DEPARTAMENTO DE HUMANIDADES
UNIVERSIDAD NACIONAL DEL SUR

Tesina de Licenciatura en Letras

Herramientas tecnológicas para la comunicación profesional.
El caso de *Google Translate*, *Grammarly* y *ChatGPT*

Diamela Frapolli
L.U. 114394

BAHÍA BLANCA 2025 ARGENTINA

Prefacio:

Esta tesina se presenta como trabajo final para obtener el título de Licenciada en Letras de la Universidad Nacional del Sur. Contiene el resultado de la investigación desarrollada por Diamela Frapolli, en la orientación Lingüística, bajo la dirección de la Dra. Lorena M. A. de-Matteis y la Mg. Mariela K. Starc.

1. Introducción

Esta tesis explora el impacto del uso de recursos digitales para la traducción y corrección en producciones escritas en lengua extranjera (LE) de textos académico-profesionales a partir de un diseño experimental. Para ello, se analizan traducciones alternativas al inglés de fragmentos de un informe de laboratorio, realizadas por futuros egresados de carreras de Ciencias Naturales de la Universidad Nacional del Sur (UNS).

Se evalúa el modo de apropiación de diferentes tecnologías digitales como soporte para la traducción y corrección de textos a través del análisis de los rasgos característicos del género, el número de errores presentes en las producciones y las operaciones lingüísticas ejecutadas por los participantes. En función de estos aspectos, se examina el empleo de tres herramientas tecnológicas específicas: *Google Translate*, *Grammarly*[1] y *ChatGPT*. Los primeros dos sistemas fueron seleccionados por su consolidación como referentes en sus respectivos fines de aplicación. En primer lugar, *Google Translate* es uno de los traductores automáticos más reconocidos y utilizados a nivel mundial. En el caso de *Grammarly*, esta es una plataforma líder en corrección gramatical y estilo, que informa más de 30 millones de usuarios activos a la fecha. En lo que respecta a *ChatGPT*, a pesar de haber sido desarrollado y lanzado al mercado hace menos tiempo que los otros dos sistemas, su popularidad ha crecido de forma vertiginosa y se ha convertido en una plataforma adoptada por diversos sectores sociales para múltiples tareas vinculadas a la producción textual, lo que justifica su inclusión en este estudio. Para complementar el análisis, se consideran las percepciones de los participantes sobre la actividad realizada y las opiniones de jueces expertos en el área disciplinar y en lengua inglesa.

1.1. Contextualización y problema de investigación

Esta investigación se realiza en un contexto formativo en el que los planes de estudio de las carreras de Ciencias Naturales que cursan los participantes de la actividad organizada para esta tesina no contemplan de manera obligatoria el aprendizaje de inglés como LE orientada a la producción escrita, si bien su formación incluye evaluaciones de suficiencia en inglés enfocadas a la lectura de bibliografía especializada. Esta situación plantea interrogantes sobre la preparación de estos estudiantes para producir textos escritos en LE que se ajusten a los géneros comunicativos requeridos en su futura labor profesional. Esto cobra particular relevancia en el caso de quienes aspiren a insertarse en entornos laborales con proyección internacional o vinculados a instituciones extranjeras. En el grupo de

[1] Originalmente, se había iniciado a trabajar con las dos primeras tecnologías en el marco de una beca CIN. Sin embargo, a raíz de la aparición de *ChatGPT*, se decidió incluir este tercer recurso debido a su rápida difusión y a su potencial para ser utilizado como herramienta asistente en diferentes tareas relacionadas con la corrección y traducción de textos.

estudiantes seleccionados, esto se refleja en contextos como empresas o laboratorios con sedes en distintos países, equipos de trabajo con jefaturas extranjeras, o proyectos de investigación que demanden la escritura de informes, presentaciones o publicaciones en inglés.

Como es sabido, el impacto de procesos globales ha transformado las dinámicas de inserción profesional en múltiples disciplinas. En particular, la globalización —definida por Al-Rodhan y Stoudmann (2006, p. 5) como un proceso que "(…) encompasses the causes, course, and consequences of transnational and transcultural integration of human and non-human activities"— ha reconfigurado la comunicación en diversos ámbitos, incluido el laboral. Este fenómeno ha impulsado a empresas y organizaciones de todo el mundo a expandir sus actividades más allá de sus fronteras nacionales, y esto ha generado una creciente demanda de comunicación efectiva en lenguas extranjeras. En este escenario, el inglés ha adquirido un estatus particular debido a su valor simbólico y funcional en el mercado global. Tal como plantea Calvet (2006):

> the notions of 'value' and 'prestige' have to do as much with representations as with realities, but that these representations foster and reinforce the realities. It is because we attribute a 'market value' to English that the vast majority of pupils take it as their first language at school and thereby increase its value. (p. 3)

Según este autor, la prominencia actual del inglés como lengua franca global está vinculada con su difusión en los ámbitos empresariales y académicos, impulsada por la hegemonía cultural y económica de los países angloparlantes. Aun así, como advierte Bein (2003), el valor atribuido a esta lengua en contextos laborales y educativos muchas veces se presenta de forma acrítica, como un "fetiche", como si se tratara de una cualidad intrínseca de esta lengua y no del resultado de procesos históricos, sociales y políticos determinados:

> la utilidad de la lengua puede parecer un hecho objetivo porque en cierto momento histórico es condición necesaria pero no suficiente para conseguir trabajo, sin que se perciba que se trata de una situación histórica determinada igualmente por variables socioeconómicas, políticas y culturales. Como a cualquier otro, al fetiche lingüístico se le atribuyen cualidades mágicas: se deposita en él la virtud de conseguir empleo, o la de reunificar una comunidad, o la de hacer perdurar una religión. (p. 36)

Pese a ello, no puede negarse que, en el marco de una economía globalizada, la capacidad de comunicarse en inglés se percibe —al menos desde el punto de vista de las representaciones sociales— como una competencia clave para la inserción y el desarrollo profesional en múltiples sectores. A título ilustrativo, Ne'Matullah et al. (2023) realizan una revisión bibliográfica sobre la relación entre

empleabilidad[2] y el uso de esta lengua, y encuentran que se emplea con frecuencia en sectores como finanzas, tecnología, ciencia y el ámbito académico (p. 169). Asimismo, al analizar encuestas realizadas en organizaciones malayas de los sectores público y privado, los autores informan que el 84.9% de los empleadores manifiesta preferencia por candidatos con dominio de inglés (p. 173).

La tecnología, por otra parte, ha sido otro factor central en esta transformación global. Esto se debe a que ha facilitado nuevas modalidades de comunicación interna y externa de las organizaciones que han permitido a los equipos de trabajo mantenerse en contacto en tiempo real, incluso si están ubicados en diferentes lugares o países. La disponibilidad de recursos digitales relacionados con el lenguaje ha crecido como respuesta a las demandas surgidas de la situación laboral actual y al desarrollo del sector informático. En este contexto, los traductores automáticos y los correctores lingüísticos emergen como recursos posibles, aunque no infalibles, para sortear los desafíos que implica la comunicación en una lengua extranjera. Los primeros emplean algoritmos para traducir textos entre diferentes idiomas sin intervención humana y son una de las tecnologías del lenguaje más representativas. A su vez, los recursos de corrección gramatical, ortográfica y estilística ofrecen soporte para la producción textual, ya sea como funcionalidades integradas en procesadores de texto o como plataformas en línea especializadas. Algunas de estas tecnologías incluso permiten detectar patrones discursivos más complejos y formular sugerencias orientadas a mejorar la cohesión, el estilo y la claridad del texto. En suma, estos sistemas permiten a los usuarios detectar y corregir errores comunes, y optimizar así su proceso de escritura en LE.

Es relevante señalar que, aunque en sus primeros años de desarrollo estas herramientas presentaron diversas limitaciones, y aún pueden mostrar algunas deficiencias, con el tiempo han mejorado de forma considerable en lo referente a su precisión. En particular, su desempeño se ha visto potenciado por la integración progresiva de sistemas de inteligencia artificial (IA), un campo que ha experimentado un desarrollo vertiginoso en los últimos años. Su incorporación ha ampliado de forma notable las capacidades de las herramientas de traducción y corrección, en especial en lo relativo al procesamiento del lenguaje natural mediante grandes modelos de lenguaje (LLM), que se fundan en el análisis estadístico complejo de las estructuras lingüísticas para realizar este tipo de tareas.

No obstante, su creciente sofisticación también plantea nuevos desafíos que van más allá de su utilidad inmediata. Si bien estas tecnologías pueden facilitar la producción textual en lenguas extranjeras, su impacto sobre la corrección gramatical y adecuación pragmática de los textos generados,

[2] La *empleabilidad* se entiende como la capacidad para conseguir y mantener un empleo satisfactorio. De manera más amplia, implica la habilidad de desplazarse de manera autónoma en el mercado laboral para desarrollar el propio potencial mediante un empleo sostenible. Esta capacidad depende del conocimiento, habilidades y actitudes que posea el individuo, de cómo las utilice y las presente a los empleadores, así como del contexto personal y del entorno del mercado laboral en el que busca trabajo (Hillage y Pollard, 1998, p. 2).

así como en los procesos lingüísticos implicados en su elaboración, aún requiere ser explorado en profundidad. En este sentido, cabe preguntarse: ¿hasta qué punto quienes las utilizan evalúan de forma crítica las decisiones lingüísticas que estos sistemas toman por ellos? ¿Qué nivel de control ejercen los hablantes en estos procesos, cada vez más automatizados? ¿En qué medida las intervenciones que los hablantes realizan sobre los textos generados o modificados por estas plataformas contribuyen a subsanar las deficiencias lingüísticas o discursivas que estas puedan producir?

Dado este escenario, resulta pertinente volver la mirada a los futuros profesionales de las Ciencias Naturales de la UNS como un caso ilustrativo, a fin de considerar cómo interactúan con estas tecnologías en su formación académica. En este contexto, la problemática que se propuso abordar este trabajo puede formularse en términos de una tensión entre tres factores. Esto es, las necesidades de producción escrita en contextos académico-profesionales globalizados, la formación lingüística en lengua extranjera de los futuros profesionales, y el uso de tecnologías como mediadoras entre las necesidades y formación antedichas.

El interrogante que integra las preguntas de investigación planteadas, entonces, es si los estudiantes de las áreas académicas consideradas se apropian de estas tecnologías de forma crítica para producir textos del género informe adecuados en LE. La cuestión no se limita a considerar los fallos o aciertos puntuales en sus traducciones, sino que implica examinar el tipo de decisiones lingüísticas que toman, cómo negocian entre sus propios conocimientos y las sugerencias tecnológicas, y qué impacto tiene esto sobre los textos producidos.

Así, el presente trabajo se inscribe en una problemática que articula el desarrollo de competencias comunicativas profesionales para la escritura en inglés con la creciente mediación de tecnologías digitales. Su aporte se orienta a ofrecer un diagnóstico del modo en que los futuros profesionales de Ciencias Naturales interactúan con herramientas digitales en situaciones de producción escrita en LE. Esta caracterización puede contribuir a la reflexión sobre posibles enfoques pedagógicos más ajustados a sus necesidades y a las prácticas comunicativas reales de las que tendrán que participar. Además, esta investigación propone un aporte metodológico que puede servir como referencia para estudios similares, presentando una forma particular de abordar y analizar las dinámicas de uso de la tecnología en procesos de producción escrita. Por último, el análisis de los usos, fortalezas y limitaciones de estos recursos en contextos auténticos puede resultar de interés tanto para el ámbito educativo como para el desarrollo y evaluación de tecnologías lingüísticas.

1.2. Objetivos

En consonancia con la problemática expuesta, el propósito de esta tesina es indagar cómo los estudiantes de carreras de Ciencias Naturales de la UNS utilizan herramientas digitales de traducción y

corrección automática para la producción escrita en inglés, en el marco de situaciones comunicativas que anticipan su futuro desempeño profesional. A partir de ello, se delinearon los siguientes objetivos generales y específicos:

1.2.1. Objetivos generales

- Aportar a la comprensión de los modos de apropiación de las tecnologías digitales en el ámbito del lenguaje entre estudiantes universitarios de carreras de Ciencias Naturales.
- Contribuir al estudio de las competencias comunicativas profesionales a partir de un examen del impacto de las nuevas tecnologías sobre el desempeño comunicativo de los hablantes.
- Determinar si el uso de recursos digitales ofrece ventajas para la comunicación profesional escrita y, en caso afirmativo, identificarlas.

1.2.2. Objetivos específicos

- Determinar qué operaciones lingüísticas son las más realizadas por los hablantes al usar estos sistemas (en particular, aquellas vinculadas a los rasgos del discurso académico).
- Analizar si el empleo secuencial y conjunto de *Google Translate*, *Grammarly* y *ChatGPT* conduce a una mejora cualitativa en sus producciones en inglés.
- Determinar si una de las tres herramientas examinadas resulta de mayor utilidad que otra para los participantes y si su uso conjunto, en una secuencia específica, supone mayores beneficios que su uso por separado.
- Examinar la percepción de los hablantes de habla hispana[3] (estudiantes de carreras universitarias de la UNS) respecto del uso de este tipo de sistemas en el marco de las tareas de comunicación que deberán ejercer en sus labores profesionales diarias.
- Contrastar estas percepciones con juicios de expertos del área en cuestión y de idioma inglés.

1.3. Hipótesis

De acuerdo con los objetivos planteados, este trabajo se apoyó en un conjunto de hipótesis que orientaron la investigación. Estas se derivan de observaciones no sistemáticas que sugieren que, aunque los estudiantes del área de las Ciencias Naturales de la UNS no requieren certificar las cuatro habilidades comunicativas en inglés, conocen y, en algunos casos, emplean herramientas digitales de asistencia lingüística —ya sea las seleccionadas para este estudio u otras de funcionamiento similar[4]—. Asimismo, se presupone que, en sus futuras prácticas profesionales, podrían recurrir a ellas tanto por

[3] Si bien no es el objeto de la tesina, es posible suponer que estas opiniones pueden variar para hablantes de otras lenguas que utilicen estas mismas herramientas.

[4] Entre otras, se pueden mencionar *DeepL*, *Microsoft Editor*, *Hemingway Editor*, *LanguageTool*, *ProWritingAid*, *Reverso*, *Google Gemini*, *Claude*, *DeepSeek*, *Perplexity* y *Microsoft Copilot*.

iniciativa propia como en respuesta a políticas institucionales que promueven su uso en entornos laborales globalizados.

Dentro de este marco, se proponen las siguientes hipótesis:

- Cuando el hablante posee un dominio al menos intermedio de la lengua inglesa, el uso de *Google Translate*, *Grammarly* y *ChatGPT*, en particular si se hace en una secuencia específica y de forma conjunta, contribuye a la mejora de las producciones propias en inglés, generando resultados textuales más adecuados a los objetivos comunicativos.
- Las mejoras en las producciones pueden observarse tanto en el nivel léxico y sintáctico de los textos producidos con apoyo de estos recursos, como en la satisfacción con la producción y en la medida en que los participantes experimentan diferentes grados dentro del continuo entre seguridad e inseguridad lingüística.
- Los hablantes seleccionados son conscientes de las ventajas que supone el empleo de estas herramientas y poseen la capacidad para usarlas críticamente.

2. Marco teórico

En un sentido amplio, entonces, esta investigación se vincula con la perspectiva de la Lingüística Aplicada, entendida como un campo orientado al diagnóstico y abordaje de problemas reales de comunicación (Lacorte, 2007), en tanto se propone analizar una situación problemática relacionada con el lenguaje en un contexto de uso verosímil. En este sentido, el trabajo se centra en el análisis de cómo los estudiantes universitarios de Ciencias Naturales de la UNS se apropian de recursos digitales para producir textos en LE. Este examen se realiza a partir de una actividad de producción textual en un entorno controlado y diseñado para emular prácticas discursivas propias del ámbito académico-profesional real.

En este marco, esta investigación requiere articular conceptos provenientes de distintos enfoques funcionales de la lingüística y de sus áreas de aplicación, tales como la etnografía de la comunicación, la sociolingüística, la alfabetización académica, la enseñanza de LE, la traductología, la correctología y la lingüística de corpus. Los apartados que siguen sintetizan los principales conceptos teóricos que subyacen a esta propuesta.

2.1. Discurso y género

En el marco de esta investigación, resulta necesario considerar la noción de *discurso*. Este concepto ha sido objeto de amplio tratamiento y desarrollo a lo largo del tiempo (Stubbs 1983; Van Dijk, 1985; Calsamiglia y Tuson, 1999; Charaudeau y Maingueneau, 2002). En el caso de Stubbs (1983), el discurso se concibe como lenguaje en uso, producido en contextos reales, que conforma unidades coherentes por encima de la oración o la frase, como la conversación o el texto escrito. El autor destaca que el lenguaje y la situación social en la que se produce son inseparables, ya que toda instancia discursiva implica una red de conocimientos compartidos, actos de habla y funciones comunicativas que dependen del contexto (pp. 17-18). Según Van Dijk (1985), por su parte, puede entenderse como una forma real del lenguaje, más allá de la gramática, que cumple funciones comunicativas específicas dentro de la interacción social (pp. 1-2). En este sentido, el discurso puede ser entendido como una forma de interacción social (p. 3)

En estrecha relación con esta noción, Swales (1990) sostiene que los miembros expertos de distintas comunidades discursivas comparten propósitos comunicativos que se manifiestan a través de eventos específicos denominados *géneros*. Esta categoría, definida por ejemplo por Bajtín (1992, p. 248), resulta central y operativa en este trabajo, y alude a tipos de textos producidos en diferentes ámbitos sociales, con características formales y funcionales específicas, que permiten a los participantes reconocer y producir mensajes adecuados dentro de un marco comunicativo determinado. Para este autor, los géneros emergen de la relativa estabilidad de los enunciados en distintas esferas de la actividad

humana. Este concepto también ha sido ampliamente trabajado por otros autores, como Bathia (1993), quien retoma la definición de Bajtín y añade la relevancia de los factores psicológicos. El autor explica que un análisis psicolingüístico de los géneros es útil para comprender las estrategias individuales que el escritor utiliza con el objetivo de volver a su escritura más efectiva en un entorno sociocultural específico. Además, explica que deben tenerse en cuenta factores como el público objetivo, el medio o prerrequisitos organizacionales para abarcar o comprehender su naturaleza (pp. 19-20).

En particular, en la esfera educativa, puede hablarse de un tipo específico de discurso constituido por una variedad de géneros que comparten expresiones, estructuras y convenciones lingüísticas propias del ámbito universitario, al que se conoce como *discurso académico*. Diversos autores han trabajado en relación a los géneros que lo integran desde múltiples enfoques, tales como Swales (1990), Halliday y Martin (1993), Hyland (2000), o Banks (2008). Swales (1990), en particular, uno de los referentes clave en el análisis de los géneros académicos, propone que estos se configuran según los propósitos comunicativos compartidos por la comunidad discursiva, científica o universitaria (p. 58).

Los aportes teóricos sobre el discurso y los géneros académicos, como los mencionados en el párrafo anterior, se traducen en estudios de aplicación, enmarcados en la perspectiva de la *alfabetización académica*, que están orientados a estudiar cómo se lleva a la práctica su enseñanza-aprendizaje. Dentro de este campo, existe una extensa tradición de trabajos, de los que mencionamos a Ciapuscio (1992), Koutsantoni (2007), Cubo de Severino et al. (2007), Cassany (2007), Carlino (2010), Hall y López (2011), Camps Mundó y Montserrat Castelló (2013), y Navarro (2017), a modo de ilustración.

A los fines de este trabajo, destacamos la distinción entre textos académicos y científicos de Carlino (2010). Entre los primeros se encuentra el tipo de informe trabajado en esta tesina, y son utilizados en el ámbito universitario con fines formativos, mientras que los segundos circulan en la comunidad de investigadores (p. 87).

Por otro lado, Navarro (2017) considera que "la escritura académica no es una sino muchas" (p. 8), y explica que existen importantes diferencias entre los diferentes tipos textuales que se producen dentro de este ámbito. Al igual que Carlino, distingue entre textos orientados específicamente a la formación y aquellos que forman parte de la comunicación científica especializada. Además, subraya la existencia de diferencias relevantes entre las producciones propias de las ciencias exactas y naturales y las de las ciencias humanas y sociales (p. 9). Asimismo, destaca que la escritura académica cumple múltiples funciones que deben considerarse en su enseñanza. Estas incluyen permitir a los estudiantes revisar, apropiarse y reelaborar los contenidos mediante la escritura; aprender formas específicas de comunicación disciplinar; acreditar sus conocimientos; posicionarse críticamente frente a los saberes de su campo, e incluso desarrollar una voz propia en la expresión escrita (pp. 9-12).

Tanto los aportes teóricos sobre discurso *y* los géneros en contextos académicos como los estudios de alfabetización académica permiten caracterizar de forma sistemática estas prácticas discursivas y sus formas textuales asociadas, lo que resulta fundamental para desarrollar materiales específicos y estrategias didácticas adecuadas para su enseñanza, y que así los estudiantes se familiaricen con las convenciones discursivas propias de su disciplina.

Entre las producciones que se inscriben en esta línea, podemos mencionar manuales concretos orientados a la adquisición de diversos géneros propios del ámbito universitario, como Botta y Warley (2002), Montolío (2003) o Barcia y Barcia (2021). Este tipo de materiales constituyen recursos tanto para docentes como estudiantes en lo referente a la enseñanza y aprendizaje de las habilidades necesarias para producir estos tipos textuales, tema que se desarrolla en el apartado siguiente.

En esta misma línea, Cubo de Severino et al. (2007) llevan adelante una tarea de caracterización de diferentes géneros académicos a partir de un trabajo de corpus, en el que clasifican y describen de forma detallada diferentes géneros académicos, e identifican varias características que definen su estilo y estructura. De forma general, se señala el predominio de estilos impersonales y desagentivados, en la mayoría de los géneros analizados.

En términos más específicos, se destaca a la voz pasiva como un recurso presente en diversos géneros académicos, tales como el artículo de investigación (p. 73), el resumen o *abstract* (p. 104), el informe de investigación (p. 312) y los manuales universitarios (p. 332). Este recurso contribuye a la construcción de textos desagentivados, característica que refiere a la omisión o minimización de la presencia del agente en la construcción oracional, lo que favorece un enfoque en los procesos, hechos o resultados más que en los sujetos que los llevan a cabo.

Las nominalizaciones, por su parte, también contribuyen a la escritura de textos impersonales y desagentivados, ya que permiten condensar información compleja en estructuras más simples. Las autoras también registran el uso de nominalizaciones en múltiples géneros académicos, como por ejemplo el artículo de investigación (p. 70), el resumen o abstract (p. 105), la ponencia (p. 137), la conferencia académica (p. 205), la monografía (p. 229), el proyecto de investigación (p. 291), los manuales universitarios y de nivel medio (pp. 332 y 349) y los documentos de cátedra (p. 380). Halliday y Matthiessen (2004), desde la Lingüística Sistémica Funcional, explican que este recurso permite transformar elementos verbales, adjetivales o de otras categorías en grupos nominales, de modo que puedan cumplir diferentes funciones dentro de la estructura de la cláusula. Los autores la clasifican como un tipo de metáfora gramatical ideacional, frecuente en discursos especializados como los de la ciencia, la educación, la burocracia y el derecho ya que permite que los procesos y las cualidades se construyan como entidades (p. 636-637).

A la luz de estas definiciones y caracterizaciones, el informe de laboratorio con el que se trabaja en este caso se considera un género específico del discurso académico que, como los restantes, posee propósitos comunicativos concretos, estructuras retóricas convencionales y exigencias de adecuación terminológica propias del ámbito en el que circula. Ezpeleta Piorno (2008) lo describe como un género que se inscribe en contextos empresariales, industriales o científicos, y cuya principal función es exponer de manera escrita el progreso o los resultados de un estudio, una investigación o un problema. La autora señala que estos textos se caracterizan por una serie de parámetros que abarcan dimensiones comunicativas, sociales, formales, convencionales y psicolingüísticas (pp. 5-6). Así, el informe técnico —y por extensión, el de laboratorio— cumple una función comunicativa precisa en contextos académico-profesionales, y exige un dominio temático y la capacidad de adecuar la producción textual a convenciones discursivas específicas.

2.2. Competencias comunicativas

A la luz de las definiciones y caracterizaciones de la sección anterior, el informe de laboratorio del que proviene el material con el que se trabaja en esta investigación se considera un género específico del discurso académico, cuya producción demanda tanto un conocimiento explícito de sus rasgos y convenciones por parte del grupo de estudiantes seleccionado, como el desarrollo de habilidades para emplearlos de forma adecuada en contextos académicos y profesionales. Por lo tanto, es de relevancia la noción de *competencia comunicativa* (Hymes, 1972), proveniente de la etnografía de la comunicación, que comprende los aspectos socioculturales de la comunicación, y la capacidad de los hablantes de usar el lenguaje de forma efectiva en diferentes contextos comunicativos y culturales. El autor explica que esta competencia va más allá de reconocer la corrección gramatical, ya que incluye también la adecuación comunicativa:

> a normal child acquires knowledge of sentences, not only as grammatical, but also as appropriate. He or she acquires competence as to when to speak, when not, and as to what to talk about with whom, when, where, in what manner. [...] This competence, moreover, is integral with attitudes, values, and motivations concerning language, its features and uses, and integral with competence for, and attitudes toward, the interrelation of language with the other codes of communicative conduct. (p. 277)

Desde su formulación inicial, la noción de competencia comunicativa se propuso en reacción a la definición de *competencia lingüística* de Chomsky (1965), que se limitaba a la capacidad ideal del hablante-oyente para generar oraciones gramaticales en una lengua. Esta concepción dejaba de lado los aspectos situacionales o sociales del uso del lenguaje.

Frente a esta perspectiva, Hymes propuso un enfoque que integra la gramaticalidad con la adecuación contextual. Este concepto de competencia comunicativa fue abordado desde distintas perspectivas por diversos autores, entre ellos Wiemann (1977), Spitzberg y Cupach (1984), Greene y Burleson (2003), y Rickheit y Strohner (2008). Estos desarrollos responden a enfoques comunicativos y a necesidades surgidas en distintos ámbitos.

Así, una de las áreas que se han beneficiado especialmente de los aportes que entraña el concepto de competencia comunicativa es la enseñanza de LE. En este sentido, la inclusión de los aportes desarrollados en este campo se justifica en tanto la formación universitaria —y, como parte de ella, la enseñanza de LE— apunta a desarrollar, además de conocimientos disciplinares, capacidades comunicativas vinculadas a prácticas profesionales. En esta línea, dado el contexto actual de globalización descrito en la introducción, se vuelve necesario para los hablantes insertarse en instituciones con proyección internacional desarrollen competencias comunicativas también en una LE, en este caso, el inglés. Por ello, se retoman en este trabajo contribuciones que orientan la enseñanza-aprendizaje de segundas lenguas, como el *Marco Común Europeo de Referencia* (MCER - Consejo de Europa, 2001), que distingue, dentro de las competencias comunicativas, las *competencias lingüísticas*, *sociolingüísticas* y *pragmáticas* (p. 106). Las *competencias lingüísticas* incluyen la competencia léxica, gramatical, fonológica, ortográfica, semántica y ortoépica (p. 107). Las *competencias sociolingüísticas* refieren al "conocimiento y las destrezas necesarias para abordar la dimensión social del uso de la lengua", e implican saberes sobre "los marcadores lingüísticos de relaciones sociales, las normas de cortesía, las expresiones de la sabiduría popular, las diferencias de registro, el dialecto y el acento" (p. 116). Las *competencias pragmáticas*, por último, abarcan la competencia discursiva, que es el conocimiento sobre la organización, estructura y ordenación de los textos, la *competencia funcional*, vinculada al uso de los textos para realizar actos de habla concretos, y la *competencia organizativa*, relativa a la secuenciación de mensajes según esquemas de interacción y transacción (p. 120). El *MCER* considera estos tres conjuntos de competencias en relación con los seis niveles de dominio: A1 y A2 (usuario básico), B1 y B2 (usuario independiente), y C1 y C2 (usuario competente), lo que permite caracterizar las habilidades comunicativas en los distintos estadios del aprendizaje de una LE. En el marco de este trabajo, el empleo en inglés de las nominalizaciones y estructuras pasivas, así como el análisis de los errores gramaticales en las traducciones, fueron considerados como indicadores de las competencias lingüísticas de los participantes. Sin embargo, también puede proponerse que constituyen indicadores de las competencias discursivas de los estudiantes, en tanto el uso correcto o incorrecto de los rasgos propios del género puede dar cuenta del grado de conocimiento que los participantes tienen sobre sus convenciones discursivas.

De acuerdo con estos antecedentes, la escritura de informes, tanto en LE como en español, constituye una práctica que requiere de —y en esta etapa formativa, también desarrolla— una competencia comunicativa específica en el ámbito académico y profesional. Ahora bien, dada la especificidad de las situaciones comunicativas que requieren de la escritura del género informe, se toma como punto de partida la categoría general de competencia comunicativa, articulada con una noción más acotada y pertinente al objeto de estudio de esta tesina: la *competencia comunicativa profesional* (CCP). Esta noción remite a los saberes vinculados a los géneros discursivos propios del ámbito laboral. Aguirre Raya (2005), tras una revisión bibliográfica sobre el tema, define a este tipo de habilidades de la siguiente manera:

> La competencia comunicativa del profesional es la potencialidad que tiene el sujeto de lograr una adecuada interacción comunicativa a partir del dominio e integración en el ejercicio profesional de los conocimientos acerca del proceso de comunicación humana, habilidades comunicativas, principios, valores, actitudes y voluntad para desempeñarse en su profesión eficientemente así como para tomar decisiones oportunas ante situaciones complejas o nuevas, que faciliten el logro de los objetivos trazados o propuestos en diferentes contextos y en las dimensiones afectivo-cognitiva, comunicativa y sociocultural. (p. 1)

Estas CCP, por último, se articulan con el desarrollo de competencias comunicativas en LE mediante la enseñanza de lenguas para fines específicos, cuyos programas contemplan el abordaje de géneros y prácticas comunicativas propias de ámbitos académicos y profesionales. Esta área permite atender a los requerimientos comunicativos de contextos profesionales, en particular en entornos globalizados donde el uso de una LE debe ser tanto efectivo como adecuado. En esa línea, la actividad realizada por los participantes en el contexto de este trabajo propuso una simulación de una situación de comunicación profesional verosímil, similar a como podría presentarse en su futuro laboral.

2.3. El error lingüístico

De este modo, y en función de lo expuesto sobre competencias comunicativas, es posible establecer una relación entre el nivel de desarrollo de estas competencias y la calidad del desempeño comunicativo en LE y géneros académicos. En este marco, puede entenderse que a menor nivel de CC, mayor es la probabilidad de que se produzcan errores lingüísticos, en tanto manifestaciones de dificultades en la adecuación o en la conformación gramatical, léxica o discursiva de los textos. A partir de esta relación, resulta pertinente señalar que una forma de aproximarse al análisis del desarrollo de la CC es a través del estudio de las dificultades que se evidencian en la escritura de ciertos géneros o lenguas en particular. Por ello, otra noción que opera en este trabajo es la de *error lingüístico*, que ha

sido abordada por diferentes disciplinas como el aprendizaje de LE, la traductología, la correctología y la posedición[5]. En este caso, se consideran aportes de los tres primeros campos, en tanto ofrecen insumos para describir, clasificar e interpretar los errores observados en las producciones textuales, así como para comprender su origen, función e impacto comunicativo.

En primer lugar, el tratamiento del *error* desde el enfoque del aprendizaje de LE resulta pertinente en vista de que los textos analizados fueron producidos por hablantes no expertos en inglés. En este contexto, Corder (1981) sostiene que los errores de los aprendientes no deben interpretarse como simples fallos, sino como manifestaciones del desarrollo de la interlengua del aprendiz (p. 10). En este caso, el análisis de los fallos presentes en las traducciones permite indagar cómo estas tecnologías impactan en las distintas versiones del texto. Esta incidencia se vincula con el desarrollo de CCP en LE de los participantes, en tanto revela su nivel de dominio de los recursos lingüísticos y las dificultades que enfrentan al producir textos especializados en inglés.

Por otro lado, si bien los participantes que realizan la actividad propuesta para la tesina no son estudiantes de traducción o corrección, se incorpora la conceptualización del *error* desde la traductología y la correctología, ya que la tarea que origina los textos del estudio implica operaciones sucesivas tanto de corrección como de traducción de un fragmento de un informe de laboratorio. Además, desde una perspectiva analítica, la noción de error de estas perspectivas permite describir y sistematizar las operaciones realizadas y los fallos detectados, y facilita así una categorización más precisa de los fenómenos observados.

Desde la traductología, Cruces Colado (2001) ofrece un análisis basado en corpus de traducciones realizadas por estudiantes, con el propósito de elaborar una taxonomía de este fenómeno lingüístico y explorar sus posibles causas. Su enfoque busca contribuir a la prevención de errores en el proceso formativo, mediante la identificación de los factores lingüísticos, culturales y contextuales que los generan. De forma general, la autora define el error de traducción como una ruptura de las reglas de coherencia del texto meta, ya sea gramatical, léxica, semántica o referida al conocimiento del mundo y la experiencia compartida (p. 2). Esta concepción permite considerar el error no sólo como una reformulación inadecuada, sino también como una manifestación de la incapacidad para reconocer que se ha producido una ruptura (p. 4). La autora distingue entre dos grandes tipos de errores: relacionados con el sentido y de naturaleza formal. En el primer grupo se incluyen los casos en que el texto meta

[5] La posedición, por su parte, comparte con la corrección de textos el objetivo de garantizar la claridad y precisión del mensaje final, pero se centra en el refinamiento de textos generados mediante recursos de traducción automática. En este sentido, de Almeida (2013) señala que, si bien estas herramientas han ganado relevancia en los procesos modernos de traducción, "human linguists are still required to verify and correct machine-translated documents. This is done through a process called post-editing (PE), in which the raw output of MT is corrected in order to make documents as accurate and readable as possible" (p. 1).

transmite un sentido diferente al del texto fuente, lo que puede deberse a una redacción dificultosa, al manejo incorrecto de los recursos expresivos en la lengua meta, a una interpretación errónea de las estructuras gramaticales del texto origen o a un uso literal de expresiones que ignora las convenciones discursivas e idiomáticas. El segundo grupo corresponde a desviaciones formales vinculadas al desconocimiento de las reglas ortográficas y/o gramaticales de la lengua meta. Según la autora, muchos de estos casos podrían evitarse mediante una revisión atenta del texto, aunque esta instancia suele pasarse por alto. A su vez, señala que, cuando el estudiante carece de competencias lingüísticas suficientes en la lengua origen y/o en la lengua meta, difícilmente la revisión pueda funcionar como estrategia eficaz para evitarlos (p. 10).

Este enfoque resulta útil para este trabajo, ya que permite considerar los errores de traducción en función de su origen y naturaleza específica, y no solo como desviaciones normativas. Asimismo, brinda herramientas conceptuales para analizar el impacto —positivo o negativo— de los recursos digitales en las sucesivas correcciones asistidas de las traducciones, y evaluar en qué medida las intervenciones de los participantes contribuyen a resolver —o a generar— fallos de sentido o de forma.

Desde la correctología, por otro lado, resultan relevantes los aportes de García Negroni y Estrada (2006), quienes identifican y clasifican las competencias necesarias para abordar distintos tipos de error lingüístico. Aunque estas competencias fueron pensadas pensando en la figura del corrector profesional, su sistematización permite observar con mayor precisión los aspectos lingüísticos que intervienen en la producción de textos especializados. Las autoras distinguen entre *competencias enciclopédicas*, *gramaticales* y *textuales*, cada una vinculada a fallas lingüísticas específicas cuya corrección requiere habilidades particulares. Las *competencias enciclopédicas*, en primer lugar, refieren al conjunto de conocimientos sobre el mundo que posee el corrector, que permiten "corregir textos especializados, pero también tomar las decisiones necesarias sobre sus aspectos más generales" (p. 29). Las *competencias gramaticales*, divididas en los subniveles fonemático[6], morfológico y sintáctico, se relacionan con el conocimiento de las normas que rigen la estructura del lenguaje, e incluyen desde cuestiones ortotipográficas hasta problemas de concordancia o tiempos verbales (pp. 5-9). Las *competencias textuales*, por su parte, refieren a la organización del discurso y permiten detectar problemas de coherencia y cohesión, como la ausencia de conectores o la repetición innecesaria de

[6] Según García Negroni y Estrada (2006, p. 6), el nivel fonemático se subdivide en dos subniveles: fónico y gráfico. En el contexto de la correctología aplicada a textos escritos, se trabaja sobre el nivel gráfico, que abarca aspectos como la ortografía, la acentuación y otros elementos vinculados a la representación visual del lenguaje. Cabe destacar que, aunque algunos autores consideran al ortográfico como un nivel independiente, las autoras la integran en el gramatical, dentro del subnivel fonemático.

ideas (pp. 9-11). En particular, este trabajo se apoya en las dos últimas categorías para sistematizar los tipos de errores considerados, y establecer criterios claros para cuantificarlos.

Ahora bien, para establecer estos parámetros con rigor, también resulta necesario considerar ciertas nociones que permiten matizar el concepto de "error", y que permiten entender que no toda desviación textual implica necesariamente un fallo desde el punto de vista comunicativo o normativo. Entre estas nociones, adquieren especial relevancia la *gramaticalidad*, la *aceptabilidad* y la *corrección*, conceptos que han sido ampliamente desarrollados en distintas tradiciones lingüísticas y que permiten pensar la relación entre forma, norma y uso desde perspectivas complementarias.

El concepto de *gramaticalidad* se inscribe originalmente en la tradición de la gramática generativa, en particular en la noción de competencia lingüística formulada por Chomsky (1965), entendida como la capacidad ideal del hablante-oyente para generar oraciones bien formadas en su lengua desde el punto de vista del sistema lingüístico, sin considerar necesariamente su uso en contextos reales. En línea con esta tradición formal, Bosque Muñoz y Gutiérrez-Rexach (2008) definen la gramaticalidad como la propiedad de una secuencia de estar bien formada según las reglas combinatorias del sistema lingüístico al que pertenece. Es decir, este concepto es interno al sistema y no depende de factores externos (p. 28).

En cambio, la noción de *aceptabilidad* surge como una ampliación a esta perspectiva, ya que alude a la manera en que los hablantes perciben, juzgan o aceptan las construcciones lingüísticas en situaciones reales de comunicación. Esta categoría se vincula con el concepto de competencia comunicativa propuesto por Hymes (1972), en tanto introduce la dimensión social y pragmática al uso del lenguaje, al considerar no solo lo que es posible desde lo formal, sino también lo que es apropiado y eficaz en función del contexto, los interlocutores y las normas del discurso.

La *corrección*, por su parte, se sitúa en un espacio intermedio, ya que refiere tanto a la adecuación gramatical como a factores extralingüísticos, como normas estilísticas, sociales o discursivas, que regulan el uso de las formas lingüísticas (Bosque Muñoz y Gutiérrez-Rexach, 2008, pp. 28-29). En esta línea, Paredes García (2009) amplía la reflexión y explica que la corrección es un concepto polisémico, cuyos límites se entrelazan con los de *aceptabilidad*, *adecuación* y *apropiación*. Un enunciado puede ser aceptable para los hablantes aun si quebranta ciertas normas gramaticales, y puede ser adecuado o inadecuado dependiendo del contexto comunicativo, la relación entre interlocutores y el género discursivo en juego. Otros conceptos como la *apropiación* remiten a la precisión en la selección léxica, mientras que la noción de *norma* implica una serie de preferencias compartidas en una comunidad lingüística que no siempre se ajustan a la frecuencia de uso (pp. 16-21). Estos aportes teóricos resultan fundamentales para comprender cómo operan tanto las plataformas como los propios estudiantes al momento de producir y revisar textos en LE, y permiten fundamentar los criterios utilizados en este

trabajo para analizar la eficacia de las correcciones y las tensiones que se generan entre lo normativo, lo aceptable y lo comunicativamente eficaz.

2.4. Variables lingüísticas y sociolingüísticas

En consonancia con las categorías y perspectivas ya mencionadas, este trabajo se apoya también en la noción de *variable* de los estudios sociolingüísticos. Introducida por Labov (1966), refiere a una unidad lingüística que presenta más de una forma de realización —las variantes— cuya distribución puede correlacionarse de manera sistemática con factores tanto lingüísticos como sociales. En sus estudios, las *variables lingüísticas* (por ejemplo, la presencia o ausencia de /r/ en inglés neoyorquino) funcionan como dependientes, mientras que las *sociales* (como clase social, edad o género) actúan como independientes, y permiten identificar patrones de estratificación social del lenguaje.

En el marco de este trabajo, las elecciones lingüísticas de los participantes, que implican mayor o menor adecuación textual y gramatical, pueden relacionarse a variables sociolingüísticas. El análisis contempla distintas operaciones de traducción y corrección, exitosas o no, que se expresan con variantes propias del discurso académico. Así, se consideran las variantes léxico-gramaticales seleccionadas en la traducción de nominalizaciones, voz pasiva e impersonales con "se" del texto fuente en español mediante estructuras análogas en inglés o mediante estructuras alternativas. Adicionalmente, se analiza el grado de especialización léxica o el uso de otros recursos discursivos del registro académico. En el estudio de las variantes seleccionadas y su frecuencia se proponen relaciones con algunas de las variables socioacadémicas consideradas, tales como el grado de dominio de inglés —establecido según los seis niveles de competencia definidos por el *MCER*—, la carrera que cursan los participantes, el grado de avance académico, o el conocimiento previo de los sistemas utilizados (v. 4.1.5.1).

La relación entre variables socioacadémicas y lingüísticas, en este sentido, constituye un eje central para la interpretación de los datos experimentales, ya que permite organizar los resultados de acuerdo con los parámetros del experimento (Hernández Campoy y Almeida, 2005, p. 46).

2.5. Inseguridad lingüística

Otro concepto operativo en este trabajo es el de *inseguridad lingüística*, introducido por Labov (1972) y trabajado posteriormente por Bretegnier (1998) y Calvet (2006). Bretegnier explica que:

> le sentiment d'IL apparaît comme lié à la perception, par un (groupe de) locuteur(s), de l'illégitimité de son discours en regard des modèles normatifs à l'aune desquels, dans cette situation, sont évalués les usages ; et partant, à la peur que ce discours ne le délégitime à son tour, ne le discrédite, ne le prive de l'identité, à laquelle aspire, de membre de la communauté qui véhicule ce modèle normatif. C'est ainsi que l'on parle

> de l'insécurité linguistique comme expression d'un sentiment d'exclusion, d'extériorité, d'exogénéité, comme quête d'admission, de communauté, de légitimité linguistique et identitaire. (p. 9)

Esta noción alude a la percepción de que las propias producciones lingüísticas —ya sea a nivel de pronunciación, léxico, gramática o estilo— son inadecuadas o insuficientes frente a variedades consideradas más legítimas o prestigiosas. Tal percepción puede derivar en una sensación de desventaja comunicativa, inhibición o autocensura. Como es esperable, estas experiencias tienden a intensificarse en contextos donde se utiliza una LE. En este sentido, el concepto resulta clave para comprender las elecciones lingüísticas observables en los textos producidos y las percepciones que los hablantes tienen sobre su propio desempeño y el de las herramientas utilizadas. En el marco de esta investigación, la inseguridad lingüística permite interpretar las respuestas obtenidas en las encuestas de percepción realizadas al finalizar la actividad. En dichas encuestas, se consultó a los participantes sobre su grado de satisfacción con los textos que elaboraron, así como sobre la facilidad, la fiabilidad y la utilidad que atribuyen a estos recursos. En particular, se indagó en qué medida estas tecnologías inciden —positiva o negativamente— en sus sensaciones de competencia o (in)seguridad a la hora de producir textos escritos en inglés en un contexto académico-profesional simulado.

2.6. Noción de corpus

Esta investigación se apoya también en la noción de *corpus*, proveniente del enfoque de la Lingüística de Corpus, en tanto permite concebir el conjunto de textos recopilados no solo como un simple insumo, sino como una unidad organizada de datos para el análisis lingüístico. Si bien este concepto ha sido trabajado y abordado por diversos autores como Biber (1988), Sinclair y Carter (1991), Stubbs (1996), Baker et al (2006), McEnery y Hardie (2011) o Yatsko (2014), en el caso de esta investigación, se recurre a la definición de Hunston (2002), que junto con la taxonomía que ofrece aporta un marco útil para describir las características del conjunto de datos construido. La autora define al *corpus digital* como una colección planificada de ejemplos de texto —desde breves hasta grandes conjuntos de datos— diseñada con fines lingüísticos específicos y accesible en formato electrónico (p. 2). A diferencia de archivos o bibliotecas, un corpus *digital* permite un análisis tanto cuantitativo como cualitativo. Si bien el corpus construido para este trabajo posee una configuración particular, dados sus fines específicos, comparte rasgos con distintas clases de corpus digitales tipificados por Hunston (p. 15)[7].

[7] Más razones por las que el corpus de este trabajo comparte características con estos tipos se detallan en la sección 4.2, dedicada a la descripción de sus características y organización. Cabe aclarar que, si bien el conjunto de datos recogido es accesible en formato electrónico y fue trabajado digitalmente, no se trata de un corpus digital en sentido estricto, ya que no fue diseñado con una interfaz y el conjunto de algoritmos necesarios para su análisis.

Por un lado, se vincula con los *corpus de aprendientes*, definidos por la autora como aquellos compuestos de textos producidos por usuarios no nativos con el objetivo de analizar en qué aspectos su desempeño lingüístico difiere del de hablantes expertos o nativos. Este tipo de corpus digital es utilizado en investigaciones sobre aprendizaje y enseñanza de segundas lenguas y LE, y para la evaluación del desarrollo de competencias lingüísticas. En este caso, esta vinculación se explica ya que los textos del corpus fueron producidos por usuarios no nativos del inglés, la mayoría de ellos con niveles de competencia entre intermedio e intermedio bajo (B1 a B2), y sin formación especializada en traducción. Dicha configuración sirve para identificar errores lingüísticos para analizar el desempeño de los estudiantes en situaciones próximas a las demandas comunicativas de su campo disciplinar, cuando estos utilizan recursos digitales como asistentes de corrección y traducción.

Por otro lado, presenta también elementos propios de los *corpus comparables*, que suelen incluir versiones traducidas de los mismos documentos en dos o más lenguas, o bien distintas versiones de un texto en la misma lengua producidas en contextos distintos. Estas colecciones de textos son empleadas en estudios de lingüística comparativa, traducción y procesamiento del lenguaje natural, y permiten el análisis de los patrones de variación entre versiones. El corpus recogido para este trabajo se relaciona con los *comparables* ya que está compuesto por múltiples versiones de traducción de un mismo texto fuente. Si bien todos los textos están redactados en inglés, corresponden a traducciones alternativas realizadas por los participantes bajo condiciones controladas y con diferentes combinaciones de herramientas. Esto permite comparar las modificaciones lingüísticas entre versiones y analizar de qué modo determinados factores o comportamientos —como el sistema utilizado o la intervención del participante— inciden en los resultados.

De este modo, el corpus creado se inscribe dentro de una tradición metodológica que combina el análisis empírico del lenguaje en uso con el de las tecnologías para su estudio.

Por último, este trabajo se vincula también con las Humanidades Digitales (HD). Tal como ha sido definido por autores como Berry (2011) y Rojas Castro (2013), este campo no constituye una disciplina cerrada, sino un espacio transdisciplinar en el que se articulan saberes provenientes de las ciencias humanas y sociales con metodologías propias de la informática. Desde esta perspectiva, las HD ofrecen un marco de reflexión para examinar cómo las tecnologías digitales intervienen en la mediación cultural y en la configuración de nuevas formas de producir, representar y analizar el conocimiento. En este sentido, esta investigación se centra en el análisis de una práctica comunicativa situada, mediada por tecnologías digitales, y en la que, además, se recurre a la noción de corpus digital para estructurar las producciones de los participantes que se analizan con el apoyo de tecnologías digitales que permiten la sistematización y el tratamiento de los datos (v. 4.5.).

3. Antecedentes

En este apartado se presentan antecedentes relevantes para este estudio por abordar fenómenos similares o emplear metodologías comparables. La organización de estos trabajos procura mantener una correspondencia con la organización de las categorías desarrolladas en el marco teórico, ya sea por afinidad temática o por su pertinencia metodológica para el enfoque propuesto.

3.1. Estudios sobre el género informe y sus rasgos característicos

En el apartado teórico, se define al informe como un género discursivo académico con funciones comunicativas y convenciones específicas. Diversos estudios lo han abordado, tanto por su relevancia en la formación universitaria como por sus particularidades discursivas en el marco de la escritura académica. Sin embargo, es relevante destacar que dentro de estos trabajos se emplean diferentes denominaciones y criterios de clasificación según el enfoque o la perspectiva adoptada. A pesar de estas diferencias terminológicas, los estudios que siguen permiten precisar algunas de las características distintivas del informe, así como propiedades que comparte con otros géneros académicos.

En primer lugar, es relevante ubicar el texto utilizado en este trabajo en relación a su función dentro del contexto específico del que proviene. De esta manera, Botta y Warley (2002) establecen una tipología que diferencia el informe universitario del informe de investigación. Este último es un texto científico mediante el que los investigadores comunican los resultados de sus trabajos, y puede concebirse como un género profesional propio de esa comunidad. En cambio, el informe universitario se inscribe en el campo académico y cumple una función formativa, pues permite a los estudiantes aprender a observar, recolectar y organizar datos, así como aplicar nociones teóricas en la interpretación sistemática de los hechos. Además, constituye un ejercicio de entrenamiento para investigaciones más complejas (pp. 20-21). En este sentido, el material con el que se trabajó en esta tesina se sitúa en este último grupo.

En relación con la estructura, Tapia-Ladino y Burdiles Fernández (2009) analizan un corpus de 208 informes escritos por estudiantes universitarios, con el fin de "identificar la superestructura, distinguir los formatos e identificar los tipos de informes en diferentes disciplinas" (p. 17). Las autoras distinguen seis formatos de informe, entre los que se encuentra el de observación. Esta clase, en la que se ubica el informe de laboratorio que nos ocupa, se caracteriza por seguir un protocolo preestablecido en el que los estudiantes deben registrar fenómenos observados directamente, con el fin de comunicar datos específicos. Su función principal es documentar eventos reales, como puede ser la observación de una clase o de un fenómeno en un entorno de laboratorio (p. 39).

En cuanto a las características formales del informe, Ezpeleta Piorno señala que los informes técnicos presentan una macroestructura organizada en apartados que responden a un esquema común

—aunque no siempre se incluyan todos—, con un alto grado de densidad terminológica, y el uso de frases completas, verbos en forma activa y tercera persona; en inglés, además, es frecuente el uso de la voz pasiva (pp. 6-7). En una línea similar, Bosio (2014) caracteriza al informe de investigación académico-científica por su estilo objetivo y preciso, al que contribuyen recursos como el uso de léxico disciplinar específico, construcciones impersonales con "se", la tercera persona, expresiones concisas y la presencia de gráficos de diversa índole (pp. 318-319). Por último, estudios como García Negroni et al. (2005), Olmo (2006) y Gelbes (2016) abordan rasgos característicos de este género tales como la objetividad, la precisión y el uso de las nominalizaciones, así como la desagentivación mediante el uso de la voz pasiva e impersonales con "se". En función de su relevancia para el análisis que aquí se propone, se retoman especialmente estudios que abordan en profundidad estos últimos aspectos en el discurso académico.

En el plano de las selecciones léxico-sintácticas específicas, los rasgos señalados se comparten con otros géneros propios del ámbito académico: numerosos estudios han señalado la centralidad de recursos como la nominalización y la voz pasiva en la construcción de textos propios de este ámbito. Entre ellos, destacamos a Banks (2023), quien analiza la frecuencia de nominalizaciones en artículos científicos según autor y época, con una tendencia actual de un caso cada 10 a 14 unidades léxicas, frente a 1 cada 22 en textos más antiguos (p. 11). Además, este recurso es más frecuente en ciertas disciplinas, donde contribuye a un discurso cargado de información, propio de la escritura técnica. Holtz (2009) observa que las nominalizaciones son más comunes en textos de biología e ingeniería mecánica que en los de informática o lingüística (p. 18). Este dato resulta pertinente, dado que el corpus analizado proviene de una materia de Bioquímica, vinculada estrechamente con la biología, y optativa en la carrera de Licenciatura en Ciencias Biológicas. En el caso de la voz pasiva, por su parte, se puede mencionar a Banks (2017), quien a partir de un estudio diacrónico de artículos de las *Philosophical Transactions of the Royal Society*, identifica un aumento en su frecuencia dentro de las ciencias biológicas durante el siglo XX, asociado al paso de un enfoque descriptivo a uno experimental (p. 100). Sin embargo, Subagio et al. (2019) señalan una disminución de esta estructura en las primeras décadas del siglo XXI, impulsada por la búsqueda de mayor claridad y concisión. A partir de un análisis de corpus, reportan que la voz pasiva aparece en un 30 % de las cláusulas (p. 7). No obstante, una encuesta realizada por los investigadores a estudiantes de posgrado en Singapur muestra que muchos de ellos desconocen esta tendencia: el 90 % considera que su uso sigue siendo necesario y lo asocia con profesionalismo y objetividad (p. 2). En este sentido, aunque su frecuencia ha disminuido, la voz pasiva se mantiene como una característica relevante en el análisis de textos científicos como los que integran el corpus de este trabajo.

3.2. Estudios sobre competencias comunicativas profesionales (CCP)

En consonancia con lo desarrollado en la sección teórica sobre CCP, a continuación se presentan investigaciones que, si bien se sitúan en otros campos disciplinares, comparten con esta propuesta la preocupación por el desarrollo de habilidades comunicativas vinculadas a la práctica profesional durante la formación universitaria.

En primer lugar, Włoszczak-Szubzda y Jarosz (2012) analizan, a partir de una revisión bibliográfica, el desarrollo de las CCP en estudiantes y trabajadores del área de enfermería en Polonia. A partir de esta revisión, los autores sostienen que el conocimiento teórico adquirido en materias como psicología y psicoterapia no se traduce de forma espontánea en competencias comunicativas efectivas en la práctica profesional, lo que lleva a la necesidad de implementar estrategias sistemáticas de formación complementaria (p. 187). Este antecedente aporta una perspectiva crítica sobre la brecha entre los saberes disciplinares y su transferencia al ámbito comunicativo profesional, una problemática presente en nuestro objeto de estudio.

También en un contexto de formación profesional, Khudyakova y Filatova (2013) estudian la formación de las CCP en estudiantes de psicología en Rusia, a través de un diseño experimental en tres etapas. Las autoras parten de una evaluación diagnóstica, desarrollan una serie de actividades formativas, y evalúan nuevamente el desempeño de los estudiantes. Concluyen que la formación de las CCP se ve favorecida por ciertas condiciones psicológicas, educativas y actitudinales, así como por la realización de tareas que simulan situaciones propias del futuro ejercicio profesional. Este trabajo resulta relevante en tanto propone un abordaje experimental para observar el desarrollo de habilidades comunicativas específicas en estudiantes universitarios, enfoque que dialoga con el diseño metodológico adoptado en nuestra investigación.

En el contexto local, Starc y Martino (2022) indagan en las percepciones de estudiantes de la UNS acerca de su preparación para comunicarse en situaciones profesionales, tanto en su lengua materna como en una extranjera. En particular, observan que un número significativo de los encuestados manifiesta sentirse poco preparado para desenvolverse de forma oral en inglés en contextos laborales. Este trabajo refuerza la relevancia de considerar el desarrollo de las CCP en una LE como un desafío relevante en la formación universitaria actual. En este sentido, los resultados obtenidos por Starc y Martino subrayan la necesidad de seguir explorando las condiciones que favorecen el desarrollo de las CCP en LE, en especial en áreas donde el uso del inglés adquiere un papel relevante en situaciones profesionales específicas, como la abordada en esta tesina.

3.3. Tratamiento de las intervenciones lingüísticas sobre los textos

Son también relevantes para nuestra investigación los aportes de la correctología y la posedición, ya que la actividad de la que surgen los datos para este trabajo involucra la revisión y corrección de traducciones producidas con asistencia de herramientas digitales. Esto plantea la necesidad de analizar las modificaciones entre diferentes versiones de una misma producción escrita, con el objetivo de identificar transformaciones lingüísticas específicas realizadas por los participantes. Tanto en corrección como en posedición, se han desarrollado trabajos que analizan las modificaciones aplicadas en las sucesivas versiones de un texto.

En el ámbito de la corrección de textos, Wates y Campbell (2007) examinan las diferencias entre la versión original de los autores y la versión final publicada. Consideran factores como el número de consultas editoriales, los cambios tipográficos, las modificaciones sugeridas e ignoradas, y las alteraciones realizadas por los propios autores (pp. 123-125). Por su parte, Robblee (2017) analiza los enfoques que los editores adoptan al comentar propuestas de subvención, los tipos de comentarios que realizan y las ediciones efectuadas. La autora clasifica las intervenciones editoriales en categorías como ediciones de: a) contenido, b) gramaticales y mecánicas, c) de puntuación, diseño y estilo (p. 36).

En el campo de la posedición, el estudio de Koponen, Salmi y Nikulin (2018) examina las correcciones introducidas por poseditores en textos generados por tres sistemas diferentes de traducción automática. Las modificaciones identificadas incluyen: a) sustituciones léxicas, b) cambios en la morfología verbal, c) reordenamientos sintácticos, d) adiciones, e) eliminaciones (p. 63). Estos hallazgos permiten establecer categorías para clasificar las transformaciones lingüísticas observadas en las producciones analizadas.

Si bien esta investigación no adopta directamente las taxonomías propuestas en los estudios mencionados, dichos aportes resultan valiosos por evidenciar un interés sostenido en clasificar y analizar las intervenciones lingüísticas en distintas etapas de los procesos de producción textual. En nuestro caso, la categorización utilizada fue construida a partir del análisis de las intervenciones realizadas por los participantes, con el propósito de desarrollar una taxonomía adecuada a los objetivos específicos de este estudio (v 4.1.4.1). La coincidencia entre algunas de las categorías identificadas y las presentes en investigaciones previas permite respaldar la validez de nuestra propuesta y situarla en diálogo con estos campos de trabajo.

3.4. Tecnologías aplicadas a la comunicación

Otro conjunto relevante de antecedentes para este plan de investigación está constituido por los estudios sobre tecnología aplicada a la comunicación, en particular aquella orientada a la asistencia lingüística de los hablantes. Llisterri (2003) señala que el desarrollo de estas tecnologías —hoy cada

vez más extendidas— involucra no solo a informáticos e ingenieros, sino también a profesionales de la lingüística, cuyo rol específico analiza en su artículo. Además, propone una clasificación de estas tecnologías en dos grandes grupos, ambos dependientes de una serie de recursos lingüísticos que exceden el campo de la informática, como corpus, diccionarios y gramáticas (p. 2).

El primer conjunto, denominado "tecnologías del habla", abarca aquellas herramientas informáticas que procesan la lengua oral y permiten que una computadora reciba, ofrezca y procese información hablada (p. 3). El segundo grupo, "tecnologías del texto", comprende aquellas centradas en el uso escrito de la lengua que subdivide, a su vez, entre las que se utilizan para procesar el lenguaje escrito y las que se destinan al desarrollo de *software* (p. 19). De este modo, los recursos cuyo empleo nos proponemos estudiar, como los correctores ortográficos y gramaticales integrados en la mayoría de los procesadores de documentos y los correctores estilísticos, como *Grammarly*, forman parte de las tecnologías del texto, pues asisten a las prácticas de escritura (p. 19). También se incluyen aquí las tecnologías de generación de lenguaje (p. 25), entre las que se destacan los recientes avances en sistemas de IA como *ChatGPT*, y las herramientas de traducción automática (p. 26), como *Google Translate*.

En relación con su uso en tareas de traducción y corrección, se destacan algunas investigaciones relevantes para esta propuesta. Se abordan aquí dos estudios representativos sobre *Google Translate*: uno enfocado en la comunicación profesional y otro en la comunicación académica. Patil y Davies (2014) analizan su eficacia en situaciones profesionales específicas, como la comunicación entre pacientes y médicos que no comparten una lengua. El estudio —realizado durante la etapa de traducción estadística de *Google Translate*[8]— y basado en combinaciones del inglés con otras 26 lenguas, reveló que solo un 57.7% de las traducciones fueron consideradas correctas. Por este motivo, los autores desaconsejan su uso en contextos médicos, salvo en situaciones extremas en las que no sea posible recurrir a un traductor humano y se requiera una intervención urgente. Un aspecto destacado del estudio es la diferencia significativa en la precisión de las traducciones según el grupo de lenguas: las traducciones hacia lenguas europeas occidentales alcanzaron un 74% de aciertos, mientras que en lenguas africanas descendieron al 45%. Este hallazgo señala la existencia de un posible sesgo en el sistema, al menos durante la primera década de funcionamiento de esta plataforma.

El segundo trabajo que se destaca es el de Kol et al. (2018), que indaga los posibles beneficios del uso de este traductor en cursos de escritura académica en inglés. Mediante un experimento, se solicitó a estudiantes que completaran dos tareas de escritura: una sin la ayuda de *Google Translate* y otra con su asistencia. Los resultados mostraron mejoras en los textos generados con el traductor automático: mayor legibilidad y un léxico más complejo. Los autores aclaran que no se espera que los estudiantes

[8] V. sección 4.1.2.1

incorporen de forma inmediata todo el vocabulario nuevo. Sin embargo, resaltan que sí se vieron expuestos a nuevos recursos léxicos, lo que constituye un primer paso hacia la ampliación de su vocabulario. Ambos estudios muestran que, si bien la herramienta presenta limitaciones y sesgos, puede favorecer aspectos de la producción escrita y brindar apoyo significativo a usuarios con un conocimiento intermedio de la lengua, en línea con las hipótesis de este trabajo.

El surgimiento de *Grammarly* y sus posibles aplicaciones también han despertado el interés de la comunidad científica. En particular, Nova (2018) analizó sus ventajas y desventajas en la evaluación de la escritura académica de estudiantes de inglés como lengua extranjera en Indonesia. A partir de entrevistas con tres usuarios, concluyó que se trata de una plataforma útil que ofrece retroalimentación beneficiosa para mejorar la escritura y detectar errores. También valoró su disponibilidad gratuita, la rapidez con que genera sugerencias y la posibilidad de utilizarla de manera inmediata, sin configuraciones complejas ni costos asociados (p. 85). No obstante, señaló limitaciones, como correcciones que no siempre se ajustan a la intención comunicativa del usuario, una corrección excesiva en secciones como las referencias, y ocasionales deficiencias en el tratamiento de información contextual. Estos aspectos fueron considerados en este trabajo al evaluar el desempeño de dicha herramienta en la corrección de diferentes fenómenos lingüísticos.

Por su parte, O'Neill y Russell (2019) investigaron la percepción del uso de *Grammarly* combinado con asesoramiento académico. En un experimento con dos conjuntos de estudiantes universitarios australianos, uno de ellos (grupo experimental) recibió correcciones del programa y de asesores humanos, mientras que el otro (grupo de control) solo recibió correcciones humanas. El primer conjunto de participantes expresó una respuesta más positiva a las correcciones, atribuido a que *Grammarly* ofreció más correcciones gramaticales que los asesores, quienes tienden a enfocarse en la macroestructura. Los autores atribuyen tal respuesta a que este sistema puede complementar el trabajo humano, aunque aún requiere supervisión profesional, dado que no es completamente precisa para un uso autónomo (p. 1). Estos hallazgos permiten dimensionar tanto las posibilidades de mejora léxica y sintáctica que ofrece esta tecnología, como la importancia de su uso acompañado por una reflexión crítica por parte del usuario, en línea con las hipótesis de este trabajo, que abordan la contribución de estas tecnologías a la adecuación textual y la conciencia crítica de los participantes respecto de su uso.

A pesar de su reciente aparición, ya se han publicado estudios que exploran los usos potenciales de *ChatGPT* en distintos ámbitos de la comunicación. Por ejemplo, Lechien et al. (2023) evaluaron la capacidad de *ChatGPT*-4 para revisar manuscritos en el área de otorrinolaringología y concluyeron que la herramienta tiende a proponer correcciones inapropiadas, en especial en cuestiones léxicas y en el uso de preposiciones. Otro caso es el de Barrio Andrés (2023), quien analiza sus posibles aplicaciones en el campo jurídico para la generación y revisión de documentos, la transformación de textos y la

interacción con clientes. Si bien reconoce cierto grado de utilidad, también advierte limitaciones que restringen su empleo en contextos profesionales.

Uno de los trabajos relevantes en el ámbito de la comunicación académica es el de Castillo-González et al. (2022), que analiza la utilidad de tecnologías de IA generativa en tareas de revisión de textos. Los autores subrayan que el editor científico debe ser capaz de identificar errores, proponer mejoras y asegurar el cumplimiento de normas de estilo y formato, por lo que contar con recursos adecuados —incluidas las basadas en IA— resulta esencial (p. 2). En este marco, reconocen el potencial de *ChatGPT* para asistir en la revisión de manuscritos producidos en inglés por autores cuya lengua materna es otra, una situación habitual en la comunicación científica internacional.

En esta misma línea, Kim (2023) sostiene que la herramienta puede ser una alternativa más accesible y eficiente frente a los servicios de edición pagos, capaz de generar párrafos editados en segundos y de forma gratuita. Según el autor, *ChatGPT* produce oraciones más refinadas que los textos originales sin corregir, lo que contribuye a mejorar la claridad y precisión de la escritura (p. 2). No obstante, advierte sobre la necesidad de validar cualquier idea generada por el sistema, debido al riesgo de producir información falsa.

Por su parte, Lopezosa (2023) destaca tanto los beneficios como los desafíos asociados a la incorporación de *ChatGPT* en la práctica investigativa. Entre las ventajas que supone su aplicación en el ámbito de la escritura, el autor destaca que:

> es capaz de escribir marcos teóricos (con mayor o menor acierto), proponer resúmenes y palabras clave a partir de un manuscrito, mejorar la redacción de un texto, e incluso eliminar reiteraciones y párrafos para mejorar el flujo narrativo del texto final (p. 19)

Asimismo, bajo la consideración de que existen preocupaciones sobre las implicaciones éticas del uso de este sistema, propone una serie de lineamientos para su uso responsable en la comunicación científica y ejemplifica su aplicación en investigaciones recientes. En particular, considera que esta IA (y otras similares) deben ser consideradas asistentes y no co-autores en una investigación, ya que a estos no se les puede atribuir responsabilidad por los resultados. Además, el empleo de estos modelos debe ser completamente transparente, y la responsabilidad final siempre recae en el investigador.

Estos aportes permiten dimensionar el papel que recursos como *ChatGPT* pueden asumir en prácticas comunicativas profesionales especializadas, y ofrecen un punto de partida útil para evaluar sus efectos en la producción escrita en inglés por parte de hablantes no nativos, en consonancia con los objetivos de esta tesina.

4. Metodología

De acuerdo con las hipótesis y los objetivos de esta investigación, se diseñó una metodología que incluye tareas orientadas a la conformación del corpus y su análisis. En esta sección se detalla el diseño metodológico, que comprende el análisis de producciones escritas generadas en un experimento con herramientas tecnológicas, la consideración de las percepciones de los participantes sobre su experiencia con ellas y la evaluación de las producciones por jueces expertos.

4.1. Justificación y descripción del diseño experimental

Con el propósito de explorar el desempeño de usuarios cuya formación no se centra en la lengua como objeto de estudio, se optó por una carrera perteneciente al campo de las Ciencias Naturales. Esta decisión se sustenta en la hipótesis de que los estudiantes provenientes de disciplinas que no abordan las prácticas escriturales como parte de su reflexión académica podrían enfrentar un doble desafío al producir textos escritos en inglés con fines profesionales. Por un lado, afrontan la producción de un género cuyos rasgos estilísticos no suelen ser abordados como objeto de estudio específico en su disciplina[9], y por otro, deben hacerlo en una LE, sin que se les exija en su formación el dominio de las prácticas escriturales en dicha lengua, más allá de la comprensión escrita. En este marco, se seleccionó la carrera de Bioquímica dictada en la UNS, por tratarse de una disciplina con fuerte presencia de géneros discursivos que requieren precisión terminológica y claridad expositiva. Se eligió, específicamente, la asignatura Bioanalítica II, correspondiente al tercer año de la carrera, por considerarse un momento intermedio en la formación de los estudiantes, en el que ya han adquirido cierta familiaridad con los géneros propios de su campo, sin haber concluido aún su recorrido formativo. En esta asignatura, se requiere a los estudiantes la redacción de informes escritos como parte de diversas actividades prácticas. El texto en español de uno de los informes de laboratorio elaborados en esta materia fue cedido para su utilización en el presente estudio y funcionó como insumo para el diseño de la actividad experimental.

Para recoger los datos, entonces, se diseñó una actividad experimental que consistió en la elaboración de traducciones al inglés de fragmentos seleccionados del informe de laboratorio en español mencionado por parte de 20 estudiantes pertenecientes a carreras del área de las Ciencias Naturales, asistidos por herramientas digitales. La elección de este género se justifica en función de su frecuencia en la formación académica y en el ámbito profesional de las disciplinas (v. 2.1.). Para la elaboración de las traducciones, se diseñó una secuencia de trabajo en la que los participantes realizaron una traducción

[9] Si bien los estudiantes de esta área practican la escritura de informes de laboratorio, no se abordan de manera específica aspectos formales como el uso de la voz pasiva o las nominalizaciones como recursos desagentivadores.

autónoma inicial del texto fuente a partir de sus propios conocimientos de inglés y con la única asistencia de un diccionario electrónico bilingüe inglés-español[10], en una computadora que les fue provista. Luego, los estudiantes alternaron entre el uso individual y combinado de *Google Translate*, *Grammarly* y *ChatGPT*, así como en la realización de intervenciones manuales en los textos, para producir versiones alternativas.

Antes del experimento, se solicitó a cada participante que completara un perfil sociolingüístico y académico a través de un cuestionario de *Google Forms* (v. anexo 8.3.), elaborado a partir de los lineamientos propuestos por Hernández Campoy y Almeida (2005). Se relevaron variables sociodemográficas —edad y género—, así como datos vinculados a su trayectoria académica, nivel de conocimiento de inglés y experiencia previa con las herramientas a ser utilizadas. En los casos en que no se registró familiaridad con alguna de ellas antes de comenzar el experimento, se ofreció una explicación previa de sus funcionalidades y modos de uso, como parte del protocolo de trabajo. Esta encuesta permitió relevar información acerca de variables socioacadémicas y tecnológicas de interés sobre los participantes y su formación antes del experimento (v. 4.1.5.1).

Como parte del compromiso ético de las investigaciones en Ciencias Sociales (Wiles et al., 2005), se garantizó el anonimato de los participantes y de sus respuestas. Asimismo, se les brindó la posibilidad de retirar sus datos de la investigación en cualquier momento, si así lo desearan. La actividad se llevó a cabo en el marco de una observación controlada, en consonancia con lo propuesto por Hernández Campoy y Almeida (2005). Como resultado, cada participante generó una traducción inicial autónoma, cinco versiones producidas mediante el uso de estas tecnologías, y cinco versiones corregidas manualmente a partir de esas salidas.

4.1.1. El texto en español

Dado que traducir el informe completo hubiera resultado en un experimento demasiado extenso para el tiempo disponible en las sesiones con los participantes —y, por lo tanto, poco viable para su implementación—, se optó por emplear fragmentos representativos. Si bien esto implicó trabajar con un texto segmentado, se procuró mantener la coherencia discursiva general del documento, de modo tal que los fragmentos conservan una conexión temática y funcional que permite reconstruir el sentido del informe original. Los fragmentos utilizados (v. anexo 8.1.) corresponden a las secciones de introducción, discusión y conclusión, que fueron seleccionadas por poseer una mayor densidad de rasgos lingüísticos relevantes para los objetivos del estudio, así como por su representatividad dentro del género informe.

[10] El diccionario electrónico utilizado corresponde al *Cambridge Dictionary* (s.f).

Estos fragmentos (v. anexo 8.2.) fueron intervenidos con el propósito de concentrar fenómenos lingüísticos que, si bien ya estaban presentes en el informe original, resultaban especialmente relevantes para este estudio por su frecuencia en el *género* y su complejidad para los procesos de traducción automática y para usuarios no expertos. En particular, se incorporaron de forma deliberada construcciones de nominalización, así como oraciones en voz pasiva, tanto casos de pasiva perifrástica como construcciones con "se" (pasivas reflejas e impersonales). Este enfoque buscó orientar el análisis hacia la manera en que los participantes emplean las herramientas frente a problemáticas concretas y recurrentes en este tipo de producciones.

Además, durante la preparación del texto fuente para el experimento, se decidió mantener algunos errores del texto original, como imprecisiones de significado y problemas de puntuación, con el objetivo de observar si los estudiantes y los sistemas los detectaban y eran capaces de corregirlos a la hora de hacer las traducciones. En concreto, se destaca la conservación de la unidad léxica *altercados* en una de las cláusulas[11] del texto original. En este caso, se trata de un error de selección léxica de los propios autores, ya que el término —que alude a discusiones o enfrentamientos— fue utilizado en lugar de una unidad más adecuada como, por ejemplo, *complicaciones*.

4.1.2. Tecnologías utilizadas

Las tres plataformas utilizadas se seleccionaron por su amplia difusión y relevancia en contextos sociales y profesionales. En primer lugar, *Google Translate* es uno de los traductores automáticos más reconocidos y empleados a nivel global (Google Corporation, 2023). *Grammarly*, por su parte, se ha consolidado como una de las plataformas líderes en corrección gramatical y estilística, con más de 30 millones de usuarios activos al momento del estudio (Grammarly Inc, 2023). Por último, si bien *ChatGPT* fue desarrollado y lanzado con posterioridad a las otras dos herramientas anteriores (OpenAI, 2022), su adopción ha crecido de forma vertiginosa y ha alcanzado una presencia destacada en ámbitos como la educación, el trabajo y la producción de contenidos digitales. Esta rápida expansión, junto con su potencial para asistir en tareas de escritura y corrección, fundamenta su inclusión en este análisis. Cabe señalar que tanto *Grammarly* como *ChatGPT* ofrecen versiones gratuitas y de pago, estas últimas con funcionalidades ampliadas. En este trabajo, se optó por analizar el uso de las versiones gratuitas, con el objetivo de simular un escenario de uso verosímil para hablantes nativos de español sin formación lingüística específica. Esta decisión responde a la necesidad de reflejar condiciones de acceso realistas en contextos donde la adquisición de versiones *premium* puede no estar al alcance de muchos usuarios, incluso en ámbitos académicos o profesionales nacionales. De este modo, se buscó reproducir una

[11] La cláusula en cuestión es: *Si bien se produjeron altercados y resultados que no habían sido esperados (ver conclusión) los objetivos generales fueron cumplidos* (Informe original en español).

situación de revisión textual con los recursos efectivamente disponibles para todos los participantes, neutralizando en el escenario experimental el impacto de una posible brecha digital.

4.1.2.1. Google Translate

La primera de las tecnologías utilizadas en esta actividad, *Google Translate*, es un sistema que permite traducir textos de distintas extensiones entre diferentes idiomas. En su desarrollo pueden identificarse dos etapas. La primera, iniciada el 28 de abril de 2006, corresponde a un enfoque de traducción meramente estadística, basada en frases predefinidas y probabilidades de ocurrencia. La segunda, en la que se inscribe el desarrollo del experimento, fue implementada a partir del 15 de noviembre de 2016, y marca la transición hacia la traducción neuronal, que incorpora técnicas de aprendizaje automático, entre ellas, redes neuronales, para generar traducciones más precisas y contextualmente adecuadas (Schuster et al., 2016). Este enfoque aún utiliza conceptos probabilísticos, pero lo hace de forma más avanzada y adaptativa, lo que permite un procesamiento más sofisticado de las relaciones lingüísticas y semánticas entre el texto fuente y el texto meta.

Un aspecto importante a considerar es que, dado que la recolección de datos se extendió durante varios meses (desde el 5 de abril de 2024 hasta el 6 de marzo de 2025), el resultado generado por *Google Translate* para la versión T1 no fue idéntico para todos los experimentos, incluso cuando tradujeron el mismo texto fuente. Esta variabilidad responde al carácter dinámico del sistema, que se actualiza con mejoras de forma constante. En consecuencia, para los participantes E00 a E06, la plataforma produjo una versión con mayor cantidad de errores que para los participantes E07 a E19, en cuyo caso dicha cantidad se redujo de forma significativa. Por último, en el caso del último experimento (E21), que completó la actividad más de cinco meses después, el resultado de la traducción volvió a modificarse, aunque sin que ello implicara un aumento o disminución en el número total de fallas. Estos aspectos contextuales ilustran una de las complicaciones inherentes al trabajo con tecnologías en evolución constante.

4.1.2.2. Grammarly

Por su parte, *Grammarly* es una herramienta digital que sugiere correcciones gramaticales, ortográficas, de puntuación y mejoras estilísticas en inglés. Cabe señalar que la versión gratuita permite el acceso diario a dos sugerencias *premium* sin necesidad de pago. Dado que estas se presentan al usuario marcadas en el texto con los mismos colores que las gratuitas —y solo se identifican como *premium* una vez agotado ese límite—, resultaba poco realista impedir su uso tras haber sido visualizadas por los participantes. Por este motivo, se decidió permitir a los estudiantes utilizarlas en el marco de la actividad. Esta decisión busca reflejar una situación de uso en la que los usuarios

aprovechan las funcionalidades que las propias empresas ofrecen como parte de sus estrategias de promoción y captación de usuarios.

4.1.2.3. ChatGPT

Por último, *ChatGPT* es un *chatbot* basado en el modelo de lenguaje GPT-3.5[12], desarrollado por *OpenAI*, y entrenado para generar texto en respuesta a una amplia variedad de entradas. Este recurso puede ser utilizado para múltiples fines, como asistente en procesos de escritura, corrección o traducción de textos, aunque en este estudio se empleó sólo para la corrección de textos. Es importante considerar su notable flexibilidad tanto en los tipos de solicitudes que puede recibir como en las respuestas que puede generar. Si bien esta característica puede ser una fortaleza, también plantea un desafío metodológico: la variabilidad en las salidas puede dificultar la comparación entre producciones corregidas para diferentes participantes. A fin de minimizar este factor, se instruyó a todos los participantes a utilizar una única instrucción (*prompt*) al solicitar la corrección de sus traducciones: "Revisar y corregir los siguientes fragmentos de un informe de laboratorio de bioquímica, en inglés. La gramática, la sintaxis y el estilo deben ser precisos y apropiados para un contexto técnico. Mantener el lenguaje técnico y científico en estos fragmentos. Los fragmentos que requieren corrección son los siguientes: *[Fragmentos]*", en la que *[Fragmentos]* es reemplazada por el texto traducido por los estudiantes. Esta instrucción fue diseñada con el objetivo de estandarizar, en la medida de lo posible, el tipo de respuesta generada por la herramienta, y permitir así un análisis comparativo más consistente. No obstante, se reconoce que esta medida no elimina por completo la variabilidad inherente a la generación de texto por parte del modelo. En futuras investigaciones centradas exclusivamente en el uso de *ChatGPT* o un modelo de lenguaje de funcionamiento similar, podría resultar relevante explorar cómo las diferencias en las instrucciones afectan los resultados generados, así como indagar en la posible existencia de patrones o regularidades dentro de esa variabilidad. Cabe señalar, además, que existe la posibilidad de que el historial de uso haya influido en las respuestas ofrecidas por las herramientas, en consecuencia, se considera un factor potencial que podría haber incidido en los resultados.

4.1.3. Etapas del experimento

Luego de la instancia de relevamiento de datos personales (v. 4.1.), se dio inicio a la etapa de producción textual. En ella, y a partir de la selección de estos fragmentos en español, que denominaremos versión TE (texto en español), los participantes elaboraron un total de once traducciones alternativas cada uno. Estas traducciones se agrupan en dos categorías principales: las

[12] Actualmente, la versión gratuita de *ChatGPT* utiliza el modelo GPT-3.5, aunque es posible acceder a una versión basada en GPT-4 mediante una suscripción paga.

versiones *T* son aquellas generadas con asistencia tecnológica, y las *V* son aquellas que implicaron una intervención activa por parte de los participantes. La primera traducción autónoma también se incluye en este segundo grupo (V1).

De acuerdo con el diagrama de flujo presentado en la figura 1, el procedimiento se inició con la traducción de TE al inglés por parte de los participantes, sin asistencia de sistemas digitales, pero con acceso al diccionario electrónico bilingüe español-inglés (disponible también en las etapas posteriores). Esta primera traducción fue designada V1.

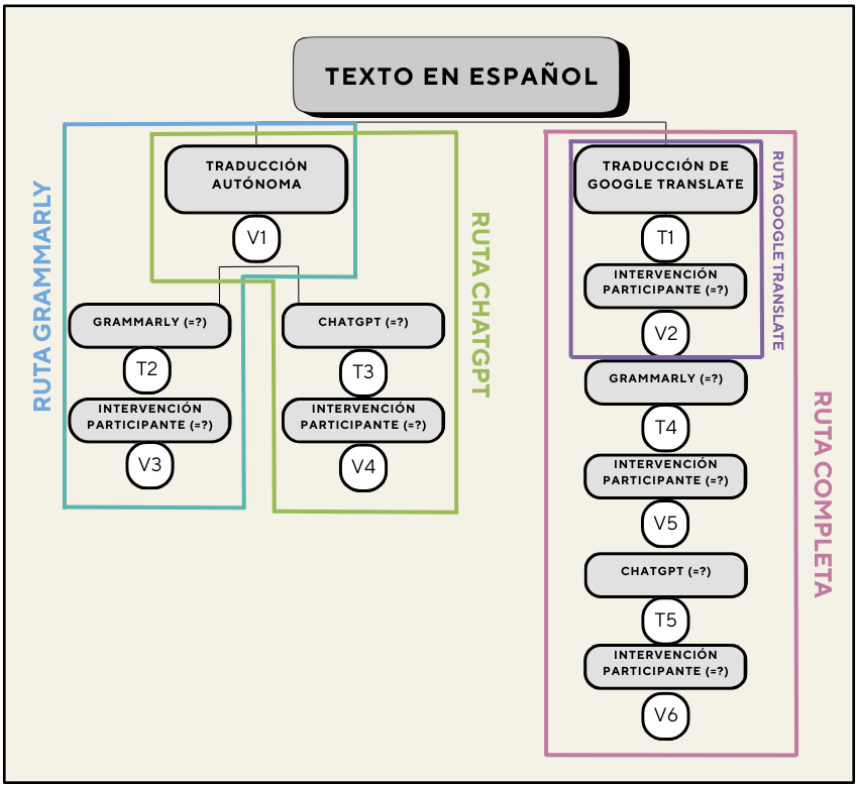

Figura 1: Diseño del experimento

A continuación, a partir de TE, se generó una versión traducida por *Google Translate*: T1. Este texto fue luego revisado por los participantes en base a su propio juicio, para dar lugar a la versión intervenida V2. En una tercera etapa, los participantes editaron V1 con *Grammarly* y *ChatGPT*. Los resultados de estas ediciones se denominaron T2 (para *Grammarly*) y T3 (para *ChatGPT*). Luego, cada una de estas versiones fue revisada e intervenida nuevamente por los participantes, para producir V3 (a partir de T2) y V4 (a partir de T3).

Hasta este punto, cada herramienta ha sido utilizada una única vez, de forma aislada, como puede observarse en las ramas verdes del diagrama de flujo. Posteriormente, los participantes continuaron a partir de V2, y generaron una secuencia de traducción que integró el uso de las tres plataformas de manera secuencial (representada mediante una línea punteada en el diagrama). En esta última fase, V2 fue editada con *Grammarly*, lo que dio lugar a T4, que fue luego intervenida por los participantes para

obtener V5. A continuación, V5 fue editada con *ChatGPT*, para dar lugar a T5, que también fue intervenida, y produjo la versión final V6.

En algunos casos, los participantes decidieron no modificar sus versiones T. Esta ausencia de intervención fue considerada como una operación lingüística en sí misma, ya que refleja conformidad con el resultado ofrecido por el sistema. En estos casos, la versión T se tomó como equivalente a la versión V correspondiente. Esta situación está representada en el diagrama mediante los símbolos "(=?)".

4.1.4. Categorías de análisis y procesamiento del corpus

En esta sección se presenta el procedimiento utilizado para sistematizar y registrar los datos obtenidos en el estudio. Se puntualiza cómo se aplicaron las categorías vinculadas con las variables lingüísticas de interés para el análisis que operaron en la organización y clasificación de las unidades analizadas, así como a las estrategias empleadas para el análisis de los textos y las herramientas que facilitaron el registro de los datos.

4.1.4.1. Categorías de análisis

Las categorías empleadas para el análisis lingüístico de los textos producidos por los participantes incluyen estructuras relevantes en el discurso académico, como la voz pasiva, las impersonales con "se" y las nominalizaciones, así como errores gramaticales y textuales. Además, se clasifican y describen los tipos de intervenciones realizadas por los participantes.

Voz pasiva e impersonales con "se"

Aunque este trabajo se centra en la identificación de estructuras pasivas en las traducciones al inglés, estos textos se originan a partir de un informe de laboratorio en español. Por lo tanto, se consideran primero los casos de voz pasiva en el texto fuente en español, y la posibilidad de que estos sean traducidos manteniendo la pasividad en el texto en inglés. Cabe aclarar que, si bien algunas construcciones en español analizadas corresponden a oraciones impersonales con "se" (por ejemplo: "se llegó a la conclusión de que…"), estas fueron consideradas dentro del análisis de voz pasiva por su equivalencia funcional con pasivas en inglés académico, como "it was concluded that…". Como fue explicado en el apartado teórico y el de antecedentes, en este tipo de discurso, tanto en español como en inglés, se prefiere un estilo impersonal y objetivo, y tanto la voz pasiva como las impersonales son de los recursos más utilizados para ello. Por esta razón, se incluyó en esta categoría analítica a las impersonales con "se" que se esperaba que fueran traducidas mediante estructuras en voz pasiva en inglés.

En este trabajo se considera para el análisis de la voz pasiva solo aquellas construcciones en inglés que incluyen el verbo "*to be* (auxiliar) + participio", excluyendo otras posibilidades de formación de la

voz pasiva. Esta decisión responde al interés por evaluar no solo en qué medida los participantes recurren a una construcción característica del discurso académico, sino también cómo resuelven su formulación en inglés, ya que este tipo de estructuras suele presentar dificultades tanto para usuarios no expertos como para los traductores y correctores automáticos utilizados.

Dado que se trata de textos en inglés traducidos desde el español por hablantes no expertos en LE, se asume la posibilidad de desviaciones de la norma gramatical en los textos, lo que incluye las construcciones pasivas. Esto implica que se consideran tanto las estructuras pasivas gramaticalmente correctas como aquellas que, aunque presenten errores, evidencian la intención de utilizar estas construcciones. Se adopta así un análisis funcional, que permite observar tanto los logros como las dificultades en la producción escrita de los participantes.

Nominalizaciones

De manera similar, se consideran primero las nominalizaciones presentes en el texto fuente en español, y la posibilidad de que se traduzcan como nominalizaciones al inglés. Para ello, se toma como referencia la *Nueva gramática de la lengua española de la Real Academia Española* (2009, pp. 337-411), que aporta una guía sistemática en relación a los sufijos de derivación morfológica característicos de estos sustantivos en español.

Si bien es posible nominalizar distintos elementos del lenguaje (Halliday y Matthiessen (2004), este trabajo se enfoca solo en el uso de nominalizaciones deverbales y deadjetivales, ya que son las más frecuentes en los estudios sobre discurso académico. Se excluyen las denominales y estructuras complejas a partir de cláusulas como "what the duke gave to my aunt" (p. 69), debido a que, en los fragmentos analizados, su aparición es escasa o nula, lo que vuelve poco productivo su abordaje en el marco del presente análisis. En particular, se atiende a las deverbales de proceso, por su función en la desagentivación de los textos científicos. Como plantea Halliday, este tipo de metáforas gramaticales añade significado al centrar el foco en el proceso más que en el agente (p. 638). Se incluyen también algunos casos de nominalizaciones instrumentales (Comrie y Thompson, 2007, p. 1), como *reagent*, por su importancia terminológica en los textos analizados, y algunos casos ambiguos entre proceso y resultado, como *measurements* en el ejemplo citado por Banks (2023, p. 12), donde ambos sentidos resultan inseparables.

Con igual criterio que para la voz pasiva, en el caso de las nominalizaciones se considera toda unidad léxica que cumpla en el texto en inglés una función nominal[13], independientemente de si presenta un error morfogramatical en su construcción. Así, en la lista de nominalizaciones detectadas de forma manual, el estudio prioriza un enfoque funcional, más que de corrección gramatical, debido a

[13] O debería de cumplirla, habiendo solucionado dicho error.

la necesidad de constatar también los casos que presenten problemas, que pueden ofrecer información valiosa sobre las estrategias y dificultades de los participantes en la escritura académica en inglés.

Errores

Para la clasificación de errores, se adoptó la tipología de competencias gramaticales y textuales propuesta por García Negroni y Estrada (2006), como se explicita en el marco teórico. Esta taxonomía permite analizar los fallos presentes en el corpus, así como identificar sus tipos más frecuentes y la capacidad de los participantes para resolverlos con la asistencia de las plataformas. En cuanto a los gramaticales, esta categoría contempla los errores que afectan la conformación interna de las unidades lingüísticas y su estructuración en la oración, y fue organizada en tres niveles según el enfoque propuesto por las autoras: fonemático, morfológico y sintáctico.

1. Errores fonemáticos: corresponden a casos vinculados con la representación gráfica de los fonemas, cuya manifestación puede afectar la inteligibilidad del texto. Incluyen dificultades de ortografía, acentuación y uso inapropiado de mayúsculas.
2. Errores morfológicos: se refieren a fallas en la construcción de unidades léxicas o en la flexión de sus componentes. Incluyen errores de concordancia, uso inadecuado de artículos, y en general, fallas en la estructura interna de los elementos léxicos.
3. Errores sintácticos: engloban las construcciones que se desvían de las reglas sintácticas del inglés. Esto incluye problemas de concordancia sujeto-verbo, uso indebido de preposiciones, uso incorrecto de los pronombres relativos, entre otros.

En lo referente a los errores textuales, estos comprenden aquellos que afectan la organización global del texto, su adecuación al contexto comunicativo y la eficacia en la transmisión del sentido, y fueron clasificados en tres subcategorías en función de las características observadas en el corpus:

1. Errores de coherencia y cohesión (CC): incluyen el uso incorrecto de conectores, artículos, preposiciones y referencias pronominales, así como repeticiones innecesarias. También contempla la conexión ilógica entre ideas, aunque este último tipo de error no se registró en los textos analizados.
2. Errores de informatividad, intencionalidad e intertextualidad (III): abarcan estructuras o expresiones que resultan en la omisión o la sobrecarga de información relevante, la falta de claridad en el contenido a pesar de que la estructura gramatical sea correcta, alteraciones del sentido general de una oración traducida en relación al texto fuente y selecciones léxicas inadecuadas que afectan el contenido o matices de las ideas expresadas.

3. Errores de situacionalidad y aceptabilidad (SA): refieren a la selección de expresiones léxicas que, si bien no modifican el contenido, resultan inapropiadas para el registro requerido, y al uso de contracciones, que no son adecuadas en el discurso académico.

Intervenciones

Se consideran las intervenciones realizadas en las distintas versiones V con el objetivo de observar con mayor precisión el modo de apropiación de las herramientas por parte de los estudiantes. Estas se clasifican en distintas categorías, lo que permite capturar la dinámica del proceso de reescritura y comprender qué operaciones son más frecuentes en este contexto. A continuación, se describen y ejemplifican las transformaciones consideradas:

- Inserciones: incorporación de unidades léxicas sin alterar la sintaxis que rodea al grupo insertado.
  - Versión T (E10T1[14]): *that is not within the calibration curve*
  - → Versión intervenida (E10V2): *that is not within the limits of the calibration curve*
- Eliminaciones: supresión de unidades léxicas, sin alterar la sintaxis que rodea al grupo eliminado.
  - Versión T (E09T5): *This could be due to the following reasons:*
  - →Versión intervenida (E09V6): *This could be due to ~~the following reasons~~:*
- Reordenamientos: cambios en el orden de las unidades léxicas dentro de una oración, sin afectar de forma sustancial la sintaxis.
  - Versión T (E11T4): *an Abs was obtained that is not within the calibration curve*
  - →Versión intervenida (E11V5): *an Abs that is not within the calibration curve was obtained*
- Reemplazos: sustitución de elementos léxicos por otros.
  - Versión T (E02T5): *This study aimed to purify*
  - →Versión intervenida (E02V6): *This work aimed to purify*
- Reformulaciones: modificaciones a nivel de frase que implican una reestructuración sintáctica significativa, combinando reemplazos, reordenamientos e intervenciones gramaticales. Se clasifican como reformulaciones cuando la frase resultante es completamente diferente en estructura a la original.
  - Versión T (E05T4): *Although some altercations and results had not been expected*
  - →Versión intervenida (E05V5): *Although some altercations and unexpected results*

[14] Las notaciones EXXVX y EXXTX se utilizan para identificar cada experimento individual junto con la versión de la que se está hablando, donde EXX indica el identificador del experimento y V/T señala la versión específica del texto asociada. Por ejemplo, E10T1 refiere al experimento E10, versión T1.

- Alteraciones gráfico-morfológicas: modificaciones en la forma de una unidad léxica sin que se produzca un reemplazo completo, como cambios ortográficos o morfológicos.

  ○ Versión T (E00T5): *To determine the concentrations*

  →Versión intervenida (E00V6): *To determined the concentrations*

- Alteraciones de puntuación: modificaciones que afectan únicamente a los signos de puntuación. Estas alteraciones pueden modificar la relación entre cláusulas.

  ○ Versión T (E01T5): *partially achieved. Despite potential errors*

  →Versión intervenida (E01V6): *partially achieved and despite potential errors*

- Reformulaciones/reemplazos + reordenamiento: intervenciones complejas en las que se combinan reformulaciones o reemplazos léxicos con reordenamientos sintácticos. A diferencia de las reformulaciones o reemplazos simples, en este tipo de intervención el elemento modificado no permanece en su posición original, sino que es desplazado dentro del enunciado, y altera la organización del contenido.

  ○Versión T (E19T3): *Additionally, we achieved the specific objectives, at least partially*

  →Versión intervenida (E19V4): *Additionally, we partially achieved the specific objectives*

- Recuperaciones: intervenciones en las que el participante vuelve a introducir un segmento del texto previo a la corrección de la herramienta que había sido eliminado o modificado por el sistema. Esta categoría general incluye casos en los que la recuperación se realiza mediante diferentes operaciones (inserciones, reordenamientos, reformulaciones, etc.), siempre que el objetivo sea restituir una parte del enunciado original[15].

  ○ Versión original (E09V5): *With the completion of the work in the laboratory, it can be concluded that the main objectives were achieved*

  →Versión T (E09T5): *The laboratory work successfully achieved the main objectives*

  ⇒Versión intervenida (E09V6): *It can be concluded*[16] *that the laboratory work successfully achieved the main objectives*

- Recuperaciones con reemplazo: intervenciones en las que se restituye un elemento del texto previo a la corrección realizada por la herramienta mediante un reemplazo léxico. En el caso del

[15] Esta categoría agrupa todos los casos en que se recupera un elemento del texto previo a la corrección de la herramienta, excepto cuando la recuperación implica un reemplazo léxico. En esos casos, se clasifica de manera separada como recuperación con reemplazo. Este tipo de recuperación fue diferenciada de las recuperaciones generales con el objetivo de observar con más detalle el comportamiento de los participantes en relación al léxico, ya que se detectaron en las encuestas de percepción varios casos en los que los participantes manifestaron opiniones en relación a los cambios léxicos y terminológicos realizados por las tecnologías.

[16] Recupera *It can be concluded.*

ejemplo, se conserva la estructura sintáctica sugerida por el sistema, pero el participante decide reincorporar una elección léxica propia de la versión anterior.

○Versión original (E12V1): *although of posibles mistakes that it could be make unexpective results*

→Versión T (E12T3): *Despite possible errors that may have led to unexpected results*

⇒Versión intervenida (E12V4): *Despite possible mistakes that may have led to unexpected results*

4.1.4.2. Procesamiento del corpus y registro de los datos

Como fue señalado, el foco del análisis recayó sobre la detección manual de traducciones de nominalizaciones, traducciones de estructuras en voz pasiva e impersonales con "se", errores gramaticales y textuales, e intervenciones humanas entre versiones. Los primeros tres fenómenos fueron seleccionados por su frecuencia en el discurso académico y su función en la construcción de un registro impersonal. Aunque inicialmente se consideró el uso de diversas librerías de *Python* para identificar varios de estos fenómenos, estas herramientas resultaron poco eficaces, en parte debido a las características del corpus que, como se mencionó, comparte rasgos con los corpus de aprendientes, incluyendo la presencia frecuente de errores (Frapolli, en evaluación). Esta situación comprometía la precisión de los algoritmos automáticos, por lo que se optó por una revisión manual exhaustiva, con múltiples controles para asegurar la mayor fiabilidad posible en la codificación.

Las intervenciones de los participantes sobre las salidas generadas por las plataformas fueron cuantificadas con el fin último de evaluar el grado de confianza depositado en tecnologías. Este análisis permitió identificar las decisiones lingüísticas tomadas por los participantes y las estrategias desplegadas frente a las sugerencias automáticas.

Adicionalmente, si bien no estaba previsto en el proyecto de tesina, se calculó el *Lexical Frequency Profile* (*LFP* - Laufer y Nation, 1995) de las traducciones con el objetivo de analizar el uso de léxico académico y evaluar la incidencia del uso de recursos digitales sobre la complejidad del vocabulario. El *LFP* permite estimar la proporción de unidades léxicas empleadas en un texto según su frecuencia de aparición en listas de vocabulario definidas, y diferenciar así, por ejemplo, entre léxico cotidiano y académico. Así, se calculó esta métrica con las listas de vocabulario[17] *General Service List* (*GSL* - West, 1953) y *Academic Word List* (*AWL* - Coxhead, 2002). La primera reúne las palabras más frecuentes del

[17] Si bien los trabajos que dieron lugar a estas listas de vocabulario se encuentran correspondientemente referenciados en la sección de bibliografía, se indican aquí también dos hipervínculos que permiten acceder a ellas con mayor facilidad, que corresponden además a los recursos utilizados por la plataforma que usamos para calcular la métrica:
General Service List: https://www.eapfoundation.com/vocab/general/gsl/frequency/
Academic Word List: https://www.eapfoundation.com/vocab/academic/awllists/

inglés general, mientras que la segunda contiene términos comúnmente usados en textos académicos, pero que no figuran entre los más frecuentes del lenguaje cotidiano. Esta métrica ha demostrado ser útil para comparar textos de aprendientes con distintos niveles de competencia lingüística. Aunque se han señalado limitaciones (Smith, 2005) es considerada válida para ciertos casos. En particular, es útil cuando se analizan diferencias entre grupos, cuando se trabaja con muestras de baja competencia y la extensión de los textos es controlada, como es el caso de este estudio.

Además, el *LFP* permite medir la riqueza léxica en función del uso correcto de vocabulario (en términos de ortografía), lo que es útil al evaluar traducciones realizadas por aprendientes de inglés. Esto es posible porque esta métrica calcula la proporción de unidades léxicas de un texto que se corresponden con las listas seleccionadas (en este caso, de vocabulario cotidiano y académico), y clasifica el resto —es decir, aquellas que no coinciden con ninguna de las dos listas— en una tercera categoría denominada *Off-list*. En consecuencia, las unidades que presentan errores ortográficos no son reconocidas por las herramientas utilizadas para calcular esta métrica y, por lo tanto, son incluidas en la *Off-list*. Al excluir términos con errores ortográficos, esta métrica proporciona una representación más precisa de la competencia léxica de los participantes, ya que el dominio efectivo de un ítem léxico implica su correcta utilización en todas sus dimensiones. Esta característica es muy relevante en las traducciones iniciales (V1), donde es esperable encontrar fallas de tipo ortográfico debido al nivel de inglés intermedio de los participantes. Sin embargo, es importante señalar que no todos los ítems incluidos en la *Off-list* corresponden a errores: si bien en versiones como V1 estos pueden tener un peso significativo, la lista también agrupa vocabulario que no forma parte ni de la *GSL* ni de la *AWL*, como números, términos muy técnicos, nombres propios o vocabulario fuera de registro. Esto es particularmente relevante para el análisis de versiones que presenten bajos niveles de errores ortográficos (véase *infra* por ejemplo versiones T1 o T5).

Por último, esta herramienta facilita también la comparación entre versiones al discriminar entre diferentes niveles de vocabulario, lo que permite observar posibles mejoras en la precisión léxica vinculadas al uso de tecnologías como *Google Translate*, *Grammarly* o *ChatGPT*.

El análisis de estos fenómenos se realizó a partir del seguimiento de las rutas definidas en el diseño experimental. Para cada participante se generaron once versiones: una versión inicial autónoma (V1), cinco producciones intermedias generadas o corregidas por recursos digitales (T1 a T5), y cinco versiones resultantes de la intervención humana sobre las salidas de dichos sistemas (V2 a V6). A partir de esta estructura, se identificaron cuatro secuencias principales de análisis, señaladas en la figura 1 con diferentes colores para facilitar su diferenciación. Tres de ellas corresponden al uso individual de cada herramienta: la ruta *Grammarly* (V1 → T2 → V3), la ruta *ChatGPT* (V1 → T3 → V4), y la ruta *Google*

Translate (T1 → V2). El cuarto recorrido combina todos los sistemas de forma secuencial (T1 → V2 → T4 → V5 → T5 → V6)[18], y se denomina ruta *completa*. Para cada uno de estos caminos, se examinó la dinámica de los fenómenos seleccionados con el objetivo de observar incrementos o reducciones en los distintos pasos. Asimismo, se realizaron comparaciones entre estos trayectos para identificar patrones recurrentes y diferencias en los efectos de cada herramienta o combinación de ellas. En el caso de las intervenciones humanas, dado que el foco estuvo puesto en los cambios introducidos por los participantes, el análisis se centró en cada versión V modificada, con el fin de clasificar los tipos de intervenciones más frecuentes. Además, se realizó una comparación global entre versiones para identificar cuáles concentran, en promedio, mayor número de intervenciones.

La identificación de los fenómenos estudiados se realizó a través de un procedimiento de marcado manual en documentos *.docx* correspondientes a cada texto del corpus. Para ello, se utilizó la opción de resaltado de color de *Google Docs*, y se asignó un color específico a cada tipo de fenómeno. Para el cálculo del *LFP* se utilizó una herramienta en línea desarrollada por la *EAP Foundation* (2024). Además, se empleó la plataforma *diff_match_patch* (Google, 2006), para realizar una revisión preliminar automática de las intervenciones de los participantes. Este recurso permitió descartar aquellos pares de textos en los que no había cambios, y facilitó la focalización del análisis en los casos con modificaciones.

Para el registro y el procesamiento cuantitativo de los datos, se emplearon hojas de cálculo de *Google Sheets*. Cada uno de los fenómenos fue registrado en planillas específicas, en las que se codificaron, para cada experimento y versión, categorías como tipo de construcción o naturaleza del error (cuando correspondía). Además, se utilizó el lenguaje de programación *Python* para la elaboración de un *script* (Frapolli, 2025) para el registro de intervenciones en las hojas de cálculo. Este procedimiento permitió registrar de forma ágil las ediciones de los participantes.[19]

[18] Nótese que la ruta *completa* incluye como etapa inicial a la *Google Translate*. Como resultado, muchas veces en la tesina estas dos secuencias se analizan en conjunto para evitar redundancias.

[19] Como parte de un proyecto personal orientado a la práctica de programación en *Python*, este *script* se desarrolló para automatizar parte del proceso de registro de las intervenciones. Este desarrollo fue posteriormente incorporado como herramienta metodológica en la tesina, con el objetivo de sistematizar la extracción de información desde documentos *.docx*. El *script* fue diseñado para identificar fragmentos de texto resaltados en los documentos y asociarlos al color correspondiente. Para ello, se utilizó la biblioteca *python-docx*, que permite manipular archivos de *Word* desde *Python*, y se accedió al *XML* subyacente del archivo para detectar los resaltados y obtener el código hexadecimal del color de fondo. Se estableció un diccionario que vinculaba cada color con un tipo específico de intervención (por ejemplo, "*f9cb9c*": "REEMPLAZO"). La automatización mejoró la eficiencia del registro, posibilitando la extracción rápida y precisa de grandes volúmenes de información. No obstante, debido a ciertas limitaciones técnicas inherentes al procesamiento automático —como la detección de fragmentos vacíos, la interpretación errónea de algunos colores o fallas en el formato original de los documentos—, se optó por realizar una revisión manual exhaustiva de los resultados generados. Esta revisión permitió corregir los errores detectados y asegurar que la información final volcada en la planilla fuese más fiable y representativa. Si bien el proceso no sustituyó el criterio humano en esta etapa del análisis, el *script* desarrollado resultó fundamental para agilizar la codificación y minimizar el margen de error en las tareas repetitivas.

El análisis tuvo un foco cuantitativo, complementado con estrategias cualitativas, y se basó en el cálculo de promedios por versión (y las rutas que estas componen): se sumaron los valores correspondientes a cada fenómeno en todas las producciones de una misma versión (por ejemplo, todas las V1) y se los dividió por la cantidad total de experimentos para obtener una medida de resumen que permitiera observar tendencias generales en la dinámica de los fenómenos en cada ruta y versión estudiada.

4.1.5. Los participantes

Se convocó a 20 participantes, estudiantes de carreras de Ciencias Naturales de la UNS, y se priorizaron aquellos pertenecientes a las áreas de Bioquímica, Biología o disciplinas afines. Si bien inicialmente se había previsto limitar la participación a quienes hubieran cursado Bioanalítica II, las dificultades para reunir una cantidad suficiente de voluntarios dispuestos a involucrarse en una actividad extensa obligaron a ampliar los criterios de inclusión a estudiantes de Ciencias Naturales en general.

4.1.5.1 Caracterización de la muestra recogida

En este apartado se presenta una descripción de la muestra, que permite establecer relaciones entre los perfiles y las elecciones realizadas durante la actividad.

En primer lugar, cabe aclarar que aunque el total de datos analizados corresponde a veinte participantes, la numeración llega hasta E21. Esto se debe a dos circunstancias: por un lado, los datos producidos por uno de los participantes originales fue descartado (E08), ya que no utilizó una de las herramientas conforme a las instrucciones del estudio. Por otro lado, el reemplazo previsto para ese dato (E20[20]) no completó la actividad, por lo que fue necesario convocar a una nueva participante, identificada como E21. Por ello, si bien hay veintidós identificadores, en efecto, solo veinte datos fueron válidos para el análisis.

En relación al género, predomina el femenino (n=17), y los varones representan una proporción menor (n=4). Se buscó, en una primera instancia, conformar una muestra equilibrada en cuanto a esta variable y las distintas carreras del área, pero las dificultades para reunir voluntarios dispuestos a participar hicieron que esta tarea se tornara imposible.

En cuanto a la variable edad[21], se consideraron tres rangos etarios. La franja más representada fue la de 25 a 29 años (n=11), seguida por la de 19 a 24 años (n=7) y por último la de 30 a 34 años (n=2).

[20] El reemplazo previsto corresponde a E20 porque la muestra original se numeraba desde E00 hasta E19, lo que representa veinte casos en total. Como el conteo comienza en cero (E00), E19 era el vigésimo participante.

[21] Las variables sociales de edad y género fueron relevadas con vistas a estudios futuros por dos motivos, dado que las hipótesis de trabajo de esta investigación se vinculan principalmente con otras variables socioacadémicas y técnicas, como el nivel de dominio de inglés, el área disciplinar y la familiarización previa con las tecnologías.

Con respecto al área disciplinar, la muestra se compone en su mayoría de estudiantes de Bioquímica (n=7) y de carreras de Ciencias Biológicas (n=6)[22], seguidos por los de carreras afines, esto es, Medicina (n=3), Farmacia (n=2), Enfermería (n=1) y Licenciatura en Química (n=1).

En lo referente al grado de avance en la carrera, se observa una concentración significativa de participantes en el tercer año (n=9), seguido por una distribución más dispersa entre los niveles superiores: cuarto (n=1), quinto (n=4) y sexto año (n=4). No se registraron participantes de primer año, y solo uno se encontraba cursando segundo (n=1). Además, participaron dos estudiantes de posgrado: uno en una etapa inicial del trayecto y otro en una etapa avanzada. Esta distribución indica que la muestra está compuesta en su mayor parte por estudiantes con una trayectoria universitaria ya consolidada.

Respecto al dominio del inglés, en la encuesta de perfil se brindó a los participantes una breve descripción de las competencias asociadas a cada nivel del *MCER*. En los casos en que contaban con certificados oficiales, se indicó a qué categoría correspondía el examen aprobado (por ejemplo, *First Certificate* corresponde a B2). Para quienes no disponían de certificación, se solicitó que realizaran una autovaloración en función de las descripciones ofrecidas. Si bien este tipo de medición no es precisa como un examen estandarizado, el objetivo no era establecer con exactitud el nivel de inglés de cada participante, sino obtener una referencia general sobre sus competencias y una valoración de las capacidades propias para interpretar mejor sus decisiones lingüísticas durante la actividad. En total, diez participantes afirmaron contar con algún tipo de certificación formal que acredita sus conocimientos del idioma, mientras que los once restantes realizaron una autovaloración. Esta distribución relativamente equilibrada permite tener en cuenta tanto trayectorias formales como aprendizajes no certificados dentro del análisis. En lo que respecta a los niveles declarados propiamente dichos, la mayoría de los participantes se ubicó en la franja intermedia del *MCER*[23]: ocho declararon tener un nivel B1 y nueve, un nivel B2. Solo un participante manifestó poseer un dominio avanzado (C1), mientras que tres se identificaron con competencias básicas (uno en A1 y dos en A2)[24].

[22] Incluye estudiantes del profesorado, la licenciatura y el doctorado en Ciencias Biológicas. Dado que pertenecen al mismo campo disciplinar —aunque con orientaciones distintas: la enseñanza en el profesorado y la investigación en la licenciatura y el doctorado—, se agruparon bajo una misma categoría.

[23] Cabe aclarar que si bien originalmente se había considerado solo la inclusión de estudiantes con un dominio intermedio de inglés, se ampliaron los criterios de selección a partir de la dificultad para reunir participantes dispuestos. Esta situación, sin embargo, permitió plantear posibles relaciones preliminares entre los respectivos niveles de inglés de los participantes y variables como la cantidad de intervenciones realizadas, aunque estas deben ser corroboradas en una investigación con una muestra más equilibrada

[24] El único participante que se identificó con un nivel C1 no presenta respaldo oficial. De los ocho que indicaron tener un nivel B1, cinco no contaban con certificado al momento de realizar la actividad y tres sí. En el caso de quienes señalaron un nivel B2 (nueve en total), seis poseen acreditación formal, mientras que tres no. Por último, en los niveles básicos, el participante que mencionó un nivel A1 no posee certificación, mientras que entre los dos que consignaron A2, uno tiene constancia oficial y el otro no.

Por último, se observan diferencias en cuanto a la familiaridad con los recursos digitales propuestos. Todos los participantes declararon conocer y haber utilizado *Google Translate*, lo que evidencia su amplio alcance y uso generalizado como recurso lingüístico. En contraste, sólo tres personas afirmaron haber usado *Grammarly* con anterioridad, mientras que cinco dijeron conocerla, pero no haberla utilizado, y trece indicaron no conocerla en absoluto. Esta herramienta, por tanto, se presenta como poco difundida en el grupo. Por otro lado, *ChatGPT* muestra una presencia más marcada: dieciséis participantes dijeron conocerlo y haberlo utilizado, tres lo conocían, aunque no lo habían usado, y solo dos afirmaron no conocerlo. Estos datos permiten suponer una mayor visibilidad y accesibilidad de esta plataforma en comparación con *Grammarly*, posiblemente debido a la difusión masiva de la que ha sido objeto desde que fue lanzado.

4.2. Evaluación de los participantes

El diseño de investigación previó, como instancia inmediatamente posterior, una evaluación de las percepciones y experiencias de los participantes durante la actividad experimental. Para ello, la recolección de datos incluyó una encuesta final administrada mediante *Google Forms* e incluída en el anexo 8.4. Este cuestionario se centró en sus percepciones respecto de la facilidad de uso, eficacia y fiabilidad de las plataformas utilizadas, así como en su nivel de satisfacción con los textos producidos y el grado de (in)seguridad lingüística experimentado. Se incluyeron preguntas con escala de *Likert* y opciones múltiples, así como otras abiertas que permitieran justificar las anteriores, con el fin de captar tanto tendencias generales como apreciaciones individuales.

En lo que refiere a las percepciones de los participantes, el análisis se organizó de acuerdo con la estructura de la encuesta final de percepción. Las preguntas con opciones cerradas —como las de escala de *Likert* y las de selección múltiple— permitieron relevar datos cuantificables, a partir de los que se buscó identificar patrones de uso, valoración y confianza respecto de las herramientas empleadas. Por su parte, las preguntas abiertas fueron analizadas mediante una lectura detallada y sistemática, que permitió relevar opiniones recurrentes y contrastarlas con las tendencias cuantitativas observadas.

4.3. Juicio de los expertos

La última instancia del diseño de esta investigación comprendió la contrastación del análisis cuanti-cualitativo con las evaluaciones de jueces expertos. Para seleccionar a estos jueces, se buscó contar con la participación de un especialista bilingüe en el área disciplinar en cuestión, y dos especialistas en idioma inglés. Sin embargo, diversas circunstancias externas limitaron la concreción de esta última condición. Entre ellas, pueden mencionarse los efectos derivados de la inundación ocurrida el 7 de marzo, que afectó el normal desarrollo de las actividades laborales y académicas de potenciales

colaboradores, y el reciente proceso de jubilación de varias docentes del área de inglés en la universidad, que redujo el número de contactos disponibles. Como resultado, fue posible contar únicamente con una especialista en lengua inglesa, además del experto en el área disciplinar. A pesar de esta limitación, las observaciones realizadas por los jueces participantes permitieron incorporar una mirada complementaria y enriquecer la interpretación de los hallazgos desde una perspectiva evaluativa que aporta criterios expertos sobre los usos lingüísticos, académicos y profesionales.

En este marco, a cada evaluador se le envió un subconjunto de las traducciones producidas por los participantes, junto con una breve consigna orientadora, y se les solicitó que emitieran juicios sobre aspectos vinculados al léxico, la sintaxis, la adecuación pragmática y la organización discursiva de los textos (v. anexo 8.7.). Fue parte del protocolo asegurar que los jueces no tuvieran conocimiento acerca de la ruta de traducción específica que dio origen a cada versión a evaluar, para evitar posibles sesgos en los juicios emitidos sobre ellas.

5. Análisis de los datos

Este capítulo presenta el análisis de los datos obtenidos en el experimento, seguido de la evaluación de las percepciones de los participantes y de los juicios emitidos por los jueces expertos.

5.1. Tratamiento de los rasgos típicos del género informe

En este apartado se presenta el análisis de los rasgos del *género* informe considerados para esta tesina: las nominalizaciones, las construcciones pasivas y las impersonales con "se". En el texto fuente en español se identificaron 42 casos de nominalización. Si bien algunas corresponden a la repetición de un mismo sustantivo en diferentes contextos (por ejemplo, dos instancias de *uso*, *absorción*, entre otras), se contabilizaron todas las ocurrencias. Asimismo, se registraron 24 construcciones que incluyen pasivas reflejas con "se", impersonales con "se" y pasivas perifrásticas. De ese total, 4 fueron clasificadas como impersonales con "se", pero se incorporaron al análisis general de la voz pasiva debido a que, en inglés académico, tienden a traducirse mediante construcciones pasivas formales, cumpliendo así una función equivalente en términos de desagentivación del discurso.

Tanto para la primera parte del análisis, centrada en las nominalizaciones, como para la correspondiente a las construcciones en voz pasiva e impersonales con "se", se buscó determinar cuántas de estas expresiones habían sido traducidas mediante estructuras análogas en inglés. Con este fin, se definieron traducciones esperadas (v. anexos 8.5 y 8.6) para cada una de las ocurrencias identificadas, entendidas no como únicas posibles, sino como opciones verosímiles con base en criterios lingüísticos y disciplinarios. En el caso de términos especializados del área de las Ciencias Naturales, los términos esperados fueron definidos en consulta con un juez experto bilingüe en la disciplina. Para el resto de las expresiones, se recurrió a diccionarios, a la búsqueda de traducciones frecuentes, y a la asesoría de una profesora y traductora de inglés, quien participa como co-directora de este trabajo.

5.1.1. Tratamiento de las nominalizaciones

En el análisis de las nominalizaciones, se estudia la cantidad de expresiones traducidas al inglés mediante formas análogas, la corrección y adecuación de dichas traducciones en términos de errores gramaticales y textuales, y los casos de traducciones que presentaron mayor cantidad de fallos.

<u>Nominalizaciones traducidas como tales</u>

El primer interrogante que orienta este análisis fue cuántas nominalizaciones en el texto en español fueron traducidas al inglés también mediante estructuras nominales. Esta inquietud surgió a partir de una observación inicial de los datos, al notar durante el registro de esta clase de sustantivos que, en numerosos casos, los participantes optaron por formas verbales en inglés en lugar de mantener la estructura nominalizada original. Por ejemplo, expresiones como "para la obtención de" fueron

traducidas con frecuencia como "to obtain" (E06V1), en lugar de una versión nominal como "for the obtainment". Esta elección da cuenta de procesos de reformulación que exceden la simple traducción literal y que revelan tanto la competencia lingüística (en términos del *MCER*) del hablante como las tendencias discursivas dominantes en cada lengua. Por lo tanto, explorar qué porcentaje de las nominalizaciones se mantuvo como tal ofrece indicaciones sobre las estrategias que desplegaron los participantes —y/o los sistemas utilizados— al traducir estos textos. Para calcular esto, se contabilizaron de forma manual todas las nominalizaciones identificadas en las traducciones de los participantes, y se calcularon los promedios para cada versión.

Al observar la cantidad de nominalizaciones que fueron traducidas al inglés mediante estructuras nominales, se advierte un comportamiento diferencial según la ruta (v. fig. 2). En la de *Grammarly*, los promedios se mantienen estables (V1: 36,85; T2: 36,70; V3: 36,90). Esto sugiere que *Grammarly* no introduce modificaciones significativas en lo que respecta al uso de estructuras nominales. La relativa estabilidad puede deberse a que las sugerencias de la versión gratuita de esta herramienta tienden a centrarse en aspectos gramaticales o de estilo más que en reformulaciones estructurales profundas. Asimismo, la intervención posterior de los participantes no altera de forma sustancial la proporción de nominalizaciones traducidas como tales.

Por el contrario, en la ruta *ChatGPT* se observa una ligera disminución en la cantidad de nominalizaciones traducidas como tales entre la versión inicial (V1: 36,85) y la corrección realizada por el sistema (T3: 33,65). Aunque la diferencia no es pronunciada, sí puede sugerir que *ChatGPT* tiende a reemplazar algunas estructuras nominales por otras construcciones. Por ejemplo, en el experimento E00, la versión V1 incluye la formulación "we reached the main objectives the ones that consisted of the introduction of the handling of the techniques in a purification protocol", mientras que la versión corregida por tecnología en T3 es "we achieved the main objectives of introducing the techniques used in a purification protocol". En este caso, la nominalización "the introduction of the handling" es reemplazada por el gerundio "introducing", Algo similar ocurre en el experimento E06, cuya versión original en V1 presenta la secuencia "In context to intruduce [us][25] in the handling of", que es modificada en T3 por *ChatGPT* como "this work was conducted to familiarize with". Allí, la estructura nominal "the handling of" es sustituida por la construcción verbal "to familiarize with". La intervención posterior de los estudiantes (V4: 34,00) apenas revierte esta tendencia, lo que indica cierta aceptación de las reformulaciones ofrecidas por el sistema, en lo referente a este rasgo del *género* informe.

[25] En el experimento E06, el pronombre "*us"* aparece entre corchetes porque no fue incluido por el participante en la versión original. Se añadió aquí únicamente con fines expositivos, para facilitar la comprensión de la oración del ejemplo.

En el camino de *Google Translate*, se registran los valores iniciales más altos en cuanto al mantenimiento de nominalizaciones: T1 alcanza un promedio de 39,40, y V2 apenas desciende a 39,20. Esto parece sugerir que la herramienta tiende a realizar traducciones más literales o a mantener una fuerte correspondencia estructural entre lenguas, lo que resulta en una alta tasa de conservación de estructuras nominales del texto en español original, que los participantes mantienen al intervenir esta versión.

Por último, al analizar la ruta *completa*, que incluye sucesivas intervenciones humanas y automáticas, el promedio de estos valores se mantiene estable hasta V5 (39,15), y luego se observa una disminución en la cantidad de nominalizaciones traducidas como tales en T5 (35,35) y V6 (35,95). Esta caída refuerza lo antedicho sobre *ChatGPT*: si bien el punto de partida (T1) presenta una alta similitud estructural, este recurso tiende a introducir modificaciones que reemplazan las nominalizaciones originales. Por ejemplo, en E10V5, el participante traduce "Con la realización del trabajo en el laboratorio" como "With the realization of the work in the laboratory"[26], y la herramienta reformula esta traducción en "From the work carried out in the laboratory" (T5).

Corrección y adecuación en la traducción de nominalizaciones

Una vez determinado el porcentaje de nominalizaciones que fueron traducidas como estructuras análogas en inglés, se buscó averiguar cuántas de esas nominalizaciones fueron traducidas sin errores (v. 4.1.4.1), para evaluar no solo la frecuencia del recurso, sino también la calidad de su traducción. Este enfoque se vincula con los objetivos específicos de este trabajo, en tanto permite indagar si el uso individual o combinado de las plataformas se traduce en una mejora cualitativa de las producciones y si alguna de ellas ofrece ventajas particulares frente a las demás. La figura 3 ilustra esta dinámica, al mostrar el promedio de nominalizaciones traducidas correctamente en cada etapa del proceso, según los diferentes recorridos. Se encontró que, en aquellas que inician con una versión autónoma de los estudiantes (V1), el porcentaje de nominalizaciones correctas parte de un valor bajo (72,86%) en comparación con el valor inicial de la ruta *completa* (T1: 97,46%). En la de *Grammarly*, se observa una mejora progresiva en las versiones, y se alcanza un 81,98% en V3. En la vía *ChatGPT*, el incremento es aún más notorio: la corrección automática (T3) alcanza el 96,14%, y aunque se registre una leve caída, los estudiantes logran mantener ese nivel en su intervención posterior (94,56% en V4).

Por otro lado, la ruta *completa* (que incluye la *Google Translate*) presenta desde el inicio porcentajes elevados de corrección (97,46% en T1), que se mantienen estables y tienden a incrementarse de forma ligera hasta alcanzar un 99,03% en la versión final (V6). Sin embargo, también en esta secuencia se advierte un fenómeno recurrente: la intervención de los participantes no siempre

[26] Traducción esperada en este caso: "Upon the completion of the work in the laboratory"

mejora los resultados provistos por la tecnología, y en algunos casos produce una ligera disminución en la proporción de nominalizaciones correctas. Este comportamiento sugiere que, si bien tanto las herramientas como los estudiantes estabilizan estos valores, la intervención humana conlleva cierto margen de error.

Tipos de errores en nominalizaciones por ruta

Una vez identificado el porcentaje de nominalizaciones traducidas sin errores en general, el análisis se orientó a desglosar aquellos presentes en las traducciones que no alcanzaron una formulación adecuada. Para ello, se distinguieron dos grandes grupos de acuerdo a la taxonomía elegida: errores gramaticales (v. fig. 4), que incluyen problemas sintácticos, fonemáticos o morfológicos, y textuales (v. fig. 5), vinculados a dimensiones tales como la situacionalidad, la aceptabilidad, la informatividad, la intencionalidad y la intertextualidad. No fueron identificados problemas de coherencia y cohesión en las traducciones de las nominalizaciones. En conjunto, las fallas gramaticales tienden a disminuir de manera más marcada que las textuales en todos los trayectos.

De acuerdo al análisis cuantitativo, puede observarse que en la vía *Grammarly* los errores gramaticales descienden de 9,25 en la versión inicial (V1) a 4,35 en la final (V3), mientras que los textuales se mantienen más estables, con una leve suba de 3,60 a 4,00. Este patrón sugiere que la plataforma es efectiva en la corrección de estructuras, pero tiene un impacto limitado sobre aspectos más complejos de la textualidad, como la adecuación textual y su posible relación con intenciones comunicativas.

La secuencia *ChatGPT* muestra una tendencia similar, aunque más pronunciada: los fallos gramaticales bajan de forma sustancial de 9,25 en V1 a solo 0,50 en T3, con un pequeño incremento posterior en V4 (1,15), tras la intervención de los estudiantes. Los textuales, en cambio, bajan de 3,60 a 2,05 con este sistema, pero vuelven a aumentar a 2,45 en la versión intervenida por los estudiantes (V4). Este comportamiento refuerza lo observado con anterioridad: los participantes no siempre mejoran los resultados de la herramienta, y en algunos casos reintroducen errores.

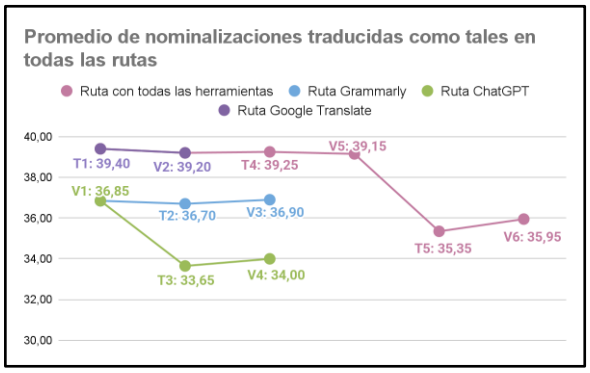

Figura 2

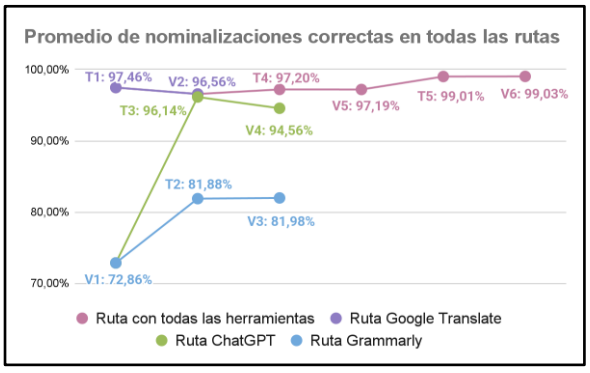

Figura 3

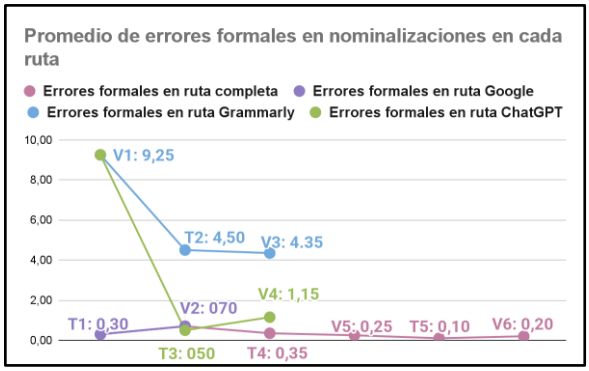

Figura 4

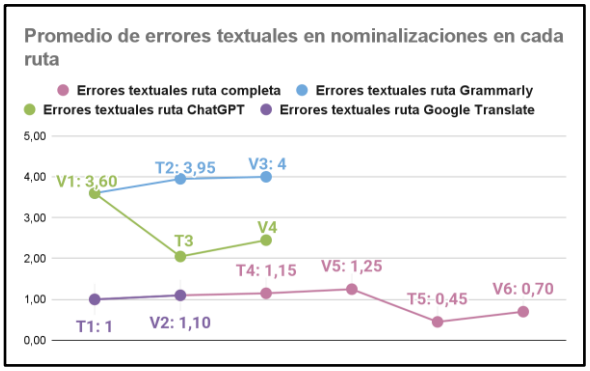

Figura 5

En cuanto a la ruta *completa*, que incorpora la de *Google Translate* en sus dos instancias iniciales, los errores gramaticales son muy bajos (0,30 en T1) y tienden a desaparecer de forma progresiva (0,10 en T5). Los textuales también son escasos, pero se mantienen más constantes en las primeras versiones, con una leve disminución hacia el final (de 1 en T1 a 0,45 en T5), tras la intervención de *ChatGPT*. Lo interesante aquí es que las plataformas actuaron con eficacia, aun con un margen de mejora limitado, y los participantes no generaron retrocesos significativos. Es posible que esto se deba a que partían de una base con pocos problemas que no requería grandes reformulaciones.

Hasta aquí, los datos permiten observar que el uso de recursos digitales tiende a disminuir el número de errores lingüísticos en las traducciones de nominalizaciones. Se puede sugerir que estos recursos son más efectivos para resolver dificultades gramaticales, ya que las textuales parecen presentar cierta resistencia a la corrección automática, al depender en mayor medida de factores como el contexto y la intención comunicativa. Por otro lado, en lo que refiere a la dificultad de corrección humana de este tipo de fallos, esta se vincula principalmente con las competencias individuales de quien realiza la revisión, así como con la interpretación de esos mismos factores extralingüísticos. Asimismo, se observa que las distintas etapas de la ruta *completa* ofrecen resultados con menor promedio de errores, lo que sugiere una ventaja respecto de los recorridos que emplean solo una herramienta y comienzan con la versión autónoma del participante. Además, queda de manifiesto que la intervención humana en las versiones T no garantiza una mejora, y en muchos casos los participantes introducen cambios que impactan de forma negativa en la cantidad de fallos en el texto.

Casos de nominalizaciones con más errores

Tras examinar esta dinámica general, el análisis se centró en detectar cuáles fueron los casos específicos de nominalizaciones que presentaron mayores dificultades en su traducción. Para ello, se identificaron aquellos casos que concentraron la mayor cantidad de fallas gramaticales y textuales. Este recorte permitió observar con mayor detalle qué aspectos podrían haber influido en las fallas de traducción, y así aportar elementos para una interpretación más precisa del desempeño tanto de las herramientas como de los participantes.

El caso más frecuente de problemas gramaticales en nominalizaciones fue *calibration*, que los participantes utilizaron habitualmente para traducir el participio *calibrado* en la expresión "curvas de calibrado" del texto fuente. En la mayoría de las instancias erróneas (99 en todo el corpus), el término fue traducido como *calibrate* (por ejemplo, en E02T2: "the use of calibrate curves") o calibrated (por ejemplo, en E04V1: "was used calibrated curves"). Esta tendencia podría deberse a que *calibrado* funciona en español como un sustantivo derivado (equivalente a "calibración"), mientras que en inglés el término esperado es *calibration* como modificador del núcleo nominal. La similitud entre *calibrado* y *calibrated* provoca que los participantes no identifiquen correctamente la función gramatical y, por lo tanto, se generen errores morfológicos y sintácticos derivados de la transferencia inadecuada entre categorías gramaticales en ambos idiomas. Este tipo de relaciones entre las lenguas resulta una fuente común de errores en el corpus.

El segundo caso más común fue *reagent*, traducción esperada de *reactivo*, en la expresión "reactivo de Bradford", en el material en español. La gran mayoría de las veces en la que se traduce erróneamente (65 instancias), *reagent* es confundido con *reactive* (por ejemplo, en E07V1: "with NaOh or Bradford Reactive"). Es posible que esto se deba a su similitud fonológica con *reactivo*, en español. Sin embargo, *reactive* es un adjetivo en inglés, y los sustantivos correspondientes son *reagent* y *reactant*. En el contexto específico del texto, *reagent* es el término más adecuado, ya que remite a una sustancia utilizada para provocar una reacción, mientras que *reactant* suele referirse a los compuestos que reaccionan entre sí. Este patrón de error parece estar influido por la cercanía morfológica entre los términos en ambas lenguas, lo que favorece elecciones léxicas que no siempre se corresponden con la unidad léxica esperada.

Por último, el tercer error formal más frecuente se da en la nominalización *absorption*, correspondiente a *absorción* en "espectrofotometría de absorción molecular" en el texto fuente. La mayoría de las veces en las que este término es traducido con un error (37 instancias), *absorption* se traduce como *absortion*, sin la *p* (como por ejemplo en E06V1: "the molecular absortion spectrometry"). Este error ortográfico parece originarse en la estructura gráfica del español, ya que *absorción* no lleva *p*, lo que conduce a los estudiantes a adaptar la forma en inglés según patrones más familiares.

Por otro lado, entre los fallos textuales más comunes en traducción de nominalizaciones, el caso más frecuente fue el de *fractionation* (101 ocasiones), expresión con la que se traduce *fraccionamiento* en "Fraccionamiento de la actividad enzimática". En la mayoría de estas instancias, se seleccionó un término alternativo que no reflejaba con precisión el proceso descrito en el texto fuente, como *breaking up* o *division* (por ejemplo, en E17V3: "The division to the enzymatic activity with sulfate ammonium"). Aunque estas opciones pueden transmitir una idea general de lo que se desea comunicar,

carecen de la precisión técnica necesaria, lo que en un contexto científico puede comprometer la claridad del mensaje y la correcta interpretación de los resultados.

El segundo caso más común fue el de *altercations*, que se tradujo de manera errónea en 91 ocasiones. Como se anticipó, este error fue conservado de forma intencional en el texto fuente en español, donde se utilizó *altercados* en lugar de voces más apropiadas como *problemas* o *complicaciones* (v. 4.1.1.). En la mayoría de las instancias con error, el problema persistió, en especial en las versiones T1 (*Google Translate*), V2, T4 y V5 (versiones corregidas por *Grammarly* y los participantes). Sin embargo, se observó una disminución significativa en las traducciones que incluyeron la corrección de *ChatGPT* (T5, T3, V6, V4), lo que sugiere que esta herramienta tuvo un impacto positivo en la mejora de la precisión léxico-semántica.

Por último, el tercer error textual más común se presentó en la nominalización *absorbance* (o su abreviatura *abs*), que fue traducida de forma incorrecta en 41 ocasiones. En la mayoría de estos casos, *abs/absorbance* se confundió con la nominalización *absorption* (por ejemplo, en E13V5: "an absorption that is not within the calibration curve was obtained"). Esta forma es incorrecta en este contexto, ya que *absorbance* hace referencia a la medida cuantificable de la cantidad de absorción en experimentos científicos, mientras que *absorption* se refiere al proceso físico de absorber, que no implica necesariamente una medición directa. La precisión en el uso de estas dos unidades léxicas es relevante para evitar interpretaciones incorrectas de los procesos o resultados experimentales, por lo que se lo considera un error de selección léxica.

5.1.2. Tratamiento de la voz pasiva y las impersonales con "se"

Para el análisis de estas estructuras en este corpus, se siguió el mismo enfoque que en el caso de las nominalizaciones. Esto es, se identificó la cantidad de estructuras pasivas e impersonales con "se" traducidas con construcciones pasivas en las distintas versiones, la corrección y adecuación de estas traducciones en términos de fallos, y los casos de voz pasiva que presentaron más errores.

Construcciones pasivas e impersonales traducidas como pasivas

En esta etapa del análisis, se procuró determinar en qué medida las estructuras en voz pasiva (e impersonales con "se") del texto fuente en español fueron traducidas al inglés manteniendo la pasividad o, por el contrario, reformuladas mediante otras estructuras. Un ejemplo de este fenómeno es el uso de construcciones activas como "we used calibration curves" [E00V1] en lugar de "calibration curves were used" (forma esperada). También se identificaron reformulaciones mediante frases preposicionales que incorporan sustantivos abstractos, como "in the development of this paper" [E04V1], en lugar de la construcción pasiva esperada "this work was carried out". Asimismo, se observaron cláusulas de participio presente (present participle clauses), como "indicating a lack of linearity" [E00T5] en lugar

de “linearity cannot be ensured” (forma esperada). Respecto de este fenómeno, se advierte nuevamente un comportamiento diferenciado según la ruta y los sistemas utilizados (v. fig. 6).

En el camino *Grammarly*, los valores promedio se mantienen en torno a 12 casos de voz pasiva traducidas como tales en las tres versiones sucesivas que componen la secuencia (V1: 11,85; T2: 11,35; V3: 11,8). Este comportamiento sugiere, como en lo relativo a nominalizaciones, que esta plataforma no introduce grandes modificaciones a nivel estructural. La voz pasiva se encuentra poco representada desde el inicio (V1: 11,85 en contraste con los 24 casos en el texto fuente en español) en este recorrido, y la intervención automática no parece influir de forma significativa en este aspecto, y los cambios realizados por los participantes en V3 tampoco alteran esta tendencia.

En el trayecto de *ChatGPT*, se observa una leve disminución de la cantidad de pasivas traducidas como tales (V1: 11,85; T3: 10,15; V4: 10,2), un patrón similar al detectado en el caso de las nominalizaciones. Si bien las diferencias no son marcadas, se perfila cierta tendencia: *ChatGPT* tiende a reformular algunas construcciones pasivas en formas activas u otras alternativas, y los participantes tienden a conservar esas reformulaciones en sus versiones finales.

En la vía *Google Translate*, los promedios resultan más altos (T1: 21; V2: 19,3), lo que indica una mayor tendencia a conservar las estructuras pasivas del texto fuente en español. Esta regularidad puede atribuirse a una traducción más literal o alineada de forma estructural con el texto fuente por parte del traductor automático, lo que también se había observado en el análisis de las nominalizaciones. La leve caída entre T1 y V2 no representa un cambio notable, y sugiere que la intervención humana posterior tampoco introduce variaciones significativas en este aspecto.

Por último, en la ruta *completa*, se evidencia una disminución progresiva en la cantidad de estructuras pasivas mantenidas, especialmente hacia las últimas versiones. El punto de partida, correspondiente al trayecto de *Google Translate*, refleja una alta tasa de conservación. Sin embargo, los valores descienden levemente en la versión corregida por *Grammarly* y la revisión de los participantes (T4: 19; V5: 19), y la caída más significativa se registra tras la intervención de *ChatGPT* (T5: 12,05), con una recuperación mínima en la versión final (V6: 13). Esta tendencia refuerza lo observado en la vía específica de *ChatGPT* que tiende a reformular estructuras pasivas con otras construcciones.

Corrección y adecuación en la traducción de las construcciones pasivas

Una vez establecida la frecuencia de traducción de las estructuras pasivas, se analizó cuántas de esas traducciones se realizaron sin errores (v. 4.1.4.1). Esta pregunta permite desplazar el enfoque desde la mera identificación del recurso hacia una evaluación cualitativa de su realización.

A diferencia del caso de las nominalizaciones, donde se analizaron tanto problemas gramaticales como textuales, el estudio de la voz pasiva se enfocó en los aspectos gramaticales vinculados a su

construcción sintáctica. Esta decisión responde a una diferencia clave en la naturaleza de las unidades analizadas: mientras que las nominalizaciones seleccionadas son unidades monoléxicas (es decir, sustantivos derivados), la voz pasiva es una construcción que se manifiesta a nivel de cláusula. Esta distinción tiene implicaciones metodológicas importantes. En el caso de las nominalizaciones, fue posible identificar, dentro de una misma unidad, fallos de distinto tipo: ortográficos, morfológicos, léxicos o de registro, ya que todos ellos afectan de forma directa a la unidad léxica portadora de sentido. En cambio, en las construcciones pasivas, estos tipos de errores (como los ortográficos en "the porcentage of purification was determined" [E00V1] o "the activity of this enzime was determined" [E01V1]) no pueden considerarse inherentes a la cláusula pasiva en sí, aun cuando afecten a alguno de sus componentes (por ejemplo, al sujeto o al complemento). Así, si bien son registrados y analizados en otras secciones del análisis, los fallos ortográficos como *porcentage* o *enzime* (en lugar de *percentage* o *enzyme*) no pueden atribuirse a la construcción pasiva como tal. En otras palabras, son problemas que ocurren en alguno de los componentes de la voz pasiva, pero no la *constituyen*. Para garantizar la coherencia del análisis y evitar una atribución arbitraria de errores a una construcción que los contiene pero no los genera, entonces, se decidió restringir el análisis de la voz pasiva a los fallos que afectan exclusivamente su formación sintáctica —esto es, a aquellos que alteran su estructura según las reglas del sistema lingüístico—, y considerar el resto de los que afectan alguno de sus componentes dentro de las categorías generales correspondientes (fonemáticos, morfológicos, de selección léxica, de registro, etc.).

La figura 7 presenta los porcentajes de traducciones correctas de estructuras pasivas en las distintas secuencias. Así, en el caso de la ruta *Grammarly*, que comienza con la traducción autónoma de los participantes (V1), los valores iniciales de pasivas traducidas sin errores son relativamente bajos (60,34%). Sin embargo, tras la intervención de la tecnología (T2), se observa una mejora significativa que alcanza un 81,94% de corrección, y que llega a un 83,05% en la versión intervenida por los estudiantes (V3). Esta progresión indica que *Grammarly* no solo mantiene las estructuras pasivas, sino que también es una herramienta efectiva para ajustarlas hacia formas correctas en inglés. Además, se evidencia en este recorrido que los participantes logran sostener y consolidar esa mejora a partir de las correcciones del sistema y en combinación con sus conocimientos del idioma.

En la vía *ChatGPT*, la tendencia es aún más notoria. Se parte del mismo valor inicial que en la ruta alternativa (60,34% en V1), pero la corrección en T3 alcanza el 100%, lo que implica que todas las pasivas conservadas en esta etapa fueron reformuladas sin errores. Si bien en la versión revisada por los participantes (V4) se registra una leve caída (98,04%), el desempeño se mantiene alto. Cabe señalar que, si bien se ha observado que *ChatGPT* tiende a reformular las pasivas con otras construcciones, cuando opta por conservarlas, su capacidad de corrección supera a la de *Grammarly* (T2: 81,94% en

contraste con T3: 100%). Por otro lado, en ambos caminos se advierte que las modificaciones de los participantes provocan cambios mínimos en los niveles de corrección, lo que sugiere una tendencia a aceptar las sugerencias provistas por los recursos en lo que respecta a este tipo de estructuras.

En lo referente a la secuencia de *Google Translate*, tanto la traducción automática (T1) como la posterior revisión humana (V2) muestran porcentajes altos y estables de corrección (98,33% y 98,45%). Estos valores sugieren que el traductor automático, además de tener una tendencia a traducir de forma literal y conservar las estructuras lingüísticas del texto fuente, logra resultados con una alta tasa de corrección gramatical. Además, se encontró que los estudiantes no realizaron grandes cambios respecto al uso de la voz pasiva en esta etapa. Esta tendencia había sido también observada en el análisis de las nominalizaciones.

La ruta *completa* mantiene también niveles muy altos de corrección. Parte del trayecto de *Google Translate* y presenta oscilaciones leves (T4: 99,21% y V5: 98,42%) hasta alcanzar un 100% de corrección en las dos últimas versiones (T5 y V6). En este tramo, se destaca el impacto de *ChatGPT*, que eleva la corrección a los valores máximos, así como se había observado en la vía específica de esta tecnología. A diferencia de lo observado en las nominalizaciones, en este caso no se registra retroceso alguno en la última etapa, por lo que es posible asumir que los participantes no realizan cambios que generen muchos problemas en lo referente a las estructuras pasivas.

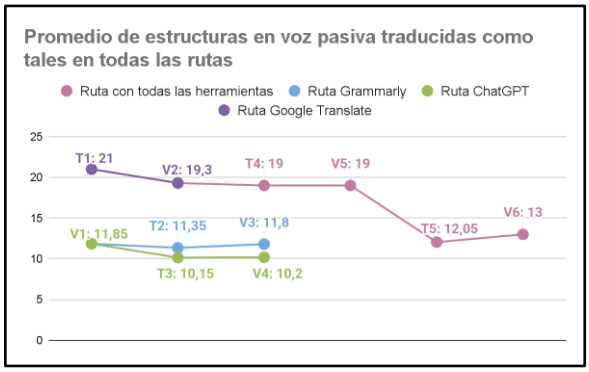

Figura 6

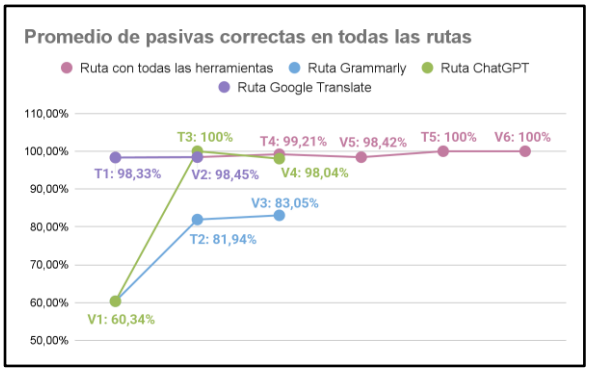

Figura 7

En conjunto, los datos hasta el momento indican que la tecnología —en particular *ChatGPT* y *Google Translate*— pueden favorecer a una traducción más precisa de estructuras pasivas. El desempeño de *ChatGPT* parece superar el de las otras herramientas en la traducción correcta de estas construcciones. No obstante, el hecho de que se hayan alcanzado niveles de corrección idénticos tanto en la ruta *ChatGPT* como en la *completa* indica que su eficacia no varía significativamente según se utilice sola o en combinación con otras tecnologías. En la mayoría de los casos, los participantes tienden a mantener los niveles de corrección alcanzados tras el uso de las herramientas, y solo se registra un leve aumento de errores en una de las rutas (a diferencia de lo que ocurría con las nominalizaciones). Esto sugiere una alta capacidad crítica de los estudiantes para evaluar las modificaciones realizadas en

estructuras pasivas. De ello puede inferirse que la capacidad de uso consciente de las herramientas no es uniforme, sino que posiblemente depende del grado de familiaridad de los usuarios con las estructuras lingüísticas implicadas. Es probable que esto se deba a que, normalmente, la voz pasiva constituye un contenido central en muchos programas de inglés como LE. En cambio, las nominalizaciones pueden estudiarse como parte del vocabulario de una disciplina o, incluso, de un género específico, en relación con los procesos de formación de palabras, no necesariamente como un contenido en sí mismo.

Casos de construcciones pasivas con más errores

Luego de examinar los porcentajes de traducciones correctas de estructuras pasivas en las distintas etapas, se procedió a identificar las construcciones pasivas que presentaron mayores dificultades, de las que se analizan las tres formas pasivas que concentraron la mayor cantidad de errores en el corpus.

En primer lugar, se encontró una dificultad recurrente en la traducción de "fue realizado el siguiente trabajo". Uno de los problemas estructurales más frecuentes en la traducción de esta frase al inglés es el orden incorrecto de los constituyentes en la voz pasiva. En lugar de seguir la estructura canónica del inglés (paciente + auxiliar + participio pasado), muchos participantes produjeron secuencias como "was developed this work" (E11V1), en lugar de "this work was developed". Este tipo de error puede explicarse a partir de diferencias entre el español y el inglés en cuanto al orden de los constituyentes: mientras que en español es común desplazar el sujeto hacia el final para enfatizar la acción o el agente, en inglés el sujeto gramatical de la pasiva se mantiene al inicio de la oración.

Un segundo error recurrente relacionado con esta frase es la formulación de la pasiva mediante la introducción de la preposición *in* donde no corresponde, lo que genera estructuras como: "In context to introduce the handling of techniques used in a purification protocol, the molecular absorption spectrometry and the adequate use of the calibrate curves were realized ***in*** the present work" (E06T2). Este tipo de formulación, que responde a la traducción errónea de la cláusula del texto fuente "En contexto de introducirnos en: el manejo de las técnicas utilizadas en un protocolo de purificación, la espectrofotometría de absorción molecular y el adecuado uso de las curvas de calibrado fue realizado el presente trabajo" altera la estructura de la oración. En lugar de mantener como sujeto paciente a "el presente trabajo", el uso de *in* hace que el sujeto gramatical pase a ser la enumeración previa de conceptos: "the management of the techniques used in a protocol of purification…". La forma resultante en inglés resulta inadecuada tanto desde el punto de vista gramatical como semántico: elementos que en el texto fuente eran parte del marco temático son transformados en sujetos gramaticales de una acción. Esto genera ambigüedad informativa y evidencia una comprensión insuficiente de las funciones sintácticas en la voz pasiva del inglés.

El segundo caso de estructura pasiva que presentó más errores estructurales en su traducción es "se determinó el porcentaje de purificación de la misma", que aparece en el siguiente contexto en el texto fuente: "A su vez la actividad de dicha enzima fue sondeada mediante una reacción de tipo colorimétrica y se determinó el porcentaje de purificación de la misma". Esta estructura es típica de los textos científicos en español, donde se utiliza la pasiva refleja (*se determinó*) para expresar una acción sin agente explícito. La traducción más directa al inglés sería una pasiva con sujeto paciente en primera posición, como "the percentage of purification was determined". Sin embargo, en las producciones analizadas, muchos estudiantes recurren a una inversión sintáctica que antepone el verbo pasivo al sujeto paciente, y produce estructuras como "was determined the percentage of purification" (E03T2) o "it was determined the purification percentage" (E06V3). Este patrón puede explicarse por la influencia del español, donde el orden verbo + sujeto es posible y aceptado en ciertas construcciones impersonales. En cambio, en inglés, la estructura pasiva exige que el sujeto gramatical ocupe la posición inicial. Además de la inversión, aparece con frecuencia el agregado innecesario del pronombre *it* al comienzo de la oración, como en "it was determined its purification percentage" (E01V1). Esta estructura resulta redundante y confusa: el pronombre *it* sugiere un sujeto anticipado que no se realiza en la oración, y la presencia simultánea de *its* como posesivo introduce ambigüedad sobre a qué se refiere cada elemento. Asimismo, se registran dificultades en la selección del pronombre posesivo, con casos como *her*, *your*, en construcciones como "by her percentage of purification", que indican una confusión tanto en la asignación de referentes como en la estructura de la pasiva. Otros casos menos representados en esta cláusula incluyen el uso de formas morfológicas incorrectas como "was determinated" (E03V1) en lugar de "was determined" o el empleo de tipos de gramaticales inadecuados, como "its purify percentage was *quantifier*" (E12T2 - *quantifier* por *quantified*).

El tercer grupo de fallos estructurales frecuentes se observa en la frase "fueron utilizadas curvas de calibrado", que aparece en el siguiente contexto: "Para la obtención de los datos que corresponden a concentraciones tanto de proteínas así como de P-Nitrofenol (p-NP) fueron utilizadas curvas de calibrado". En español, esta oración emplea una estructura pasiva perifrástica con orden verbo + sujeto. Una traducción esperable al inglés sería "calibration curves were used", donde el sujeto paciente (*calibration curves*) debe ocupar la posición inicial, seguido de la forma pasiva del verbo (*were used*). Sin embargo, en los textos analizados se observa con frecuencia la inversión de este orden, junto con el uso del verbo auxiliar en singular, con estructuras como "was used calibration curves" (E03V1), "was used curves of calibration" (E17V1), o "was used calibrated curves" (E04V1). Este tipo de error puede explicarse por la influencia del español, donde el verbo puede preceder al sujeto sin que esto se considere incorrecto. En inglés, en cambio, esta inversión rompe con la estructura gramatical esperada de la pasiva y puede generar ambigüedad, en especial cuando el sujeto es plural y el verbo está en

singular. Este último punto conduce al segundo error estructural recurrente en este segmento: la inadecuada concordancia entre el sujeto y el verbo en la forma pasiva. El orden inusual de los componentes de la pasiva y la falta de concordancia entre el verbo auxiliar y el sujeto paciente sugiere que la estructuración sintáctica de la pasiva representa una dificultad persistente. En algunos casos, los problemas se agravan por la presencia de formas verbales mal formadas como *was utilizated* (E11V1) por *was used*, o por combinaciones léxicas como "curves calibrates was used" (E12V1), que reflejan tanto una formación incorrecta del adjetivo (*calibrated* o *calibration*, según el caso) como una inversión de sujeto y verbo.

5.2. Errores generales

En las secciones anteriores se analizaron dos estructuras lingüísticas propias del registro académico y los problemas gramaticales y textuales asociados a ellas. En esta sección se presenta un análisis global de estos dos tipos de errores identificados en todo el corpus (v. fig. 8), incluyendo los estudiados en las secciones precedentes, pero incorporando otros de carácter general.

5.2.1. Errores gramaticales

Para el análisis de los fallos gramaticales, se evalúa su cantidad promedio en cada recorrido de traducción y revisión, así como el subtipo mayoritario en cada uno de ellos.

Cantidad promedio de errores gramaticales por ruta

Este apartado analiza la distribución y variación de los errores gramaticales en cada camino (v. 4.1.4.1), para evaluar si el uso combinado y secuencial de las herramientas conduce a una mejora cualitativa en las producciones escritas en inglés, y determinar si alguna de ellas resulta más efectiva.

En la vía *Grammarly* se observó una disminución notable en la cantidad promedio de fallos gramaticales entre la versión inicial (V1: 66,25) y la versión corregida por el sistema (T2: 21,50), lo que indica que el recurso cumple de forma eficaz su función de corrector gramatical. En la versión final (V3), tras la intervención de los participantes sobre el texto corregido, el promedio presenta un leve incremento (23,50). Este aumento puede deberse a la reintroducción de errores o a nuevas formulaciones por parte de los estudiantes, así como al hecho de que sus CCP en inglés escrito no resulten suficientes para identificar problemas que *Grammarly* no había señalado.

En la vía de *ChatGPT* se observa una reducción de errores gramaticales más significativa que en la de *Grammarly*. El promedio desciende de 66,25 en la versión inicial (V1) a 1,20 en la versión de la herramienta (T3), lo que señala resultados muy eficaces desde el punto de vista gramatical. En la versión final (V4), producida luego de la intervención de los participantes sobre el texto corregido por el

sistema, se registra un leve aumento en la cantidad de fallas (5,10), aunque el promedio se mantiene muy por debajo del valor inicial.

La ruta *Google Translate* muestra desde el comienzo un promedio de errores gramaticales menor al de la versión inicial de las otras rutas (5,75 en T1). Sin embargo, tras la intervención de los participantes (V2), se observa, una vez más, un leve aumento en la cantidad de fallas (8,60), lo que indica que algunas de las modificaciones introducidas generaron problemas que no estaban presentes en la traducción automática.

Este patrón recurrente se mantiene también en la secuencia *completa*, que incorpora de manera secuencial las tres plataformas. Las versiones automáticas (T4, T5) tienden a reducir la cantidad de errores, mientras que las intervenciones humanas posteriores (V5, V6) producen incrementos leves. No obstante, el promedio general disminuye en la secuencia, y alcanza su punto más bajo en T5 (1,45). La última versión producida por los participantes (V6) presenta un promedio (2,80) menor al registrado en las etapas iniciales, tanto de este trayecto como de los demás, aunque levemente mayor al promedio registrado para la versión anterior (T5).

Tipos de errores gramaticales por ruta

El examen de la distribución de los diferentes subtipos de errores gramaticales en cada vía muestra que en la ruta *Grammarly* se observan diferencias claras en la cantidad promedio de errores fonemáticos, morfológicos y sintácticos en las versiones que se inician con V1 y culminan en V3 (v. fig. 9). En la versión autónoma inicial (V1), predominan los fonemáticos (41,95), seguidos por los sintácticos (19,8). Los morfológicos, aunque presentes, son los menos frecuentes (4,5). Esto refleja que, en la fase inicial, los estudiantes tienen más dificultades relacionadas con la ortografía (por ejemplo, en la traducción de unidades léxicas como *porcentaje* que en varios experimentos se traduce como "porcentage" en lugar de "percentage") y el uso de mayúsculas, así como problemas con la estructura de las oraciones, como en "Despite of altercation" (E16V1) en lugar de "Despite altercations". Tras la intervención del sistema (T2), se observa una reducción significativa en los tres tipos de fallos. La reintroducción de errores gramaticales en la versión V3 se corresponde, en este caso, con fallos fonemáticos y de construcción sintáctica.

También se observan variaciones notables en el promedio de errores fonemáticos, morfológicos y sintácticos en las distintas versiones del trayecto *ChatGPT* (v. fig. 10). En esta ruta, la intervención de la herramienta (T3) produce una reducción aún más sustancial en los tres subtipos de fallos, en especial en los sintácticos, que casi desaparecen (0,05), siendo "this work was conducted to familiarize[27] with

[27] "To familiarize" es un verbo transitivo que requiere de un objeto directo que, en este caso, no está presente. Una forma correcta hubiera sido "this work was conducted to familiarize ourselves with molecular absorption spectrometry".

molecular absorption spectrometry" (E06T3) el único caso encontrado. En V4, producido tras la intervención de los participantes sobre T3, se registran incrementos leves, aunque los valores se mantienen muy por debajo de los observados en la versión inicial. Al igual que en el camino *Grammarly*, la reintroducción de errores por parte de los estudiantes se concentra en los fonemáticos (3,20) y sintácticos (1,5), mientras que los morfológicos casi no varían.

Las diferencias entre versiones resultan menos marcadas en el caso de la secuencia *Google Translate*, y de la *completa* (v. fig. 11), lo que parece obedecer al hecho de que el texto generado por el traductor (T1) presenta bajos promedios iniciales de errores de los tres subtipos: 5,15 fonemáticos, 0,3 morfológicos y 0,3 sintácticos. A partir de T1, los cambios posteriores muestran sólo variaciones bajas. Como se observó en las rutas *Grammarly* y *ChatGPT*, hay una mínima tendencia a la reintroducción de errores tras las modificaciones humanas, sobre todo en el plano fonemático: aumentan en V2 (6,85), se reducen en T4 (4,05), pero vuelven a incrementarse en V5 (4,25), y aunque descienden de forma notable tras la intervención de *ChatGPT* en T5 (1,45), se incrementan en la versión final V6 (2,40). Los errores sintácticos, por su parte, sólo experimentan incrementos leves en dos momentos: tras la intervención de los estudiantes en V2 (de 0,3 a 1,55) y en V5 (de 0,4 en T4 a 0,65). En cuanto a los fallos morfológicos, la trayectoria es más estable: se mantienen cercanos a cero en todas las versiones, con un pequeño aumento en V6 (0,2) tras haber desaparecido en T5. En conjunto, esta vía muestra una progresión más gradual y una reintroducción de fallos menos pronunciada que en las anteriores.

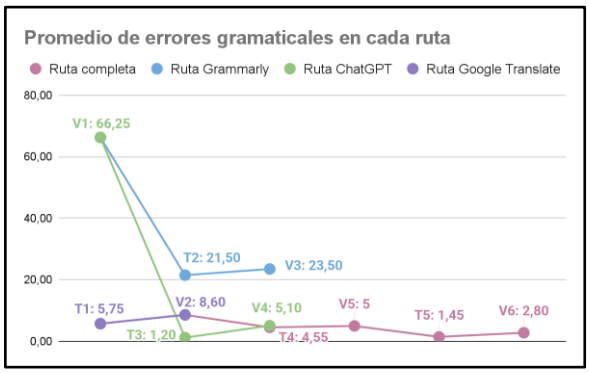

Figura 8

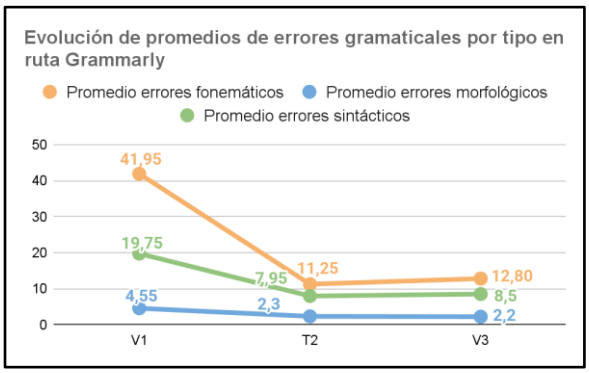

Figura 9

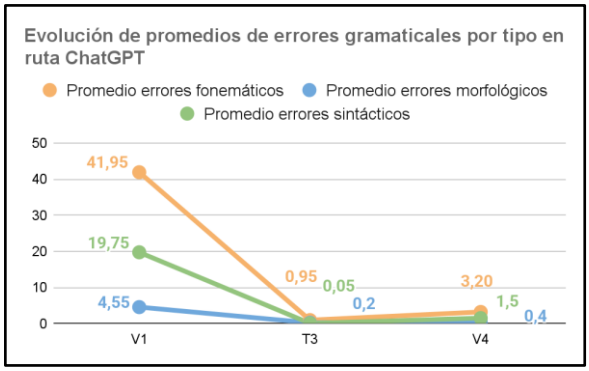

Figura 10

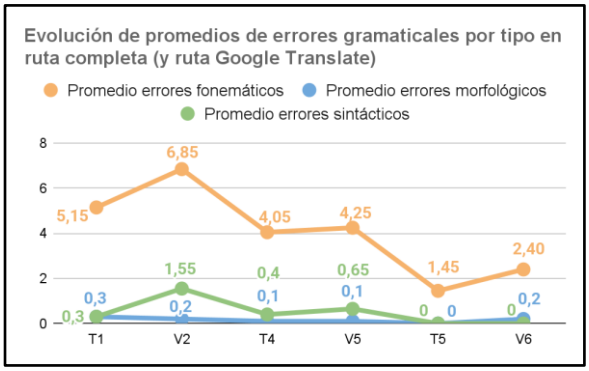

Figura 11

Estos resultados evidencian una mejora gradual en la corrección gramatical en todos los recorridos, y muestran que las ediciones humanas, si bien reintroducen ciertos errores, no revierten el efecto de las tecnologías de forma considerable. Por otro lado, la ruta *ChatGPT* presenta la mayor corrección de fallos entre la versión inicial (V1) y la automática (T3). En comparación, la de *Grammarly*, aunque también efectiva, muestra una reducción menor. La combinación secuencial, por su parte, permite iniciar con pocos errores, reducirlos, y mantener su promedio con cierta estabilidad. Sin embargo, si se contrasta la intervención de *ChatGPT* (T3) con la traducción de *Google Translate* (T1), también se observa una diferencia significativa a favor de la IA generativa. Además, *ChatGPT* presenta un rendimiento superior a *Grammarly* en los tres subtipos de fallos gramaticales, y también demuestra ser más eficaz que *Google Translate* respecto de los errores fonemáticos y sintácticos.

5.2.2. Errores textuales

Para el análisis de los problemas textuales, se siguen los mismos lineamientos que para el análisis precedente de los gramaticales.

Cantidad promedio de errores textuales por ruta

Al analizar la distribución y variación de fallos textuales (v. 4.1.4.1), la dinámica del recorrido *Grammarly* difiere de la observada en el análisis de errores gramaticales (v. fig. 12). Como se vio en las secciones precedentes, en ese caso la herramienta logró una disminución considerable de los fallos, en cambio, aquí la reducción es mínima: el promedio desciende apenas de 14,20 en la versión inicial (V1) a 13,35 en la corregida por el sistema (T2). Tras la intervención de los participantes, se registra un leve aumento en el promedio: 13,60 en la versión final (V3). Estos datos sugieren que al menos en su versión gratuita, la plataforma no corrige de manera sustancial esta clase de fallo. Por otro lado, los participantes tienden a reintroducir o generar estos errores durante su intervención.

En el recorrido de *ChatGPT*, el comportamiento de los problemas textuales muestra un descenso más significativo que en el de *Grammarly*, ya que baja de 14,20 en la versión inicial (V1) a 3,20 en la corregida por el *chatbot* (T3). No obstante, si se compara esta disminución con los resultados obtenidos en el análisis de errores gramaticales —entre los que el promedio cayó de 66,25 en V1 a 1,20 en T3—, se advierte que la eficiencia de *ChatGPT* en la reducción de problemas de índole textual, aunque considerable, es menos pronunciada. Al igual que en otras vías, los participantes tienden a reintroducir o generar fallos textuales durante su intervención posterior: se registra un aumento de 3,20 promedio en T3 a 4,05 en V4, aunque la cantidad se mantiene muy por debajo de la versión inicial.

En comparación con las versiones iniciales de otros trayectos, en el de *Google Translate* la cantidad promedio de errores es baja: 3,50 en T1. Tras la intervención de los participantes (V2), se observa también un leve aumento, con un promedio de 4,85. Este patrón se mantiene en la vía *completa*, donde

se integran progresivamente las tecnologías restantes. Las versiones T4 y T5 tienden a presentar menor cantidad de fallas (4,65 y 2,20), mientras que las ediciones humanas (V5 y V6) vuelven a registrar pequeños aumentos (4,55 y 2,75). A pesar de estas fluctuaciones, el promedio desciende gradualmente en la secuencia, se mantiene en niveles inferiores a los observados en las versiones iniciales de todas las rutas, y alcanza su valor más bajo en T5 (2,20).

Los resultados muestran una tendencia general hacia la mejora de la corrección textual en todos los caminos, aunque de manera menos pronunciada que en el caso de la gramatical. Esto sugiere que los errores textuales presentan mayores dificultades, tanto para las tecnologías del texto como para los participantes humanos. En el caso de los recursos digitales, la dificultad puede vincularse a la necesidad de interpretar el contexto comunicativo y la intención del emisor, aspectos que parecen exceder sus capacidades actuales. Para los participantes, en cambio, podría deberse a limitaciones en sus niveles de CCP en lengua inglesa, lo que afectaría su capacidad para identificar y subsanar este tipo de problemas en las producciones escritas.

Tipos de errores textuales por ruta

También se examinó la distribución de errores textuales por subtipos en cada camino. En la versión V1 predominan los problemas de informatividad, intencionalidad e intertextualidad (III), con un promedio de 6,47, seguidos por los de situacionalidad y aceptabilidad (SA), con un promedio de 5,63, y, en menor medida, los de coherencia y cohesión (CC), que registran un promedio de 2,84.

En el caso de la vía *Grammarly*, se registra una diferencia en la efectividad de corrección de cada subcategoría, con una mayor capacidad para corregir errores de CC (v. fig. 13). Tras la intervención automática (T2), los problemas de este tipo se reducen de 2,84 a 1,63, mientras que los de III experimentan un leve aumento (de 6,47 a 6,89) y los de SA se mantienen casi constantes (de 5,63 a 5,53). Este aumento en los fallos de III constituye un caso atípico, ya que es el único en el que se registra un incremento en un subtipo de error tras una intervención automática. Una posible explicación radica en su naturaleza, pues suponen la consideración de aspectos informativos que una plataforma de corrección automática como *Grammarly* no puede evaluar de manera efectiva. Además, es posible que, a partir de la corrección de errores ortográficos en V1 (analizada en la sección precedente), la herramienta haya generado problemas de selección léxica. Al corregir la ortografía en T2, en varios casos *Grammarly* reemplazó unidades léxicas mal escritas por otras con formas similares, pero con un significado diferente. Por ejemplo, en el caso de E11, la traducción del participante en V1 incluye el texto "possible mistakes that could conduced to inesperated results". Al corregir esta sección, la herramienta generó la siguiente sugerencia: "possible mistakes that could conduce to desperate results". Mientras que la corrección de "conduced" resultó correcta, el intento de corregir "inesperated" (que

podría haber sido corregido como “unexpected”) resultó en la selección de una unidad léxica con un significado diferente del deseado. Así, aunque la forma ortográfica resultó correcta, el contenido del texto se vio alterado, lo que originó nuevos problemas de III. La reintroducción de fallos en la versión intervenida por los participantes (V3), por su parte, incide especialmente sobre los errores de SA, cuyo promedio aumenta a 5,79, mientras que los de CC se incrementan sólo de forma leve (a 1,74) y los de III disminuyen de forma marginal (a 6,79).

Al recorrer la secuencia *ChatGPT*, se detectan también cambios significativos en el promedio de errores textuales por subtipo en las sucesivas versiones (v. fig. 14). La intervención automática (T3) exhibe una reducción notable en los problemas de CC, que disminuyen de 2,84 a 0,11, y en los de SA, que descienden de 5,63 a 0,26. Si bien no se registra un aumento en el caso de los de III, como sucede en la vía *Grammarly*, sí se observa que la reducción es menor en comparación con las otras categorías (6,47 a 3,00). Algunos ejemplos de errores de III remanentes que *ChatGPT* no logró corregir son de selección léxica como “spectroscopy” (en lugar de “spectrophotometry” - E03T3) o el uso de la sigla “F.E.” (E02T3). Este último caso resulta particularmente interesante, ya que proviene de un acrónimo en español que quiere decir “Fracción de la enzima”. Muchos participantes optaron por traducir la expresión manteniendo la expresión fuente, lo que generó un problema informativo: en inglés, el significado de “F.E” no resulta evidente. Al procesar estos textos, *ChatGPT* tendió a conservar la sigla, reproduciendo así esta ambigüedad. Por otro lado, también en este caso, tras la intervención de los participantes (V4), se registran leves incrementos en las tres clases de fallos, aunque con valores por debajo de los observados en la versión inicial (V1).

En el caso de las rutas *Google Translate* y *completa*, las diferencias entre las dinámicas de cada categoría resultan menos marcadas (v. fig. 15). Esto parece vincularse al hecho de que el texto generado por el traductor (T1) presenta ya bajos promedios de errores textuales en los tres subtipos: no se registran fallos de CC, y los de III y SA se ubican en 3,00 y 0,68. A partir de esta base, los cambios posteriores muestran variaciones moderadas. Como se observó en las vías *Grammarly* y *ChatGPT*, se detecta una tendencia a la reintroducción de errores tras las ediciones humanas, en este caso, con más incidencia en el subtipo SA: el promedio aumenta en V2 (1,58), disminuye en T4 (1,42) y se estabiliza en V5 (1,42). Ejemplos de este subtipo son el uso recurrente de contracciones, como es el caso de “doesn`t” en “It doesn't wait 4 minutes before reading the absorbance at 595nm.” (E03V5). Luego, tras la intervención de *ChatGPT* (T5), se produce una disminución notable (0,11), aunque en la versión final (V6) se registra un leve incremento (0,26). Los fallos de III, por su parte, parten de un promedio de 3,00 en T1, experimentan aumentos leves en V2 (3,21) y T4 (3,32), y luego se estabilizan en torno a 3,26 en V5. Al igual que en el camino *Grammarly*, resulta significativo que los errores de III se incrementen de manera leve tras emplear la herramienta (T4). Como se señaló, este fenómeno podría estar vinculado al

intento de corrección ortográfica que reemplaza unidades léxicas mal escritas por otras de forma similar, pero con un significado diferente. Por último, estos descienden en T5 (2,16) y vuelven a incrementarse en V6 (2,47). En cuanto a los problemas de CC, los cambios son muy pequeños: el promedio aumenta de 0 en T1 a 0,32 en V2, disminuye a 0,16 en T4, baja a 0,11 en V5, y vuelve a incrementarse a 0,16 en V6. En general, en esta secuencia todas las fluctuaciones de los valores resultan ínfimas, lo que puede estar vinculado a los bajos promedios iniciales de fallos textuales.

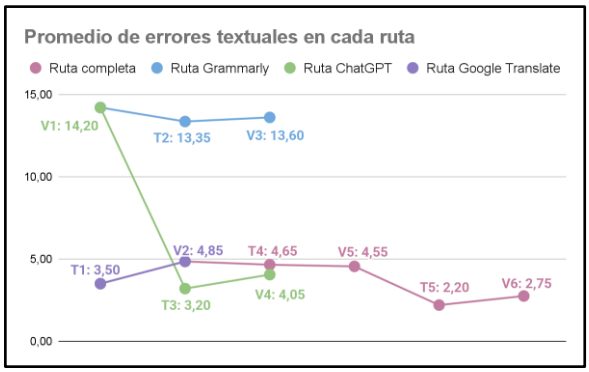

Figura 12

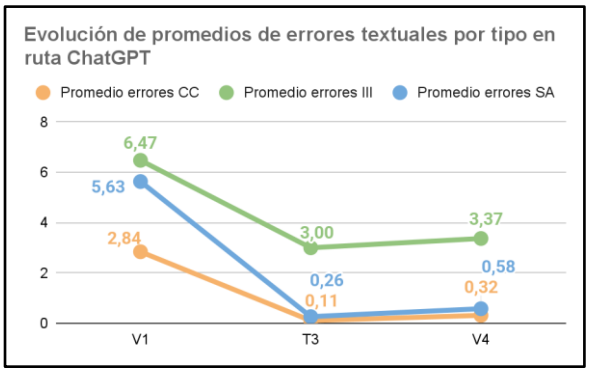

Figura 14

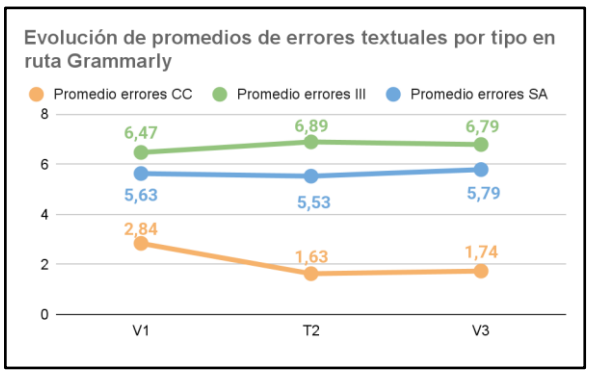

Figura 13

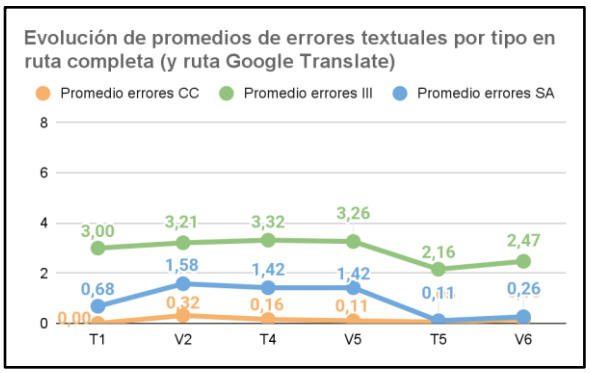

Figura 15

Las tendencias observadas en este apartado son similares a las del nivel gramatical, pero las mejoras son más limitadas. Se advierte que *ChatGPT* resulta también el sistema más eficaz para la reducción de errores textuales. No obstante, si se comparan los descensos relativos, se observa que su eficiencia es menor que la evidenciada en la corrección gramatical. Además, si bien las intervenciones humanas tienden a incrementar levemente los errores, las versiones finales de cada recorrido mantienen niveles de fallo inferiores a los registrados en un principio, lo que indica que el proceso de corrección asistido impacta de modo positivo sobre la calidad textual general. En lo referente a los subtipos de errores, la eficacia de *ChatGPT* también destaca, ya que logra una disminución significativa en todos ellos, en especial al corregir problemas de CC y SA. En cambio, si bien presenta una mínima mejora en los errores de CC, *Grammarly* no logra reducir los de las otras dos categorías, y muestra un leve incremento en los de III, lo que indica que la eficacia de los sistemas para intervenir en todos los aspectos difiere. Por otro lado, el desempeño inicial de *Google Translate*, que ya presenta bajos niveles de

problemas textuales en T1, parece explicar la menor magnitud de las fluctuaciones observadas en la vía *completa* que lo incluye.

5.3. Intervenciones de los participantes

En este apartado se analizan las modificaciones que los participantes realizaron sobre las distintas versiones de los textos generadas o corregidas por las herramientas con el fin de comprender su proceso de apropiación de las tecnologías en la tarea de traducción y corrección planteada experimentalmente. Las intervenciones se clasificaron según las categorías detalladas en la sección de metodología. En primer lugar, el análisis se centró en identificar las modificaciones más frecuentes, para luego considerar las que predominan en cada una de las versiones intervenidas (V2, V3, V4, V5 y V6). A diferencia de las secciones anteriores, donde el análisis se organizó por rutas, en este caso el foco recae sobre las modificaciones realizadas a las versiones T.

Intervenciones más frecuentes en el corpus

En primer lugar, se observa que las operaciones más frecuentes fueron las recuperaciones (12,45 en promedio), seguidas por los reemplazos (8,2 en promedio). Un ejemplo del primer grupo es el caso de "this value should not be taken into account" (E07V5), que *ChatGPT* reformuló en T5 como "this value should be disregarded" (E07T5); sin embargo, el participante volvió a la versión original en V6. En el caso de los reemplazos, puede citarse el ejemplo de "the activity of said enzyme was probed" (E10T1), que fue modificado en V2 como "the activity of this enzyme was analysed", realizando así dos reemplazos: *said* por *this* y *probed* por *analysed*. Por su parte, las recuperaciones con reemplazo[28] presentaron un promedio de 5,1. Estos datos indican que los participantes aceptaron algunas de las modificaciones propuestas por los sistemas, con tendencia a recuperar y ajustar el vocabulario según sus preferencias originales. En menor medida, realizaron operaciones como alteraciones gráfico-morfológicas[29] (7), reformulaciones (3,65) y eliminaciones (3,5), lo que evidencia intervenciones adicionales en los niveles estructural y morfológico. Las ediciones menos frecuentes fueron las inserciones (3,2), los reordenamientos (1) y las alteraciones de puntuación (0,8), mientras que las combinaciones de reformulación/reemplazo + reordenamiento fueron las menos comunes en promedio (0,65). Estos resultados sugieren que, en general, los participantes se enfocaron en recuperar o ajustar las unidades léxicas y frases, mientras que prestaron menos atención a modificaciones estructurales o sintácticas más complejas (v. fig. 16).

[28] Esta categoría fue creada para identificar los casos en los que los participantes sustituyeron una unidad léxica que había sido modificada por la plataforma por su versión original.

[29] Las categorías "alteración gráfico-morfológica", "alteración de la puntuación" y "reformulación/reemplazo + reordenamiento" han sido denominadas ALT_GM, ALT_PUNT y RR_REORD, respectivamente, en la figura 16, por cuestiones de espacio.

Intervenciones más frecuentes en cada ruta

En cuanto a los tipos de intervenciones más comunes por versión (v. fig. 17), se observa que, cuando los participantes intervienen sobre T1, generada por *Google Translate,* los reemplazos fueron la operación predominante (V2: 5,74 casos promedio). Esto revela que se concentraron en realizar cambios léxicos, con una tasa menor de modificaciones estructurales (3,74 alteraciones gráfico-morfológicas y 2,63 reformulaciones, en promedio). De hecho, las encuestas de percepción (v. 5.5.) revelan que los participantes consideran que el traductor tiende a ser demasiado literal, lo que podría explicar su tendencia a modificar más el léxico que otros aspectos. Por otro lado, no se registran recuperaciones en V2, ya que esta versión es la intervención de la versión inicial de la vía *Google Translate* y de la *completa*, y no hubo nada que "recuperar" de una versión previa.

En V3, V4, V5 y V6, las recuperaciones se convirtieron en la operación más frecuente. Esto sugiere que, en estos casos, los participantes tendieron a revertir ciertos cambios más que a proponer alternativas de traducción no consideradas en los textos previos. En particular, el alto promedio de las recuperaciones en V4 (3,26) y V6 (6,16) puede estar relacionado con el hecho de que los participantes consideran que el *chatbot* realiza correcciones demasiado profundas, al reescribir el texto en su totalidad. Así, los resultados cuantitativos, en conjunto con las percepciones de los participantes, sugieren que las correcciones de *ChatGPT* les resultan excesivas, lo que los llevó a recuperar más elementos de su versión original que al revisar las correcciones de *Grammarly*.

Cantidad de intervenciones por versión

En cuanto a la cantidad de cambios por versión (v. fig. 18), se observa que la que presenta mayor promedio es V2 (16,70). Esto puede sugerir que los participantes tienden a desconfiar más de *Google Translate* que de las otras herramientas. A continuación, con una frecuencia algo menor se ubican las versiones V4 (9,10) y V6 (12,00), correspondientes a las correcciones realizadas sobre los textos generados por *ChatGPT*. En este caso, los valores podrían vincularse con la percepción de que sus correcciones son excesivas o invasivas.

Por último, las intervenciones de V3 (6,15) y V5 (1,60), ambas realizadas sobre correcciones de *Grammarly*, son las menos numerosas. En particular, el caso de la versión V5 podría explicarse en relación al análisis de los errores de V2, que presentó pocos fallos gramaticales o textuales, lo que sugiere que *Grammarly* tuvo un margen reducido para proponer mejoras significativas. Como resultado, y ante una menor cantidad de correcciones automáticas, los participantes pueden no haber considerado la necesidad de realizar más modificaciones sobre los textos revisados por este sistema.

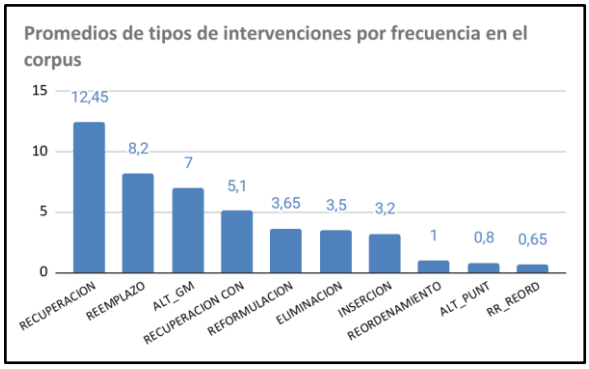

Figura 16

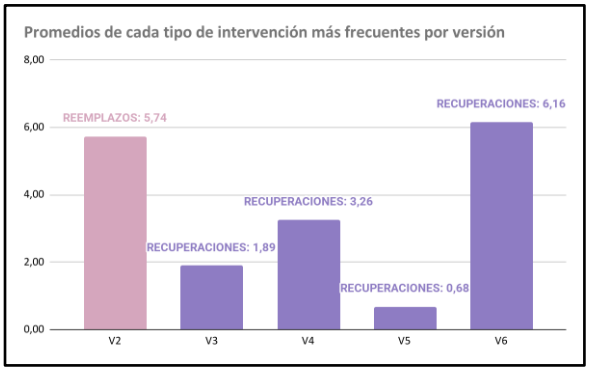

Figura 17

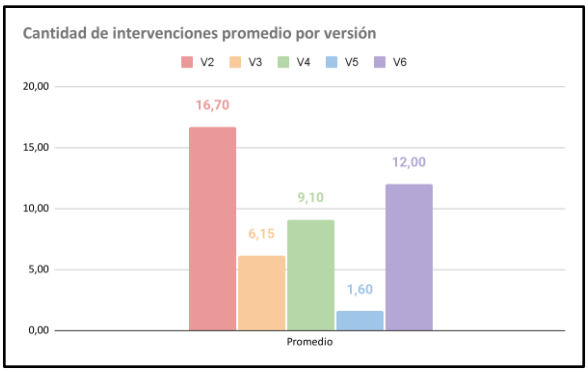

Figura 18

5.4. Perfil de vocabulario

Tal como se anticipó en la sección de metodología, se calculó el *Lexical Frequency Profile* (*LFP*) de las traducciones obtenidas, utilizando las listas *Academic Word List* (AWL) y *General Word List* (GSL) para generar perfiles de vocabulario de cada una (v. figs. 19-21) e identificar la proporción de vocabulario de uso cotidiano y académico presente en cada versión. El objetivo de este análisis fue detectar su posible relación con el uso de las plataformas utilizadas en cada ruta. A partir del análisis realizado, logramos observar ciertas tendencias con respecto a las listas de vocabulario *GSL* y *AWL*. En primer lugar, en la vía *Grammarly*, advertimos un leve aumento en el uso de vocabulario académico (en la *AWL*), al pasar de un promedio de 6,28 en la versión autónoma (V1) a un 7,91 en la versión corregida por la herramienta (T2). Esto puede vincularse a la corrección ortográfica realizada por *Grammarly*, que permite la identificación de unidades léxicas de la *AWL*, que en versiones con errores podrían haber sido clasificadas como fuera de las listas (*Off-list)*. Asimismo, aparece una disminución del promedio de elementos en la *Off-list* entre V1 (21,31) y T2 (17,28), por lo que es posible que, al menos en este recorrido, *Grammarly* actúe más como un corrector ortográfico que como un sistema que proponga modificaciones léxicas sustanciales. En la siguiente versión (V3), que incluye intervenciones de los participantes, los valores se mantienen estables, lo que indica que no realizan demasiados cambios en el léxico o sobre la corrección ortográfica de *Grammarly*.

En el recorrido de *ChatGPT*, en cambio, se observa un incremento más notorio del vocabulario de la *AWL*: la proporción de unidades léxicas de esa lista se eleva del 6,28 inicial al 11,33 en T3, y se mantiene en niveles altos tras la intervención de los participantes (V4: 10,85). Esto parece mostrar que *ChatGPT* no solo corrige fallos ortográficos, sino que también introduce vocabulario perteneciente a registros más académicos. Paralelamente, se advierte un leve descenso en la frecuencia de vocabulario de la *GSL* en T3, lo que sugiere una sustitución parcial de léxico cotidiano por uno más especializado. La *Off-list*, por su parte, también se reduce, lo que podría atribuirse, al menos parcialmente, a la corrección de problemas ortográficos.

El camino de *Google Translate*, que contempla una traducción producida por el traductor (T1) y una versión intervenida por los participantes (V2), presenta un patrón distinto. El vocabulario de la *AWL* se mantiene casi estable (7,98 en T1 a 7,80 en V2), al igual que el de unidades léxicas en la *GSL* (75,65 en T1 a 75,27 en V2). No obstante, el valor de *Off-list* se incrementa de forma notable (de 7,98 en T1 a 16,92 en V2), lo que puede vincularse a un aumento de fallos ortográficos o a la introducción de léxico no contemplado por las listas empleadas tras la intervención humana.

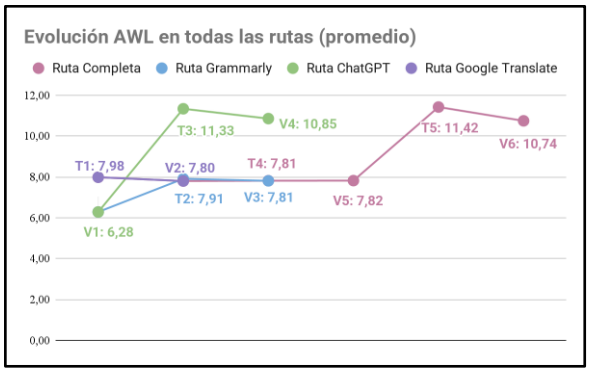

Figura 19

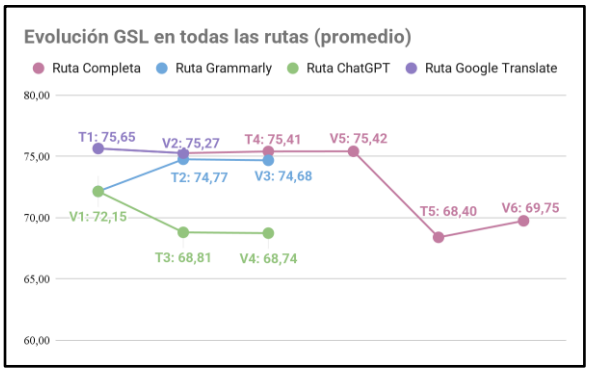

Figura 20

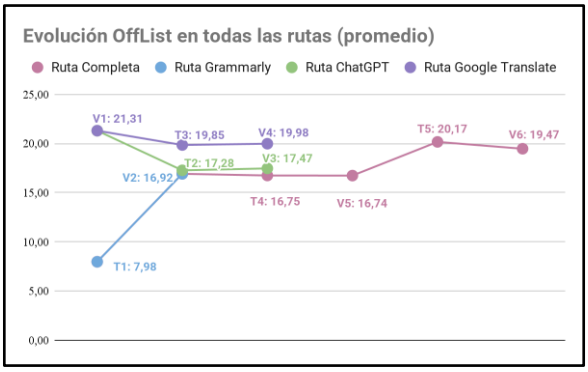

Figura 21

Por último, la proporción de vocabulario académico se mantiene estable durante las primeras etapas (T1 a V5) de la ruta *completa*, con valores cercanos al 7,8, pero se incrementa significativamente luego de la corrección de *ChatGPT* (T5: 11,42; V6: 10,74). En paralelo, el uso del vocabulario de la *GSL* desciende (T1:75,65 en T1; T5:68,40), mientras que la proporción de *Off-list* aumenta (de 7,98 a

19,47). Esto refuerza la tendencia observada en la vía *ChatGPT*: la herramienta introduce léxico académico, pero también puede contribuir al uso de vocablos no contemplados por las listas.

En general, los datos sugieren que *Grammarly* tiende a reducir la proporción de vocabulario en las *Off-list* a partir de la corrección ortográfica, mientras que *ChatGPT* parece reformular el texto con mayor profundidad y, en consecuencia, se eleva la proporción de vocabulario en la *AWL*, así como el léxico que no se encuentra en ninguna de las dos listas utilizadas. *Google Translate*, por su parte, ofrece una base de vocabulario similar a la lograda a partir del uso de *Grammarly* sobre las traducciones autónomas (V1), que los participantes tienden a mantener con muy leves modificaciones. La incorporación de *Grammarly* en etapas posteriores, cuando el texto ya contiene pocos fallos ortográficos, no parece tener un impacto significativo sobre el léxico académico empleado.

5.5. Percepciones de los participantes

Además del análisis precedente de las producciones escritas, el diseño metodológico contempla una instancia de evaluación subjetiva por parte de los participantes implementada mediante una encuesta de percepción. Esta encuesta tuvo como finalidad explorar la valoración que los estudiantes hacen del uso de recursos digitales en tareas de producción escrita en inglés, tanto de manera individual como combinada. En esta sección se presentan las principales tendencias cuantitativas observadas, junto con el análisis cualitativo de las opiniones expresadas por los participantes.

Percepciones sobre las herramientas

Se consultó a los participantes sobre el nivel de dificultad que experimentaron con el uso de cada recurso digital (v. fig. 22), que se midió mediante una escala de *Likert* de 5 puntos (1 = Muy sencillo; 5 = Muy complicado). *Google Translate* fue el que recibió mejores valoraciones ya que para el 80% de los estudiantes resultó "muy sencillo" utilizar esta herramienta. El 70% consideró también que es "muy sencillo" usar *ChatGPT* y el 20% seleccionó la siguiente opción en nivel de dificultad (2). Las respuestas sobre *Grammarly*, en cambio, mostraron una mayor dispersión. Si bien el 40% indicó que su uso fue "muy sencillo", el 25% eligió el nivel 2 de dificultad, el 30% eligió el nivel intermedio (opción 3), y el 5% (un participante) seleccionó la opción 5 ("Muy complicado"). *Grammarly* parece haber generado más dudas o exigido una mayor familiarización previa. Esto puede deberse a que, si bien la herramienta presentaba una interfaz similar a la de otros programas conocidos como *MS Word*, resultó ser la herramienta que menos participantes conocían al iniciar la actividad.

Se les consultó si consideraban que cada uno de los recursos utilizados ofrecía algún beneficio para traducir un informe de laboratorio al inglés. El 90% respondió afirmativamente en relación con *ChatGPT*, el 80 % respecto de *Grammarly* y el 60 % en el caso de *Google Translate*. Ninguno consideró que todos los sistemas carecieran de utilidad (v. fig. 23).

Al indagar cuál de las plataformas resultaba más útil, *ChatGPT* volvió a destacarse: fue seleccionada por el 65 % de los encuestados. *Grammarly* fue elegida por el 30 %, mientras que *Google Translate*, contrariamente a lo que podría suponerse, solo recibió el 5 % de las preferencias (v. fig. 24). Nadie optó por la opción "todas son igual de útiles", lo que sugiere una diferenciación clara en cuanto al aporte de cada recurso.

También se consultó por la fiabilidad de cada plataforma (v. fig. 27). En el caso de *Google Translate*, el 85% (17 de 20) consideró que puede ser confiable para traducir textos académicos o profesionales al inglés, pero solo si el usuario revisa los resultados para asegurar su calidad. Solo una persona (5%) manifestó plena confianza en el traductor e indicó que no considera necesaria una revisión posterior, mientras que el 10% (2) expresó que no la considera fiable en ningún caso. La fiabilidad de *Google Translate* como traductor automático continúa siendo objeto de duda, en especial en términos de precisión y adecuación, según se puede apreciar cuando justifican esa valoración.

Muchos estudiantes señalaron que este sistema tiende a realizar traducciones literales, sin contemplar el contexto, lo que da lugar a textos poco naturales, rígidos o incoherentes. Así, E09 afirmó: "Hay ciertas cosas que las traduce literalmente y a veces, como ya mencioné, se pierde la intención o la información que se quiere informar. Además, algunas oraciones no las traduce adecuadamente". Mencionaron también percibir en estas versiones la presencia de errores ortográficos, fallas en la interpretación de expresiones idiomáticas, y una falta de sensibilidad hacia el registro del texto fuente: "Porque a veces no es preciso y no traduce de la mejor manera. Con respecto a la ortografía muchas veces no la tiene en cuenta y el texto tiene errores. A veces también no traduce bien una oración y cuando se lee es confusa, hablando en cuanto a tiempos verbales." (E01). Algunos participantes basaron su opinión en experiencias negativas previas con el sistema, y en la recomendación de sus docentes de inglés: "Particularmente porque me encontré con traducción mal escritas de google translate y mi profesora de ingles nos recomienda que dudemos de este traductor" (E17). Por otro lado, resulta interesante destacar que algunas opiniones parecen haberse basado más en percepciones generales de desconfianza hacia el sistema que en una evaluación de su desempeño en algún aspecto concreto. Por ejemplo, E15 comentó: "A pesar de haber funcionado de una forma mucho más satisfactoria que la que esperaba, es probable que siempre haya al menos una palabra u oración que no haya sido traducida correctamente", mientras que E10 expresó: "No me da confianza, creo que el programa no está preparado para eso". Estas percepciones podrían vincularse con la imagen negativa que *Google Translate* tuvo durante años, en especial antes de la implementación del sistema de traducción neuronal en 2016. Aunque el rendimiento del traductor ha mejorado significativamente desde entonces, es posible que muchos usuarios, en particular quienes no siguen de cerca la evolución de estas tecnologías, tengan una percepción desactualizada de su funcionamiento.

Al preguntarles sobre la fiabilidad de *Grammarly*, todos respondieron que la consideraban al menos parcialmente fiable. La mayoría de los estudiantes (17 de 20) consideró que es fiable, siempre y cuando el usuario revise los resultados para asegurar su calidad, y tres personas manifestaron plena confianza en el corrector, señalando que no consideran necesaria una revisión posterior. Aunque *Grammarly* es percibida como un poco más precisa y confiable que *Google Translate*, también se considera necesario un control humano para garantizar la calidad del texto final. En la justificación de esta valoración, los estudiantes destacaron que *Grammarly* resulta eficaz para corregir problemas ortográficos, gramaticales y de puntuación, y valoraron que sus sugerencias no suelen alterar de forma significativa el significado original: "En general esta herramienta revisa gramática y ortografía. En principio, el significado del texto no cambiaría, por lo que me parece muy buena para optimizar la escritura (incluso aunque esté bien traducida)" (E15).

No obstante, también señalaron limitaciones, como que no siempre capta errores de contexto, ni ofrece las mejores sugerencias en términos de estilo o sintaxis. Algunos participantes expresaron que las correcciones pueden resultar erróneas, como cuando se trata de corregir errores de ortografía: "Hay palabras que no conoce y no puede traducir y por otro lado solo corrige gramática pero tal vez toma la palabra erróneamente. Corrige la palabra pero al estar mal escrita entiende que es otra palabra" (E03). Además, mencionaron restricciones respecto a la versión gratuita: "Grammarly solo corrige errores gramaticales y puntuaciones en su version gratuita, entonces puede que algun error a nivel sentido de la oración puede quedar sin que el programa lo note" (E00).

Con respecto a la fiabilidad de *ChatGPT* para corregir textos académicos o profesionales en inglés, la mayoría de las opiniones (19 de 20) indica que, si bien el *chatbot* es fiable, es imprescindible revisar sus salidas para garantizar su calidad. Solo un participante expresó plena desconfianza en este sistema de IA. Esta distribución, similar en los tres casos, muestra que, aunque se reconoce el potencial de todas las plataformas, persiste una valoración generalizada sobre la necesidad de control por parte del usuario. Las justificaciones ofrecidas a esta pregunta refuerzan esta percepción: "Lo mismo que lo explicado con grammarly, creo que confio mas que el traductor porque es mas abierto y no te traduce palabra por palabra pero igualemente hay que controlar los resultados porque se trata de una inteligencia aritficial y capaz no se confia tanto" (E13). Además, valoraron características específicas de este recurso: "Siento que ChatGPT es la mejor de estas herramientas para corregir y traducir, puesto que puede hacer análisis mucho más complejos (no sólo corrige o traduce, es capaz de considerar el contexto e incluso responder a necesidades específicas por parte del usuario)" (E15).

Sin embargo, los estudiantes también señalaron sus limitaciones, como que "ChatGPT reescribe el texto completamente y puede que algunos datos que el usuario considere pertinentes se pierda." (E00). Además, se subraya que: "Puede tener errores de interpretación, por eso le debes brindar un

adecuado PROMPT" (E19). En general, se percibe que la posibilidad de producir textos gramatical y estilísticamente correctos no garantiza que el resultado sea adecuado en términos de intención: "Porque a veces reduce mucho el texto original y deja afuera información importante para uno mismo. También puede ser xq cambie mucho las ideas del texto original y que quede otra cosa completamente diferente" (E01).

Percepciones sobre el modo de usar las herramientas

Por otro lado, se preguntó a los participantes si el uso combinado de los sistemas había facilitado o dificultado la realización de la tarea. El 95% de los encuestados afirmó que se les hizo más fácil trabajar con las tres plataformas en conjunto, mientras que el 5% (un participante) expresó lo contrario. Sin embargo, este último señaló que "Aunque se me haya hecho más complicado, reconozco que es muy diferente (se llega a un mejor resultado) al usar las tres herramientss[30] y poder ir comparando palabras y sintaxis que usando una sola o solo el diccionario" (E10).

Al justificar sus opiniones, muchos estudiantes compararon las fortalezas de cada sistema. *Google Translate* fue vista como útil para lograr una primera versión rápida: "Es más sencillo empezar con algo ya armado y a partir de ahí ir cambiando" (E12) y *Grammarly*, para corregir aspectos formales: "Gramarly ayuda a encontrar fallas en la gramática qué tal vez por escribir rápido o por saber palabras por su pronunciación no se como se escriben." (E03). Por último, *ChatGPT* fue considerado útil para mejorar la naturalidad del texto: "ChatGPT hizo una muy buena traducción a partir de la mía, haciendola mas natural con respecto a la de Google" (E02). También destacaron que el uso conjunto facilitó y aceleró el proceso, al ofrecer puntos de partida y opciones de corrección más accesibles que una traducción hecha desde cero. Otros subrayaron el aporte al enriquecimiento del léxico, con mejoras en el vocabulario técnico y expresiones más variadas. Se consideró que las plataformas son útiles para detectar errores y afinar la escritura, aunque algunos mencionaron limitaciones puntuales. Por ejemplo, el participante E02 señaló: "Noté que las herramientas hicieron la traducción más fácil, sin embargo el traductor de Google es bastante literal". Otro ejemplo destacable es el del participante E05, quien destaca que "ChatGPT suele tomarse ciertas libertades en la redacción que requieren que la persona sepa inglés para corroborar que el contenido del texto esté preservado". En general, sin embargo, los participantes coincidieron en que la combinación permitió lograr un mejor resultado.

Ante la consulta sobre si el uso combinado y secuencial de estos recursos ofrecía mayores beneficios que su empleo por separado, el 85% respondió que sí, frente al 15% que consideró que no había diferencia (v. fig. 25). Esta tendencia refuerza la idea de que los estudiantes no sólo valoran las

[30] Las respuestas citadas fueron transcritas tal como fueron escritas por los participantes, sin corregir errores de ningún tipo.

distintas funciones en cada sistema, sino que también perciben una mejora en la calidad de la traducción al integrarlas. En este sentido, al solicitarles que explicaran el motivo de su respuesta anterior, se observó una tendencia clara a valorar la complementariedad: los participantes destacan que cada una cumple una función distinta, lo que permite obtener un texto más claro, preciso y adecuado al contexto. Por ejemplo, un participante afirmó: "Creo que el uso secuencial de las herramientas es muy beneficioso al realizar la traduccion porque estas se centran sobre puntos distintos, y dan asi respuestas disintas que se pueden complementar para generar una traduccion mas verdadera" (E13). Así, la combinación de herramientas se percibe como una forma de mejorar el resultado final, ya que el pasaje del texto por ellas permite detectar errores, filtrar los problemas introducidos por las propias herramientas, corregir construcciones problemáticas y refinar la expresión.

Al mismo tiempo, sin embargo, varios estudiantes advirtieron sobre los posibles riesgos del empleo encadenado de herramientas. Por ejemplo, uno de los participantes explicó que: "el uso consecutivo de las herramientas me dio una ventaja a la hora de ir realizando las correcciones, porque tenía en cuenta las correcciones anteriores. Digamos que "iba mejorando". Sin embargo, resulta fundamental tener en cuenta que el texto va siendo manipulado cada vez más, y es seguro que haya diferencias a corregir con respecto al texto original (español). Más allá de la traducción, que a esa altura resulta perfecta, puede (probablemente) haber modificaciones indeseadas con respecto al significado original (español)." (E15). El participante E09, por su parte, señaló: "Siento que si se usa secuencialmente sin ningún tipo de corrección o intervención en el medio, la intención o cierta información del texto se pierde".

También aparece la idea de que procesar el texto con más de un sistema brinda mayor seguridad y confianza: "Suelo desconfiar dé traducciones de google o a veces de que tan correcto es chat gpt, el uso en conjunto me asegura que lo que esta escrito es entendible y me da la tranquilidad de saber que pasa por mas de un "filtro", me da mas seguridad a la hora de hacer una traducción" (E17). Además, algunos participantes mencionan como ventaja adicional la posibilidad de acceder a sinónimos o alternativas de expresión que, al surgir del uso combinado de diferentes tecnologías, amplían las opciones disponibles para su selección y facilitan una mayor adecuación del texto a los requerimientos discursivos del informe.

Por otro lado, se les preguntó qué plataformas (o combinaciones de ellas) recomendarían a un colega de su misma área que necesitara escribir un texto en inglés con fines profesionales, con el objetivo de relevar de forma indirecta qué herramientas consideran más adecuadas para este tipo de tarea. La opción más elegida fue recomendar el uso conjunto y secuencial propuesto en el experimento (60% de los encuestados), frente al uso individual de *ChatGPT* (40%), *Grammarly* (35%) y *Google Translate* (20%). Todos los participantes recomendaron al menos una de las herramientas utilizadas en la actividad. Además, tres personas (15%) brindaron respuestas alternativas que revelan un enfoque

flexible y estratégico: una de ellas sugirió combinar *Grammarly* y *ChatGPT* para generar un borrador y explorar opciones; otra propuso comparar resultados a partir de pequeños fragmentos traducidos por diferentes medios, y una tercera recomendó adaptar la secuencia de herramientas según el nivel de inglés de quien escribe, eligiendo entre las combinaciones *Google Translate-ChatGPT* o *Grammarly-ChatGPT*. Para estos participantes, el uso de recursos digitales para escribir y traducir es una práctica adaptable a cada situación y nivel del idioma de dominio de la LE.

Percepciones sobre los requerimientos para usarlas

Por otro lado, se consultó a los participantes si consideraban que la calidad del texto en español influía en la calidad de las traducciones generadas con herramientas digitales[31]. El 90% respondió que sí: 18 de los 20 encuestados creen que un texto en español bien escrito mejora los resultados de la traducción, mientras que solo el 5% consideró que no influye, y otro 5% expresó dudas. Esta amplia mayoría sugiere que se comparte una fuerte percepción de que el contenido original determina, en gran medida, la precisión, coherencia y fidelidad de la versión traducida, incluso si esta es realizada con la asistencia de recursos digitales. Al elaborar este punto, muchos destacaron que una escritura clara en español permite que las plataformas interpreten con mayor eficacia el contenido y produzcan traducciones más adecuadas. Una de las respuestas que permite introducir esta cuestión con mayor claridad es la de E00, quien señala: "Muchas personas del ámbito de las exactas no saben redactar correctamente, entonces al llevarlo a una traducción a otro idioma puede haber problemas para el entendimiento de los textos". Esta observación revela un nivel significativo de conciencia lingüística en torno a las particularidades de la escritura en el campo de las Ciencias Naturales. Como se detalló en la sección de Metodología, el texto fuente consistía en fragmentos de un informe de laboratorio en los que se conservaron errores del original y se incorporaron otros nuevos, con el objetivo de observar si las herramientas digitales, así como los propios participantes, lograban detectarlos y corregirlos. Si bien el análisis de las correcciones mostró que en varios casos los participantes no advirtieron algunos errores específicos o mantuvieron problemas de puntuación, las justificaciones analizadas en esta sección indican que sí fueron capaces de reconocer que el texto fuente presentaba errores y no estaba redactado de manera completamente adecuada. Las opiniones recogidas destacan que "Mientras peor este escrito en español, mas dificil va a ser psra la herramienta traducirlo ya que posiblemente exprese otra cosa que vos no estas queriendo decir" (E07). Varios participantes señalaron que frases mal construidas, uso inadecuado del vocabulario o expresiones redundantes tienden a trasladarse a la traducción. Otros

[31] Como se explicó en la sección de Metodología, se conservaron ciertos errores de puntuación y selección léxica del texto fuente para el experimento con el objetivo de observar su capacidad de detectarlos y traducirlos de forma correcta. La pregunta de la encuesta de percepción sobre la calidad del texto en español responde también a este respecto, ya que se buscó constatar si los participantes detectaron estos problemas en el texto fuente o no.

mencionaron que ciertos sistemas —como *ChatGPT*— tienen mayor capacidad para mejorar el texto fuente antes o durante la traducción, lo que podría mitigar, en parte, los efectos de un texto escrito de forma deficiente: "Desde luego que influye. Sin embargo, es posible que una herramienta como ChatGPT (y, tal vez, Grammarly) pueda detectar deficiencias en el texto original y corregirlas. Pero evidentemente el texto original (español) debe estar redactado lo mejor posible para asegurar la correcta traducción" (E15).

Percepciones sobre los resultados

Otra pregunta situó a los estudiantes en un escenario hipotético: se les planteó la idea de que los fragmentos traducidos correspondían a su propio informe de laboratorio y que debían presentarlo en inglés por motivos laborales. A partir de ello, se les consultó si consideraban que ese texto en inglés cumplía con los objetivos de comunicación profesional. El 100% de los encuestados respondió afirmativamente, lo que evidencia una percepción común de que la traducción obtenida mediante la secuencia propuesta es adecuada y funcional en contextos reales de uso profesional.

Por otro lado, se les consultó si se sentían más seguros al escribir o traducir un texto utilizando herramientas como las presentadas en la actividad o sin recurrir a ellas. El 85% indicó sentirse más seguros utilizándolas (17 participantes), mientras que el 15% manifestó experimentar el mismo nivel de seguridad independientemente de usar tecnologías de asistencia. Ningún participante eligió la opción que indicaba mayor seguridad sin dichas herramientas. Estos resultados sugieren que los sistemas digitales tienen un efecto positivo en la percepción de seguridad lingüística de quienes los utilizan, en relación con la escritura de este tipo de textos en LE (v. fig. 26).

Además, se buscó explorar la existencia de una posible relación con el nivel de inglés declarado antes del experimento, en particular, si quienes afirmaron sentirse igual de seguros sin utilizar herramientas digitales correspondían a niveles intermedio o avanzado. Sin embargo, se observó que los tres participantes en cuestión reportaron niveles diversos: A2, B2 y C1. Este hecho, sumado a la baja representatividad de niveles no intermedios en la muestra, impide sugerir una relación concluyente entre el nivel de inglés y la seguridad percibida en relación con el empleo de herramientas digitales para la realización de este tipo de tareas. No obstante, la indagación sobre esta posible asociación queda abierta para futuras investigaciones que cuenten con una muestra más amplia y equilibrada en términos de nivel de competencia lingüística en LE.

La encuesta finalizó con el pedido de que seleccionen cuál de todas las versiones de traducción obtenidas durante la actividad elegirían para publicar en un contexto profesional, y que justifiquen por

qué. Se les ofrecieron las 11 opciones correspondientes a las distintas etapas del proceso de traducción[32]. Se encontró que la opción más elegida fue V6 (9 menciones), es decir, la versión final que resultó de un proceso de múltiples correcciones automáticas e ediciones humanas. En segundo lugar, se ubicó V4 (6 menciones), producto de un trabajo combinado entre *ChatGPT* y la producción y revisión del propio participante. En conjunto, las versiones intervenidas por los participantes (V) representaron 19 de las 22 menciones totales (86%), lo que indica una clara preferencia por las traducciones que integran la acción humana con el apoyo de sistemas automáticos. Una de las justificaciones ofrecidas por los participantes para elegir V6 destaca "que es la que mejor traducida esta en conjunto con mi traduccion y aplicada las correcciones con las herramientas. Es la mas completa en cuestion de gramatica, cohesion y es la que mejor refleja la version en español" (E13). Otro participante explica que elegiría esta versión "porque es la más completa. Paso por varias correcciones" (E16). Estas valoraciones de "completitud" y "calidad acumulada", sugieren que los estudiantes confían en que un proceso iterativo contribuye a una mayor precisión en la traducción.

Por otro lado, las traducciones automáticas (las opciones T) fueron poco elegidas, y algunas —como T1, T3 y T4— no recibieron ningún voto. T2 fue elegida una vez y T5 fue elegida en dos ocasiones, lo que sugiere que, aunque se reconoce el impacto positivo de *Grammarly* o *ChatGPT*, su uso aislado no se percibe como suficiente para garantizar una versión final adecuada para contextos profesionales. De hecho, uno de los participantes afirmó que "No elegiría ninguna de las T porque no están chequeadas por mi, elegiría alguna de las V, pero recuerdo que varias me parecían igual de válidas" (E10).

Una observación relevante es que las versiones V5, T5 y V6 —todas derivadas de una misma ruta de traducción que inicia con *Google Translate* y culmina en V6— concentraron 13 de las 22 menciones totales (59%). Esto sugiere que dicho trayecto (o al menos ciertas instancias) fue considerado como el más efectivo por la mayoría. Estos resultados refuerzan la valoración del trabajo humano en la producción de textos de calidad profesional, aun cuando se empleen herramientas de asistencia lingüística, ya que estas se consideran un apoyo, pero no una solución definitiva.

[32] Cabe aclarar que, al tratarse de una pregunta abierta, algunos participantes mencionaron más de una versión como favorita. En estos casos, todas las opciones indicadas fueron tomadas en cuenta para el análisis, por lo que la cantidad de versiones elegida supera a la cantidad de participantes.

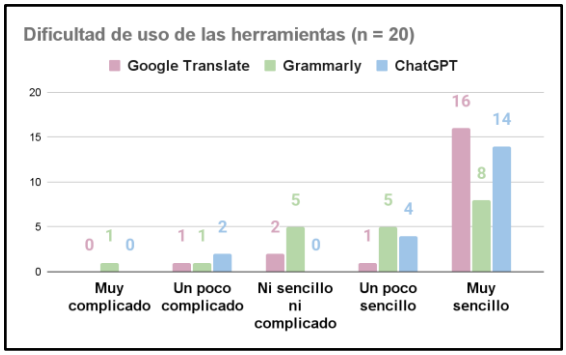

Figura 22

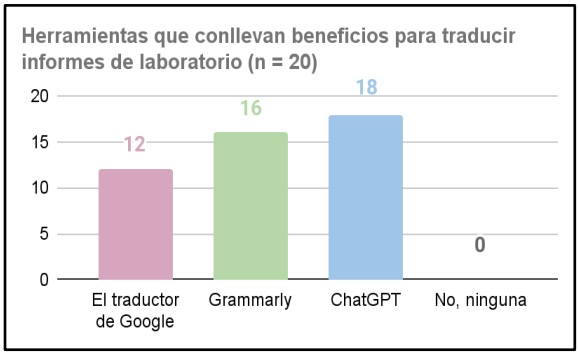

Figura 23

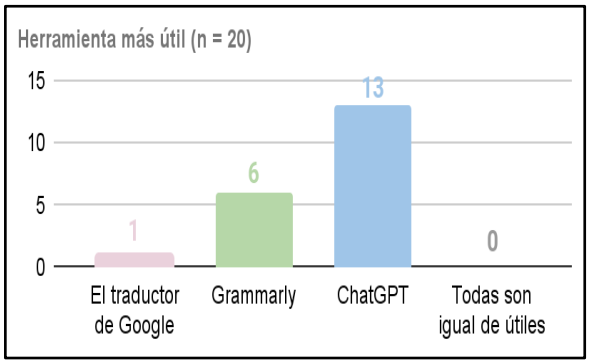

Figura 24

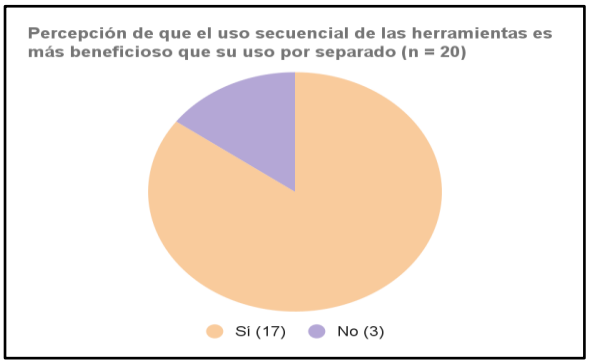

Figura 25

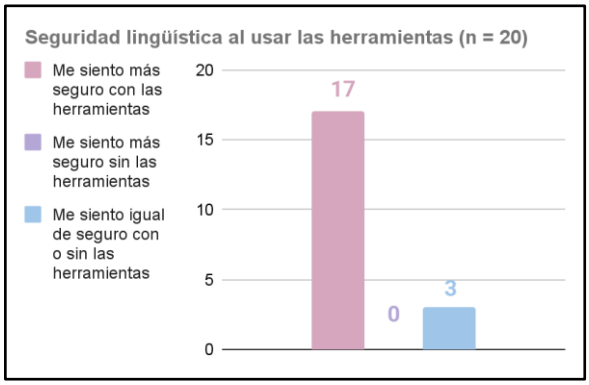

Figura 26

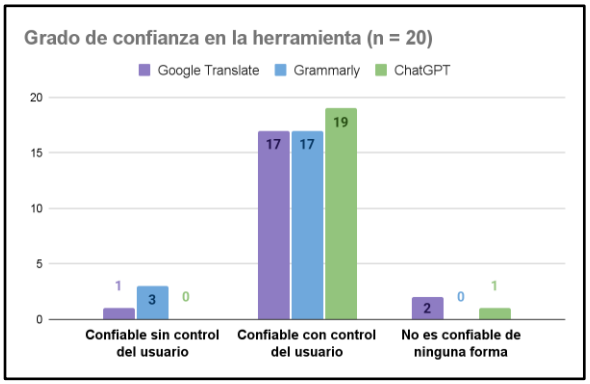

Figura 27

5.6. Juicios de expertos

Con el objetivo de obtener una evaluación experta de las producciones, se solicitó la colaboración de dos jueces, uno del área disciplinar, familiarizado con la escritura académica y profesional en inglés en este ámbito (juez 1), y una lingüística, esto es profesora y traductora de inglés con trayectoria docente universitaria (jueza 2).

Calificación de las versiones por parte de los jueces

Ambos evaluadores calificaron de manera independiente un subconjunto de textos correspondientes a cuatro versiones seleccionadas por su valor representativo dentro del diseño metodológico: la versión V1 (producción inicial autónoma (V1), sin tecnologías lingüísticas de asistencia, solo con un diccionario bilingüe inglés-español) y las versiones V3, V4 y V6, que

corresponden a las producciones resultantes de las tres principales secuencias del experimento. Estas son las rutas *Grammarly*, *ChatGPT*, y *completa* (que incluye la *Google Translate*). Las calificaciones fueron realizadas en una escala del 1 al 10 a partir de cuatro descriptores definidos en un formulario de evaluación: léxico, sintaxis, pragmática y discurso. Estos aspectos fueron seleccionados en consonancia con las competencias lingüísticas y pragmáticas definidas por el *MCER*. Según este marco, la competencia lingüística incluye elementos como la precisión léxica y la corrección gramatical, mientras que la competencia pragmática abarca tanto la adecuación funcional del texto a su propósito comunicativo como su coherencia discursiva. En concreto, a los efectos de esta evaluación, se entiende por léxico, la precisión de la selección de unidades léxicas en un texto científico en inglés; por sintaxis, la corrección de las estructuras gramaticales empleadas; por pragmática, la efectividad general del texto para cumplir su función comunicativa y por discurso, el ajuste al género discursivo de los informes científicos.

Los resultados obtenidos muestran diferencias marcadas entre las distintas versiones, que se corresponden con las herramientas utilizadas en cada ruta de producción textual (v. figs. 28-31). En términos generales, las versiones que integran recursos digitales combinados tienden a recibir puntuaciones más altas en todas las dimensiones evaluadas, en comparación con aquellas que se apoyan en recursos más limitados.

La versión V1, escrita solo con los conocimientos del participante y el diccionario bilingüe, obtuvo las calificaciones promedio más bajas en todas las categorías. En léxico, recibió un 1,95 por parte del juez 1 y un 2,76 por la jueza 2. En sintaxis, las puntuaciones fueron 2,40 y 2,57 respectivamente. En cuanto a la dimensión pragmática, ambos evaluadores otorgaron valores muy similares: 2,25 y 2,57. Por último, en discurso, el juez 1 asignó un puntaje de 2,35 y la jueza 2 de 2,65.

La versión V3, que incorpora la intervención de *Grammarly* sobre la versión inicial (V1), seguida de una revisión del participante, mostró un incremento leve pero consistente en las puntuaciones promedio. Para léxico, se asignaron 4,10 (juez 1) y 3,86 (jueza 2); en sintaxis, 3,86 y 4,43; en pragmática, 4,14 y 3,79; y en discurso, 3,86 y 4,06.

En cambio, la versión V4, donde se introduce la asistencia de *ChatGPT* antes de la revisión humana, fue evaluada de manera significativamente más favorable. Las puntuaciones en léxico fueron 6,76 (juez 1) y 7,48 (jueza 2); en sintaxis, 6,76 y 8,24; en pragmática, 6,95 y 7,59; y en discurso, 6,67 y 7,94.

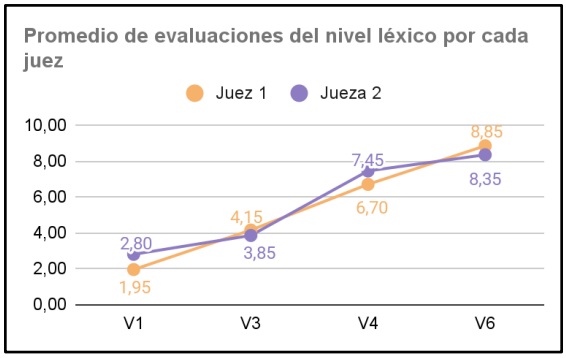

Figura 28

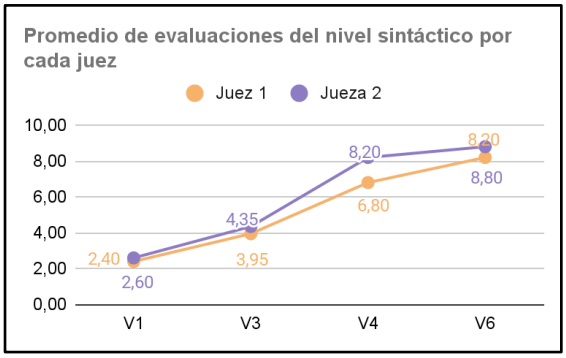

Figura 29

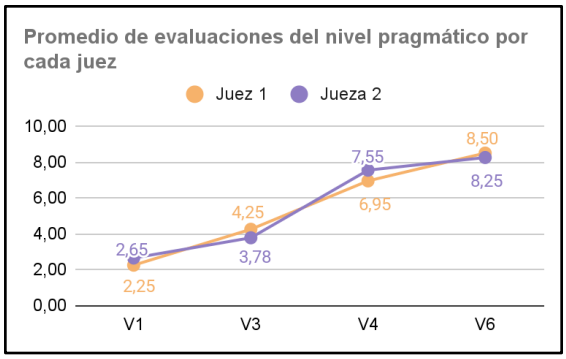

Figura 30

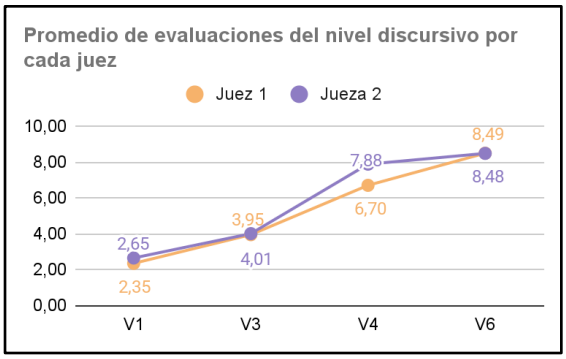

Figura 31

Por último, la versión V6, que combina el uso de *Google Translate*, *Grammarly*, *ChatGPT* y múltiples instancias de edición por los participantes, obtuvo las puntuaciones más altas en todas las dimensiones evaluadas. En léxico, recibió 8,86 (juez 1) y 8,33 (jueza 2); en sintaxis, 8,19 y 8,81; en pragmática, 8,48 y 8,24; y en discurso, 8,49 y 8,48.

Mejor versión en general

Por otro lado, se solicitó a los jueces que indicaran, por cada participante, cuál consideraban la versión más lograda para explorar la percepción global de la calidad textual. Los resultados muestran una preferencia por las versiones producidas con múltiples sistemas y edición humana. El juez 1 eligió la versión V6 en 19 de los 20 casos, y V4 en uno. Por su parte, la jueza 2 también manifestó una preferencia por la versión V6 (12 casos), seguida por la versión V4 (8 casos). Estos resultados sugieren que la combinación estratégica de las herramientas estudiadas, junto con la intervención humana en diferentes momentos del proceso, puede generar textos mejor evaluados por expertos tanto en dimensiones específicas como de forma global.

6. Integración de los datos

La integración de los distintos aspectos analizados posibilita una interpretación más profunda de los fenómenos lingüísticos observados. Este capítulo explora, entonces, cómo se interrelacionan los datos observados para ofrecer una visión más articulada y comprensiva del fenómeno estudiado.

6.1. Relación entre errores y rasgos típicos del género informe

Se analizó, en primer lugar, la relación entre los errores presentes en los textos y los dos rasgos distintivos del *género* informe estudiados —las nominalizaciones y la voz pasiva— para determinar qué proporción de los fallos gramaticales y textuales en cada versión de los experimentos puede vincularse con dichos rasgos.

Para ello, se calcularon los promedios de errores generales en cada texto en el corpus (gramaticales o textuales, según el caso) y los promedios de errores vinculados a nominalizaciones o voz pasiva. A partir de estos valores, se obtuvieron los porcentajes correspondientes mediante la siguiente fórmula:

$$Porcentaje = \left(\frac{Promedio\ de\ errores\ en\ rasgos\ del\ género\ informe}{Promedio\ de\ errores\ totales}\right) x\ 100$$

Este cálculo se aplicó por separado para los errores gramaticales en nominalizaciones y voz pasiva, y para los fallos textuales en nominalizaciones. Para facilitar la interpretación de los gráficos, es importante señalar que los porcentajes presentados en los gráficos de líneas (figs. 32, 36 y 40) fueron calculados a partir de los valores promedio de errores representados en los correspondientes gráficos de barras (fig. 33 a 35, 37 a 39, y 41 a 43). En estos últimos, las barras naranjas reflejan los promedios de errores vinculados a rasgos del género informe (es decir, el numerador en la fórmula), mientras que las barras azules indican los promedios de errores totales (el denominador).

Errores gramaticales totales y en nominalizaciones

En relación con los fallos gramaticales vinculados al uso de nominalizaciones, se observan comportamientos diferentes según el camino analizado (v. fig. 32). En particular, en la ruta *Grammarly*, los porcentajes de errores gramaticales que corresponden a fallos de nominalizaciones se incrementan: van desde un 13,96 % en V1 a un 20,70 % en T2 y un 18,51 % en V3 (v. fig. 33). En la vía *ChatGPT*, la tendencia es similar: el porcentaje parte de un 13,96% en V1, asciende a un 41,67 % en T3 y baja a un 22,55 % en V4 (v. fig. 34). En el caso de la ruta *Google Translate* y la *completa*, se observan porcentajes más bajos desde el comienzo: 5,22 % en T1, 8,14% en V2, 6,59% en T4, 5% en V5, 6,90% en T5 y 7,14% en V6. Estos valores iniciales se vinculan con los bajos promedios de errores gramaticales generales y en nominalizaciones observados en los textos. Además, la tendencia creciente

en los porcentajes de problemas gramaticales vinculados a nominalizaciones en las versiones T, así como su disminución en las versiones V, se vuelve a evidenciar en estas dos secuencias (v. fig. 35).

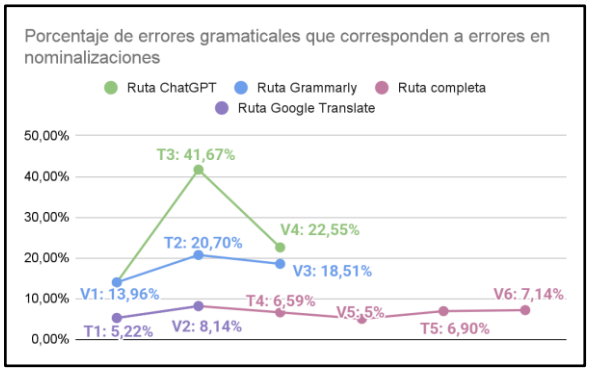

Figura 32

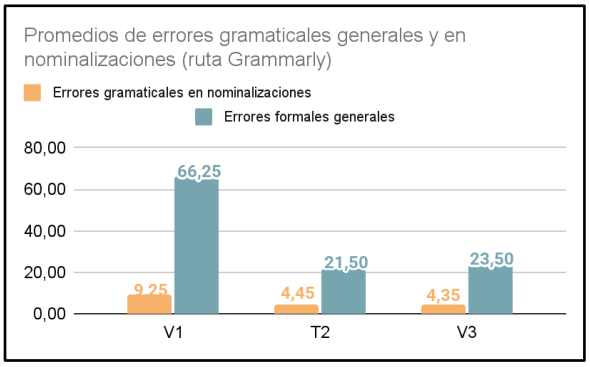

Figura 33

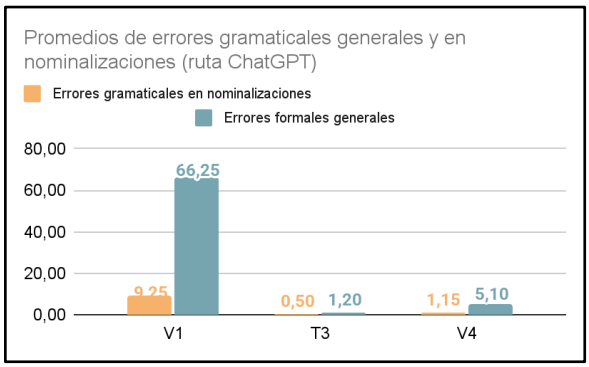

Figura 34

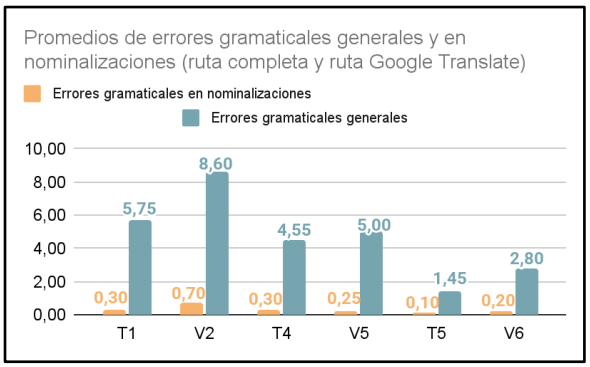

Figura 35

Estos aumentos porcentuales en las versiones T en todos los caminos no implican necesariamente un incremento en el número de nominalizaciones problemáticas, sino que pueden responder a una reducción más significativa de otros errores gramaticales generales, lo que modifica la proporción al bajar el denominador. Esta dinámica puede observarse también en el análisis de los errores gramaticales en relación con la voz pasiva y de los errores textuales en relación a las nominalizaciones. Asimismo, la disminución de estas proporciones en las versiones V puede estar más relacionada con la introducción o reintroducción de fallos gramaticales generales que con la corrección de los propios de las nominalizaciones, en consonancia con los comportamientos para cada tipo de error examinados en las secciones previas del análisis. Estos incrementos y disminuciones sugieren que, si bien las herramientas contribuyen a disminuir problemas gramaticales en general y los vinculados a nominalizaciones en particular, los problemas relacionados con este rasgo persisten en mayor medida, y adquieren así un peso relativo mayor dentro del conjunto.

Errores gramaticales totales y en voz pasiva

En cuanto a fallos gramaticales vinculados con el uso de la voz pasiva, también se evidencian comportamientos diferentes según la vía analizada, aunque con particularidades propias (v. fig. 36).

Se observa un incremento progresivo en los porcentajes: del 7,09% en V1, a un 9,53% en T2 y a un 8,51% en V3 en la ruta *Grammarly*. Esta tendencia, que refleja un mayor número de revisiones gramaticales generales que de correcciones vinculadas a rasgos propios del género informe, reproduce el patrón previamente identificado en relación con las nominalizaciones (v. fig. 37). En la vía *ChatGPT*, en cambio, el comportamiento se aparta del observado con anterioridad. A pesar de partir del mismo valor inicial (7,09%), el porcentaje de fallos gramaticales de voz pasiva desciende de forma notable a 0 en T3 y luego asciende a 3,92% en V4. La ausencia total de este tipo de problemas en la versión generada por el sistema sugiere gran efectividad de *ChatGPT* para corregir construcciones incorrectas de voz pasiva, en contraste con los resultados observados en el caso de las nominalizaciones, donde persistían ciertos problemas (v. fig. 38). Por su parte, en el trayecto *Google Translate* y en la ruta *completa* los porcentajes de errores gramaticales que corresponden a fallos de voz pasiva son bajos desde el comienzo: 6,09% en T1 y 3,49% en V2. Esta tendencia de valores reducidos iniciales también fue observada en el caso de las nominalizaciones. En este caso, el porcentaje baja mínimamente en T4 (3,30%) y vuelve a aumentar en V5 (6%). Por último, en T5, el porcentaje vuelve a descender a 0, manteniéndose en ese valor en V6, lo que refuerza la hipótesis de que *ChatGPT* es muy eficaz para eliminar problemas de voz pasiva (v. fig. 39).

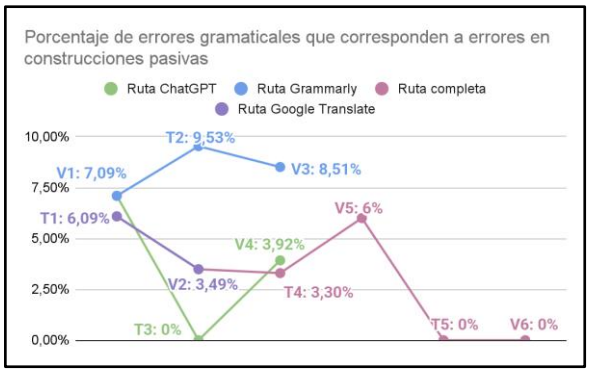

Figura 36

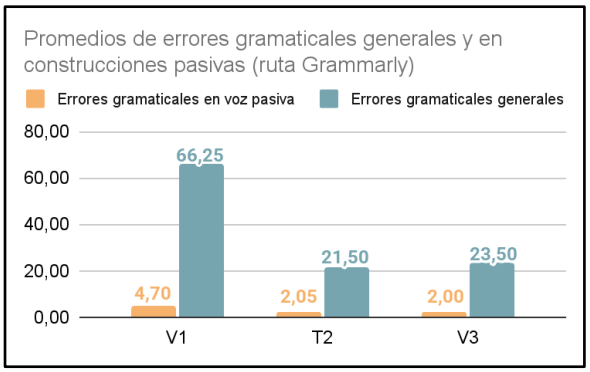

Figura 37

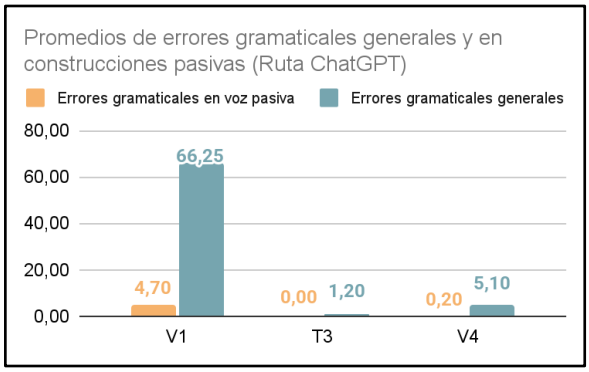

Figura 38

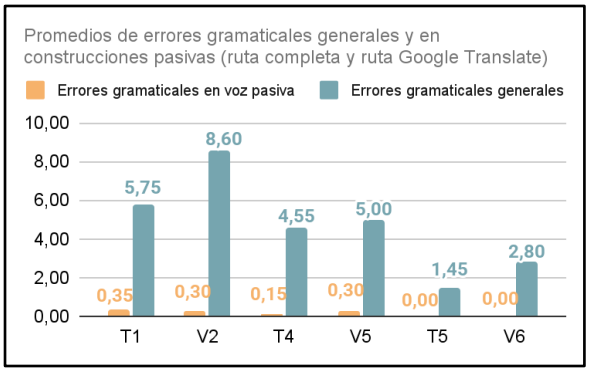

Figura 39

Los resultados permiten observar que, si bien los errores gramaticales vinculados con fenómenos léxico-sintácticos característicos del género representan una proporción relativamente baja del total de fallos registrados en las versiones iniciales de cada recorrido, su presencia adquiere un peso mayor en

términos relativos en las versiones intermedias y finales. Por ejemplo, en el caso de T1, 0,30 errores promedio de nominalizaciones en 5,75 errores promedio generales tiene un peso proporcional menor que 0,20 errores promedio de nominalizaciones en 2,80 errores promedio generales, en V6. Esta tendencia sugiere que los problemas relacionados con estos rasgos del género no constituyen el núcleo principal de las dificultades gramaticales de los estudiantes en las primeras etapas del proceso de traducción, pero sí parecen ser más resistentes a la corrección en comparación con otros tipos de problemas. Al mismo tiempo, los errores gramaticales en voz pasiva parecen ser menos resistentes a la corrección que las nominalizaciones por parte de *ChatGPT*. Esto sugiere que algunos rasgos del discurso académico pueden ser más fácilmente identificables y corregibles en el proceso de traducción asistida por tecnología. Por otro lado, en lo que refiere a las otras secuencias, se vuelve a observar que los errores gramaticales en este fenómeno particular presentan una mayor resistencia a la corrección.

Errores textuales totales y en nominalizaciones

Por último, en relación a los errores textuales que se corresponden con el uso inadecuado de nominalizaciones, se observa que la variación entre porcentajes en todas las versiones es mucho menor que en los gramaticales. Esto se vincula con el análisis realizado de fallos textuales generales y en nominalizaciones, en el que se pudo observar que estos presentan una mayor resistencia a la corrección tanto por parte de las herramientas como de los participantes (v. fig. 40).

En la vía *Grammarly*, los porcentajes muestran un leve aumento, del 17,09% en V1 al 21,43% en T2 y al 21,87% en V3. La proporción creciente de este tipo de errores reproduce el patrón observado para la relación entre fallos gramaticales generales y gramaticales en nominalizaciones, ya que se evidencia una mayor corrección de errores textuales generales que de los vinculados a este rasgo del género (v. fig. 41). El recorrido *ChatGPT* presenta el mismo patrón de aumento, pero con saltos más notables entre los porcentajes: parte del mismo valor inicial en V1, y se observa un aumento al 62,69% en T3, seguido de un descenso a 50,56% en V4. En este caso, la mayor corrección de errores textuales generales en comparación con la de fallos en nominalizaciones es incluso más marcada (v. fig. 42). En la ruta *Google Translate*, la proporción de errores textuales en nominalizaciones se mantiene muy estable: 16,26% en T1 y 17,46% en V2, y continúa de la misma manera en la ruta *completa*: T4 (19,01%) y V5 (20,66%). Sin embargo, vuelve a presentarse un aumento moderado en el porcentaje de fallos textuales correspondientes a nominalizaciones luego de la corrección realizada por *ChatGPT*, en T5 (25,71%). Como muestra el gráfico, el *chatbot* reduce significativamente los problemas textuales generales, mientras que la reducción en nominalizaciones resulta menor, aunque eficaz. El porcentaje vuelve a descender en V6 (21,43%), lo que se presenta

como un aumento general de los errores textuales en esta versión, debido a las operaciones de recuperación observadas en el análisis previo (v. fig. 43).

Los datos revelan que estos fallos en nominalizaciones constituyen una proporción significativa dentro del conjunto de los errores textuales, en especial si se los compara con los gramaticales. Sin embargo, el progresivo aumento observado en todas las vías revela que, así como sucedía para los gramaticales, los asociados a las nominalizaciones presentan una mayor resistencia a la corrección que otros fallos textuales.

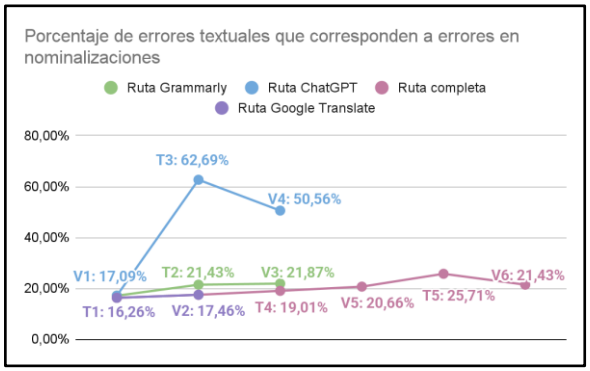

Figura 40

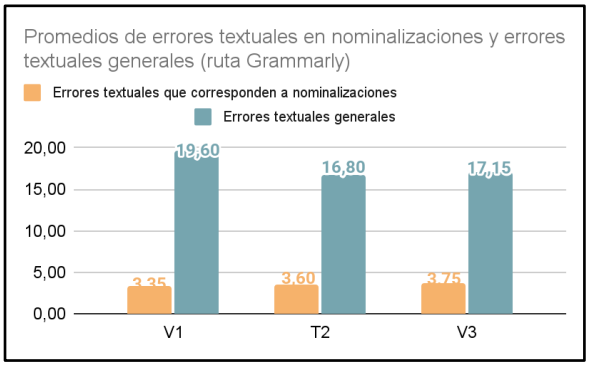

Figura 41

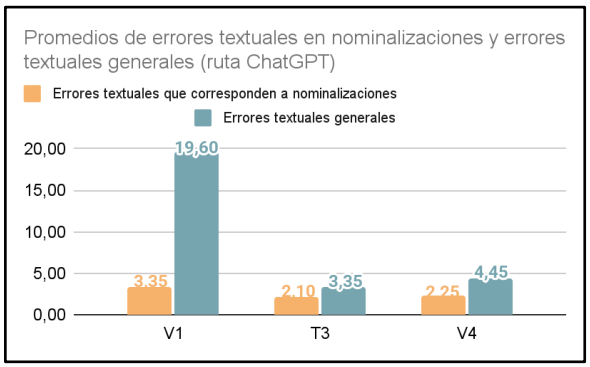

Figura 42

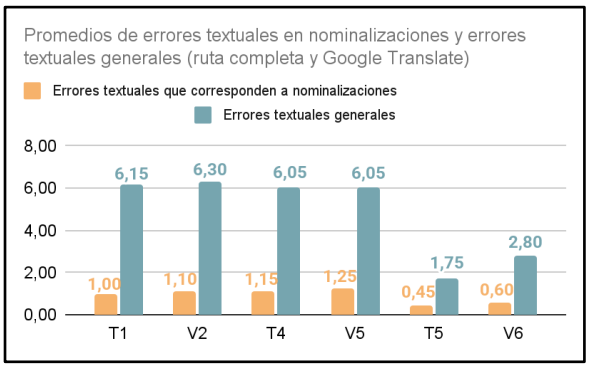

Figura 43

6.2. Relación entre errores e intervenciones de los participantes

La segunda relación entre los datos del experimento que se buscó explorar fue la influencia que las intervenciones realizadas por los participantes habría tenido en el aumento registrado de errores gramaticales y textuales promedio de las versiones V. Este análisis permitió aproximar en qué proporción sus modificaciones generaron o reintrodujeron fallos de ambos tipos y, por lo tanto, el nivel de incidencia de las transformaciones de los participantes en la corrección de los textos.

Errores gramaticales e intervenciones

En primer lugar, se observa que, tras la alta reducción de errores gramaticales que logra *Grammarly* en T2 (de 62,55 a 19,9 en promedio), estos aumentan levemente en V3 a 21,85 tras 6,15 ediciones humanas promedio. Los cambios son más numerosos que los fallos introducidos, lo que sugiere que muchas de las modificaciones no afectaron la calidad gramatical. A partir del número de errores

generados y la cantidad de intervenciones, es posible afirmar que alrededor del 31% de ellas generaron fallos gramaticales en este recorrido, mientras que el resto no lo hizo (v. fig. 44).

En lo que respecta a la vía *ChatGPT*, luego de la significativa reducción de problemas gramaticales lograda por la herramienta en T3 (de 62,55 a 1,75), estos vuelven a subir a 5,45 a partir de las 9,1 modificaciones promedio de los participantes en V4. En este sentido, alrededor del 40% de las ediciones en esta ruta generan o reintroducen errores gramaticales. No obstante, este patrón presenta casos extremos: en E03 y E19 se registran aumentos significativos tanto de modificaciones como de fallos (E03: 62 cambios y 28 errores; E19: 25 cambios y 32 errores). Estos valores afectan la media y podrían exagerar la percepción de que las intervenciones generan muchos fallos, ya que no son la norma.

Otro aspecto a destacar es que, pese a la eficacia de *ChatGPT* para reducir errores, los participantes realizan más ediciones que en el camino *Grammarly*. Esta tendencia podría reflejar una mayor desconfianza hacia el sistema, que parece no estar relacionada con su desempeño objetivo sino con su modo de operación. En las preguntas abiertas de la encuesta, varios participantes manifestaron que *ChatGPT* no solo corrige, sino que reescribe en su totalidad las traducciones, y que se aleja del estilo o de las decisiones tomadas por el participante. Esta sensación de "pérdida de control" sobre el texto podría explicar por qué, a pesar de la baja cantidad de errores en T3, los participantes realizan numerosas modificaciones. A este factor se suma que *ChatGPT* es la herramienta más nueva entre las evaluadas, lo que también podría generar cierta resistencia o escepticismo por parte de los usuarios, que todavía no confían (plenamente) en ella, a diferencia de *Grammarly*, que, si bien la mayoría de los participantes no la conocía antes de la actividad, presenta una interfaz y un tipo de correcciones que resultan más familiares. Sistemas como *Microsoft Word*, que son muy conocidas y utilizadas, utilizan sistemas de sugerencias y marcas similares a las de *Grammarly*, lo que podría haber facilitado su aceptación y disminuido la percepción de extrañeza frente a sus correcciones. En cambio, *ChatGPT* no solo es menos familiar, sino que además interviene de una forma menos controlable (v. fig. 45).

En el caso de *Google Translate* (v. fig. 46), la cantidad de errores promedio se duplica al pasar de la versión automática (T1: 4,25) a la versión intervenida (V2: 8,3). A su vez, la cantidad promedio de ediciones realizadas por los participantes es de 16,7, es decir, casi cuatro veces la cantidad promedio de fallos que presentaba T1. Este dato permite realizar una estimación aproximada sobre el impacto de las modificaciones en la introducción de nuevos errores. Si se considera que estos aumentan en promedio 4,05, y que se realizaron alrededor de 16,7 intervenciones por versión, se puede inferir que cerca de un 24% de ellas generaron un nuevo error gramatical. Cabe destacar que, en casi todos los experimentos, las modificaciones introducidas no redujeron el número de fallos originales, sino que los incrementaron. Solo en el caso de E02 se observa una leve reducción a partir de la intervención sobre la T1 (de 5 a 4).

Este fenómeno puede interpretarse como una combinación de varios factores. Por un lado, es posible que el nivel intermedio de competencia en inglés (B1 y B2) de los participantes no sea suficiente para detectar de manera sistemática los fallos presentes en la versión automática o los introducidos durante su intervención. Así, merece destacarse que las ediciones realizadas en V2 son significativamente mayores que las realizadas en V3 (ruta *Grammarly*). Esto resulta llamativo si se tiene en cuenta que las versiones corregidas por *Grammarly* presentaban más del doble de errores gramaticales que las traducciones de *Google Translate*. Esta aparente contradicción puede explicarse, en parte, por la confianza diferencial que los participantes depositan en cada sistema y en sus propias traducciones. En el caso de *Google Translate*, se parte de una traducción automática, mientras que *Grammarly* actúa sobre una traducción inicial realizada por los propios participantes (V1). Esta diferencia, junto con las percepciones previamente señaladas por los participantes, parece haber influido en la intensidad de las modificaciones.

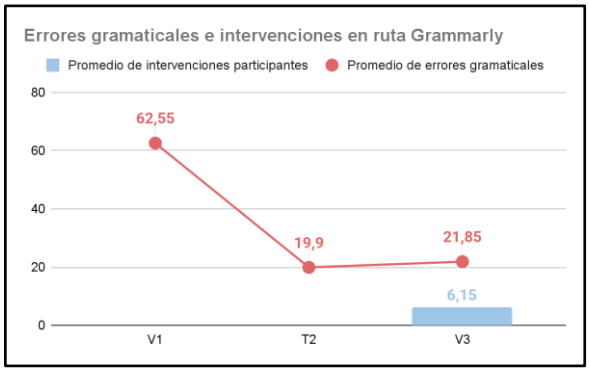

Figura 44

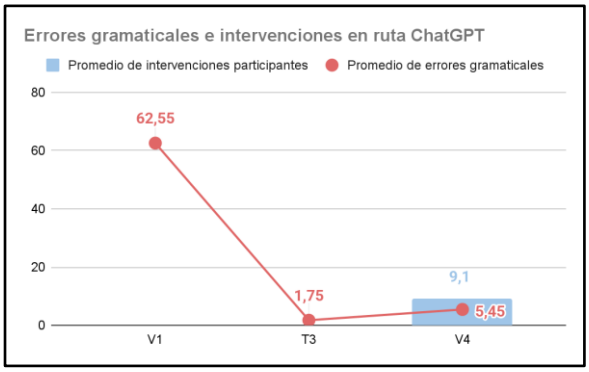

Figura 45

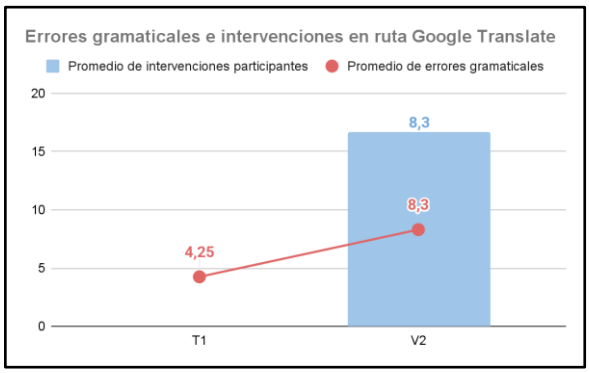

Figura 46

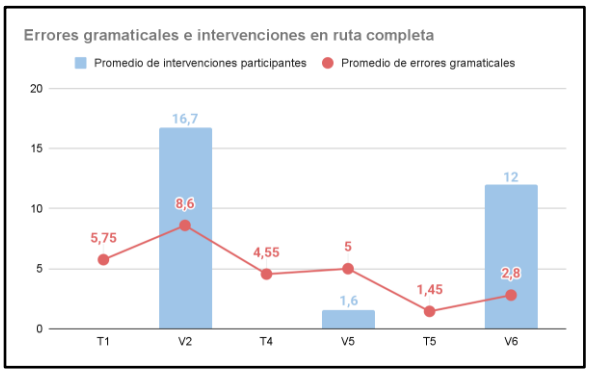

Figura 47

En el caso del resto de la ruta *completa*, que incluye la de *Google Translate*, se puede observar que tras la reducción de errores producida por *Grammarly* en T4 (de 8,6 a 4,55), estos vuelven a aumentar de forma leve en V5 (a 5 en promedio). Esta variación mínima es consistente con la cantidad promedio de intervenciones observadas en V5 (1,6), lo que representa que aproximadamente un 28% tuvo un impacto negativo. No obstante, dado que el aumento en la cantidad de fallos es inferior a medio error por versión, este porcentaje debe interpretarse con cautela, ya que se basa en cifras muy bajas. Por último, luego de la significativa disminución de problemas gramaticales lograda por *ChatGPT* (1,45 en

T5), estos aumentan levemente en V6 (2,8). En este último caso, puede observarse que el porcentaje de ediciones promedio que generan errores es mínimo, ya que de las 12 modificaciones promedio registradas, solo un 11% produce nuevos fallos gramaticales (v. fig. 47).

Errores textuales e intervenciones

En el caso de los errores textuales, el recorrido *Grammarly* (v. fig. 48) exhibe un patrón recurrente de una leve reducción de fallos seguida de un pequeño incremento. En términos cuantitativos, el promedio pasa de ser 15,05 en V1 a 14,4 en T2 luego de las ediciones del corrector automático. Tras las 6,15 ediciones promedio, este valor aumenta a 14,65 en V3. A partir de estos datos, puede estimarse un aumento porcentual de errores cercano al 3,9%, lo que sugiere que la mayor parte de ellas no afecta negativamente los aspectos textuales[33].

En la vía *ChatGPT*, se observa una reducción significativa en la cantidad de errores textuales tras el uso del *chatbot*. El promedio pasa de 15,05 en V1 a 3,35 en T3. Aunque no tan pronunciada como la registrada en los problemas gramaticales, esta reducción continúa siendo considerable. Luego de las ediciones humanas en V4, el número de fallos asciende a 4,2. Alrededor del 9,3% de las modificaciones en esta etapa provocaron nuevos problemas textuales (0,85 errores nuevos / 9,1 intervenciones). Como se mencionó, los participantes tienden a intervenir más en esta ruta que en la de *Grammarly*. Es posible que esto se deba a que *ChatGPT* genera una sensación de pérdida de control sobre el contenido. Sin embargo, a diferencia de lo que ocurre con los errores gramaticales, el impacto negativo de estas modificaciones sobre la calidad textual es menor. No obstante, los datos revelan algunos casos extremos en los que la cantidad de fallos aumenta de forma más considerable, con mayor un número de intervenciones. Por ejemplo, el participante E03 pasa de 4 errores en T3 a 9 en V4, con un total de 62 cambios, mientras que E19 pasa de 4 a 12 errores, con 25 cambios. Si bien estos valores son atípicos, muestran que, en ciertos casos, los participantes parecen haber reincorporado problemas previos o introducido nuevos al intentar recuperar fragmentos que percibían como más fieles a su versión original, sin advertir que estos contenían fallos textuales. En líneas generales, la mejora observada tras el uso de *ChatGPT* parece estar asociada con su capacidad para interpretar mejor el contexto, identificar términos especializados y reformular pasajes con mayores ajustes a nivel textual. (v. fig. 49).

En la secuencia *Google Translate* (figura 50) los fallos textuales aumentan más notablemente que en los otros tramos de análisis. Sin embargo, el promedio sigue siendo bajo (2,85). En términos proporcionales, la relación entre intervenciones y nuevos errores textuales (17,06% aproximadamente) es menor que en la dimensión gramatical, lo que sugiere que, aunque el número de modificaciones es

[33] Este valor, así como el porcentaje estimado para la vía *ChatGPT*, deben ser tomados con precaución, ya que los incrementos de errores textuales luego de la intervención humana responden a menos de un error promedio.

elevado (16,7), solo una fracción reducida de estas genera fallos de este tipo. Esto puede indicar que muchas modificaciones responden a otras necesidades más allá de la corrección de errores. Si se considera que el promedio combinado de fallos (gramaticales + textuales) es de alrededor de 11, pero las ediciones promedio superan las 16, se infiere que los participantes intervinieron más allá de los errores detectados. Varias razones podrían explicar este comportamiento: la percepción de que las traducciones automáticas son poco confiables, aun cuando el texto no contenga problemas graves; la búsqueda de un estilo más personal; o una tendencia general a intervenir como parte del cumplimiento de la actividad experimental. En conjunto, aunque *Google Translate* produce textos con fallos limitados, la herramienta no parece generar suficiente confianza en los usuarios como para evitar una modificación extensa de los textos.

En lo que refiere a la ruta *completa* (v. fig. 51), la intervención de *Grammarly* (T4) produce una leve reducción en el número de errores, que pasa de 5,05 a 4,85. Tras las intervenciones de los participantes en V5, el promedio desciende apenas a 4,75. Este pequeño cambio se acompaña de una cantidad promedio muy baja de ediciones (n=1,6), lo que es consistente con lo observado en la secuencia *Grammarly*, donde los participantes tendían a intervenir poco. El mayor descenso en la cantidad de fallos textuales se produce en T5, tras el uso de *ChatGPT*, donde estos bajan en promedio a 2,25. No obstante, en V6 se observa un aumento leve a 2,9, acompañado por un incremento considerable en la cantidad de modificaciones (n=12).

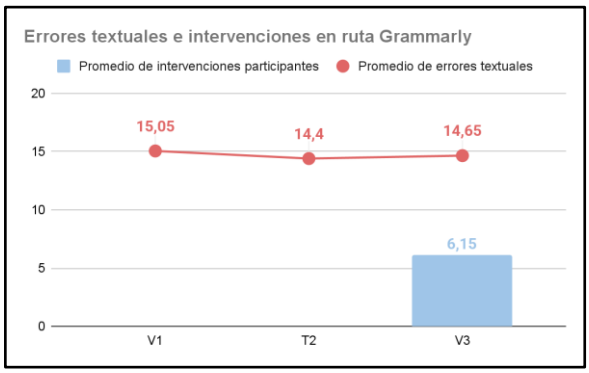

Figura 48

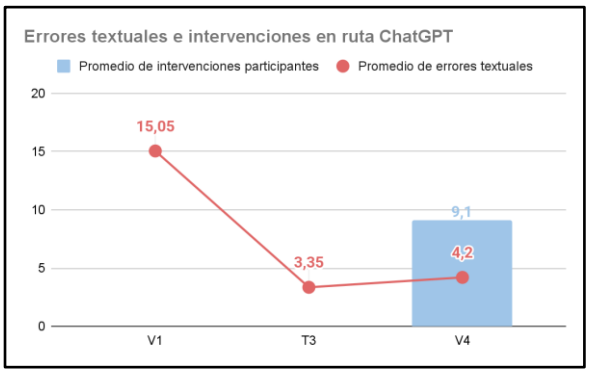

Figura 49

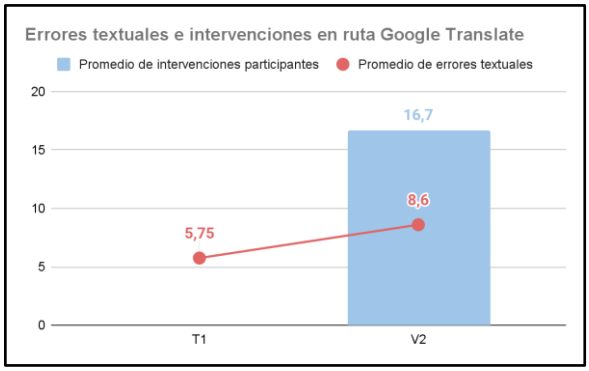

Figura 50

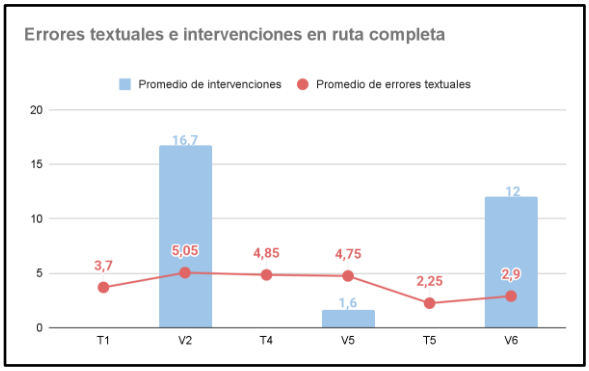

Figura 51

6.3. Relación entre errores gramaticales y perfil de vocabulario

Este análisis busca establecer una relación entre las variaciones del perfil de vocabulario (calculado a partir del *LFP*) en cada versión y el número promedio de errores fonemáticos que estas presentan. La elección de este enfoque se debe a que los problemas fonemáticos incluyen, en su mayoría, fallos ortográficos (aunque también abarcan cuestiones como el uso incorrecto de mayúsculas, tildes, puntuación, abreviaturas, entre otros). El *LFP* solo clasifica dentro de la *GSL* y la *AWL* aquellas unidades léxicas que no presentan errores ortográficos. Si bien la categoría *Off-list* también incluye unidades léxicas escritas correctamente que no pertenecen a ninguna de las dos listas, se parte de la hipótesis de que muchos elementos que, en las primeras versiones, terminan en la categoría *Off-list*, lo hacen por estar mal escritos, lo que impide que la calculadora del *LFP* las reconozca. Por ello, se busca analizar en qué medida la disminución de errores ortográficos afecta el comportamiento de los puntajes del perfil de vocabulario en las sucesivas versiones, en cada camino.

En el caso de la ruta *Grammarly*, se observa que el promedio de unidades léxicas en la *Off-list* es del 21,31% en V1, mientras que el promedio de fallos fonemáticos alcanza los 41,95. En cuanto a las listas de frecuencia, el porcentaje del léxico clasificado dentro de la *AWL* se sitúa en 6,28% y dentro de la *GSL*, en 72,15%. Tras la corrección de la herramienta, el promedio de elementos en la *Off-list* baja a 17,28% y los errores fonemáticos descienden a 11,25 en T2. Es posible que esta importante reducción haya permitido que varias unidades léxicas previamente mal escritas —y por tanto clasificadas como *Off-list*— pasaran a ser reconocidas dentro de las listas académicas o de alta frecuencia. Este cambio se ve respaldado por un aumento en el porcentaje del léxico clasificado dentro de la *AWL* (que sube a 7,91%) y dentro de la *GSL* (que sube a 74,77%). Sin embargo, cabe señalar que la disminución de errores fonemáticos (una reducción del 73,2%) es mucho más pronunciada que los cambios en las listas léxicas (que fluctuaron entre 1,63 y 2,62 puntos porcentuales), lo que indica que la relación entre estos factores no es estrictamente proporcional. En V3, se observa una leve inversión de esta tendencia: el promedio de elementos en la *Off-list* sube ligeramente a 17,47% y los fallos fonemáticos aumentan a 12,8. Esta reintroducción o generación de errores podría haber contribuido a esta variación. Además, el porcentaje de unidades léxicas clasificadas dentro de la *AWL* desciende levemente a 7,81% y el de la *GSL* baja a 74,68%, lo que refuerza parcialmente esta hipótesis (v. fig. 52).

Por su parte, en la secuencia *ChatGPT*, se observa que el promedio de elementos en la *Off-list* es del 21,31% en V1, mientras que los fallos fonemáticos alcanzan un promedio de 41,95. El porcentaje de unidades clasificadas dentro de la *AWL* se sitúa en 6,28% y dentro de la *GSL*, en 72,15%. En T3, tras la edición del *chatbot*, el promedio de ítems en la *Off-list* desciende a 19,85% y los errores fonemáticos bajan de forma considerable a 0,95. Al igual que con *Grammarly*, es posible que esta

importante corrección haya permitido que unidades léxicas antes mal escritas y, por ende, fuera de las listas, pasaran a ser reconocidas dentro de la *AWL* y la *GSL*. Este efecto se refleja en el marcado incremento del porcentaje de vocabulario clasificado dentro de la *AWL*, que sube a 11,33% —más alto que en el camino *Grammarly*—. Sin embargo, el porcentaje de elementos clasificados dentro de la *GSL* desciende a 68,81%, lo que podría indicar que algunas unidades léxicas comunes fueron sustituidas por variantes de registro más académico. Esto también podría explicar el aumento más pronunciado en la *AWL*. Cabe destacar que la disminución de errores fonemáticos en esta etapa (una reducción del 97,7%) es mucho mayor que las fluctuaciones observadas en las listas léxicas (que varían entre 3,34 y 5,05 puntos porcentuales), lo que sugiere, una vez más, que la relación entre estos factores no es estrictamente proporcional. En V4, se registra que el promedio de ítems en la *Off-list* sube de forma leve a 19,98% y los fallos fonemáticos aumentan a 3,2. La reintroducción o generación de nuevos problemas podría haber influido en este cambio. Por otro lado, el porcentaje de léxico clasificado dentro de la *AWL* desciende levemente a 10,85% y el de la *GSL* también baja a 68,74%, lo que refuerza el significado provisional de esta interpretación (v. fig. 53).

En la ruta *Google Translate*, se observa que el promedio de unidades léxicas clasificadas dentro de la *Off-list* es del 16,34% en T1, mientras que los errores fonemáticos alcanzan un promedio de 5,15 casos. El porcentaje de elementos clasificados dentro de la *AWL* se sitúa en 7,98% y dentro de la *GSL*, en 75,65%. Luego de la intervención de los participantes, se registra un leve aumento tanto en el porcentaje de ítems en la *Off-list*, que alcanza el 16,92% en V2, como en la cantidad de fallos fonemáticos (6,85). Es posible que algunas unidades antes reconocidas dentro de la *AWL* o de la *GSL* pasaran a la *Off-list* por estar mal escritas. Esta hipótesis parece respaldarse en la leve disminución del porcentaje de elementos clasificados dentro de la *AWL* (7,80%) y de la *GSL* (75,27%). A diferencia de las vías anteriores, en este caso el aumento de errores fonemáticos no supera ampliamente las fluctuaciones en las listas léxicas, lo que sugiere una relación más equilibrada entre ambos fenómenos.

En T4, versión correspondiente a la secuencia *completa* (que continúa la intervención tras la etapa con *Google Translate*), se observa una reducción tanto en el porcentaje de ítems fuera de lista (16,75%) como en los fallos fonemáticos (4,05). Este patrón es similar al de las rutas *Grammarly* y *ChatGPT*: la disminución de errores parece facilitar el reconocimiento de unidades léxicas mal escritas con anterioridad y favorecer su inclusión en las listas léxicas, lo que se evidencia en el aumento del porcentaje de elementos clasificados dentro de la *AWL* (7,81%) y de la *GSL* (75,41%). En V5, los valores se estabilizan: el porcentaje de ítems fuera de lista desciende apenas a 16,74%, mientras que el porcentaje clasificado dentro de la *AWL* sube de forma leve a 7,82% y el de la *GSL* a 75,42%. Los fallos fonemáticos aumentan ligeramente (4,25), pero sin alterar de forma significativa los porcentajes de las listas. En T5, sin embargo, se produce una disminución considerable de problemas fonemáticos (de

4,25 a 1,45), acompañada de un crecimiento marcado de elementos en la *Off-list* (hasta 20,17%). Aunque los errores disminuyen, este fenómeno, también observado en la vía *ChatGPT*, podría deberse a que la herramienta utilizada introduce vocabulario que no forma parte de ninguna de las listas (*AWL* o *GSL*), lo que lleva a que dichas unidades sean clasificadas como *Off-list*. Esta hipótesis se ve respaldada por el aumento del porcentaje de ítems en la *AWL* (11,42%) y una baja sustancial en los que se encuentran en la *GSL* (68,40%), lo que sugiere una sustitución de vocabulario frecuente por términos de registro más académico o disciplinarmente específico. En esta etapa, los fallos fonemáticos parecen tener menor incidencia en las fluctuaciones de las listas. Es posible que esto se deba a que la mayor parte de los errores ortográficos ya fueron corregidos en fases anteriores, y los remanentes estén más vinculados a cuestiones menos relevantes para las listas, como el uso de mayúsculas o abreviaturas. En V6, etapa final de la secuencia *completa*, los fallos fonemáticos aumentan de forma leve (2,4), pero esta suba no se refleja en un incremento de los elementos de la *Off-list* que, por el contrario, disminuyen a 19,47%. A su vez, el porcentaje del léxico dentro de la *GSL* sube a 69,75% y el de la *AWL* baja a 10,74%. Esto refuerza la idea de que, en esta instancia, las variaciones en las listas léxicas están determinadas por factores distintos de los errores fonemáticos (v. fig. 54).

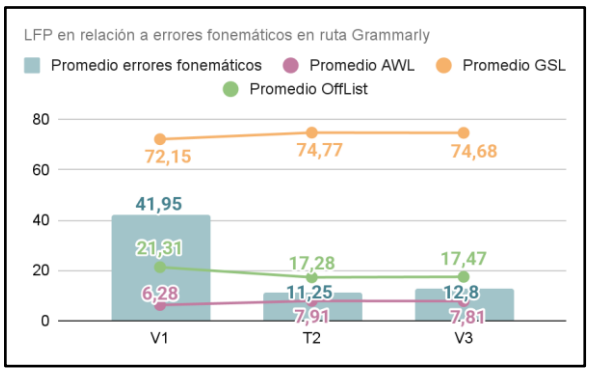

Figura 52

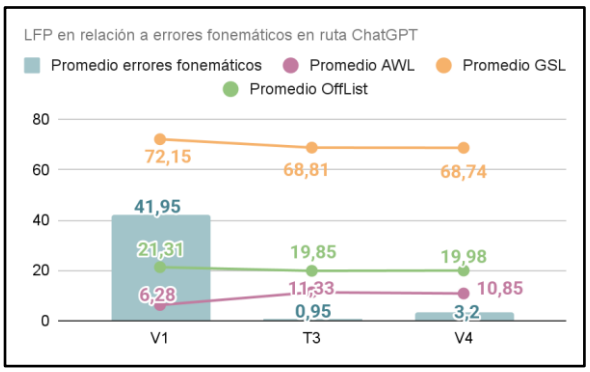

Figura 53

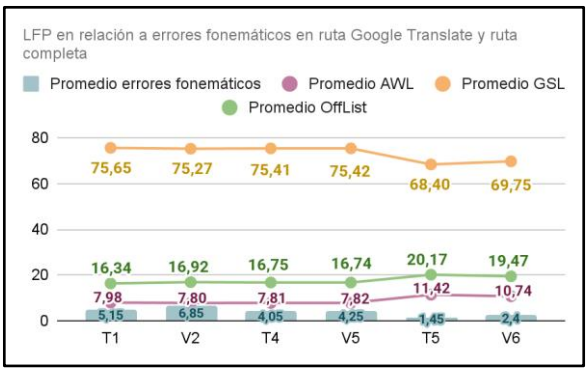

Figura 54

En síntesis, los datos analizados sugieren que la disminución de errores fonemáticos influye en las fluctuaciones del perfil léxico hasta cierto punto, en especial en etapas iniciales (V1), donde la incidencia de problemas ortográficos es alta. A medida que estos son corregidos, se observa una progresiva reubicación de unidades léxicas desde la *Off-list* hacia la *GSL* y la *AWL*. En etapas más

avanzadas (V5 o V6 en la vía *completa*), la incidencia de estos errores sobre el perfil de vocabulario se reduce de forma considerable, ya que se encuentran corregidos la mayoría de los fallos ortográficos, y las fluctuaciones en las listas parecen responder a otros factores, como las elecciones específicas de vocabulario por parte de las plataformas o los participantes. Además, en las secuencias *Grammarly* y *ChatGPT*, la cantidad de fluctuaciones en las listas fue mucho menor que la cantidad de errores corregidos entre V1 y T2 o T3. Esta disparidad sugiere que la incidencia de los fallos fonemáticos en las fluctuaciones de las listas es menos significativa de lo que podría suponerse.

6.4. Relación entre juicios de expertos y otros datos

Esta sección analiza el vínculo entre los juicios realizados por el panel de expertos y dos dimensiones clave: por un lado, la presencia y el tipo de errores detectados en las producciones escritas de los participantes, y por otro, las percepciones declaradas por los propios participantes respecto al uso de los recursos digitales. En particular, se considera la relación entre la consulta realizada tanto a participantes como a jueces que conlleva la elección de una versión preferida entre un grupo de opciones dadas.

Relación de juicios de expertos y errores

En esta sección se analizan asociaciones entre distintos tipos de errores y las dimensiones evaluadas por los jueces, para observar cómo varían las calificaciones asignadas a las distintas versiones del texto en relación con la frecuencia de ciertas categorías de fallos.

En primer lugar, se focaliza en el grupo de problemas vinculados a la intencionalidad, la informatividad y la intertextualidad —que comprenden fallas en la selección del léxico y la alteración, omisión o sobrecarga de información respecto al texto fuente— y su posible impacto en las evaluaciones de los jueces en las dimensiones de léxico y pragmática (v. figs. 55-56). Aunque no se establecerá aquí una correlación estadística formal, es posible observar una tendencia consistente: a medida que disminuyen los problemas de este grupo, aumentan las puntuaciones otorgadas por los jueces en estos aspectos.

En la versión V1 se registró un promedio de 6,47 errores del grupo III. Esta versión recibió las calificaciones más bajas tanto en léxico como en pragmática: el juez 1 asignó un promedio de 1,95 en la dimensión léxica, mientras que la jueza 2 otorgó un 2,80. En cuanto a la dimensión pragmática, el juez 1 evaluó con un 2,25 y la jueza 2 con un 2,65. En la versión V3, donde interviene *Grammarly* seguida de una revisión por parte del participante, los errores del grupo III se mantuvieron en un nivel alto (aunque cercano a V1), con un promedio de 6,79 casos. No obstante, se observa una leve mejora en las calificaciones: en léxico, el juez 1 asignó un 4,15 y la jueza 2 un 3,85; en pragmática, el juez 1 otorgó un 4,25 y la jueza 2 un 3,78. La versión V4, que incorpora la intervención de *ChatGPT* antes de

la revisión humana, muestra una reducción significativa en los errores del tipo III, con un promedio de 3,37. Esta disminución se corresponde con un aumento considerable en las puntuaciones asignadas por los jueces. En la dimensión léxica, el juez 1 evaluó con un 6,70 y la jueza 2 con un 7,45, mientras que en la dimensión pragmática se asignaron valores de 6,95 y 7,55 respectivamente. Por último, en la versión V6, que integra el uso combinado de *Google Translate*, *Grammarly*, *ChatGPT* y múltiples etapas de edición por parte del participante, los fallos del grupo III descienden aún más, alcanzando el valor más bajo con un promedio de 2,47 casos. Esta versión recibió las puntuaciones más altas en ambas dimensiones: en léxico, 8,85 por parte del juez 1 y 8,35 por parte de la jueza 2; en pragmática, 8,50 y 8,25 respectivamente.

Si bien los datos muestran una tendencia general en la que la reducción de errores de III se asocia con una mejora en las evaluaciones de léxico y pragmática, esta correspondencia no se presenta de forma absoluta en todos los casos. Un ejemplo es la versión V3, donde, pese a un leve incremento en los fallos del grupo III respecto de V1, las puntuaciones otorgadas por los jueces en ambas dimensiones evaluadas son más altas. Esta aparente contradicción puede deberse a varios factores. En primer lugar, debe considerarse que mientras el análisis de errores de este trabajo es cuantitativo, la evaluación de los jueces tiene un componente cualitativo: si bien se les proporcionaron dimensiones específicas y una escala del 1 al 10 para realizar la puntuación, su juicio se sustenta en una percepción global de la calidad textual y no en un recuento de errores. En segundo lugar, es posible que mejoras en cuestiones vinculadas con aspectos como la corrección sintáctica, que no están contemplados en la definición de lo que es un error de III (pero que evidentemente tienen un impacto), hayan influido positivamente en la percepción. Esto pudo haber favorecido la valoración de la formulación léxica y la efectividad comunicativa, aun cuando los fallos vinculados a la intencionalidad o la fidelidad informativa no hayan disminuido. Por lo tanto, aunque existe una relación observable entre ciertos tipos de errores y las evaluaciones, está mediada por múltiples factores textuales y evaluativos que complejizan una interpretación unívoca.

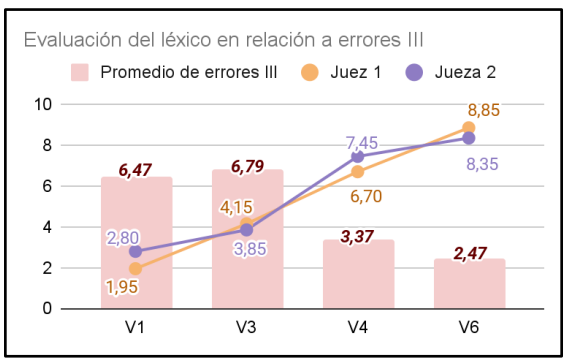

Figura 55

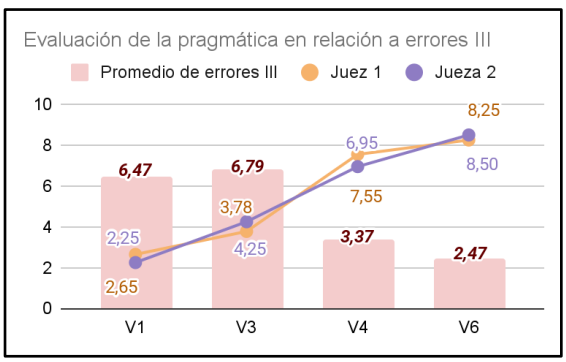

Figura 56

En segundo lugar, se examina la relación entre los fallos sintácticos y las puntuaciones asignadas por los jueces en la dimensión de sintaxis (v. figura 57). La versión V1 presenta el promedio más alto de errores sintácticos, con 19,75 por texto. En consonancia con este resultado, recibió las calificaciones más bajas en la dimensión correspondiente: el juez 1 asignó un promedio de 2,40 y la jueza 2, un 2,60. En la versión V3, que incluye la intervención de *Grammarly* seguida de una revisión por parte del participante, la cantidad de errores sintácticos desciende de forma significativa a 8,5. Esta mejora se refleja en las puntuaciones otorgadas por los jueces, que ascienden a 3,95 (juez 1) y 4,35 (jueza 2). Con la versión V4, en la que interviene *ChatGPT* antes de la revisión humana, se registra un descenso aún más marcado en los fallos sintácticos, con un promedio de apenas 1,5 por texto. Esta mejora se corresponde con una suba considerable en las calificaciones: 6,80 y 8,20 respectivamente. Por último, en la versión V6 —que integra el uso combinado de múltiples sistemas y etapas de intervención—, no se registran errores sintácticos. En este caso, se observan también las puntuaciones más altas en la dimensión evaluada: 8,20 por parte del juez 1 y 8,80 por parte de la jueza 2[34].

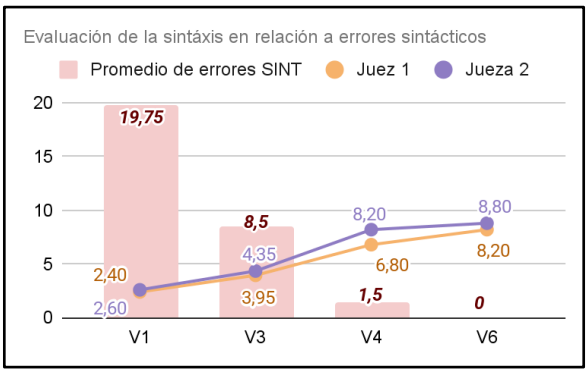

Figura 57

Por último, se explora la relación entre la dimensión discursiva evaluada por los jueces y dos grupos de errores textuales: aquellos vinculados a la coherencia y cohesión (CC) y los relacionados con la situacionalidad y aceptabilidad del texto (SA). Al tratarse de una dimensión que remite a la organización global del texto, a la adecuación al propósito comunicativo y a la consistencia interna de las ideas, resulta esperable que esté influida por múltiples niveles de análisis (v. figs. 58-59).

La versión inicial (V1) obtuvo una calificación de 2,35 por parte del juez 1 y 2,65 por parte de la jueza 2, en la dimensión discursiva, y en esta versión se identificaron en promedio 2,84 errores de CC y 5,63 errores de SA. Si bien los errores de coherencia y cohesión no aparecen en una cantidad particularmente elevada, esto se debe en parte a las dificultades estructurales de esta versión, que

[34] El hecho de que estas calificaciones no alcancen el valor máximo, a pesar de la ausencia de errores identificados, puede explicarse por el enfoque adoptado en la evaluación de los jueces sobre la dimensión sintáctica. Esta no se limita exclusivamente a la presencia o ausencia de errores gramaticales, sino que incluye también criterios como la claridad, la adecuación y la complejidad estructural. De este modo, una construcción puede ser formalmente correcta y, sin embargo, resultar menos adecuada desde una perspectiva sintáctica si su organización interna es oscura o dificulta la comprensión.

dificultan la identificación de fenómenos discursivos, como el uso incorrecto de conectores o referencias, debido a que partes de algunos fragmentos de estas traducciones autónomas (V1) no son del todo comprensibles. La elevada cantidad de errores de situacionalidad y aceptabilidad refuerza la baja calificación, ya que indica que el texto no logra adecuarse a las condiciones comunicativas de la tarea. En cambio, la cantidad de fallos de coherencia y cohesión no es particularmente alta, pero esto no parece reflejar una fortaleza del texto, sino más bien una consecuencia de sus fallas estructurales: los problemas sintácticos impiden que emerjan ciertos recursos de cohesión, lo que limita incluso la posibilidad de registrar errores en esta dimensión. En V3, la evaluación discursiva mejora levemente: 3,95 (juez 1) y 4,01 (jueza 2). Este incremento es consistente con una reducción de errores de CC a 1,74, aunque los de SA se mantienen en un nivel similar al de V1 (5,79 en promedio). La mejora parcial en CC parece estar reflejada en la ligera subida de las calificaciones. Sin embargo, la persistencia de fallos de adecuación contextual podría explicar por qué el aumento en la calificación no es más significativo. Es decir, aunque el texto mejora en cuanto a la estructuración de las ideas, sigue sin responder adecuadamente a los requerimientos del *género discursivo*. La relación entre puntuaciones y errores vuelve a ser coherente: la intervención (*Grammarly* + revisión del participante) tiene impacto estructural limitado, pero no aborda la dimensión discursiva del texto. La versión V4 presenta un salto significativo en la evaluación: 6,70 (juez 1) y 7,88 (jueza 2). Esta mejora se corresponde con una disminución marcada de fallos en ambas categorías: apenas 0,32 en CC y 0,58 en SA.

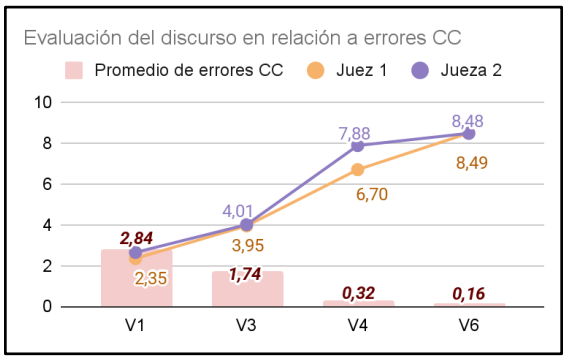

Figura 58

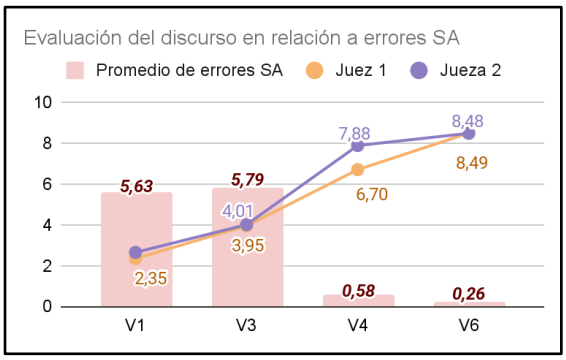

Figura 59

La versión V6 recibió las calificaciones más altas: 8,49 (juez 1) y 8,48 (jueza 2). Se identificaron en promedio 0,16 errores en CC y 0,26 en SA, los valores más bajos registrados en todas las versiones, por lo que la correspondencia entre evaluación y errores es evidente: los jueces valoran mejor un texto que casi no presenta fallas en la organización ni en la adecuación situacional. La versión parece tener relaciones cohesivas mejor establecidas y una construcción que responde más a las convenciones del *género discursivo*.

Relación de juicios de expertos y percepciones de participantes

Tanto los estudiantes como los evaluadores mostraron una tendencia a privilegiar las versiones que integran múltiples etapas de intervención, especialmente aquellas que combinan el uso estratégico de herramientas digitales con la revisión y edición humana. Sin embargo, mientras que en los jueces se observa una marcada concentración en torno a la versión V6, en los participantes las elecciones fueron más diversas, aunque con una preferencia igualmente destacada por esa misma versión. Esta dispersión puede interpretarse a la luz de las percepciones individuales sobre los beneficios del uso combinado de tecnologías, en tanto muchos estudiantes valoran la posibilidad de complementar funciones entre sistemas, ajustarlas a sus propias necesidades y realizar un control final sobre el producto textual. Así, la elección de V6 por parte de casi la mitad de los participantes y su predominio en la evaluación experta refuerzan la idea de que los procesos iterativos que articulan automatización e intervención humana tienden a ser percibidos como los más eficaces para lograr textos adecuados a contextos profesionales.

Si bien las motivaciones detrás de las elecciones pueden diferir —los participantes valoraron la acumulación de mejoras, mientras que los jueces evaluaron dimensiones como el léxico, la sintaxis, la pragmática y el discurso—, los resultados apuntan en la misma dirección: las traducciones valoradas como más eficaces surgen de un proceso de revisión iterativo y sostenido, que combina distintas formas de asistencia automática con la intervención crítica humana. Esta coincidencia sugiere que los criterios de calidad de quienes producen los textos tienden a alinearse con los criterios de los evaluadores expertos.

6.5. Otras posibles relaciones entre los datos

En este apartado se presentan algunas posibles relaciones entre los datos relevados en el corpus, formuladas en forma de proyecciones exploratorias. Dado que en varios casos se observa una marcada desproporción entre los grupos comparados, no es posible establecer interpretaciones concluyentes. No obstante, se consideró pertinente indagar preliminarmente estas posibles asociaciones y presentar de forma somera los patrones observados. La exploración de estas relaciones queda abierta para futuros estudios, con una muestra más amplia y equilibrada.

6.5.1. Errores y perfiles de los participantes

En primer lugar, se exploró una posible relación entre el nivel de inglés de los participantes y la cantidad de errores de cada tipo en los textos. Si bien se identificaron algunas relaciones preliminares, estas deben considerarse con reserva, ya que la muestra cuenta con solo un participante de nivel A1, dos de nivel A2 y uno de nivel C1, lo que impide establecer tendencias firmes.

En el caso de los problemas gramaticales, se observa una disminución progresiva en la cantidad promedio a medida que el nivel de inglés aumenta (A2: 21,45; A1:19,45; B1: 15,16; B2: 9,64; C1: 6,18), lo que sugiere una posible correlación entre el dominio lingüístico y la corrección gramatical en la realización de tareas de traducción asistida. En cuanto a los errores textuales, la relación con el nivel de inglés resulta menos definida. Si bien existe una ligera disminución en el promedio de fallos textuales entre los niveles B1 (8,22) y B2 (5,47), los promedios en otros niveles no siguen un patrón consistente. El único dato disponible para el nivel C1, por ejemplo, muestra un promedio de 5,18, cercano al observado en el nivel B2. Sin embargo, el comportamiento observado podría indicar que, para los participantes, los fallos textuales son más difíciles de identificar o corregir que los gramaticales. Para evaluar con mayor precisión esta relación, sería necesario contar con una muestra más amplia y equilibrada en términos de niveles de competencia lingüística.

También se exploró la posible influencia del área disciplinar de los participantes sobre la cantidad de fallos textuales vinculados al contenido del texto, es decir, aquellos clasificados en las categorías de informatividad, intertextualidad e intencionalidad (III). Dado que el texto fuente era un informe de laboratorio de una asignatura de las carreras de Bioquímica y Biología, se hipotetizó que quienes provenían de áreas más alejadas de estas disciplinas podrían presentar más errores relacionados con el contenido y con la terminología específica.

Sin embargo, los resultados no permiten confirmar esta hipótesis. Aunque los participantes de otras áreas presentan un promedio levemente más alto de errores de contenido (4,14), en comparación con los de Bioquímica (3,97) y Biología (4,00), la diferencia no resulta lo suficientemente significativa como para atribuirla a la distancia disciplinar. Una posible explicación es la gran diversidad dentro del grupo de disciplinas no afines, ya que los siete participantes restantes se distribuyen entre cuatro carreras distintas, lo que dificulta establecer patrones claros en una muestra tan heterogénea.

6.5.2. Errores y percepciones de los participantes

Otro aspecto explorado fue la posible relación entre la confianza que los participantes expresaron hacia las tecnologías y la cantidad de errores cometidos al corregir las versiones T. Además, se examinó si la que los participantes calificaron como preferida coincide con la que presenta menos fallos.

Relación entre la confianza en las herramientas y cantidad de errores

En la muestra, la mayoría de los participantes consideraron que cada recurso es confiable con supervisión humana, mientras que solo unos pocos (de 1 a 3, según el caso) manifestaron que las consideraban confiables sin supervisión, o no confiables en absoluto. Esta disparidad limita las comparaciones posibles, por lo que las tendencias observadas a continuación se presentan como proyecciones a estudiar a futuro con una muestra más equilibrada.

Si bien *Grammarly* tiende a reducir los fallos gramaticales en T2, la intervención humana en V3 no siempre supone una mejora. La mayoría, quienes consideraron que *Grammarly* es confiable con supervisión humana, incrementó el promedio de errores de 23,59 (T2) a 26,12 (V3). En cambio, entre quienes confiaban en el sistema sin necesidad de supervisión (3), estos se redujeron de 9,67 a 8,67 en promedio. Si se desglosa el comportamiento dentro del segundo grupo, se observa que, en un caso (E05), la cantidad de se mantiene constante en 7; en otro (E09), aumenta de 7 a 8; y en el tercero (E15), desciende de 15 a 11. Aunque solo uno de los tres participantes muestra una reducción en la cantidad de errores, las variaciones en los otros dos casos son mínimas, lo que podría sugerir una mayor estabilidad en la intervención humana sobre el texto, aspecto a corroborar con una muestra más amplia.

En el recorrido *Google Translate* (T1 y V2) se advierte un patrón similar. La mayoría (17), que consideró necesaria la supervisión, aumentó la cantidad de fallos tras la intervención de T1 (de 6,06 a 8,88). Lo mismo ocurrió con quienes expresaron desconfianza total (4 a 8,05). El único participante que declaró confiar de forma plena no modificó la cantidad de errores.

En el trayecto *ChatGPT* (V1, T3 y V4), los resultados son menos consistentes. Los participantes que consideraron que el *chatbot* es confiable con supervisión humana (19) aumentaron el promedio de fallos de 1,58 a 5,44, mientras que el único participante que desconfiaba por completo del sistema logró reducirlos de 5 a 4. Esto contradice en parte las tendencias observadas en las otras rutas, ya que a mayor desconfianza no aumentan los fallos, como sí ocurre en los otros casos.

En cuanto a la relación entre la confianza en los sistemas y la variación de errores textuales, se observa un panorama menos claro que en el caso de los gramaticales. En el camino *Grammarly*, el desempeño presentó una ligera diferencia en cada grupo en la versión intervenida (V3): en el grupo de confianza plena, el promedio de fallos textuales descendió de 11,00 (T2) a 9,33 (V3), mientras que en el grupo que declaró confianza con supervisión humana, el promedio subió de 15,00 (T2) a 15,59 (V3).

En la vía *Google Translate* (T1 y V2), se observan en los errores textuales patrones similares a los registrados para los gramaticales. La mayoría de quienes consideraron que el traductor es confiable con supervisión aumentaron los fallos tras la intervención (de 3,65 a 5,12). Los dos participantes que consideraron al sistema no confiable también los incrementaron de 4,00 a 5,00. El participante que confió de forma plena en el traductor mantuvo la cantidad constante.

Por último, en la ruta *ChatGPT*, los 19 participantes que consideraron a esta plataforma confiable con supervisión aumentaron los errores de 3,32 (T3) a 4,00 (V4), y el único que expresó desconfianza hacia *ChatGPT* también incrementó sus errores de 4,00 (T3) a 8,00 (V4).

En algunas rutas, quienes confiaron más en los sistemas tendieron a mantener o reducir la cantidad de errores tras su intervención. Sin embargo, dado lo reducido del grupo que confió plenamente en las

herramientas sin supervisión, estos indicios deben considerarse proyecciones exploratorias, abiertas a un estudio más profundo en trabajos con muestras más equilibradas.

Versión preferida vs versión con menos errores

Se exploró si existía una correspondencia entre las versiones más correctas —en términos de errores gramaticales y textuales— y las versiones preferidas por los participantes en la última pregunta de la encuesta de percepción. En solo 4 casos se observa una coincidencia entre la versión con menos errores y la preferida por el participante. En la mayoría, en cambio, se tiende a elegir versiones V (intervenidas por ellos) como las favoritas, mientras que las versiones con menos fallos suelen ser las T, para ambos tipos de errores. Esta situación sugiere que los participantes perciben que sus intervenciones mejoran el texto, más allá de si en efecto reducen los errores o los incrementan. Esta tendencia puede estar vinculada a sus opiniones sobre las tecnologías, en especial en relación con lo expresado sobre la necesidad de supervisión humana. Esta preferencia podría reflejar una necesidad de mayor control sobre el proceso de corrección, incluso si las versiones intervenidas no son las más correctas en términos gramaticales.

Además, el nivel intermedio de inglés (B1 y B2) de la mayoría de los participantes podría estar relacionado con el hecho de que, aunque no siempre son capaces de detectar ciertos problemas gramaticales en los textos, confían más en los resultados cuando pueden intervenir en el proceso. Las versiones V podrían ofrecer un texto más alineado con las expectativas del participante o un estilo que les parece más adecuado en relación al nivel de inglés que manejan, lo que genera una sensación de mejora a partir de su intervención.

7. Conclusión

Las siguientes conclusiones se organizan en tres apartados. El primero ofrece una interpretación de los hallazgos, en función de los objetivos e hipótesis planteados en la "Introducción". El segundo delinea proyecciones en relación a la aplicabilidad de dichos hallazgos en distintos ámbitos. Por último, se exponen reflexiones de carácter metodológico que destacan tanto los aspectos del diseño que resultaron eficaces y susceptibles de ser replicados, como aquellos que podrían optimizarse en futuras investigaciones con datos u objetivos similares.

7.1. Interpretación de los hallazgos

En primer lugar, puede concluirse que el uso de *Google Translate*, *Grammarly* y *ChatGPT* para la traducción y corrección de textos tiende a mejorar la calidad de las producciones escritas en inglés, y se logran resultados más adecuados a los objetivos comunicativos del género informe de laboratorio que los obtenidos sin su empleo. Asimismo, estas mejoras se pueden observar tanto en los planos léxico y sintáctico, así como en los niveles de (in)seguridad lingüística manifestados por los participantes.

Estas interpretaciones son respaldadas por los resultados de las diferentes secciones del análisis. Por un lado, la cantidad promedio de errores gramaticales y textuales, tanto en la traducción de nominalizaciones y voz pasiva, como en el caso de fallos lingüísticos generales en las traducciones, se reduce considerablemente en las versiones corregidas por las herramientas, respecto de la realizada de forma autónoma (V1). En particular, la reducción de errores sintácticos dentro del nivel gramatical, así como la mejora en la dimensión del vocabulario, evidenciada por la disminución de fallos de III (que contienen en su mayoría errores de selección léxica), corroboran la segunda hipótesis, referida a la mejora en los planos léxico, sintáctico y en los niveles de (in)seguridad lingüística.

Asimismo, se observa que la versión V6, culminación de la ruta *completa* que implica el uso combinado de las tres plataformas junto a la intervención de los participantes, ofrece resultados con menor cantidad de fallos promedio que V2, V3, y V4. Esto sugiere una ventaja inicial respecto de los caminos que emplean sólo un sistema y comienzan con la versión autónoma del participante (V3 y V4), y de la versión que consta de la intervención de los participantes sobre la versión traducida por *Google Translate* (V2). Esta tendencia respalda la hipótesis de que el uso combinado y secuencial de las tecnologías del texto propuestas mejora las producciones en mayor medida que su aplicación individual.

Por otro lado, entre los hallazgos que respaldan estas dos hipótesis, se destaca la dinámica observada en el perfil de vocabulario, calculado a través del *LFP*, que muestra un aumento en el porcentaje de vocabulario académico asociado al uso de la tecnología, especialmente de *ChatGPT*. En conjunto, los datos indican que *Grammarly* tiende a reducir la proporción de vocabulario en la *Off-list* a partir de la corrección ortográfica en T2 y, por lo tanto, aumenta la proporción de léxico de uso

cotidiano y académico. *Google Translate*, por su parte, ofrece en T1 una base de vocabulario académico similar a la lograda a partir del uso de *Grammarly* sobre las traducciones autónomas, que los participantes tienden a mantener con muy leves modificaciones. La incorporación de *Grammarly* en la ruta *completa* (T4), en la que el texto suele contener pocos errores ortográficos desde el inicio, no parece tener un impacto significativo sobre el léxico académico empleado. Sin embargo, *ChatGPT* tiende a reformular el texto con mayor profundidad y elevar considerablemente la proporción de vocabulario académico, tanto en su trayecto de uso individual como en el *completo* (T3 y T5, respectivamente).

Las hipótesis planteadas también quedan avaladas por los datos recogidos sobre las percepciones de los participantes con respecto al uso de las plataformas. Estos consideran, en general, que el uso de estas tecnologías tiene efectos positivos en la producción de este tipo textual en LE. Los recursos digitales fueron considerados útiles para detectar errores y afinar el resultado del proceso de escritura, si bien algunos mencionaron limitaciones puntuales, como la literalidad de *Google Translate* o el estilo, a veces demasiado libre, de *ChatGPT*. En general, sin embargo, los participantes opinaron que la combinación de herramientas les permitió lograr un texto más completo, claro y coherente. Por su parte, cuando se les consultó sobre qué sistemas recomendarían a un colega de su misma área que necesitara escribir un texto profesional, la mayoría optó por elegir la secuencia propuesta en el experimento. Asimismo, ante la consulta sobre qué versión sería más apropiada para publicar en un contexto profesional, la mayoría de ellos eligió la V6, versión final de la vía *completa*. Por último, cabe destacar también que la mayoría de los participantes indicó sentirse más seguro al redactar o traducir un texto con tecnologías como las presentadas en esta actividad. Este conjunto de validaciones de los participantes indican que el enfoque secuencial no solo mejora la calidad gramatical y textual, sino que también satisface sus exigencias comunicativas y les ayuda a sentirse más seguros sobre los resultados de sus producciones escritas.

Por su parte, el análisis de las evaluaciones de los jueces expertos también se encuentra en línea con las hipótesis formuladas. En general, estas valoraciones indican que la combinación de los recursos usados, junto con la intervención humana en diferentes etapas del proceso, puede ayudar a producir textos mejor evaluados por los expertos que aquellas producciones escritas sin asistencia tecnológica, en particular en las dimensiones léxica, sintáctica, pragmática y discursiva.

Con respecto a la hipótesis que plantea que los hablantes son conscientes de las ventajas que implica el empleo de estas tecnologías y que poseen la capacidad para usarlas de forma crítica, se proponen las siguientes observaciones. En relación con la primera parte de la hipótesis, el análisis de las respuestas de la encuesta de percepción indica que, en efecto, existe una percepción clara de los beneficios que conlleva el uso de estas herramientas. Por otro lado, los participantes consideran que estas deben ser utilizadas siempre con supervisión humana, lo que puede interpretarse como una actitud

crítica hacia ellas, en tanto no se limitan a aceptar de forma pasiva la versión sugerida por el sistema. Sin embargo, en muchos casos, al intervenir los textos generados por el sistema, los participantes terminaron por incrementar el número de errores gramaticales y textuales (aunque, en general, este aumento fue leve). Esto permite plantear interrogantes en torno a si el nivel de competencia en inglés de los participantes —en su mayoría intermedio o intermedio bajo— resulta suficiente para que exista un verdadero empleo de las tecnologías que resulte efectivo. Es decir, si ese uso consciente de sus limitaciones y del rol activo del usuario logra generar resultados adecuados, o incluso superiores a los que las herramientas podrían producir por sí solas o con mínima intervención. En este sentido, sería interesante replicar este experimento o desarrollar uno con características similares, con un grupo de estudiantes que posean niveles más altos de competencia en inglés.

A partir de lo expuesto, además, se puede concluir que se cumplieron los objetivos generales propuestos: por un lado, comprender con mayor profundidad los modos en que los estudiantes se apropian de estas tecnologías, y por otro, contribuir al estudio de las CCP mediante el análisis del impacto que dichos sistemas tienen en su desempeño lingüístico. Asimismo, se determinó que el uso secuencial de estas tecnologías ofrece ventajas para la comunicación profesional, entre las que se identifican mejoras en el léxico, la sintaxis y la seguridad lingüística de quienes las emplean.

Con respecto a los objetivos específicos, se observó que las operaciones lingüísticas más frecuentes al utilizar estas tecnologías son, en el caso de *Google Translate*, los reemplazos léxicos, y en el caso de *Grammarly* y *ChatGPT*, la recuperación de formas utilizadas en instancias anteriores. En lo referente al objetivo de determinar si una de las tres plataformas resulta más útil que las otras, diversos resultados indican que *ChatGPT* presenta ventajas significativas en varias dimensiones. En términos de errores gramaticales y textuales —tanto en relación con los rasgos del *género* como en aspectos generales—, se observa que esta herramienta corrige de forma más eficaz que *Grammarly*. Sumado a esto, en los casos en que el texto es generado a partir de *Google Translate*, *ChatGPT* logra reducir el número de fallos presentes en esas versiones iniciales. En lo que respecta al léxico, además, el análisis del perfil de vocabulario muestra que las traducciones resultantes de la ruta *ChatGPT* y de la *completa*, en especial luego de la intervención del *chatbot*, tienden a incrementar el uso de vocabulario académico. En cuanto a las percepciones de los estudiantes, *ChatGPT* también es valorado como más útil que las otras dos tecnologías. Finalmente, en las evaluaciones realizadas por los jueces, la versión V4 (corrección de *ChatGPT* sobre la versión autónoma) obtuvo valoraciones notablemente superiores a la versión V3 (corrección de *Grammarly* sobre la misma base).

Como proyección para futuros trabajos, sería pertinente indagar también en las preocupaciones sociales y posibles aspectos negativos asociados al uso de estas (u otras) tecnologías del texto. A modo ilustrativo, a la fecha de entrega de esta tesina, ha llegado a nuestro conocimiento el estudio de Kosmyna

et al. (2025) que explora las consecuencias neuronales del uso de modelos de lenguaje en la escritura de ensayos, y cuyos hallazgos advierten sobre posibles costos cognitivos vinculados a su empleo sostenido. En este sentido, sería valioso desarrollar investigaciones que identifiquen tanto los beneficios de estos recursos, como sus posibles efectos negativos a largo plazo, con el objetivo de fomentar un uso crítico, informado y reflexivo, en particular, en contextos educativos y profesionales.

7.2. Proyecciones

Desde la perspectiva de Lingüística Aplicada en la que se enmarca este trabajo, se considera que las observaciones previas pueden ser de utilidad para el diseño de materiales y estrategias pedagógicas, tanto para la enseñanza de LE con fines específicos como para la formación de CCP en lengua materna. Estas prácticas podrían enriquecerse a partir de los hallazgos del presente estudio, en tanto ofrecen insumos para incorporar el uso de recursos digitales de asistencia para la escritura analizados (u otros similares) en los procesos de escritura académica y profesional, a partir de un uso crítico que potencie y desarrolle las competencias de los estudiantes.

En segundo lugar, los resultados también pueden ser relevantes para estudiantes interesados en conocer el potencial de estos sistemas en la producción de textos que deberán enfrentar en contextos académico-profesionales. Si bien este trabajo no se propuso analizar sistemáticamente sus desventajas, algunos hallazgos permiten, asimismo, advertir ciertos aspectos que los usuarios de estas tecnologías deberían tener en cuenta. Por ejemplo, en determinadas situaciones, las intervenciones sobre los textos generados por estas herramientas pueden introducir o reintroducir errores, lo que subraya la necesidad de una lectura atenta y una participación activa del usuario que valore de forma justa sus CCP en LE, sin subvalorarlas ni sobrevalorarlas. Esto implica que el uso de estas tecnologías no reemplaza el estudio del idioma inglés ni garantiza por sí solo la elaboración de textos adecuados.

Además, este trabajo ofrece elementos que pueden ayudar a revisar ciertas percepciones frecuentes sobre algunas plataformas. Por ejemplo, se identificaron valoraciones negativas relacionadas con la fiabilidad de *Google Translate* que no siempre se condicen con los resultados obtenidos: por el contrario, la traducción automática seguida de una intervención del estudiante generalmente produce textos con menos errores que aquellos elaborados de forma autónoma y corregidos luego con *Grammarly*. En este sentido, se espera que este estudio contribuya a una toma de decisiones más y mejor informada respecto del uso de estas tecnologías que le permita al usuario reconocer tanto posibilidades como límites en función de los contextos y sus objetivos comunicativos.

En este mismo sentido, y en línea con su empleo crítico y complementario que aquí se promueve, y en consonancia con la <transparencia por la que aboga Lopezosa (2023), cabe destacar que se utilizó *ChatGPT* como asistente de corrección durante la escritura de esta tesina. Su uso se orientó a la

búsqueda de sinónimos, la mejora de la cohesión textual, y la revisión de formulaciones propias a partir de borradores comentados con indicaciones integradas a las consultas realizadas al sistema. Por ejemplo, una de las oraciones de esta tesina, se formuló originalmente de la siguiente manera:

> Se puede sugerir que estos recursos son más efectivos para resolver dificultades gramaticales, ya que las textuales parecen presentar cierta resistencia a la corrección automática y humana, al depender en mayor medida del contexto y la intención comunicativa.

Esta formulación nos resultó problemática, y enviamos el texto al *chatbot* seguido de una instrucción (*prompt*) con una indicación puntual[35]:

> """
>
> Se puede sugerir que estos recursos son más efectivos para resolver dificultades gramaticales, ya que las textuales parecen presentar cierta {resistencia}* a la corrección automática y {humana}*, al depender en mayor medida del contexto y la intención comunicativa.
>
> """
>
> ** Necesito reformular esta parte: que ciertos errores presenten "resistencia" a la corrección automática tiene sentido, ya que un sistema puede tener problemas para corregir ciertos tipos de errores, por ejemplo, a causa de la falta de información contextual, pero en el caso de la corrección humana, esta tiene más que ver con las competencias y capacidades de cada persona y cómo interpreta esa información contextual.*

A partir de esta consulta, el modelo devolvió la siguiente respuesta:

> *Claro, aquí te dejo una reformulación que aclara ese punto, evitando la ambigüedad sobre la "resistencia" de los errores a la corrección automática y humana:*
>
> *Se puede sugerir que estos recursos son más efectivos para resolver dificultades gramaticales, mientras que las correcciones de errores de tipo textual resultan más complejas. Esto se debe a que la corrección automática puede verse limitada por la falta de información contextual suficiente, y la corrección humana depende, en gran medida, de las competencias y capacidades individuales de quien realiza la revisión, así como de la interpretación del contexto y la intención comunicativa.*

[35] Se utilizan triples comillas como separador entre el texto a corregir y la instrucción, así como marcadores gráficos (llaves o asteriscos), para indicar al modelo de lenguaje en dónde se encuentra el texto que debe considerar para resolver el pedido del usuario. Este tipo de prácticas en la formulación de instrucciones —recomendadas en cursos de ingeniería de *prompts*— ayuda a obtener respuestas más precisas del modelo.

Evaluamos esta respuesta y, si bien decidimos no integrarla en su totalidad, las formulaciones que se encuentran subrayadas nos parecieron adecuadas, y fueron utilizadas para corregir nuestra expresión original de la siguiente manera:

> Se puede sugerir que estos recursos son más efectivos para resolver dificultades gramaticales, ya que las textuales parecen presentar cierta resistencia a la corrección automática, al depender en mayor medida de factores como el contexto y la intención comunicativa. Por otro lado, en lo que refiere a la dificultad de corrección humana de este tipo de fallos, esta se vincula principalmente con las competencias individuales de quien realiza la revisión, así como con la interpretación de esos mismos factores extralingüísticos. (p. 53)

En diálogo con la de los participantes, nuestra experiencia muestra que apelar a la asistencia lingüística de estos sistemas requiere, por un lado, un aprovechamiento estratégico de la tecnología, basada en la reflexión metalingüística y, por el otro, el control consciente del proceso de escritura.

En otro sentido, se espera que los resultados de este trabajo —así como los de otros que abordan fenómenos similares— no solo contribuyan a reflexionar sobre el uso de las tecnologías del texto, sino que también ofrezcan insumos valiosos para futuras mejoras en el desarrollo de estas herramientas. Si bien este estudio no se centra en los aspectos técnicos de su funcionamiento, aporta información empírica sobre uno de los contextos en que su desempeño puede resultar más eficaz, así como sobre aquellos en los que presenta limitaciones. Entre los aportes positivos, se destacan la baja tasa de errores que arrojan las traducciones de *Google Translate*, la capacidad de *Grammarly* y de *ChatGPT* para corregir los fallos de los participantes, y la eficacia de este último sistema para mejorar aspectos como el vocabulario y la adecuación al registro. Entre las limitaciones, aparecen el carácter literal de las traducciones producidas por el sistema automático de *Google*, el alcance limitado de *Grammarly* en lo relativo al procesamiento de información contextual para realizar mejoras más amplias en el nivel textual, y la cantidad excesiva de correcciones introducidas por *ChatGPT*, que pueden generar en los usuarios la sensación de pérdida de control.

7.3. Conclusiones metodológicas

Por último, se destacan aportaciones de tipo metodológico para aquellos investigadores que deseen realizar estudios que aborden problemáticas afines, ya sea por su temática o por presentar dinámicas entre usuario y tecnologías comparables a las analizadas en este trabajo. En línea con esto, en este apartado, se detallan aquellos aspectos que pueden ser replicables, y aquellos que pueden mejorarse para investigaciones futuras.

Sobre la recogida de datos, se destaca que demandó más tiempo del previsto, debido a una sobreestimación de la cantidad de estudiantes dispuestos a participar en una actividad prolongada. En este sentido, se sugiere que futuras investigaciones consideren estrategias más eficaces de convocatoria de participantes, la ampliación de los criterios de inclusión, o la selección de otros procedimientos de recolección de datos, en especial en los casos en los que se disponga de plazos de ejecución acotados.

En relación con el material utilizado como texto fuente, la elección de fragmentos breves pero densos en fenómenos lingüísticos —en lugar de un informe de laboratorio completo— permitió una observación detallada del modo en que los participantes abordaban estos desafíos con la asistencia de las herramientas. Esta decisión metodológica resultó especialmente pertinente si se considera el tiempo limitado disponible para la interacción con los participantes.

Como se explicó en el apartado de metodología, el procesamiento y registro de datos consistió casi en su totalidad de un análisis manual, complementado por las estrategias automatizadas explicitadas. Contar con el análisis manual y con el de los procesos automatizados permitió realizar múltiples verificaciones y contrastes internos con el fin de minimizar posibles errores o inconsistencias. Si bien todas las discrepancias que pudieron ser detectadas durante estas instancias fueron corregidas, se considera pertinente que futuras investigaciones incorporen sistemas más avanzados de análisis de consistencia, con el objetivo de maximizar la precisión y fiabilidad de los resultados. Por otro lado, el uso de *diff_match_patch* para identificar las versiones sin modificaciones de los participantes fue un aporte importante que agilizó el proceso de detección de intervenciones al focalizar solo en aquellos textos que habían sido editados. Además, la implementación de un *script* en *Python* para registrar las ediciones marcadas manualmente en *Google Docs*, también contribuyó a agilizar el procesamiento de datos. Por otra parte, no fue posible incorporar recursos adicionales que habrían permitido automatizar otras fases del análisis, como el uso de librerías en *Python* para la identificación de nominalizaciones, voz pasiva y ciertos tipos de errores, debido a las limitaciones que estas aún presentan en términos de precisión y aplicabilidad a este tipo de corpus, como fue especificado en metodología.

Para finalizar, en cuanto a la incorporación de evaluadores expertos para complementar el análisis de los datos experimentales, se constata que sus observaciones incorporaron una mirada adicional que enriqueció la interpretación de los hallazgos, al acercarla a una perspectiva más cercana a los criterios reales de uso académico y profesional.

8. Anexos

Este anexo reúne los materiales que respaldan el desarrollo del estudio. Se incluye el texto fuente en español que los participantes debieron traducir, una posible traducción, las encuestas de perfil y percepción que completaron, el formulario dirigido a los jueces expertos para la evaluación de las versiones enviadas, y las traducciones esperadas de nominalizaciones y construcciones pasivas e impersonales con "se".

8.1. Texto fuente en español

Fraccionamiento de la actividad enzimática con sulfato de amonio

Introducción
En contexto de introducirnos en: el manejo de las técnicas utilizadas en un protocolo de purificación, la espectrofotometría de absorción molecular y el adecuado uso de las curvas de calibrado fue realizado el presente trabajo en el cual se buscó purificar la enzima fosfatasa ácida vegetal a través del fraccionamiento de la actividad enzimática con sulfato de amonio. A su vez la actividad de dicha enzima fue sondeada mediante una reacción de tipo colorimétrica y se determinó el porcentaje de purificación de la misma.
Si bien se produjeron altercados y resultados que no habían sido esperados (ver conclusión) los objetivos generales fueron cumplidos.

Resultados
Para la obtención de los datos que corresponden a concentraciones tanto de proteínas asi como de P-Nitrofenol (p-NP) fueron utilizadas curvas de calibrado que fueron realizadas con Solución Estándar de Albúmina y Solución Estándar de P-Nitrofenol.

[…]
Dado que para el TUBO 2 se obtuvo una Abs que no se encuentra dentro de la curva de calibrado no se puede asegurar la linealidad por lo cual dicho valor no debe tomarse en cuenta.

[…]
En el caso del TUBO de 2 gr para la cuantificación de proteínas dado que se obtuvo una Abs menor a la que se emplea en la curva de calibrado, la cual fue realizada con solución estándar de albúmina, no puede utilizarse en los cálculos posteriores a todo el protocolo. Esto pudo deberse a:

- El punto anterior mencionado
- Se realizó una dilución de la F.E ya sea agregando volúmenes mayores a los requeridos tanto de NaOH 0,1 N o Reactivo de Bradford
- No se dejó en reposo los 4 minutos necesarios antes de leer la absorbancia a 595 nm.

Con respecto a uno de los objetivos del trabajo práctico, el cual se refería a determinar la concentración óptima de sal para conseguir una purificación eficaz, se llegó a la conclusión de que el tubo de 0,75 gr de sal es el de mayor actividad enzimática, por lo que corresponde al tubo con el valor óptimo para la precipitación de la Proteína Enzimática.
Se considera que este tubo contiene el valor óptimo de sal para la precipitación de la proteína enzimática debido a que tiene la mayor cantidad de producto formado respecto del total de proteínas que precipitaron, es decir de todo lo que precipitó, la mayor parte corresponde a proteína enzimática.

Conclusión
Con la realización del trabajo en el laboratorio se puede concluir que se alcanzaron los objetivos principales los cuales consistieron en la introducción en el manejo de las técnicas de un protocolo de purificación, el uso de la espectrofotometría de absorción molecular y la utilización de las curvas de calibrado.
Por otro lado, se alcanzaron parcialmente los objetivos específicos, pese a los posibles errores que pudieron llevar a los resultados inesperados, el protocolo fue seguido y se trabajó correctamente.

8.2. Traducción posible del texto fuente

Fractionation of enzymatic **activity** with ammonium sulfate

Introduction
In the context of introducing ourselves to: the **management** of the techniques used in a **purification** protocol, molecular **absorption** spectrophotometry and the appropriate **use** of **calibration** curves, this work was carried out, in which it was sought to purify the enzyme plant acid phosphatase through of the **fractionation** of the enzymatic **activity** with ammonium sulfate. In turn, the **activity** of said enzyme was probed through a colorimetric **reaction** and its **purification** percentage was determined.
Although there were **altercations** and results that had not been expected (see **conclusion**), the general objectives were met.

Results
For the **obtainment** of data corresponding to concentrations of both proteins and P-Nitrophenol (p-NP), **calibration** curves were used which were created with Albumin Standard Solution and P-Nitrophenol Standard Solution.

[...]
Given that for TUBE 2 an **absorbance** value was obtained that is not within the **calibration** curve, **linearity** cannot be ensured, which is why said value should not be taken into account.

[...]
In the case of the 2 g TUBE for protein **quantification**, given that a lower Abs was obtained than the one used in the **calibration** curve, which was created with standard albumin solution, it cannot be used in subsequent **calculations** in the protocol. This could have been due to:
○ The previous point mentioned
○ A **dilution** of the enzymatic fraction was carried out either by adding larger volumes than required of either 0.1 N NaOH or Bradford **Reagent**.
○ The solution was not allowed to rest the necessary 4 minutes before reading the **absorbance** at 595 nm.

Regarding one of the objectives of the practical work, which referred to determining the optimal **concentration** of salt to achieve effective **purification**, it was concluded that the tube with 0.75 g of salt is the one with the highest enzymatic **activity**, and therefore it corresponds to the tube with the optimal value for the **precipitation** of the enzymatic protein. It is considered that this tube contains the optimal value of salt for the **precipitation** of the enzymatic protein because it has the largest amount of product formed with respect to the total proteins that precipitated. That is, of everything that precipitated, the majority corresponds to enzymatic protein.

Conclusion
Upon the **completion** of the work in the laboratory, it can be concluded that the main objectives were achieved, which consisted in the **introduction** in the **management** of the techniques of a **purification** protocol, the **use** of molecular **absorption** spectrophotometry and the **use** of the **calibration** curves.
On the other hand, the specific objectives were partially achieved, despite possible errors that could have led to unexpected results, the protocol was followed and the work in the laboratory was carried out correctly.

8.3. Encuesta de perfil

Encuesta para relevar perfil
La presente encuesta tiene como objetivo general recolectar datos sobre su perfil sociolingüístico y académico, así como su conocimiento previo sobre las herramientas a utilizar en la siguiente actividad. Esta investigación tiene como objetivo contribuir al campo de las competencias comunicativas profesionales, permitiendo una comprensión más profunda de las percepciones de los participantes sobre el uso de las herramientas digitales en la actividad.
Su participación en esta encuesta, así como en las siguientes actividades, es completamente confidencial y voluntaria, pudiendo retirar su participación cuando así lo desee. La información que proporcione se utilizará únicamente con fines académicos y de investigación, y todos los datos recopilados se mantendrán de manera anónima, sin revelar su identidad en ningún informe o publicación resultante.
Si, en cualquier momento, decide retirar sus datos de esta investigación, puede hacerlo contactándose a través del siguiente mail: diamelafrapolli@gmail.com

Código de experimento

_____

Género
◯Femenino
◯Masculino
◯Otro/prefiero no decirlo

Edad

____

Actualmente, ¿qué carrera/s estás cursando?

_____

Actualmente, ¿qué año de tu carrera estás cursando?
(Si estás cursando materias de diferentes años, seleccioná una opción en función de las materias que sean del grado de avance más elevado de la carrera)
◯Primero
◯Segundo
◯Tercero
◯Cuarto
◯Quinto
◯Sexto
◯Posgrado (trayecto inicial)
◯Posgrado (trayecto medio)
◯Posgrado (trayecto avanzado)

¿Cuál es tu nivel de dominio del inglés?
◯No tengo conocimientos en lengua inglesa
◯A1 (Mis conocimientos de inglés son todavía muy básicos. Puedo entender y usar expresiones simples en situaciones concretas, como presentarme o pedir información básica.)
◯A2 (Mi nivel de inglés es elemental. Puedo comunicarme en tareas simples y cotidianas, como dar direcciones, hacer compras o describir mi entorno.)
◯B1 (Tengo un nivel intermedio bajo de inglés. Puedo entender y comunicarme en situaciones familiares y cotidianas, como hablar de mis intereses o contar experiencias pasadas.)
◯B2 (Mi nivel de inglés es intermedio. Puedo participar en conversaciones más complejas sobre temas variados y expresar opiniones. Además, puedo entender textos más elaborados.)
◯C1 (Tengo un nivel avanzado de inglés. Puedo comunicarme fluidamente en situaciones exigentes, debatir sobre temas abstractos y entender discursos más complejos.)
◯C2 (Mi inglés es casi como el de un hablante nativo. Puedo comprender y expresarme con facilidad en cualquier situación, incluso en contextos especializados y abstractos.)

¿Tenés algún certificado que compruebe tu nivel de inglés? Si es así, ¿cuál?

_____

¿Conocés y/o has usado el Traductor de Google?
◯Conozco el Traductor de Google y lo he usado antes.
◯Conozco el Traductor de Google pero no lo he usado antes.
◯No conozco el Traductor de Google.

¿Conocés y/o has usado Grammarly?
◯Conozco Grammarly y lo he usado antes.
◯Conozco Grammarly pero no lo he usado antes.
◯No conozco Grammarly.

¿Conocés y/o has usado Chat GPT?
◯Conozco Chat GPT y lo he usado antes.
◯Conozco Chat GPT pero no lo he usado antes.
◯No conozco Chat GPT.

Agradecimiento

Agradecemos inmensamente su colaboración al participar en la encuesta.
Recordamos, nuevamente, que puede retirar sus datos de esta investigación en el momento que desee contactándose a través de: diamelafrapolli@gmail.com

8.4. Encuesta de percepción

Encuesta de percepción sobre herramientas digitales para la comunicación profesional

La presente encuesta tiene como objetivo general recolectar datos sobre sus percepciones y experiencias relacionadas con el uso de las herramientas digitales utilizadas en la actividad previa. Como parte de esta investigación esta encuesta busca contribuir al campo de las competencias comunicativas profesionales, permitiendo una comprensión más profunda de las percepciones de los participantes sobre el uso de las herramientas digitales en su actividad anterior.
Su participación en esta encuesta, así como en la previa actividad, es completamente confidencial y voluntaria, pudiendo retirar su participación cuando así lo desee. La información que proporcione se utilizará únicamente con fines académicos y de investigación, y todos los datos recopilados se mantendrán de manera anónima, sin revelar su identidad en ningún informe o publicación resultante.
Si, en cualquier momento, decide retirar sus datos de esta investigación, puede hacerlo contactándose a través del siguiente mail: diamelafrapolli@gmail.com

Código de experimento

_____

¿Cómo te resultó usar el Traductor de Google para esta actividad?[36]
¿Cómo te resultó usar el Grammarly para esta actividad?
¿Cómo te resultó usar el Chat GPT para esta actividad?

Al usar las herramientas en forma conjunta, ¿se te hizo más fácil o más difícil realizar esta traducción?
◯Se me hizo más fácil
◯Se me hizo más difícil
◯No noté ninguna diferencia

Por favor, detallá un poco mejor tu respuesta en la pregunta anterior.

_____

¿Te parece que las herramientas utilizadas conllevan algún beneficio para alguien que necesite traducir un informe de laboratorio al inglés? Seleccioná todas las que correspondan
☐ Sí, el Traductor de Google
☐ Sí, Grammarly
☐ Sí, ChatGPT
☐ No, ninguna de ellas

Sí respondiste sí a la pregunta anterior ¿alguna de las herramientas te parece más útil que las otras?
◯Sí, el Traductor de Google
◯Sí, Grammarly
◯Sí, ChatGPT
◯No, todas son igual de útiles

Sí respondiste que al menos algunas de las herramientas conllevan algún beneficio ¿te parece que el uso en conjunto y secuencial con ellas resulta más beneficioso que su uso por separado para lograr una mejor traducción?
◯Sí
◯No
Por favor, explica el motivo de tu respuesta anterior:

_____

[36] Tres preguntas de escala lineal con opciones del 1 al 5, en donde 1 significa "Muy sencillo" y 5 significa "muy complicado"

Imaginá ahora que el texto que te dimos a traducir era tu propio informe de laboratorio, y que debías presentarlo en inglés por razones laborales. ¿Considerás que el texto obtenido al usar la secuencia propuesta cumple tus objetivos de comunicación profesional?
◯Sí
◯No

¿Te sentís más seguro/a al redactar/traducir un texto con o sin herramientas como las presentadas en esta actividad?
◯Me siento más seguro/a al redactar/traducir un texto utilizando herramientas como las presentadas en esta actividad.
◯Me siento más seguro/a al redactar/traducir un texto sin utilizar herramientas como las presentadas en esta actividad.
◯Me siento igual de seguro/a al redactar/traducir un texto utilizando herramientas como las presentadas en esta actividad que al redactarlo/traducirlo sin ellas.

Si algún conocido dentro de tu área te comentara que tiene que escribir un texto en inglés para una situación profesional:
☐ Recomendaría el Traductor de Google.
☐ Recomendaría Grammarly.
☐ Recomendaría ChatGPT
☐ Recomendaría usar las tres herramientas en forma conjunta y secuencial
☐ No recomendaría ninguna de las herramientas usadas en esta actividad.
☐ Otra…
_____

¿Considerás que Google Translate/Grammarly/ChatGPT es una herramienta fiable para traducir textos académicos/profesionales al inglés?
◯Sí, y no es necesario ningún control por parte del usuario para obtener buenos resultados.
◯Sí, pero es necesario que el usuario revise los resultados para asegurar su calidad.
◯No, no la considero fiable de ninguna forma.

¿Podrías explicar por qué?[37]
_____

¿Considerás que la calidad del texto en español influye en la calidad de la traducción realizada con herramientas digitales?
◯Sí
◯No
◯Tal vez

Por favor, detallá más tu respuesta:
_____

De todas las versiones de traducción obtenidas durante la actividad, ¿cuál elegirías para publicar en un contexto profesional, y por qué?
Las opciones:
- V1 (la que hiciste únicamente con el diccionario online)
- T1 (la versión de Google Translate)
- V2 (la que creaste a partir de intervenir la versión generada por Google Translate)
- T2 (La corrección de Grammarly sobre V1)
- V3 (La intervención que hiciste sobre T2)
- T3 (La corrección de ChatGPT sobre V1)
- V4 (La intervención que hiciste sobre T3)
- T4 (la corrección de Grammarly sobre V2)
- V5 (la intervención que hiciste sobre T4)
- T5 (la corrección de ChatGPT sobre V5)
- V6 (la intervención que hiciste sobre T5, o sea la que quedó al final)

_____
Agradecemos inmensamente su colaboración al participar en la encuesta.
Recordamos, nuevamente, que puede retirar sus datos de esta investigación en el momento que desee contactándose a través de: diamelafrapolli@gmail.com

[37] Estas dos preguntas fueron repetidas de forma idéntica para cada una de las tres herramientas (*Google Translate*, *Grammarly* y *ChatGPT*). En el anexo se presentan de manera condensada por motivos de espacio.

8.5. Formulario enviado a los jueces

Evaluación de versiones de traducción del participante X[38]

Bienvenido/a y gracias por su participación en esta evaluación. En esta encuesta, se le solicitará evaluar distintas versiones de la traducción de un informe de laboratorio, considerando aspectos clave como la elección de palabras, la estructura gramatical, la efectividad comunicativa y la adecuación al estilo de los informes científicos en inglés. Su opinión es fundamental para valorar la calidad de las traducciones y obtener información sobre los aspectos que contribuyen a una mejor versión.

Le pedimos que responda cada pregunta con la mayor precisión posible y, al finalizar, elija la versión que considere mejor en términos generales, justificando su elección.

Ante cualquier duda, no dude en ponerse en contacto.

En primer lugar, le pedimos que responda una serie de preguntas para cada versión de la traducción del informe de laboratorio de este participante. Cada pregunta se responde en una escala del 1 al 10, donde 1 indica una calidad muy baja y 10 una calidad excelente.

Versión x:

Léxico

Para esta versión, ¿cómo evalúa la elección de palabras en términos de precisión y adecuación, en una escala del 1 al 10? Considere si los términos empleados son correctos dentro del ámbito del informe de laboratorio, si reflejan fielmente el contenido original y si son apropiados para un texto científico en inglés.

Puntaje: ▾ (menú desplegable con números del 1 al 10)

Sintaxis

Para esta versión, ¿qué tan clara y correcta es la construcción de las oraciones? Evalúe si las estructuras gramaticales son adecuadas para el inglés, si facilitan la comprensión del texto y si mantienen la coherencia y fluidez propias de un informe científico.

Puntaje: ▾ (menú desplegable con números del 1 al 10)

Pragmática

Para esta versión, ¿qué tan efectiva es la traducción para cumplir el objetivo comunicativo del informe? Considere si el texto transmite la información de manera precisa y sin ambigüedades, y si permite que el lector acceda fácilmente a los datos y conclusiones sin dificultades de comprensión, más allá de la corrección absoluta de la gramática o el registro.

Puntaje: ▾ (menú desplegable con números del 1 al 10)

Discurso

Para esta versión, ¿qué tan bien se ajusta el texto al estilo de los informes científicos en inglés? Evalúe si el tono, la organización y el nivel de formalidad son los esperados en este tipo de textos, y si la traducción sigue las convenciones propias del género.

Puntaje: ▾ (menú desplegable con números del 1 al 10)

[38] Por razones de extensión, en este anexo se presenta un modelo simplificado del formulario de evaluación utilizado por los jueces. El formulario completo, implementado a través de un archivo de Excel, consistía en 20 pestañas, una por cada participante, y en cada pestaña se incluían las mismas preguntas para evaluar las cuatro versiones de traducción correspondientes. A modo de ejemplo, se incluye a continuación una única sección ilustrativa con las preguntas formuladas para una sola versión. Las mismas preguntas se repitieron para cada una de las cuatro versiones.

8.6. Traducciones esperadas de nominalizaciones

Caso	NOM/EN/Posible
abs/absorbancia	abs/absorbance
absorción	absorption
actividad	activity
altercados	altercations
cálculos	calculations
calibrado	calibration
calibrado	calibration
realización	completion
concentración	concentration
concentraciones	concentrations
conclusión	conclusion
dilución	dilution

Caso	NOM/EN/Posible
fraccionamiento	fractionation
introducción	introduction
linealidad	linearity
manejo	management
obtención	obtainment
precipitación	precipitation
purificación	purification
cuantificación	quantification
reacción	reaction
reactivo	reagent
uso	use
utilización	utilization

8.7. Traducciones esperadas de voz pasiva e impersonales con "se"

Caso	Pasiva/EN/Posible
fue realizado el presente trabajo	this work was carried out
se buscó purificar la enzima fosfatasa	it was sought to purify
la actividad de dicha enzima fue sondeada	the activity of said enzyme was probed
y se determinó el porcentaje de purificación de la misma	its purification percentage was determined
resultados que no habían sido esperados	results that had not been expected
los objetivos generales fueron cumplidos	the general objectives were met
fueron utilizadas curvas de calibrado	calibration curves were used
que fueron realizadas con	which were created with
se obtuvo una Abs	an absorbance value was obtained
no se puede asegurar la linealidad	linearity cannot be ensured
no debe tomarse en cuenta	said value should not be taken into account
se obtuvo una Abs menor	a lower Abs was obtained
a la que se emplea en la curva de calibrado	than the one used in the calibration curve
la cual fue realizada con solución estándar de albúmina	which was created with standard albumin solution
no puede utilizarse en los cálculos posteriores	it cannot be used in subsequent calculations
Se realizó una dilución	A dilution was carried out
No se dejó en reposo	the solution was not allowed to rest
se llegó a la conclusión de que el tubo de 0,75 gr	it was concluded
Se considera que este tubo	It is considered
se puede concluir que	it can be concluded
se alcanzaron los objetivos principales	the main objectives were achieved
se alcanzaron parcialmente los objetivos específicos	the specific objectives were partially achieved
el protocolo fue seguido	the protocol was followed
se trabajó correctamente	the work in the laboratory was carried out correctly

9. Bibliografía

Aguirre Raya, D. A. (2005). Reflexiones acerca de la competencia comunicativa profesional. *Revista cubana de Educación Médica Superior*, 19(3), 1. http://scielo.sld.cu/scielo.php?pid=S0864-21412005000300004&script=sci_abstract

Al-Rodhan, N. R. F., y Stoudmann, G. (2006). Definitions of Globalization: A Comprehensive Overview and a Proposed Definition. *Program on the Geopolitical Implications of Globalization and Transnational Security*, 6, 1-21.

Bajtin, M. M. (1992). *Estética de la creación verbal* (5a. ed.). Siglo XXI.

Baker, P., Hardie, A., y McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburgh University Press.

Banks, D (2008). *The development of scientific writing: Linguistic features and historical context*. Equinox Publishing.

Banks, D. (2017). The extent to which the passive voice is used in the scientific journal article, 1985-2015. *Functional Linguistics*, 4(12), 1-17. https://doi.org/10.1186/s40554-017-0045-5

Banks, D. (2023). Types of nominalization in scientific English. *Scientific Notes of NTSSPI. Series: History and Philology*, 1, 8-19. https://ntspi.ru/upload/2023-1.pdf

Barcia, P. L., y Barcia, M. (2021). *Los géneros comunicativos universitarios orales y escritos. Teoría y práctica*. Editorial UCALP. https://play.google.com/store/books/details?id=I5o2EAAAQBAJ

Barrio Andrés, M. (2023). ChatGPT y su impacto en las profesiones jurídicas. *Diario La Ley*. https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAMtMSbF1CTEAAmMzC0sLQ7Wy1KLizPw8WyMDI2MDU0MDtbz8lNQQF2fb0ryU1LTMvNQUkJLMtEqX_OSQyoJU27TEnOJUtdSk_PxsFJPiYSYAAOR06GhjAAAAWKE

Bein, R. (2003). Propaganda de lenguas. *Letras*, 27, 27-37. https://doi.org/10.5902/2176148511895

Berry, D. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine*, 12, 1-22. http://www.culturemachine.net/index.php/cm/article/view/440/470

Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman.

Biber, D. (1988). *Variations across speech and writing*. Cambridge University Press.

Bosio, I. V. (2014). El informe de investigación. En L. Cubo de Severino, *Los textos de la ciencia* (pp. 305-321). Comunic-arte.

Bosque Muñoz, I., y Gutiérrez-Rexach, J. (2008). ¿Qué es la sintaxis? Caracterización y bases empíricas. En *Fundamentos de sintaxis formal* (pp. 11-53). Ediciones Akal.

Botta, M. y Warley, J. (2002). *Tesis, tesinas, monografías e informes*. Biblos.

Bretegnier, A. (1998). Introduction. En A. Bretegnier y G. Ledegen (Eds.), *Sécurité / insécurité linguistique - Terrains et approches diversifiées, propositions théoriques et méthodologiques. Actes de la 5e table ronde du Moufia,* (pp. 8-33). Collection Espaces Francophones.

Calsamiglia, H., y Tusón, A. (1999). *Las cosas del decir: Manual de análisis del discurso*. Editorial Ariel.

Calvet, L.-J. (2006). *Towards an Ecology of World Languages*. Polity Press.

Cambridge University Press & Assessment. (s.f.). *Cambridge Dictionary*. https://dictionary.cambridge.org/

Camps Mundó, A., y Castelló, M. (2013). La escritura académica en la universidad. *Revista de Docencia Universitaria,* 11(1), 17-36.

Cantamutto, L., y Delfa, C. V. (2016). El Discurso Digital como objeto de estudio: De la descripción de interfaces a la definición de propiedades. *Aposta. Revista de Ciencias Sociales*, 69, 296-323.

Carlino, P. (2010). *Escribir, leer y aprender en la universidad*. Fondo de Cultura Económica.

Cassany, D. (2007). *Afilar el lapicero: Guía de redacción para profesionales*. Editorial Anagrama.

Castillo-González, W., Lepez, C. O., y Bonardi, M. C. (2022). Chat GPT: A promising tool for academic editing. *Data & Metadata*, 23(1), 1-6. https://www.researchgate.net/publication/373185669_Chat_GPT_a_promising_tool_for_academic_editing

Charaudeau, P., y Maingueneau, D. (Dirs.). (2002). *Dictionnaire d'analyse du discours*. Éditions du Seuil.

Ciapuscio, G. E. (1992). Impersonalidad y desagentivación en la divulgación científica. *LEA. Lingüística Española Actual,* 14(2), 183-206.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Comrie, B., y Thompson, S. (2007). Lexical nominalization. En T. Shopen (Ed.), *Language typology and syntactic description* (pp. 334-381). Cambridge University Press. https://doi.org/10.1017/CBO9780511618437.006

Consejo de Europa. (2001). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación* (2a. ed.). MECD - Anaya. https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press.

Coxhead, A. (2002). The academic word list: A corpus-based word list for academic purposes. *Language and Computers*, 42, 73-89.

Crucés Colado, S. (2001). El origen de los errores en traducción. En E. Real, D. Jiménez, D. Pujante y A. Cortijo (Eds.), *Écrire, traduire et représenter la fête* (pp. 813-822). Universitat de València.

Cubo de Severino, L., et al. (2007). *Los textos de la ciencia. Principales clases del discurso científico*. Comunic-arte Editorial.

de Almeida, G. (2013). *Translating the post-editor: An investigation of post-editing changes and correlations with professional experience across two Romance languages* [Tesis doctoral]. School of Applied Language and Intercultural Studies.

Frapolli, D. (2025). *docx-highlighter-processor*. GitHub. https://github.com/DiamelaFrapolli/docx-highlighter-processor

(en evaluación). *Python* para el análisis de corpus. *Actas de las Xmas. Jornadas de Investigación en Humanidades*. Departamento de Humanidades.

EAP Foundation (2024). *Vocabulary profiler*. Recuperado el 25 de noviembre de 2024, de https://www.eapfoundation.com/vocab/profiler/

Ezpeleta Piorno, P. (2008). El informe técnico. Estudio y definición del género textual. En *La traducción del futuro: mediación lingüística y cultural en el siglo XXI*,1, 429-438. https://dialnet.unirioja.es/servlet/articulo?codigo=5660335

García Negroni, M. M., Hall, B., y Marín, M. (2005). Ambigüedad, abstracción y polifonía del discurso académico: Interpretación de las nominalizaciones. *Revista signos*, 38(57), 49-60. https://doi.org/10.4067/S0718-09342005000100004

García Negroni, M. M., y Estrada, A. (2006). ¿Corrector o corruptor? Saberes y competencias del corrector de estilos. *Páginas de guarda: Revista de lenguaje, edición y cultura escrita*, 1, 26-40. https://dialnet.unirioja.es/servlet/articulo?codigo=2014997

Gelbes, S. P. R. (2016). Una pista polifónico-argumentativa para interpretar el agente en pasivas e impersonales con se. *Letras de Hoje*, 51(1), 17-28. https://doi.org/10.15448/1984-7726.2016.1.21629

Google. (2006). *diff-match-patch*. GitHub. Recuperado el 16 de junio de 2025, de https://github.com/google/diff-match-patch

Google Translate (Versión 2023). (2003). [Software]. Google Corporation. https://translate.google.com.ar

Grammarly Inc (2023). My Grammarly (2023) [Software] https://www.grammarly.com

Grammarly Inc. (2023). About. https://www.grammarly.com/about

Greene, J. O., y Burleson, B. (Eds.). (2003). *Handbook of communication and social interaction skills*. Erlbaum.

Hall, B., y López, M. I. (2011). Discurso académico: manuales universitarios y prácticas pedagógicas. *Literatura y lingüística*, 23, 167-192. https://doi.org/10.4067/S0716-58112011000100010

Halliday, M. A. K., y Martin, J. R. (1993). *Writing science: literacy and discursive power*. The Falmer Press.

Halliday, M. A. K., y Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3a. ed.). Hodder Education. (Trabajo original publicado en 1985).

Hernández Campoy, J. M., y Almeida, M. (2005). *Metodología de la investigación sociolingüística*. Comares.

Hillage, J., y Pollard, E. (1998). *Employability: Developing a framework for policy analysis* (Informe N° 85). Institute for Employment Studies. https://www.employment-studies.co.uk/resource/employability-developing-framework-policy-analysis

Holtz, M. (20-23 de julio de 2009). *Nominalisation in scientific discourse: A corpus-based study of abstracts and research articles*. En Michaela Mahlberg, Victorina González-Díaz y Catherine Smith (coords.), *Proceedings of CL2009*, p. 341. https://www.researchgate.net/publication/290488300_Nominalisation_in_scientific_discourse_A_corpus-based_study_of_abstracts_and_research_articles

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hyland, K. (2000). *Disciplinary discourses: social interactions in academic writing*. Longman.

Hymes, D. (1972). On Communicative Competence. En J. B. Pride, y J. Holmes (Eds.), *Sociolinguistics. Selected Readings* (pp. 269-293). Penguin Books.

Khudyakova, T., y Filatova, L. (2013). The Formation of a Professional Communication Competence in Psychologists. *Procedia - Social and Behavioral Sciences*, 86, 224-227. https://doi.org/10.1016/j.sbspro.2013.08.554

Kim, S.-G. (2023). Using ChatGPT for language editing in scientific articles. *Maxillofacial Plastic and Reconstructive Surgery*, 45(1). https://doi.org/10.1186/s40902-023-00381-x

Kol, S., Schcolnik, M., y Spector-Cohen, E. (2018). Google Translate in Academic Writing Courses? *The EuroCALL Review*, 26(2), 50-57. https://doi.org/10.4995/eurocall.2018.10140

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., y Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. arXiv. https://doi.org/10.48550/arXiv.2506.08872

Koponen, M., Salmi, L., y Nikulin, M. (2019). A product and process analysis of post-editor corrections on neural, statistical and rule-based machine translation output. *Machine Translation*, 33(1), 61-90. https://doi.org/10.1007/s10590-019-09228-7

Koutsantoni, D. (2007). *Developing academic literacies*. Peter Lang.

Labov, W. (1966). *The social stratification of English in New York City*. Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Lacorte, M. (2007). *Lingüística aplicada del español*. Arco/Libros.

Laufer, B., y Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322. https://doi.org/10.1093/applin/16.3.307

Llisterri, J. (2003). Lingüística y tecnologías del lenguaje. *LynX. Panorámica de Estudios Lingüísticos*, 2, 9-71.

Lechien, J. R., Gorton, A., Robertson, J., y Vaira, L. A. (2023). Is ChatGPT-4 accurate in proofreading manuscript in otolaryngology-head and neck surgery? *Otolaryngology-Head and Neck Surgery*, 0(0), 1-4. https://doi.org/10.1002/ohn.526

Lopezosa, C. (2023). ChatGPT y comunicación científica: Hacia un uso de la Inteligencia Artificial que sea tan útil como responsable. *Hipertex.net*, 26, 17-21. https://doi.org/10.31009/hipertext.net.2023.i26.03

McEnery, T., y Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Montolío, E. (Coord.). (2003). *Manual práctico de escritura académica*. Ariel.

Navarro, F. (2017). De la alfabetización académica a la alfabetización disciplinar. En R. Ibáñez y C. González (Eds.), *Alfabetización disciplinar en la formación inicial docente: Leer y escribir para aprender* (pp. 7-15). Ediciones Universitarias de Valparaíso.

Ne'Matullah, K. F., Seong Pek, L., y Roslan, S. A. (2023). Speaking English in job interviews increases employability opportunities: Malaysian employer's perspectives. *International Journal of Language, Literacy and Translation*, 6(2), 166-178. https://doi.org/10.36777/ijollt2023.6.2.085

Nova, M. (2018). Utilizing Grammarly in evaluating academic writing: A narrative research on EFL students' experience. *Journal of English Education and Applied Linguistics*, 7(1), 80-96.

Olmo, S. O. del. (2006). The Role of Passive Voice in Hedging Medical Discourse: A Corpus-based Study on English and Spanish Research Articles. *Revista de Lenguas para Fines Específicos*, 12, 205-218.

O'Neill, R., y Russell, A. (2021). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), 42-56. https://doi.org/10.14742/ajet.3795

OpenAI (2023). ChatGPT (Versión del 25 de septiembre de 2023) [Software] https://chat.openai.com/

Paredes García, F. (2009). *Guía práctica del español correcto*. Guías prácticas del Instituto Cervantes (pp. 16-24). Espasa-Calpe.

Patil, S., y Davies, P. (2014). Use of Google Translate in medical communication: Evaluation of accuracy. *The BMJ*, 349(dec15 2). https://doi.org/10.1136/bmj.g7392

Real Academia Española. (2009). La derivación nominal (I): Nombres de acción y efecto. *En Nueva gramática de la lengua española* (pp. 337-411). Espasa.

Rickheit, G., y Strohner, H. (Eds.). (2008). *Handbook of communication competence*. Walter de Gruyter.

Robblee, S. K. (2017). Identifying editing strategies for grant proposals: Results of coding editors' approaches to comments, comment types and edit types, (Doctoral dissertation, Texas Tech University).

Rojas Castro, A. (2013). Las Humanidades Digitales: principios, valores y prácticas. *JANUS*, 2, 74-99. Recuperado de http://www.janusdigital.es/articulo.htm?id=24

Schuster, M., Johnson, M., y Thorat, N. (2016, noviembre 22). Zero-Shot Translation with Google's Multilingual Neural Machine Translation System. *Google Research* https://ai.googleblog.com/2016/11/zero-shot-translation-with-googles.html

Sinclair, J. M., y Carter, R. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Smith, R. (2005). The lexical frequency profile: Problems and uses. En K. Bradford-Watts, C. Ikeguchi, y M. Swanson (Eds.), *JALT2004 Conference Proceedings*, 439-451.

Spitzberg, B. H., y Cupach, W. R. (1984). *Interpersonal communication competence*. Sage.

Subagio, U., Prayogo, J. A., e Iragiliati, E. (2019). Investigation of passive voice occurrence in scientific writing. *International Journal of Language Teaching and Education*, 3(1), 61-66. https://doi.org/10.22437/ijolte.v3i1.7434

Stubbs, M. (1983). *Discourse analysis: The sociolinguistic analysis of natural language*. Basil Blackwell.

Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Wiley-Blackwell.

Swales. (1990). The concept of genre: English in Academic and Research Settings. En *Genre Analysis* (pp. 33-67). Cambridge University Press.

Starc, M., y Martino, A. (2025). Percepciones de los estudiantes de grado acerca de la comunicación oral en su futuro desempeño profesional. En Negrin M., Randazzo M. B. y Barelli S. G. (coords), *IX Jornadas de Investigación en Humanidades*, (pp. 104-117), EdiUNS.

Tapia-Ladino, M., y Burdiles Fernández, G. (2009). Una caracterización del género informe escrito. *Letras*, 51(78), 17-49.

Van Dijk, T. A. (Ed.). (1985). *Handbook of discourse analysis. Volume 2: Dimensions of discourse*. Academic Press.

Villayandre Llamazares, M. (2010). *Aproximación a la lingüística computacional* [Tesis doctoral]. Universidad de León. https://doi.org/10.18002/10612/903

Wates, E., y Campbell, R. (2007). Author's version vs. publisher's version: An analysis of the copy-editing function. *Learned Publishing*, 20(2), 121-129. https://doi.org/10.1087/174148507X185090

West, M. (1953). *A general service list of English words*. Longman, Green and Co.

Wiemann, J. M. (1977). Explication and test of a model of communicative competence. *Human Communication Research*, 3(3), 195-213. https://doi.org/10.1111/j.1468-2958.1977.tb00518.x

Wiles, R., Heath, S., Crow, G., y Charles, V. (2005). Informed Consent in Social Research: A Literature Review (Documento de revisión metodológica n.º NCRM/001). *National Centre for Research Methods*. https://eprints.ncrm.ac.uk/id/eprint/85

Włoszczak-Szubzda, A., y Jarosz, M. J. (2012). Professional communication competences of nurses - a review of current practice and educational problems. *Annals of Agricultural and Environmental Medicine*, 19(3), 601-607.

Yatsko, V. A. (2014). Computational linguistics or linguistic informatics? *Automatic Documentation and Mathematical Linguistics*, 48(3), 149-157. https://doi.org/10.3103/S0005105514030042