considerando que el grupo de alto conocimiento en matemática tuvo tiempos de reacción menores en el contexto congruente en comparación con el grupo de bajo conocimiento en inglés. Sin embargo, no hubo diferencias significativas al comparar el grupo de alto conocimiento en matemática con el grupo de alto conocimiento en inglés.

Estos resultados pueden explicarse si se consideran dos modelos teóricos. A nombrar, el modelo de activación interactiva bilingüe (Grainger et al., 2010) y el modelo SECO (Gros et al., 2020). Como se indicó con anterioridad, en ambos modelos, la experiencia de los individuos cumple una función clave al momento de inhibir información que no es relevante. En primer lugar, el modelo

de activación interactiva bilingüe (Grainger, 2010) intenta dar cuenta de la capacidad de los hablantes bilingües para acceder correctamente a las estructuras lingüísticas necesarias para comunicarse, evitando las interferencias que puede crear las lenguas que el hablante conoce. Para evitar estas interferencias, se plantea que la selección de la lengua a utilizar se realiza de manera descendente, donde distintos nodos interactúan para aumentar la activación de la lengua objetivo e inhibir la activación de la otra lengua. Después de identificar por pistas contextuales la probabilidad de que una palabra pertenezca a una lengua X, se activan los nodos de esa lengua, manteniendo activadas las representaciones lingüísticas de esa lengua. Esto no significa que la activación de la lengua Y es suprimida por completo, debido a que ambas lenguas se encontrarán activadas, pero en distintos grados. Por lo tanto, siempre habrá un grado de interferencia por parte de la lengua no-objetivo. Esta interferencia es limitada gracias al control inhibitorio descendente que plantea este modelo.

Grainger et al. (2010) incluso plantearon una actualización a este modelo que considera la realidad de los aprendientes adultos de una lengua extranjera. En el modelo de activación bilingüe interactiva del desarrollo, se plantea que la interferencia a la que se exponen los aprendientes es mayor cuando su nivel de proficiencia es más bajo y que esta se reduce a medida que el nivel de

proficiencia aumenta y no es necesario anclar la representación semántica de una palabra en la L2 a su traducción en la L1.

Aunque los estímulos de la tarea Stroop fueron en español para todos los grupos experimentales, se puede tomar en cuenta la actualización del modelo planteado por Grainger y colegas para bilingües al momento de explicar los resultados obtenidos. Al comparar el grupo de alto conocimiento en inglés y el grupo de alto conocimiento en matemática, se puede explicar el hecho de que no haya habido diferencias significativas en los porcentajes de acierto y tiempos de reacción, ya que el grupo de alto conocimiento en inglés tiene amplia experiencia inhibiendo de manera activa su L1 al momento de expresarse en la L2. Por lo tanto, esto los podría poner a la par con el grupo de alto conocimiento en matemática al momento de inhibir estímulos irrelevantes al completar una tarea experimental, incluso si esta tarea es en español. En contraste, el grupo de bajo conocimiento en inglés aún no tiene suficiente experiencia inhibiendo estímulos irrelevantes, lo que explica por qué tuvieron tiempos de reacción significativamente mayores que los tiempos de reacción del grupo de alto conocimiento en matemática, necesitando más tiempo para inhibir la información irrelevante.

En segundo lugar, el modelo SECO (Gros et al., 2020) tiene distintos componentes que interactúan en procesos de codificación de información para

llegar a un algoritmo de solución que exitosamente resuelva un problema matemático escrito. Hay un componente y un proceso de codificación que está estrechamente relacionado con el grupo de alto conocimiento en matemática.

Primero, el componente de la semántica de la matemática se enfoca en el conocimiento matemático que tienen los individuos. En el caso del grupo de alto conocimiento en matemática, este conocimiento es amplio y no solo da ventajas para resolver problemas matemáticos. Particularmente, la semántica de la matemática permitiría inhibir información del problema matemático escrito que no es relevante para llegar al algoritmo de solución de una manera más expedita.

Segundo, esta capacidad de inhibir información que no es relevante es clave para el proceso de codificación experta, donde los individuos utilizan su conocimiento de la semántica de la matemática para omitir las interpretaciones basadas en la semántica del mundo y generar un algoritmo de solución. Particularmente, esto se traduciría en un mayor control de la función ejecutiva de inhibición. Esto explicaría por qué el grupo de alto conocimiento en matemática tuvo tiempos de reacción significativamente menores que el grupo de bajo conocimiento en inglés para el contexto incongruente. El grupo de alto conocimiento en matemática, al estar constantemente inhibiendo información que no es relevante para llegar al algoritmo de solución que buscan, se han convertido en expertos en inhibir respuestas a estímulos irrelevantes para resolver la tarea

en cuestión de manera exitosa. Por eso, tuvieron tiempos de reacción significativamente menores en el contexto incongruente de la tarea Stroop, que se relaciona con nuestra capacidad para inhibir una respuesta ante un estímulo irrelevante.

Los supuestos teóricos de estos modelos permiten explicar las diferencias en tiempos de reacción, lo que sugiere que los expertos en matemática e inglés, no requieren tanto tiempo para inhibir los estímulos irrelevantes en comparación a los hablantes que están recién iniciando su camino para convertirse en expertos. Por lo tanto, los efectos de la manipulación experimental son notorios en los tiempos de reacción y no en los porcentajes de acierto de la tarea Stroop.

Conclusiones

Esta investigación psicolingüística se enfocó en estudiar cómo estudiantes universitarios evaluaron las relaciones semánticas (categóricas) y sintácticas (funcionales) entre sustantivos en español, si la resolución de problemas matemáticos escritos en español e inglés se ve afectado por el procesamiento semántico y sintáctico y determinar cuán eficiente es la capacidad de inhibición de los participantes que resolvieron problemas matemáticos escritos.

En relación a la valoración de relaciones semánticas y sintácticas, se aplicaron dos escalas de Likert a dos grupos de estudiantes universitarios para evaluar si la relación de los pares de sustantivos era semántica, categórica de acuerdo a Bassok y colegas (1995; 1998; 2001; y 2016), o si era sintáctica, funcional de acuerdo a Bassok y colegas (1995; 1998; 2001; y 2016). Una lista evaluó si la relación era semántica, mientras que la segunda lista evaluó si la relación era sintáctica. Luego, estas valoraciones fueron analizadas con *t* de Student. El propósito de realizar este estudio, expresado como objetivo específico, fue comparar las posibles diferencias significativas en los puntajes atribuidos a pares de palabras que se encuentran en una relación matemática categórica, en comparación a los puntajes atribuidos a pares de palabras en relación matemática funcional en español para estudiantes universitarios.

A modo de resumen, los puntajes de los pares de palabras en relación semántica fueron significativamente mayores que los puntajes de los pares de palabras en relación sintáctica en la lista de evaluación semántica. En el caso de la lista de evaluación sintáctica, los puntajes de los pares de palabras en relación sintáctica fueron significativamente mayores que los puntajes de los pares de palabras en relación semántica. Por lo tanto, los participantes efectivamente perciben estas dos relaciones de manera significativamente distinta. Estos resultados permitieron crear la batería lingüística en español necesaria para tener el material lingüístico para el estudio de resolución de problemas matemáticos en

relación semántica y relación funcional. Además, los resultados obtenidos permitieron aceptar ambas hipótesis planteadas para este estudio y cumplir el objetivo planteado a cabalidad.

En este segundo estudio de esta tesis doctoral, tres grupos de participantes (alto conocimiento en matemática, bajo conocimiento en inglés y alto conocimiento en inglés) tuvieron que resolver problemas matemáticos escritos en relación semántica y relación sintáctica. La manipulación experimental se enfocó en alinear o desalinear los pares de palabras que aparecían en el problema matemático. Particularmente, los pares de palabras en relación sintáctica se encontraban alineados con la división a resolver, mientras que los pares de palabras en relación semántica se encontraban desalineados. También se incluyó un contexto control, donde pares de palabras sin una relación semántica clara se utilizaron para resolver sustracciones, una operación matemática considerada neutra (Bassok y colegas, 1995; 1998; 2001; 2016). A la vez, el grupo de alto conocimiento en inglés resolvió los problemas matemáticos en inglés, mientras que los otros dos grupos los resolvieron en español. El propósito de este estudio, expresado como objetivo específico, fue comparar la relación categórica y funcional de frases lingüísticas en términos de porcentajes de acierto y tiempos de reacción en estudiantes con distinta proficiencia lingüística y matemática.

Los resultados a destacar son a) no hubo diferencias significativas entre los distintos grupos del estudio; b) hubo diferencias significativas en los porcentajes de acierto de los distintos contextos experimentales; c) hubo diferencias significativas entre los porcentajes de acierto del contexto sintáctico y los porcentajes de acierto del contexto control; d) hubo diferencias significativas entre los porcentajes de acierto del contexto semántico y los porcentajes de acierto del contexto control; y e) no hubo diferencias significativas entre grupos del estudio ni en los contextos experimentales al considerar los tiempos de reacción.

Debido a estos resultados, la primera hipótesis fue rechazada puesto que no hubo diferencias significativas entre los distintos grupos del estudio y contextos experimentales. En relación a la segunda hipótesis, esta fue parcialmente aceptada, ya que se esperaba que el contexto sintáctico tuviese porcentajes de acierto significativamente mayores en comparación al contexto semántico y contexto control, lo que solo se vio reflejado al compararlo con el contexto control. También se esperaba tiempos de reacción menores para el contexto sintáctico por una posible facilitación por estar alineada con la división, pero no hubo diferencias significativas. A pesar de estos resultados no significativos, se logró el objetivo de comparar los porcentajes de acierto y tiempos de reacción de los participantes.

Finalmente, en el tercer estudio de esta tesis doctoral, los participantes de tres grupos (alto conocimiento en matemática, bajo conocimiento en inglés y alto conocimiento en inglés) tuvieron que completar la tarea Stroop. La manipulación experimental se enfocó en crear contextos congruentes (facilitación) e incongruente (inhibición), donde la palabra que indicaba un color y su tinta coincidían en el contexto congruente, mientras que para el contexto incongruente la palabra que indicaba un color y su tinta no correspondían. El propósito de este estudio, expresado como objetivo específico, fue analizar los efectos cognitivos de inhibición en estudiantes con diferentes proficiencias lingüísticas y matemáticas.

Los resultados principales fueron que a) no hubo diferencias significativas en los porcentajes de acierto al comparar los distintos grupos y contextos experimentales; b) hubo una interacción significativa entre grupos y contextos experimentales en los tiempos de reacción; y c) hubo diferencias en los tiempos de reacción del contexto incongruente entre el grupo de alto conocimiento matemático y bajo conocimiento en inglés.

Considerando estos resultados, se aprobó parcialmente la primera hipótesis, puesto que el grupo de alto conocimiento en matemática tuvo tiempos de reacción significativamente menores en comparación al grupo de bajo conocimiento en inglés en el contexto incongruente. Al no haber diferencias

significativas en los porcentajes de acierto ni tiempos de reacción significativamente menores que favoreciesen al contexto congruente, se rechazó la segunda hipótesis de este estudio. Al igual que en el estudio anterior, a pesar de que no se pudieron aprobar las hipótesis planteadas, se logró el objetivo ya que se pudo analizar los efectos cognitivos de inhibición en los participantes.

Contribuciones

En cuanto a las contribuciones de esta tesis doctoral, primero se destaca la batería lingüística creada para estudiar el fenómeno de desalineación semántica en español. Como se indicó anteriormente, de acuerdo a nuestro conocimiento, no existe una batería lingüística para estudiar este fenómeno en español. Junto con esto, este material lingüístico fue adaptado al español de Chile, se crearon nuevos pares de palabras como estímulos y fue validado luego de analizar los resultados de las valoraciones realizadas en escala de Likert por 80 estudiantes universitarios.

Luego, se destaca que no se habían realizado estudios sobre la resolución de problemas matemáticos en relación semántica y relación sintáctica en español y que comparasen los resultados de grupos con distinto conocimiento de inglés y alto conocimiento en matemática. A pesar de que sólo una hipótesis fue parcialmente aprobada, las relaciones lingüísticas sí tuvieron porcentajes de

acierto significativamente mayores al ser comparadas individualmente con el contexto neutral control. Esto sugiere que nuestra percepción de cuán estrecha es la relación lingüística entre pares de palabras tiene un efecto en cómo las procesamos al momento de resolver problemas matemáticos escritos. Esta evidencia podría ayudar a mejorar el material de instrucción de grupos desaventajados al momento de enfrentarse a problemas matemáticos escritos.

Finalmente, en la tarea experimental Stroop solo hubo diferencias significativas en los tiempos de reacción del contexto incongruente para el grupo de alto conocimiento matemático en comparación al grupo de bajo conocimiento en inglés, donde el grupo de alto conocimiento matemático obtuvo tiempos de reacción significativamente menores al momento de inhibir su respuesta ante el estímulo incongruente. A la vez, se destaca el rol de la experiencia al momento de inhibir información que no es relevante para completar la tarea experimental. Particularmente, los grupos de alto conocimiento en matemática y alto conocimiento en inglés no tuvieron diferencias significativas en porcentajes de acierto y tiempos de reacción, lo que sugiere que la capacidad de inhibir información irrelevante en sus propios campos de estudio tiene un efecto al momento de completar la tarea Stroop.

Limitaciones metodológicas

En cuanto a las limitaciones metodológicas de esta tesis doctoral, hay dos. La primera es la posibilidad de que el número de participantes necesarios haya sido calculado con un tamaño del efecto muy optimista, lo que afectó los niveles de significancia. Esto ocurrió particularmente en el estudio de resolución de problemas matemáticos escritos en contexto funcional y contexto categórico y en el estudio Stroop. Como ya se indicó con anterioridad, en cada estudio una hipótesis fue rechazada por completo, mientras que la otra hipótesis fue parcialmente aceptada. Esto sugiere que, de haber tenido un mayor número de participantes, se podría haber alcanzado los niveles de significancia necesarios para aceptar las hipótesis de cada uno de estos estudios.

La segunda limitación es que, debido a la pandemia, la recolección de datos fue realizada en línea para el estudio de resolución de problemas matemáticos escritos en contexto funcional y contexto categórico y en el estudio Stroop. A pesar de que en cada sesión de recolección de datos los participantes compartieron sus pantallas para monitorear que no hubiese problemas, hubo una participante cuyos datos no fueron considerados, ya que no fueron grabados por el sitio web utilizado. Junto con esto, las condiciones en las que los participantes se encontraban, iluminación, ruido ambiente, distancia de la pantalla, entre otras,

no pudieron ser controlados de la misma manera que si los datos hubiesen sido recolectados en un laboratorio.

Junto con eso, no se puede descartar el efecto techo que puede haber aparecido en esta tarea experimental ya que la mayoría de los participantes tuvo puntajes muy altos, donde la moda fue de 100 en todos los grupos experimentales, lo que dificulta la distribución normal de los datos y sus análisis (Salkind, 2010).

Futuras Investigaciones

Finalmente, para futuras investigaciones, se propone lo siguiente. En primer lugar, aumentar el número de participantes para el estudio de resolución de problemas matemáticos escritos en contexto funcional y contexto categórico y en el estudio Stroop para evaluar si esto permite alcanzar los niveles de significancia necesarios para aceptar las hipótesis. En segundo lugar, realizar los estudios de resolución de problemas matemáticos escritos y Stroop con estímulos en inglés y en español para estudiar cómo procesan estos estímulos los participantes de bajo conocimiento en inglés y de alto conocimiento en inglés. En tercer lugar, se podría realizar un estudio de electrofisiología para estudiar qué componentes se registran cuando se evalúa si la relación entre pares de palabras es semántica (categórica) o sintáctica (funcional). En cuarto lugar, se

podría realizar un estudio de electrofisiología de resolución de problemas matemáticos escritos en contexto funcional y contexto categórico con grupos de participantes con bajo conocimiento en inglés, alto conocimiento en inglés y alto conocimiento en matemática para determinar qué componentes se registran y cómo se relacionan con otros componentes que se registran en tareas de inhibición. En quinto lugar, se podría realizar un estudio de neuroimagen para estudiar qué patrones de activación cerebral existen al momento de evaluar si la relación entre pares de palabras es semántica (categórica) o sintáctica (funcional).

Bibliografía

- Abedi, J. & Lord, C. (2001). The Language Factor in Mathematics Tests. *Applied Measurement in Education*, 14(3), 219-234.
http://dx.doi.org/10.1207/S15324818AME1403_
- Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research*, 74(1), 1-28.
- Acosta, B. D., Rivera, C., & Willner, L. S. (2008). *Best practices in state assessment policies for accommodating English language learners: A delphi study*. Arlington: Center for Equity and Excellence in Education, the George Washington University
- Alvarez, J., & Emory, E. (2006). Executive Function and the Frontal Lobes: A Meta-Analytic Review. *Neuropsychology Review*, 16(1), 17-42.
<http://dx.doi.org/10.1007/s11065-006-9002-x>
- Ashcraft, M. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, 44 (1-2), 75–106.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International journal of psychological research*, 3(2), 12-28.
- Baker, D., Basaraba, D. & Polanco, P. (2016). Connecting the Present to the Past: Furthering the Research on Bilingual Education and Bilingualism. *Review*

of *Research in Education*, 40(1), 821-883.

<https://doi.org/10.3102/0091732X16660691>

Banks, K., Jeddeeni, A. & Walker, C. (2016). Assessing the Effect of Language Demand in Bundles of Math Word Problems. *International Journal of*

Testing, 16(4) 269-287. <https://doi.org/10.1080/15305058.2015.1113972>

Barbu, O. C., & Beal, C. R. (2010). Effects of linguistic complexity and math difficulty on word problem solving by English learners. *International Journal*

of Education, 2(2), 1-19.

Barendregt, J. & Doi, S. (2016). MetaXL User Guide (Version 5.3). EpiGear International Pty Ltd.

Bassok, M. (2001). Semantic Alignments in Mathematical Word Problems. In Gentner, D., Holyoak, K. and Kolinov, K. (eds). *The Analogical Mind Perspectives from Cognitive Science* (pp. 401-433). Bradford Books.

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology*, 35(2), 99–134. <https://doi.org/10.1006/cogp.1998.0675>

Bassok, M., Pedigo, S. F., & Oskarsson, A. T. (2008). Priming addition facts with semantic relations. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition, 34(2), 343–352. <https://doi.org/10.1037/0278-7393.34.2.343>

- Bassok, M., Wu, L.-L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition*, 23(3), 354–367. <https://doi.org/10.3758/BF03197236>
- Bautista, D., Mitchelmore, M. & Mulligan, J. (2009) Factors influencing Filipino children's solutions to addition and subtraction word problems. *Educational Psychology*, 29(6), 729-745.
- Bechtold, L., Bellebaum, C., Egan, S., Tettamanti, M., & Ghio, M. (2019). The role of experience for abstract concepts: Expertise modulates the electrophysiological correlates of mathematical word processing. *Brain and Language*, 188, 1-10.
- Bermúdez, J. (2014) *Cognitive Science: An Introduction to the Science of the Mind*. (2da ed). Cambridge University Press.
- Bernardo, A. (2002). Language and Mathematical Problem Solving Among Bilinguals. *The Journal of Psychology*, 136(3), 283-297. <http://dx.doi.org/10.1080/00223980209604156>
- Bernardo, A., & Calleja, M. (2005). The Effects of Stating Problems in Bilingual Students' First and Second Languages on Solving Mathematical Word Problems. *The Journal Of Genetic Psychology*, 166(1), 117-129. <http://dx.doi.org/10.3200/gntp.166.1.117-129>
- Brunellière 1, A., & Hans Frauenfelder, U. (2014). On the locus of grammatical context effects on word recognition. *L'Année psychologique*, 114(3), 447-467.

- Campbell, A., Adams, V., & Davis, G. (2007). Cognitive demands and second language learners: A framework for analyzing mathematics instructional contexts. *Mathematical Thinking and Learning*, 9(1), 3-30.
- Carter, B. T., Foster, B., Muncy, N. M., & Luke, S. G. (2019). Linguistic networks associated with lexical, semantic and syntactic predictability in reading: A fixation-related fMRI study. *NeuroImage*, 189, 224-240.
- Castelain, T., & Marín, B. (2014). *Psicología Experimental: Fundamentos y Protocolos experimentales para el estudio de procesos cognitivos. San José (CR): Universidad de Costa Rica.*
- Chikara, R. K., & Ko, L. W. (2020). Global neural activities changes under human inhibitory control using translational scenario. *Brain Sciences*, 10(9), 640.
- Chung, S., Chen, X., & Geva, E. (2019) Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics*, 50, 149-161.
<https://doi.org/10.1016/j.jneuroling.2018.01.003>
- Clementson, T., Tilbury, A., Hendra, L., Rea, D., Doff, A., & Goldstein, B. (2010). *English Unlimited placement test.* Cambridge University Press.
- Coe, R. (2002). *It's the Effect Size, Stupid. What Effect Size Is and Why It Is Important* [Paper]. The British Educational Research Association Annual Conference, Exeter, U.K.
<http://www.leeds.ac.uk/educol/documents/00002182.htm>

- College Board (2023, Diciembre 20) *SAT Suite Annual Report 2023*. College Board. <https://reports.collegeboard.org/sat-suite-program-results>
- Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion Volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Creswell, J. W. (2012). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Pearson
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics*, 50, 149-161.
- Daroczy, G., Wolska, M., Meurers, W., & Nuerk, H. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Frontiers In Psychology*, 06. <http://dx.doi.org/10.3389/fpsyg.2015.00348>
- Davis, C.J., & Perea, M. (2005) BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods* 37(4), 665–671. <https://doi.org/10.3758/BF03192738>

- Declerck, M. & Philipp, A. (2018). Is inhibition implemented during bilingual production and comprehension? n-2 language repetition costs unchained. *Language, Cognition and Neuroscience*, 33(5), 608-617. <https://doi.org/10.1080/23273798.2017.1398828>
- Declerck, M., Thoma, A. M., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition: Evidence from n-2 language repetition costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1911–1916. <https://doi.org/10.1037/xlm0000138>
- Driver, M. & Powell, S. (2017). Culturally and Linguistically Responsive Schema Intervention: Improving Word Problem Solving for English Language Learners With Mathematics Difficulty. *Learning Disability Quarterly*, 40(1), 41-53.
- Duque de Blas, G., Gómez-Veiga, I., & García-Madruga, J. (2021). Arithmetic Word Problems Revisited: Cognitive Processes and Academic Performance in Secondary School. *Education Sciences*, 11(4), 155.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191
- Federmeier, K. D., Jongman, S. R., & Szwedczyk, J. M. (2020). Examining the Role of General Cognitive Skills in Language Processing: A Window Into Complex Cognition. *Current Directions in Psychological Science*, 29(6), 575–582. <https://doi.org/10.1177/0963721420964095>

- Field, A. & Gillet, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63(3), 665-694.
<https://doi.org/10.1348/000711010X502733>
- Fricke, M., Zirnstein, M., Navarro-Torres, C., & Kroll, J. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, 22(1), 200-207.
<https://doi.org/10.1017/S1366728918000482>
- Gajewski, P. D., Falkenstein, M., Thönes, S., & Wascher, E. (2020). Stroop task performance across the lifespan: High cognitive reserve in older age is associated with enhanced proactive and reactive interference control. *NeuroImage*, 207, 116430.
- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, 42(3), 237-288.
- Gómez, DM., Jiménez, A., Bobadilla, R., Reyes, C., y Dartnell, P. (2014). Exploración de la comprensión de fracción en niños en edad escolar. En S. Oesterle, P. Liljedahl, C. Nicol, y D. Allan (Eds.), *Actas de la reunión conjunta de la PME y 38 PME-NA 36* (Vol. 3, pp. 185-192). Vancouver, Canadá: PME.
- Gómez, D., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction

comparison in middle school children. *ZDM*, 47(5), 801-811.

<http://dx.doi.org/10.1007/s11858-015-0685-4>

Grainger, J. & Dijkstra, T. 1992. On the representation and use of language information in bilinguals. In *Cognitive Processing in Bilinguals. Advances in Psychology*, R. J. Harris (ed.), 207–220. Amsterdam: North-Holland.

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). *Language acquisition across linguistic and cognitive systems*, 52, 267-283.

Gros, H., Thibaut, J. P., & Sander, E. (2020). Semantic congruence in arithmetic: A new conceptual model for word problem solving. *Educational Psychologist*, 55(2), 69-87.

Guthormsen, A. (2007). *Conceptual integration of mathematical and semantic knowledge* (Tesis Doctoral, University of Washington).

Guthormsen, A., Fisher, K., Bassok, M., Osterhout, L., DeWolf, M., & Holyoak, K. (2016). Conceptual Integration of Arithmetic Operations With Real-World Knowledge: Evidence From Event-Related Potentials. *Cognitive Science*, 40(3), 723-757. <http://dx.doi.org/10.1111/cogs.12238>

Gutiérrez-Rexach, J. (2011). The Syntax-Semantics Interface in Grammar. *Studies in Hispanic and Lusophone Linguistics*, 4(1), 209-220.

- Hackl, M. (2013). The syntax–semantics interface. *Lingua*, 130, 66–87.
<https://doi.org/10.1016/j.lingua.2013.01.010>
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of cognitive neuroscience*, 15(6), 883-899.
- Hammer, A., Jansma, B. M., Lamers, M., & Münte, T. F. (2008). Interplay of meaning, syntax and working memory during pronoun resolution investigated by ERPs. *Brain research*, 1230, 177-191.
- Hernandez, R., Fernandez, C, & Baptista, P. (2014). Metodología de la Investigación (6ta ed.). McGraw-Hill.
- Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558
- Kang, W., Hernández, S. P., Rahman, M. S., Voigt, K., & Malvaso, A. (2022). Inhibitory control development: a network neuroscience perspective. *Frontiers in Psychology*, 13, 651547.
- Kempert, S., Saalbach, H. & Hardy, I. (2011). Cognitive Benefits and Costs of Bilingualism in Elementary School Students: The Case of Mathematical Word Problems. *Journal of Educational Psychology*, 103(3), 547-561.
- Kieffer, M., Lesaux, N., Rivera, M. & Francis, D. (2009). Accommodations for English Language Learners Taking Large-Scale Assessments: A Meta-Analysis on Effectiveness and Validity. *Review of Educational Research*, 79(3), 1168-1201. <https://doi.org/10.3102/0034654309332490>

- Kieffer, M., Rivera, M., Francis, D.J. (2012). *Research-based recommendations for the use of accommodations in large-scale assessments: 2012 update*. Practical Guidelines for the Education of English Language Learners. Book 4. New York: Center on Instruction.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 46, 75–86.
- Kroll, J. & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25(5), 497-514. <https://doi.org/10.1080/20445911.2013.799170>
- Kunda, Z., & Spencer, S. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, 129(4), 522-544. <http://dx.doi.org/10.1037/0033-2909.129.4.522>
- Lager, C. (2006). Types of Mathematics-Language Reading Interactions that Unnecessarily Hinder Algebra Learning and Assessment. *Reading Psychology*, 27(2-3), 165-204.

- Lee, H. S., DeWolf, M., Bassok, M., & Holyoak, K. J. (2016). Conceptual and procedural distinctions between fractions and decimals: A cross-national comparison. *Cognition*, *147*, 57-69.
- LeFevre, J.-A., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory & Cognition*, *16*(1), 45–53. <https://doi.org/10.3758/BF03197744>
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of educational research*, *66*(4), 579-619.
- Lubin, A., Rossi, S., Lanoë, C., Vidal, J., Houdé, O., & Borst, G. (2016). Expertise, inhibitory control and arithmetic word problems: A negative priming study in mathematics experts. *Learning and Instruction*, *45*, 40-48.
- Luce, R. D. (1991). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press.
- MacLeod, C. (2007). In D. Gorfein & C. MacLeod, *Inhibition in cognition* (1st ed., pp. 3-23). American Psychological Association
- Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior* *6*(5), 729–763.
- Martin, B. & Fuchs, L. (2019). The Mathematical Performance of At-Risk First Graders as a Function of Limited English Proficiency Status. *Learning*

Disability Quarterly, 42(4), 244-251.

<https://doi.org/10.1177/0731948719827489>

McClelland, J. L., & Cleeremans, A. (2009). Connectionist models. In T. Byrne, A. Cleeremans, & P. Wilken (Eds.), *Oxford companion to consciousness*.

New York: Oxford University Press

Mertins, B. (2016). The use of experimental methods in linguistic research: Advantages, problems and possible pitfalls. *Slavic Languages in Psycholinguistics*, 15-33.

MINEDUC (2023, Diciembre 20) *Resultados SIMCE 2022*. MNEDUC.

<https://www.mineduc.cl/resultados-del-simce-2022/>

Müller, U., & Kerns, K. (2015). The development of executive function. *Handbook of child psychology and developmental science: Cognitive processes*, 2, 571-623. Willey.

Newell, A. & Simon, H. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, 19(3), 113–126.

<https://doi.org/10.1145/360018.360022>

Ní Ríordáin, M., & O'Donoghue, J. (2009). The relationship between performance on mathematical word problems and language proficiency for students learning through the medium of Irish. *Educational Studies in Mathematics*, 71(1), 43-64.

Obermeyer, J., Schlesinger, J. & Martin, N. (2020). Evaluating the Contribution of Executive Functions to Language Tasks in Cognitively Demanding

- Contexts. *American Journal of Speech-Language Pathology*, 29(1), 463-473. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0216
- Ockey, G. (2007). Investigating the Validity of Math Word Problems for English Language Learners with DIF. *Language Assessment Quarterly*, 4(2), 149-164
- Orosco, M. J., & Abdulrahim, N. A. (2018). Examining comprehension strategy instruction with English learners' problem solving: Study findings and educator preparation implications. *Teacher Education and Special Education*, 41(3), 215-228.
- Paas, F., Renkl, A. & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1-4. https://doi.org/10.1207/S15326985EP3801_1
- Pavón, V. & Cabezuelo, R. (2019). Analysing mathematical word problem solving with secondary education CLIL students: A pilot study. *Latin American Journal Of Content & Language Integrated Learning*, 12(1), 18-45. <https://doi.org/10.5294/lacilil.2019.12.1.2>
- Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, & E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, 51, 195-203. <https://doi.org/10.3758/s13428-018-01193-y>

- Pennock-Roman, M., & Rivera, C. (2011). Mean effects of test accommodations for ELLs and non-ELLs: A meta-analysis of experimental studies. *Educational Measurement: Issues and Practice*, 30(3), 10-28.
- OECD (2023), *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/53f23881-en>.
- Powell, S., Berry, K. & Tran, L. (2020). Performance Differences on a Measure of Mathematics Vocabulary for English Learners and Non-English Learners with and without Mathematics Difficulty. *Reading and Writing Quarterly*, 36(2) 124-141. <https://doi.org/10.1080/10573569.2019.1677538>
- Qiu, K. & Wang, Y. (2021). Conceptual distinctions and preferential alignment across rational number representations. *European Journal of Psychology of Education*, 36, 865–881 <https://doi.org/10.1007/s10212-020-00502-4>
- Rios, J., Ihlenfeldt, S. & Chavez, C. (2020). Are Accommodations for English Learners on State Accountability Assessments Evidence-Based? A Multistudy Systematic Review and Meta-Analysis. *Educational Measurement Issues and Practices*, 39(4), 65-75. <https://doi.org/10.1111/emip.12337>
- Ronasi, G., Fischer, M. H., & Zimmermann, M. (2018). Language and arithmetic: A failure to find cross cognitive domain semantic priming between exception phrases and subtraction or addition. *Frontiers in Psychology*, 9, 1524.

- Sadowsky, S., & Martínez-Gamboa, R. (2012). LIFCACH 2.0: Word Frequency List of Chilean Spanish (Lista de Frecuencias de Palabras del Castellano de Chile), versión 2.0. Zenodo. <http://doi.org/10.5281/zenodo.268043>
- Salkind, N. J. (Ed.) (2010). Encyclopedia of research design. SAGE Publications, Inc., <https://doi.org/10.4135/9781412961288>
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. doi:10.1016/j.neunet.2014.09.003
- Sireci, S., Li, S., & Scarpati, S. (2003). The effect of test accommodation on test performance: A review of the literature (Research Report 495). Amherst: University of Massachusetts School of Education, Center for Educational Assessment.
- Smith, E. E., & Kosslyn, S. M. (2008). *Procesos cognitivos: modelos y bases neurales* (Vol. 16). Madrid: Pearson Educación.
- Star, J.R., & Pollack, C. (2015) Inhibitory control and mathematics learning: definitional and operational considerations. *ZDM Mathematics Education* 47(5), 859–863. <https://doi.org/10.1007/s11858-015-0716-1>
- StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Swanson, H., Kong, J., Moran, A. & Orosco, M. (2019). Paraphrasing Interventions and Problem-Solving Accuracy: Do Generative Procedures Help English Language Learners with Math Difficulties? *Learning*

Disabilities Research and Practice, 34(2), 68-84.

<https://doi.org/10.1111/ldrp.12194>

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning.

Cognitive Science, 12 (2), 257–285. [https://doi.org/10.1016/0364-](https://doi.org/10.1016/0364-0213(88)90023-7)

[0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)

Trakulphadetkrai, N., Courtney, C., Clenton, J., Treffers-Daller, J., & Tsakalaki,

A. (2020). The contribution of general language ability, reading comprehension and working memory to mathematics achievement among children with English as additional language (EAL): an exploratory study.

International Journal of Bilingual Education and Bilingualism, 23(4), 473-

487. <https://doi.org/10.1080/13670050.2017.1373742>

Tyumeneva, Y., Larina, G., Alexandrova, E., DeWolf, M., Bassok, Keith. M., &

Holyoak, J. (2018). Semantic alignment across whole-number arithmetic and rational numbers: evidence from a Russian perspective. *Thinking &*

Reasoning, 24(2), 198-220, doi: 10.1080/13546783.2017.1374307

Van Dooren, W. & Inglis, M. (2015). Inhibitory control in mathematical thinking,

learning and problem solving: a survey). *ZDM Mathematics Education*,

47(5), 713–721. <https://doi.org/10.1007/s11858-015-0715-2>.

Verschaffel, L., Schukajlow, S., Star, J. & Van Dooren, W. (2020). Word problems

in mathematics education: a survey. *ZDM Mathematics Education*, 52(1),

1–16. <https://doi.org/10.1007/s11858-020-01130-4>

- Wang, Y. (2021). *Generic pronouns in Mandarin Chinese* (Tesis Doctoral, University of York).
- Wang, Y., Gao, J., Wang, T., Huang, B., Feng, H. & Si, J. (2023). The influence of semantic alignment on the performance of addition and division operation: age-related differences. *Cognitive Processing*, 24, 245–252. <https://doi.org/10.1007/s10339-023-01125-5>
- Wood, J., Mathews, A., & Dalgleish, T. (2001). Anxiety and cognitive inhibition. *Emotion*, 1(2), 166-181. <http://dx.doi.org/10.1037/1528-3542.1.2.166>

Anexos

Modelo Consentimiento Informado



Universidad de Concepción

Dirección de Postgrado
Facultad de Humanidades y Arte
Doctorado en Lingüística

El efecto de la proficiencia en la L2 y la función ejecutiva de inhibición sobre el procesamiento aritmético en oraciones con relación categórica y funcional: evidencias empíricas con estudios conductuales

Investigador responsable: Martín Flores Quiroz

Profesora guía: Dra. Mabel Urrutia Martínez– Universidad de Concepción

DOCUMENTO DE CONSENTIMIENTO INFORMADO PARA LA PARTICIPACIÓN EN EL ESTUDIO DE INVESTIGACIÓN TITULADO: “EL EFECTO DE LA PROFICIENCIA EN LA L2 Y LA FUNCIÓN EJECUTIVA DE INHIBICIÓN SOBRE EL PROCESAMIENTO ARITMÉTICO EN ORACIONES CON RELACIÓN CATEGÓRICA Y FUNCIONAL: EVIDENCIAS EMPÍRICAS CON ESTUDIOS CONDUCTUALES”. UNA INVESTIGACIÓN PARA OPTAR AL GRADO DE DOCTOR EN LINGÜÍSTICA DE LA UNIVERSIDAD DE CONCEPCIÓN.

Antes de autorizar su participación en este estudio, usted tiene el derecho de obtener toda la información relativa a los procedimientos que se utilizarán en el mismo. En estas páginas se le proporciona toda la información que deberá leer detenidamente. No dude en preguntar a la investigadora responsable si tiene alguna duda o necesita alguna aclaración bien sea antes, durante o después de leer este documento.

INVESTIGADOR RESPONSABLE:

Nombre: Flores Quiroz, Martín Andrés

RUT: 16.480.428-3

Institución: Universidad de Concepción. Facultad de Humanidades y Arte

E-mail: mfloresq@udec.cl/martin.flores@ucentral.cl

PROFESORA GUIA:

Nombre: Urrutia Martínez, Mabel Alejandra
Institución: Universidad de Concepción. Facultad de Educación
E-mail: maurrutia@udec.cl

INTRODUCCION

Se invita a Ud. a participar en el estudio titulado: “el efecto de la proficiencia en la L2 y la función ejecutiva de inhibición sobre el procesamiento aritmético en oraciones con relación categórica y funcional: evidencias empíricas con estudios conductuales”. En el marco de una investigación para optar al grado de Doctor en Lingüística de la Universidad de Concepción.

OBJETIVO DEL ESTUDIO

En este se quiere estudiar cómo la proficiencia en una segunda lengua, en este caso inglés, puede afectar la manera en que se procesan los problemas matemáticos escritos.

PROCEDIMIENTO

Si Ud. acepta participar en esta investigación, se le comunicará oportunamente el día y la hora en que debe presentarse para realizar la sesión experimental, el lugar físico donde le será aplicado el procedimiento corresponde al Laboratorio de Idiomas de la Facultad de Educación y Ciencias Sociales de la Universidad de Central de Chile. Para poder realizar la medición de respuestas conductuales, usted deberá estar sentado frente a la pantalla de un computador y que interactúe con un teclado. Luego se le solicitará que realice una tarea de operaciones matemáticas en el computador mientras sus respuestas son registradas por el computador.

BENEFICIOS

Se espera que su participación en este estudio contribuya al estudio interdisciplinar del conocimiento matemático y el procesamiento del lenguaje en segundas lenguas.

RIESGOS

La investigación se llevará a cabo mediante dos estudios conductuales, donde se registrará las respuestas conductuales ante diversos estímulos que serán presentados en forma de palabras en una pantalla de computador. El estudio en esta forma provee medios no invasivos de evaluar el funcionamiento de los mecanismos psicológicos de comprensión del lenguaje, por lo que no se identifican riesgos para su salud ni traerá alguna consecuencia negativa producto de su participación.

CONFIDENCIALIDAD

La identidad de cada participante en este estudio se mantendrá de forma confidencial, no se revelará bajo ninguna circunstancia y tampoco aparecerá su nombre en ningún informe o publicación derivada de este estudio. El investigador responsable custodiará los datos del estudio, identificando en un código con claves los nombres de cada participante y resguardando la información en su computador

COSTOS

Su participación en la investigación no tendrá costo alguno para usted.

COMPENSACIÓN

Ud. recibirá (si así lo desea) una compensación económica por su colaboración en este estudio. Usted recibirá un pago de \$2.000.- en efectivo para compensarle por gastos de locomoción y colación. Para poder recibir este dinero, necesitaremos que rellene un recibo simple su nombre, RUT y firma.

DERECHO A RETIRARSE DEL ESTUDIO

Su participación en el estudio es libre y voluntaria, teniendo derecho a retirarse de la investigación en cualquier momento que desee.

CONTACTOS

En cualquier momento, podrá solicitar información adicional al investigador responsable Martín Flores Quiroz, Magíster en Lingüística (teléfono 951378727, mfloresq@udec.cl/martin.flores@ucentral.cl) sobre cualquier duda o aclaración que necesite.

Si usted tiene alguna pregunta acerca de sus derechos como participante en esta investigación o siente vulnerados sus derechos, usted puede llamar al Presidente u otro miembro del Comité de Ética de la Universidad de Concepción (Dr. José Becerra Allende, al fono: (41) 2204302).

DECLARACIÓN DE CONSENTIMIENTO INFORMADO

Yo, D./Dña. _____,

declaro que: he leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre las características del estudio.

He recibido suficiente información sobre el estudio y he entendido dicha información.

Comprendo que mi participación es libre y voluntaria.

Comprendo que puedo retirarme del estudio en cualquier momento.

Presto libremente mi conformidad a participar de la investigación.

Mi consentimiento es libre y voluntario y no ha sido forzado ni obligado.

Fecha

Firma del participante

Firma y RUT del investigador responsable

Firma del Director del Programa de Doctorado en
Lingüística.

Escala de Likert Evaluación Funcional

Instrucciones:

En este cuestionario, tu tarea consiste en evaluar hasta qué punto existe una relación de dependencia entre los siguientes pares de palabras. Para realizar la evaluación, debes calificar de 1 a 7 los pares de palabras, donde 1 equivale a una relación de dependencia baja y 7 a una relación de dependencia alta. Por favor, marca con una X la casilla correspondiente y trata de utilizar todo el rango de calificaciones disponible.

A continuación, te daremos algunos ejemplos antes de comenzar con la evaluación:

Relación de dependencia alta: camisa y colgador. El colgador tiene una relación de dependencia alta con la camisa, ya que su única función es colgar ropa. Por lo tanto, un posible puntaje podría ser 7.

Relación de dependencia intermedia: camisa y sillón tienen una relación de dependencia intermedia, ya que en ciertos contextos el sillón podría ser utilizado para dejar ropa, pero su uso principal es para sentarse. Por lo tanto, un posible puntaje podría ser 4.

Relación de dependencia baja: fútbol y tenis tienen una relación de dependencia baja ya que ambos deportes se pueden realizar de manera independiente. Por lo tanto, un posible puntaje podría ser 1.

Palabra 1	Palabra 2	Calificación						
		menos dependencia			mayor dependencia			
maestras	alumnas	1	2	3	4	5	6	7
arqueros	defensas	1	2	3	4	5	6	7
cellos	violines	1	2	3	4	5	6	7
galletas	frascos	1	2	3	4	5	6	7
buses	trenes	1	2	3	4	5	6	7
gorriones	loros	1	2	3	4	5	6	7
tazas	bandejas	1	2	3	4	5	6	7
soldados	marinos	1	2	3	4	5	6	7
sillas	mesas	1	2	3	4	5	6	7
registros	archivos	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		menos dependencia			mayor dependencia			
platos	bols	1	2	3	4	5	6	7
pastores	templos	1	2	3	4	5	6	7
alcaldes	vecinos	1	2	3	4	5	6	7
cobras	terrarios	1	2	3	4	5	6	7
leones	cubiles	1	2	3	4	5	6	7
plumones	crayones	1	2	3	4	5	6	7
mucamas	hoteles	1	2	3	4	5	6	7
diamantes	pulseras	1	2	3	4	5	6	7
claveles	rosas	1	2	3	4	5	6	7
chefs	casinos	1	2	3	4	5	6	7
monjas	conventos	1	2	3	4	5	6	7
aviones	hangares	1	2	3	4	5	6	7
hamsters	jaulas	1	2	3	4	5	6	7
cabras	rebaños	1	2	3	4	5	6	7
naranjas	canastos	1	2	3	4	5	6	7
flautas	músicos	1	2	3	4	5	6	7
clavos	martillos	1	2	3	4	5	6	7
químicos	pipetas	1	2	3	4	5	6	7
osos	cuevas	1	2	3	4	5	6	7
láminas	fotos	1	2	3	4	5	6	7
árboles	parques	1	2	3	4	5	6	7
canoas	muelles	1	2	3	4	5	6	7
dentistas	clínicas	1	2	3	4	5	6	7
sartenes	hornos	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		menos dependencia			mayor dependencia			
ranas	pozas	1	2	3	4	5	6	7
ponies	burros	1	2	3	4	5	6	7
abejas	colmenas	1	2	3	4	5	6	7
lirios	ramos	1	2	3	4	5	6	7
autos	talleres	1	2	3	4	5	6	7
casas	pasajes	1	2	3	4	5	6	7
actores	teatros	1	2	3	4	5	6	7
chalecos	chaquetas	1	2	3	4	5	6	7
sopranos	tenores	1	2	3	4	5	6	7
libros	repisas	1	2	3	4	5	6	7
presos	gendarmes	1	2	3	4	5	6	7
preguntas	dudas	1	2	3	4	5	6	7
estatuas	pinturas	1	2	3	4	5	6	7
perros	camadas	1	2	3	4	5	6	7
rubíes	zafiros	1	2	3	4	5	6	7
tumbas	criptas	1	2	3	4	5	6	7
gansos	bandadas	1	2	3	4	5	6	7
poodles	terriers	1	2	3	4	5	6	7
choferes	turistas	1	2	3	4	5	6	7
peras	duraznos	1	2	3	4	5	6	7
carretas	tractores	1	2	3	4	5	6	7
frazadas	cojines	1	2	3	4	5	6	7
cucharas	cuchillos	1	2	3	4	5	6	7
tortas	pasteles	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		menos dependencia			mayor dependencia			
zapatos	botas	1	2	3	4	5	6	7
juristas	jueces	1	2	3	4	5	6	7
soles	lunas	1	2	3	4	5	6	7
ciervos	zorros	1	2	3	4	5	6	7
payasos	circos	1	2	3	4	5	6	7
vacas	cerdos	1	2	3	4	5	6	7

¡Muchas gracias por tu participación!

Escala de Likert Evaluación Categórica

Instrucciones:

En este cuestionario, tu tarea consiste en evaluar hasta qué punto los siguientes pares de palabras se relacionan entre sí directa o indirectamente. Para realizar la evaluación, debes calificar de 1 a 7 los pares de palabras, donde 1 equivale a una relación directa baja y 7 a una relación directa alta.

Por favor, marca con una X la casilla correspondiente y trata de utilizar todo el rango de calificaciones disponible.

A continuación, te daremos algunos ejemplos antes de comenzar con la evaluación:

Relación directa alta: fútbol y basketball tienen una relación directa alta, ya que ambos son deportes en que dos equipos se enfrentan, utilizando un balón. Por lo tanto, un posible puntaje podría ser 7.

Relación directa intermedia: fútbol y ajedrez tienen una relación directa intermedia, ya que ambos son juegos, pero las características del juego varían bastante entre ambos, en cuanto a números de jugadores, tipo de juego, etc. Por lo tanto, un posible puntaje podría ser 4.

Relación directa baja o indirecta: camisa y televisor tienen una relación directa baja, ya que la función de la camisa es vestir y la del televisor entretener. De ahí que, es difícil establecer una conexión entre ellas. Por lo tanto, un posible puntaje podría ser 1.

Palabra 1	Palabra 2	Calificación						
		Relación no directa			Relación directa			
chalecos	ganchos	1	2	3	4	5	6	7
químicos	físicos	1	2	3	4	5	6	7
galletas	brownies	1	2	3	4	5	6	7
hamsters	jerbos	1	2	3	4	5	6	7
monjas	monjes	1	2	3	4	5	6	7
vacas	granjas	1	2	3	4	5	6	7
canoas	kayaks	1	2	3	4	5	6	7
choferes	taxis	1	2	3	4	5	6	7
zapatos	closets	1	2	3	4	5	6	7
cobras	pitones	1	2	3	4	5	6	7
perros	gatos	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		Relación no directa			Relación directa			
platos	armarios	1	2	3	4	5	6	7
buses	andenes	1	2	3	4	5	6	7
tazas	vasos	1	2	3	4	5	6	7
gansos	patos	1	2	3	4	5	6	7
cabras	ovejas	1	2	3	4	5	6	7
láminas	álbumes	1	2	3	4	5	6	7
actores	actrices	1	2	3	4	5	6	7
soles	galaxias	1	2	3	4	5	6	7
tumbas	féretros	1	2	3	4	5	6	7
clavos	tornillos	1	2	3	4	5	6	7
aviones	cohetes	1	2	3	4	5	6	7
diamantes	perlas	1	2	3	4	5	6	7
soldados	patrullas	1	2	3	4	5	6	7
naranjas	manzanas	1	2	3	4	5	6	7
estatuas	galerías	1	2	3	4	5	6	7
árboles	arbustos	1	2	3	4	5	6	7
arqueros	equipos	1	2	3	4	5	6	7
registros	carpetas	1	2	3	4	5	6	7
sopranos	coros	1	2	3	4	5	6	7
plumones	cajas	1	2	3	4	5	6	7
rubíes	collares	1	2	3	4	5	6	7
payasos	magos	1	2	3	4	5	6	7
alcaldes	pueblos	1	2	3	4	5	6	7
flautas	violines	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		Relación no directa			Relación directa			
cucharas	cajones	1	2	3	4	5	6	7
ponies	establos	1	2	3	4	5	6	7
chefs	meseros	1	2	3	4	5	6	7
autos	camiones	1	2	3	4	5	6	7
abejas	avispas	1	2	3	4	5	6	7
ranas	sapos	1	2	3	4	5	6	7
gorriones	nidos	1	2	3	4	5	6	7
osos	lobos	1	2	3	4	5	6	7
poodles	caniles	1	2	3	4	5	6	7
dentistas	doctores	1	2	3	4	5	6	7
mucamas	conserjes	1	2	3	4	5	6	7
sillas	agencias	1	2	3	4	5	6	7
libros	revistas	1	2	3	4	5	6	7
peras	cuencos	1	2	3	4	5	6	7
preguntas	pruebas	1	2	3	4	5	6	7
cellos	estuches	1	2	3	4	5	6	7
pastores	curas	1	2	3	4	5	6	7
ciervos	trampas	1	2	3	4	5	6	7
lirios	violetas	1	2	3	4	5	6	7
leones	tigres	1	2	3	4	5	6	7
sartenes	ollas	1	2	3	4	5	6	7
frazadas	camas	1	2	3	4	5	6	7
presos	celdas	1	2	3	4	5	6	7
juristas	juzgados	1	2	3	4	5	6	7

Palabra 1	Palabra 2	Calificación						
		Relación no directa			Relación directa			
claveles	floreros	1	2	3	4	5	6	7
maestras	clases	1	2	3	4	5	6	7
casas	chozas	1	2	3	4	5	6	7
carretas	sendas	1	2	3	4	5	6	7
tortas	vitriñas	1	2	3	4	5	6	7

¡Muchas gracias por tu participación!