

La traducción automática en el ámbito de la atención al cliente: percepciones de los estudiantes de traducción sobre la idiomática*

Machine translation and customer service: How do translation trainees perceive idiomaticity?

AURORA MARTÍN DE SANTA OLALLA SÁNCHEZ

Universidad Complutense de Madrid. Facultad de Filología. Departamento de Estudios Románicos, Franceses, Italianos y Traducción. Edificio D. Calle Profesor Aranguren, s/n. 28040 Madrid. España.

Dirección de correo electrónico: aurosant@ucm.es

ORCID: <https://orcid.org/0000-0002-4907-302X>

CELIA RICO PÉREZ

Universidad Complutense de Madrid. Facultad de Filología. Departamento de Estudios Románicos, Franceses, Italianos y Traducción. Edificio D. Calle Profesor Aranguren, s/n. 28040 Madrid.

Dirección de correo electrónico: celrico@ucm.es

ORCID: <https://orcid.org/0000-0002-5056-8513>

Recibido/Received: 27/12/2023. Aceptado/Accepted: 31/7/2024.

Cómo citar/How to cite: Martín de Santa Olalla Sánchez, Aurora y Celia Rico Pérez, Nombre, «La traducción automática en el ámbito de la atención al cliente: percepciones de los estudiantes de traducción sobre idiomática», *Hermēneus. Revista de Traducción e Interpretación*, 27 (2025): pp. 401-445.

DOI: <https://doi.org/10.24197/jbm77c22>

Artículo de acceso abierto distribuido bajo una [Licencia Creative Commons Atribución 4.0 Internacional \(CC-BY 4.0\)](#). / Open access article under a [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](#).

Resumen: Presentamos los resultados de un estudio que analiza la percepción de la idiomática en un conjunto de oraciones traducidas mediante un motor neuronal del inglés al español, por parte de dos grupos de estudiantes de traducción. Partimos de un corpus de cincuenta oraciones y un cuestionario con una pregunta en la que los estudiantes tienen que valorar la naturalidad de las oraciones y otra en la que se les pide reformular la oración española. Las respuestas nos llevarán a un análisis de carácter cuantitativo que muestra las diferencias en la percepción de la idiomática y otro cualitativo que muestra el desempeño en una muestra de oraciones. Los

* Esta publicación es parte del proyecto de I+D+i ROBOT-TALK PID2022-140897OB-I00 financiado por MCIN/AEI/10.13039/501100011033/ y FEDER/UE.

datos globales muestran la confianza en los resultados de la traducción automática y la variabilidad a la que está sujeta la idiomaticidad.

Palabras clave: Traducción automática neuronal; atención al cliente; idiomaticidad; competencias de traducción; TA y umbrales de aceptabilidad.

Abstract: The article analyzes the perception of idiomaticity and machine translation acceptability thresholds in customer service. A corpus of fifty sentences was first translated from English into Spanish using a neural machine translation system. Then two groups of translation students reviewed this corpus and evaluated the naturalness of the translations, providing their answers in a two-question form which gathered both qualitative and quantitative data. The overall results show the confidence in the results of machine translation and how idiomaticity is subject to variable perception.

Keywords: Neural machine translation; customer service; idiomaticity; translation skills; MT and acceptability thresholds.

Sumario: Introducción; 1 El concepto de idiomaticidad en el campo de la atención al cliente; 2. Diseño de la investigación y metodología de análisis, 2.1. Corpus de análisis, 2.2. Participantes, 2.3. Instrumento para la recogida de datos, 2.4. Metodología de análisis; 3. Resultados, 3.1. Resultados sobre la percepción de la naturalidad (pregunta 1), 3.2. Resultados sobre la posibilidad de registrar una opción mejor a la de la traducción automática (pregunta 2), 3.2.1. Falta de idiomaticidad por interferencias léxicas: oraciones con calcos o traducciones literales, 3.2.2. Falta de idiomaticidad por interferencias en el nivel morfosintáctico, 3.2.3. Falta de idiomaticidad por cuestiones pragmática: la cortesía de las peticiones, 3.2.4. Falta de idiomaticidad por cuestiones culturales, 3.2.5. Falta de idiomaticidad por cuestiones léxicas relacionadas con la variedad latinoamericana (Latam); 4. Discusión; 5. Conclusiones; Referencias bibliográficas.

Summary: Introduction; 1. The concept of idiomaticity in customer service; 2. Research design and methodology, 2.1. Corpus of analysis, 2.2. Participants, 2.3. Instrument for data collection, 2.4. Methodology for analysis; 3. Results and discussion, 3.1. Results on the perception of naturalness (question 1), 3.2. Results on the possibility of registering a better option than the machine translation (question 2), 3.2.1. Lack of idiomaticity due to lexical interference: sentences with calques or literal translations, 3.2.2. Lack of idiomaticity due to interference at the morphosyntactic level, 3.2.3. Lack of idiomaticity due to pragmatic issues: politeness of requests, 3.2.4. Lack of idiomaticity due to cultural issues, 3.2.5. Lack of idiomaticity due to lexical issues related to the Spanish Latin American variety (Latam); 4. Discussion; 5. Conclusions; References.

INTRODUCCIÓN

Con el crecimiento de la comunicación digital tanto en dispositivos electrónicos como en plataformas digitales ha surgido toda una serie de nuevos géneros textuales que, de algún modo, nos llevan a pensar en una comunicación mediada por ordenador. Así, como consecuencia de la digitalización de las estrategias discursivas, cambia también «el modo en que adquieren nuevas formas o funciones los paradigmas léxicos, pragmáticos, semánticos y sintácticos» (Vargas-Sierra y Sánchez, 2022, p. 8) y se hace necesario estudiar cómo los nuevos contextos digitales reflejan las variaciones específicas que se derivan del uso en diferentes situaciones

comunicativas. Uno de los campos más productivos en este sentido es el de los contenidos generados por los usuarios en áreas como la mercadotecnia, el turismo, las tecnologías de la información o los servicios en Internet, áreas estas en las que las opiniones que los consumidores vierten en las plataformas digitales afectan directamente a los procesos de compra y superan incluso, a veces, el impacto de las campañas publicitarias (Candel-Mora, 2022, p.117). Este efecto directo de la opinión del consumidor en las decisiones de compra se muestra claramente en la atención al cliente, donde las expectativas de los usuarios con respecto a la experiencia que proporciona la empresa son casi tan importantes como los productos o servicios que esta ofrece. Así lo indica, por ejemplo, el informe de Salesforce (SalesForce, 2020) sobre la experiencia del cliente, en el que el 80 % de los clientes encuestados espera que la empresa comprenda sus necesidades y que esta adapte el modo en que se comunica a la hora de resolver dudas y problemas, de manera que se establece una relación de confianza única.

El género textual de la atención al cliente se ha abordado poco en los estudios de traducción a pesar de su creciente importancia en la industria. Cuando una empresa entra en el mercado internacional, además de identificar a sus clientes potenciales, necesita saber qué les motiva a comprar un producto o servicio. Por este motivo resulta clave que una empresa pueda ofrecer a sus clientes una atención personalizada en diferentes idiomas de modo que los clientes internacionales puedan comunicar fácilmente sus necesidades y recibir ayuda en su lengua materna. En este punto, es necesario tener en cuenta que la atención al usuario es el campo en el que más se utiliza la traducción automática (TAUS, 2020), con un porcentaje de proyectos que la usan cercano al 100 %. Se trata de un nuevo mercado de contenidos generados por los usuarios en los que la traducción automática permite una traducción instantánea de contenidos efímeros y que, por su volumen y urgencia de publicación, necesitan una traducción inmediata.

En este contexto, el estudio que planteamos tiene como objetivo analizar la percepción de los estudiantes de traducción con respecto a la idiomática de un conjunto de textos del ámbito de la atención al cliente traducidos de manera automática. Los resultados del estudio contribuyen a situar al estudiante en el centro del proceso de aprendizaje en una conceptualización transversal de la tecnología que va más allá del mero aprendizaje de las particularidades técnicas de la herramienta de traducción. Siguiendo las propuestas de Rico y González (2022), nuestro

trabajo parte de la idea de que el factor humano es la clave del proceso de traducción y que, por tanto, es necesario que los estudiantes entiendan el cambio de los procesos más tradicionales (traducción-edición-revisión) al nuevo modelo de traducción que incorpora la traducción automática y la posesición. En este marco la alfabetización en traducción automática y el conocimiento de sus posibilidades y limitaciones forman parte integrante de la competencia traductora profesional, de manera que la revisión y la posesición comparten competencias tanto desde el punto de vista conceptual como en la práctica (Konttinen, Salmi y Koponen, 2021). Además, esta investigación se enmarca en una línea de trabajo más amplia relacionada con los cambios que la traducción automática, y más específicamente la traducción automática neuronal, introduce en la docencia de la traducción. Este marco de trabajo nos lleva a plantear la necesidad de que los estudiantes analicen con espíritu crítico los textos producidos por la máquina, de manera que esa reflexión sirva para capacitar a los estudiantes, ayudándoles a definir el valor añadido que aporta el factor humano con respecto a los sistemas de traducción automática.

1. EL CONCEPTO DE IDIOMATICIDAD EN EL CAMPO DE LA ATENCIÓN AL CLIENTE

El ámbito de la atención al cliente tiene unas características peculiares que vienen determinadas por el tipo de comunicación bilateral que se da entre los llamados «agentes de atención al cliente» y los usuarios o clientes. La función de los agentes es contestar preguntas o atender a las reclamaciones de los clientes, procesar pedidos y devoluciones y proporcionar información sobre los productos y servicios. En esta tarea, los agentes deben seguir un protocolo específico marcado por la empresa en el que se puede incluir el uso de plantillas, así como de terminología específica que impiden toda variación estilística y que pueden dar como resultado la repetición de estructuras en el proceso comunicativo. A esto hay que añadir el hecho de que los agentes trabajan en un entorno de estrés, lo que los lleva a cometer errores ortotipográficos con cierta frecuencia. En el otro extremo de la cadena de comunicación, el contenido que generan los clientes no sigue ninguna estructura fija, es muy variado e incluye el uso de abreviaturas, emoticonos y expresiones idiomáticas, sin olvidar los errores derivados de escribir en el teclado del móvil (Gonçalves *et al.*, 2022). Desde el punto de vista de la intencionalidad, este contenido es

emocionalmente volátil y efímero. Los clientes suelen ponerse en contacto con el servicio de atención al cliente para quejarse de un producto o un servicio y pueden mostrar altos niveles de impaciencia y frustración, lo cual se refleja en elecciones léxicas únicas, como el uso de blasfemias, el uso variable de mayúsculas y la puntuación. Cuando la atención al cliente se da a través de un chat *online* multilingüe que presta ayuda a clientes en diferentes idiomas, la cuestión es más compleja puesto que la respuesta debe ser inmediata y el contenido queda fragmentado en los diferentes mensajes que intercambia el usuario con el agente. Es en este contexto cuando se opta por la traducción automática como un modo de responder a esta necesidad de inmediatez.

Para que el contenido que se genera en el chat de atención al cliente pueda traducirse automáticamente de manera eficaz, los agentes siguen las siguientes de recomendaciones (Graça, 2022):

- a. Usar frases cortas de no más de treinta palabras y que expresan una única idea.
- b. Evitar los modismos y la fraseología.
- c. Utilizar la voz activa antes que la pasiva.
- d. No utilizar abreviaturas.
- e. Prestar una atención especial al uso correcto de los espacios en blanco, los signos de puntuación, así como a la concordancia de número, género y tiempo.
- f. Formatear correctamente los mensajes en cuestiones como los saltos de línea y los saltos de párrafo innecesarios, que pueden confundir al algoritmo de traducción.
- g. Emplear un registro formal para dirigirse al cliente.

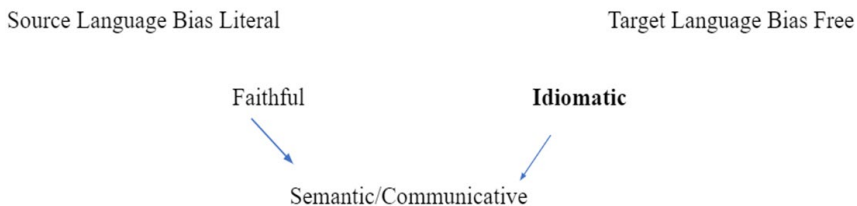
Con el fin de ilustrar el tipo de frases que manejan los agentes de atención al cliente mostramos, a continuación, algunos ejemplos en inglés tomados de nuestro corpus de análisis:

- a. The order has been cancelled.
- b. May I ask for the order number?
- c. I will arrange for a reshipment at no cost.

Ahora bien, conviene tener en cuenta que, a pesar de estas recomendaciones, la traducción automática no siempre da el resultado esperado y que, en ocasiones, tal como demuestra Toral (2019), este resultado puede interferir directamente en cómo se lleva a cabo la

posedición. Cuando esto ocurre, la huella de la traducción automática puede quedar visible en cuestiones como la densidad y variedad léxica, el uso de determinadas secuencias sintácticas o la longitud de las oraciones. Todo ello puede tener consecuencias importantes para la lengua meta a largo plazo que, por ejemplo, puede llegar a empobrecerse bien como efecto de la simplificación o por una interferencia excesiva de la lengua de partida (Daems *et al.*, 2017).

En el caso del trabajo que presentamos aquí, recogemos las percepciones de los estudiantes con respecto al resultado de la traducción automática y, concretamente, en la percepción sobre la idiomática. Nuestro punto de partida es la definición de idiomática como la cualidad de lo que se reconoce no solo como gramatical y entendible en una lengua, sino como habitual, natural o no marcado. Es este un concepto clásico en teoría de la traducción que nos remite al tradicional esquema de Newmark (1988), que sitúa la traducción idiomática en el ámbito de la traducción comunicativa del siguiente modo:



Fuente: Newmark (1988, pp. 39-41). La negrita es nuestra.

Y que define en los siguientes términos:

(...) a communicative translation is likely to be smoother, simpler, clearer, more direct, more conventional, conforming to a particular register of language, tending to undertranslate (...). A semantic translation tends to be more complex, more awkward, more detailed, more concentrated, and pursues the thought-process rather than the intention of the transmitter. It tends to overtranslate (Newmark, 1988, p. 39).

Esta definición de idiomática contrasta, en cierto modo, con la concepción que se da al término en el campo de la fraseología, donde el este cobra una mayor especificidad y se concreta en diferentes definiciones (Fan, 2018). Por una parte, Zuluaga (1980, p. 122) se refiere a la

idiomaticidad como «el rasgo semántico propio de ciertas construcciones lingüísticas fijas, cuyo sentido no se puede establecer a partir de los significados de sus elementos componentes ni del de su combinación». De manera complementaria, Corpas (1996, p. 26) se refiere a este concepto como «aquella propiedad semántica que presentan ciertas unidades fraseológicas, por la cual el significado global de dicha unidad no es deducible del significado aislado de cada uno de sus elementos constitutivos». Desde esta perspectiva, podemos decir entonces que, en el campo de la fraseología, las unidades fraseológicas pueden presentar, por una parte, un significado denotativo literal y, por otra, el significado denotativo figurativo o idiomático. Mellado (2016) recoge esta idea y la explica en relación con la divisibilidad de la cadena fraseológica y a la autonomía semántica de los constituyentes del fraseologismo, de manera que la idiomaticidad se considera como «una magnitud gradual relacionada con la composicionalidad semántica del significado fraseológico» (Mellado, 2016, p. 229). Así, por ejemplo, consideraríamos la unidad fraseológica *entrarle algo/alg. a alg. por el ojo derecho* (Coulmas, 1982, p. 19) como totalmente idiomática porque su significado idiomático no es la suma de los significados libres de los constituyentes. Tendríamos una idiomaticidad parcial, por ejemplo, en una expresión como *hablar por los codos* (Mellado, 2016, p. 238), puesto que su paráfrasis contiene un lexema que coincide con parte de la unidad fraseológica (*hablar*).

En el ámbito de la traducción automática, los estudios sobre la idiomaticidad son aún incipientes y se han ocupado de cuestiones como el trasvase de locuciones nominales del español al inglés (Ortego y Seghiri, 2019), el análisis de la fraseología en el área de la propiedad intelectual e industrial (Alemán, 2018), la traducción de paremias y locuciones verbales del ruso al español para la enseñanza de ELE (Ramírez, 2022), la definición de métricas automáticas para la evaluación de las expresiones idiomáticas (Baziotis *et al.*, 2022) o la creación de un corpus paralelo inglés-alemán de frases idiomáticas para el entrenamiento de motores de traducción automática neuronal (Marzieh Fadaee *et al.*, 2018). En todo caso, conviene recordar que el tratamiento de las expresiones idiomáticas es aún uno de los grandes desafíos a los que se enfrenta la traducción automática, en primer lugar porque los corpus de entrenamiento no suelen contener este tipo de expresiones en una cantidad suficientemente representativa y, en segundo lugar, porque los modelos de inteligencia artificial tienen dificultades para distinguir entre una expresión idiomática

y otra cuyo significado debe interpretarse de modo literal (Parmar, 2020 y Santing *et al.*, 2022).

Sin entrar necesariamente en el campo de la fraseología, nuestro trabajo orienta lo idiomático hacia el destinatario y el texto meta, a diferencia de la fidelidad, que se orienta hacia la preservación del texto original. Lo idiomático remite a lo que es más genuino, a lo más original, a lo más natural en la lengua de llegada y se refiere a cuestiones léxicas, sintácticas (la pasiva refleja es más genuinamente española que la pasiva) o pragmáticas (el español utiliza en las peticiones fórmulas directas a diferencia del inglés donde predominan las indirectas).

Lo idiomático está sujeto a una cierta variabilidad condicionada por aspectos como la edad, las variables socioculturales o incluso el contexto en el que se produce un enunciado, como veremos en esta investigación. Esta es una definición del concepto de idiomaticidad que está en consonancia con el planteamiento que hacen Laver y Mason (2018, p. 62): «*Idiomatic translation* n. an approach to translating which, while not sacrificing source-text meaning, seeks to ensure naturalness of expression in the target language».

2. DISEÑO DE LA INVESTIGACIÓN Y METODOLOGÍA DE ANÁLISIS

2. 1. Corpus de análisis

El estudio parte de un corpus de cincuenta oraciones extraídas de comunicaciones reales, totalmente anonimizadas, de diferentes empresas a sus clientes en forma de mensajes de correo electrónico, chats o FAQ. Estas oraciones se han traducido automáticamente del inglés al español mediante los motores de traducción automática neuronal de Unbabel,¹ gracias a un acuerdo de colaboración con esta empresa.

El corpus completo está disponible en acceso abierto en Rico y Martín de Santa Olalla (2024). Para nuestra investigación hemos agrupado el conjunto de las oraciones del siguiente modo: oraciones idiomáticas y oraciones no idiomáticas.²

¹ Unbabel es una multinacional portuguesa líder en la utilización de motores de traducción automática neuronal en el ámbito de la atención al cliente.

² Algunas traducciones podrían parecer *a priori* poco naturales fuera de contexto, pero debe tenerse en cuenta que estos resultados están dentro del estándar de la industria cuando se utiliza TA.

En primer lugar, el grupo de oraciones idiomáticas comprende nueve oraciones, entre las que podemos destacar las siguientes:

Original en inglés	Traducción automática al español
Thanks for getting in touch.	Gracias por ponerte en contacto con nosotros.
We're sorry to hear that you are experiencing issues when attempting to access Online features.	Lamentamos que tengas problemas al intentar acceder a las funciones en línea.
We will be happy to help you.	Estaremos encantados de ayudarte.

En segundo lugar, el grupo de oraciones no idiomáticas está compuesto por cuarenta y una oraciones con diferentes problemas relacionados todos ellos con la idiomática. La identificación del problema que causa la falta de idiomática nos ha permitido distinguir cinco subgrupos dentro de las oraciones no idiomáticas. En el grupo 1, el problema tiene que ver con la traducción de cognados, más concretamente, la preferencia por palabras que se parecen desde el punto de vista ortográfico frente a otras más alejadas ortográficamente, pero más naturales, más idiomáticas en español. Como ejemplo, obsérvese la traducción de *great*, *additional* o *rare* cases en las oraciones siguientes.

Original en inglés	Traducción automática al español
Have a great day Lucía!	¡Qué tengas un gran día, Lucía!
Can I assist with anything additional at the moment?	¿Puedo ayudarte con algo adicional en este momento?

En el grupo 2, la falta de idiomática está relacionada con cuestiones de carácter morfosintáctico, en la traducción al español se opta por seguir la literalidad de la oración en inglés: así en los siguientes ejemplos se prefiere la pasiva a la pasiva refleja o el posesivo frente al artículo definido, opciones estas dos últimas mucho más idiomáticas en español.

Original en inglés	Traducción automática al español
The payment was returned to the sender.	El pago fue devuelto al remitente.

Once we received this information, we will be able to investigate your issue further.	Una vez que recibamos esta información, podremos investigar tu problema más a fondo.
--	---

En el grupo 3, la falta de idiomática está relacionada con cuestiones pragmáticas: oraciones en las que se trasvasa a la traducción las fórmulas indirectas de cortesía del inglés en casos en los que, en español, resultaría más idiomático la utilización de una fórmula directa.

Original en inglés	Traducción automática al español
May I ask for the order number.	¿ Puedo solicitar el número de pedido?

En el grupo 4 la falta de idiomática está relacionada con cuestiones culturales: oraciones que resultan poco naturales para un hablante de español peninsular. Este tipo de expresiones están muy relacionadas con la manera de entender la atención al cliente en el ámbito anglosajón o más específicamente en los Estados Unidos (Morate, 2018).

Original en inglés	Traducción automática al español
Be safe, stay healthy and take care.	Estate seguro, mantente saludable y cuídate.
We're sorry your product arrived in this condition and we apologize for any disappointment caused.	Lamentamos mucho que su producto haya llegado en esa condición y nos disculpamos por cualquier decepción causada.
I've successfully updated the shipping address.	He actualizado con éxito la dirección de envío.

Finalmente, nuestro corpus incluye oraciones que presentan falta de idiomática por incluir fenómenos léxicos y sintácticos que son más propios de la variedad que la industria ha dado en llamar latinoamericana (Latam, de forma abreviada) en oposición al español europeo.³

³ La distinción de latinoamericano frente a español europeo es simplificadora y podría considerarse poco rigurosa. La realidad dialectológica del español es, sin duda, mucho más compleja. Prueba de ello son las diferentes propuestas de división del continuo

En este caso, la falta de idiomática tiene que ver con la extrañeza con la que puede percibir un español de la variedad centro norte peninsular el uso de léxico o de construcciones ajenas a su variedad. Ocurriría lo mismo a la inversa, para un mexicano, un colombiano o un peruano resultarían extraños ciertos usos, por diferentes a los de su variedad, sin que eso suponga ningún juicio de valor hacia una u otras variedades.

Original en inglés	Traducción automática al español
However, the serial number you've provided won't authenticate on our end.	Sin embargo, el número de serie que nos ha facilitado no se autenticará en nuestro sistema.
The refund arrived in your account today.	El reembolso llegó a tu cuenta hoy.
I am sorry to hear you're having problems with the app.	Lamento que estés teniendo problemas con la aplicación.

En el primer ejemplo, nos encontramos con el verbo, *autenticar*, que se usa, sobre todo, en América. Efectivamente, el *Diccionario panhispánico de dudas*⁴ recoge, en la entrada de *autenticar*, el significado de 'Certificar la autenticidad [de algo, especialmente un documento]' (Real Academia Española / Asociación de Academias de la Lengua, s. v. *autenticar*) y explica que esta forma es la más usada en América, mientras que, en España, se prefiere el verbo *autenticar*.

En el segundo caso, se pone de manifiesto la utilización del pretérito indefinido para referirnos a hechos recientes. Según la *Nueva Gramática de la lengua española* (Real Academia, 2009), la utilización del pretérito perfecto compuesto o del indefinido para referirnos a hechos recientes está delimitada geográficamente tanto en España como en América.

En la variedad que la industria ha dado en llamar Latam, se utiliza el pretérito indefinido, mientras que, en el llamado español europeo, se prefiere el pretérito perfecto compuesto. Finalmente, en el tercer caso, nos encontramos con la perífrasis *estar* + gerundio que es una de las más habituales del español, tanto peninsular como americano.

dialectal del español que se han realizado desde hace más de 150 años. Por lo innovador de su aproximación y por la actualidad de su análisis, recomendamos al lector interesado la lectura de Moreno y Ueda (2018).

⁴ Véase <https://www.rae.es/dpd/autenticar>

Como ha señalado García (2012), en América, al aspecto durativo de esta perífrasis se ha sumado otro de carácter puntual que es el que se pone de manifiesto en este ejemplo en el que el español europeo preferiría una forma de presente, en este caso, de subjuntivo (*Lamento que tengas problemas con la aplicación*).

2.2. Participantes

Los participantes en el estudio son alumnos del Grado de Traducción e Interpretación de la Universidad Complutense de Madrid divididos en dos grupos. El grupo 1 está integrado por treinta y cuatro estudiantes de la asignatura «Redacción y composición de textos en español».

Esta asignatura es la segunda dedicada a la expresión escrita en español y se cursa en el primer cuatrimestre del segundo curso. En ella, se abordan, entre otros contenidos, la corrección gramatical y estilística. Los estudiantes están habituados a corregir textos, pero escritos siempre en español, no traducidos al español. En este momento de sus estudios todavía no han cursado ninguna asignatura de traducción.

El grupo 2 está compuesto por treinta y siete estudiantes de la asignatura «Tecnología de la traducción asistida por ordenador», que se cursa en el primer cuatrimestre del tercer año del grado. Esta asignatura es la continuación de otras dos dedicadas a la tecnología, «Tecnologías Aplicadas a la Traducción y la Interpretación I», que se imparte en primer curso y «Tecnologías Aplicadas a la Traducción y la Interpretación II», que se imparte en segundo curso.⁵ Los estudiantes de este segundo grupo están familiarizados con la traducción automática desde el punto de vista de un usuario avanzado, conocen cómo funcionan los motores de traducción neuronal, han hecho prácticas de evaluación manual y automática y han recibido formación teórica y práctica sobre la posesición.

2.3. Instrumentos para la recogida de datos

Para el proceso de recogida de datos se crearon dos cuestionarios, uno para cada grupo, en el que se preguntaba a los participantes su percepción

⁵ El detalle completo de los objetivos, contenidos y competencias de las asignaturas del grado en Traducción e Interpretación de la Universidad Complutense de Madrid puede consultarse en el siguiente enlace: <https://www.ucm.es/estudios/grado-traduccioneinterpretacion>

sobre la idiomática de las cincuenta oraciones que componen el corpus. Los cuestionarios se diseñaron de manera que pudiéramos obtener datos cuantitativos y cualitativos, con preguntas cerradas (cincuenta, una por oración) y abiertas (cincuenta, una por oración). En el caso del grupo 1, estudiantes de la asignatura «Redacción y composición de textos en español», el cuestionario se diseñó como simulación de una tarea real, con las instrucciones de revisar las comunicaciones que los agentes del departamento de Atención al Cliente de una empresa multinacional dan al usuario a través de las respuestas en forma de correo electrónico, chats o FAQ. En esta tarea, los participantes solo tenían acceso a las oraciones traducidas automáticamente al español (es decir, no se les dio la oportunidad de ver las oraciones originales en inglés).

La tarea constaba de dos pasos. En primer lugar, los participantes debían valorar de 1 a 5 la naturalidad de las comunicaciones, siendo 1=nada natural y 5=muy natural. A continuación, se les pedía que, si lo creían necesario, escribiesen otra versión que, a su juicio, resultase más natural. De manera adicional, se les indicó que su valoración y propuesta de corrección debía estar basada en la variedad de español centro norte peninsular y que debían utilizar la segunda persona del singular, variante informal tú, para dirigirse al usuario.

El cuestionario del grupo 2, estudiantes de la asignatura «Tecnología de la traducción asistida por ordenador», constaba de dos preguntas iguales a las del grupo 1: una primera pregunta cerrada en la que se les pedía que indicaran de 1 a 5 su percepción sobre la naturalidad de una oración y otra pregunta abierta en la que se daba la posibilidad de registrar una opción mejor. La diferencia del cuestionario del grupo 2 con el del grupo 1 radica en dos aspectos. Por una parte, la tarea se plantea como un trabajo de posesión y, por otra, los estudiantes tienen acceso al texto original en inglés.

2.4. Metodología y criterios de análisis

Nuestro estudio recoge datos cuantitativos y cualitativos. En el primer caso, llevamos a cabo tres tipos de análisis con el software de tratamiento estadístico de datos SAS 9.5.⁶ La versión íntegra de estos estudios está

⁶ La parte estadística de este estudio se llevó a cabo con la inestimable colaboración de Ricardo García Mata, de los Servicios Informáticos de la Universidad Complutense de Madrid, a quien agradecemos su tiempo y su ayuda.

disponible en el repositorio de acceso abierto Zenodo (<https://zenodo.org/records/10953240>). A continuación, indicamos el detalle del tipo de análisis realizado:

1. Análisis descriptivo básico con dos tipos de resultados:
 - a. Resultados de las variables categóricas originales, esto es, las que recogen la composición de los dos grupos de estudiantes y las puntuaciones sobre su percepción de la idiomática, de 1 a 5, para cada oración.
 - b. Resultados de las variables numéricas calculadas, es decir, de las puntuaciones de las variables identificadas, a saber: puntuación de las oraciones idiomáticas, puntuación de las oraciones no idiomáticas, puntuación de las oraciones que tienen algún fenómeno relacionado con Latam. Para cada una de estas variables se proporciona la siguiente información: muestra útil, en nuestro caso, número de estudiantes (*N*), media (*Mean*), desviación estándar (*StdDev*), valor mínimo (*Min*), mediana (*Median*), cuartiles inferior y superior (*Lower/Upper Quartile*) y valor máximo (*Max*).

(Este tipo de análisis nos permite obtener los datos estadísticos básicos de nuestra investigación).

2. Análisis comparativo de las percepciones de idiomática entre los dos grupos de estudiantes para las cuatro variables descritas en el apartado anterior. A estos resultados se les aplica un *t-test* (o test de la *t* de *student*) con el fin de conocer si las diferencias son estadísticamente significativas.

(Este segundo análisis fundamenta las principales conclusiones de nuestro estudio: las diferencias, estadísticamente significativas, en la percepción de la idiomática en los dos grupos).

3. Análisis descriptivo, en forma de tabla, en la que se muestra el perfil de respuesta de cada grupo de estudiantes para cada una de las cincuenta oraciones que componen el corpus. En este estudio también se analiza el resultado de la aplicación del test no paramétrico de la Suma de Rangos de Wilcoxon que nos permitirá

ver si las diferencias entre los dos grupos, para cada una de las oraciones, son (o no) significativas.

(Este tercer análisis nos permite aportar datos cuantitativos a la aproximación cualitativa de una muestra de las oraciones que componen el corpus).

Como apuntamos en el apartado 2.3, el cuestionario para la recogida de datos contenía una segunda pregunta en la que se les pedía a los estudiantes que escribiesen otra versión de la oración en español que, a su juicio, resultase más natural. Esta propuesta de corrección debía estar basada en la variedad de español centro norte peninsular y debía utilizar la segunda persona del singular, variante informal tú, para dirigirse al usuario. Para presentar las respuestas de los estudiantes a esta pregunta, nos centramos aquí en los fenómenos de no idiomaticidad, con la selección de un conjunto de ocho oraciones que ilustran los siguientes fenómenos:

- a. Falta de idiomaticidad por cuestiones léxicas (calcos o traducciones literales).
- b. Falta de idiomaticidad por interferencias en el nivel morfosintáctico.
- c. Falta de idiomaticidad por cuestiones pragmáticas: la cortesía en las peticiones.
- d. Falta de idiomaticidad por cuestiones culturales.
- e. Falta de idiomaticidad por cuestiones léxicas relacionadas con la variedad Latam.

Para cada conjunto de estas oraciones, se mostrará:

- a. El *p*-valor, resultado de la aplicación del test no paramétrico de la Suma de Rangos de Wilcoxon.⁷ Cuanto menor sea el *p*-valor más significativa será la diferencia en los resultados del grupo 1 y del grupo 2.
- b. El porcentaje de estudiantes que detectan de manera correcta el problema que origina la falta de naturalidad y que, además, proporcionan una oración más natural (columna «desempeño»).

⁷ Recordemos que este valor nos permite determinar si las diferencias entre los dos grupos de estudiantes son o no significativas.

- c. El porcentaje de estudiantes que sobrecorrijen, entendiendo como sobrecorrección el cambio de palabras o estructuras que no plantean ningún problema en la oración original, así como los cambios de estilo no necesarios que, en ocasiones, suponen un cambio sustancial de significado en la oración original (columna «sobrecorrección»).

Con el fin de evaluar la variabilidad en la percepción de la idiomática en las respuestas de los dos grupos de estudiantes, empleamos los criterios que define el modelo MQM (Multidimensional Quality Metrics),⁸ que incluye en la categoría *estilo* la etiqueta *Unidiomatic Style*. Esta etiqueta se define en los siguientes términos: «Style that is grammatical, but unnatural, often due to interference from the source language»⁹. Esta aproximación está en consonancia con estudios similares como el realizado por Gonçalves *et al.* (2022) o Candel Mora (2022). La recogida de datos se llevó a cabo en dos sesiones consecutivas los días 24 (grupo 1 de estudiantes) y 29 de noviembre de 2022 (grupo 2 de estudiantes). Cada grupo completó los cuestionarios durante el tiempo de clase de las respectivas asignaturas con una duración total de dos horas por asignatura. Se recogieron un total de 34 respuestas en el grupo 1 y 37 respuestas en el grupo 2.

3. RESULTADOS

En este apartado presentamos los resultados y abordamos, en primer lugar, las respuestas a la primera pregunta sobre la percepción del grado de naturalidad y, en segundo lugar, las respuestas a la segunda pregunta, de carácter abierto en la que se daba a los estudiantes la posibilidad de registrar una opción mejor a la traducción automática.

3.1. Resultados sobre la percepción de la naturalidad (pregunta 1)

En la siguiente tabla hemos recogido las respuestas de cada uno de los dos grupos de estudiantes con respecto a su percepción de la naturalidad de cada tipo de oración: las idiomáticas, las no idiomáticas y, dentro de estas segundas, las de variedad de lengua Latam.

⁸ Véase <https://themqm.org/>

⁹ Véase <https://themqm.info/typology/>

GRUPO	Variable	N	Media	DT	Mín	Cuartil inferior	Mediana	Cuartil superior	Máx.
Estudiantes grupo 1	percepción de la naturalidad de las oraciones idiomáticas	34	4,05	0,55	2,43	3,75	4,22	4,44	4,78
	percepción de la naturalidad de las oraciones no idiomáticas	34	3,13	0,59	1,90	2,69	3,11	3,47	4,66
	percepción de la naturalidad sobre la variedad de lengua Latam	34	3,18	0,58	1,86	2,78	3,11	3,71	4,11
Estudiantes grupo 2	percepción de la naturalidad de las oraciones idiomáticas	37	4,36	0,36	3,33	4,11	4,44	4,56	5,00
	percepción de la naturalidad de las oraciones no idiomáticas	37	3,64	0,43	2,84	3,35	3,66	4,0	4,58
	percepción de la naturalidad sobre la variedad de lengua Latam	37	3,59	0,44	2,86	3,33	3,44	3,88	4,88

Tabla 1: Resultados de la pregunta 1: percepción de la naturalidad.

En la tabla, la columna *grupo* indica el grupo de estudiantes al que se refieren las respuestas: grupo de estudiantes 1 y grupo de estudiantes 2. Estas respuestas se refieren a tres grandes conjuntos de oraciones, recogidas en la columna *variable* y dividida en tres tipos: las oraciones idiomáticas, las oraciones no idiomáticas y las oraciones que contienen fenómenos relacionados con la variedad de lengua Latam. El resto de las columnas muestran los siguientes datos (Bryman, 2016, pp. 338-339):

1. N: número total de respuestas obtenidas para cada tipo de pregunta.
2. *Media*: media aritmética de las respuestas recibidas, esto es, la suma de todos los valores dividida entre el número total de valores.
3. DT: desviación estándar o variación en la que las respuestas individuales difieren de la media.

En la tabla hemos marcado en negrita los resultados de la media y la mediana y, en el siguiente gráfico, recogemos estos mismos valores de manera visual.

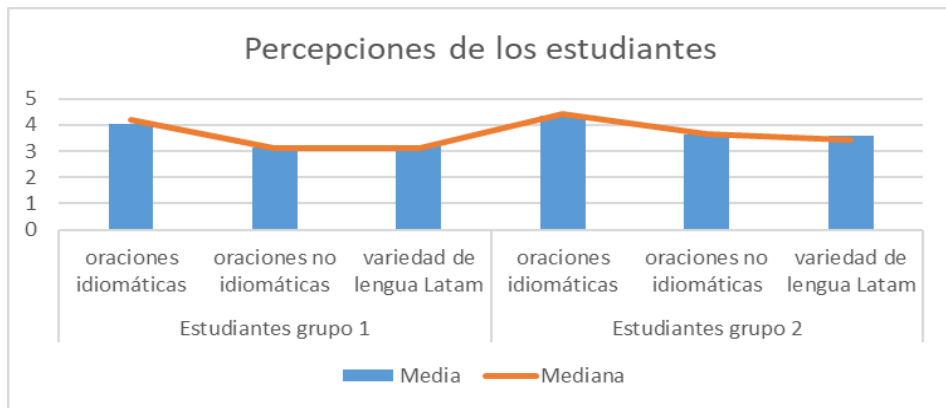


Gráfico 1: Percepción de los estudiantes: media y mediana.

Como se infiere de la tabla 1, la puntuación de la naturalidad es mayor, en todas las variables, en el grupo 2. Con el objetivo de analizar si las diferencias que se observan en la descriptiva pueden considerarse estadísticamente significativas, aplicamos un *t-test* (o test de la *t* de *student*) a cada una de las tres variables. En esta prueba, para un nivel de confianza superior al 95 % se deben obtener $p\text{-valores} < \text{Alpha} = 0,05$

(Bhattacharyya *et al.*, 1977). Como se puede apreciar en los resultados de la tabla 2, nuestros p-valores son mucho menores, lo que hace que las diferencias observadas entre los dos grupos pueden considerarse significativas al 99 %.

Variable	p-valor
Puntuación idiomáticas	<,0073
Puntuación no idiomáticas	<,0001
Puntuación variedad Latam	<,0011
Puntuación total	<,0001

Tabla 2: Resultados de la aplicación del *t-test*

El primer análisis de estos datos nos hace pensar que el acceso al texto en inglés inclina a los estudiantes hacia una valoración más positiva de los textos traducidos al español. Lo que vendría a confirmar la excesiva confianza de los estudiantes en los resultados de la traducción automática, conclusión a la que se llega también en otros estudios anteriores de carácter experimental a partir de la traducción automática neuronal (véase, por ejemplo, Rico y González, 2022; Lock *et al.*, 2022) o estadística (Jolley y Maimone, 2015).

3.2. Resultados sobre la posibilidad de registrar una opción mejor a la de la traducción automática (pregunta 2)

A continuación, presentamos los resultados agrupados por tipo de oración.

3.2.1. Falta de idiomática por interferencias léxicas: oraciones con calcos o traducciones literales

Núm. oración 11	Oración en inglés	If you no longer have the shipping box or packing material, please advise what condition it was.
	Traducción automática en español	Si ya no tienes la caja de envío o el material de embalaje, indícanos en qué condición estaba.

p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,611	5,41%	5,41%	29,73 %	10,81 %

Tabla 3: Análisis de la oración 11: *condition*.

En este primer caso (oración 11), el p-valor es de 0,611 lo que nos indica que la variación de la naturalidad no es muy diferente en los dos grupos, hecho que se confirma en los gráficos 2A y 2B. Obsérvese que el número de respuestas en los dos casos (33 en el primer y 36 en el segundo) es inferior al número de estudiantes por grupos (34 y 37, respectivamente). La razón tiene que ver con el hecho de que no todos los estudiantes responden a todas las preguntas. Estas discrepancias quedan reflejadas en en el *Frequency Missing* que aportamos debajo del número de respuestas.

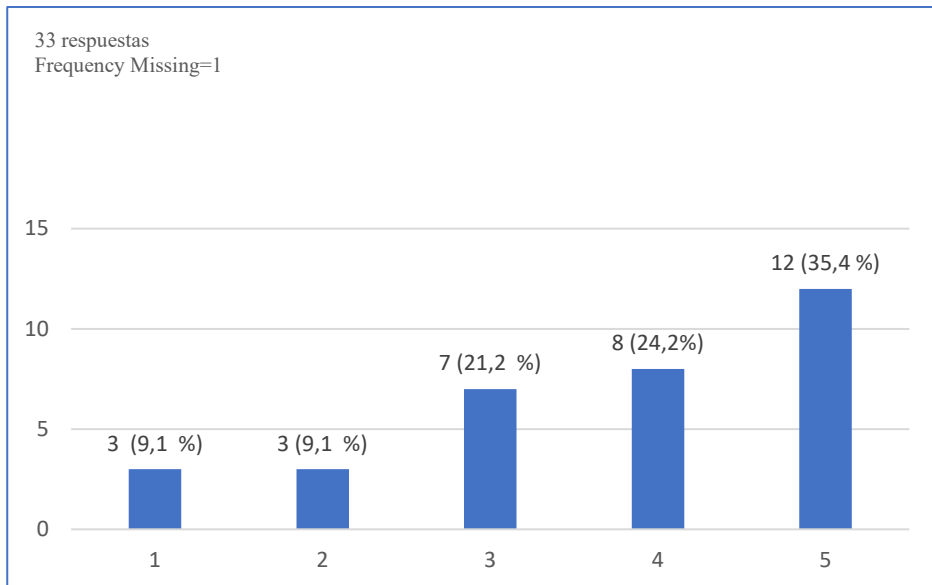


Gráfico 2A. Oración 11: valoración de la naturalidad en el grupo 1.

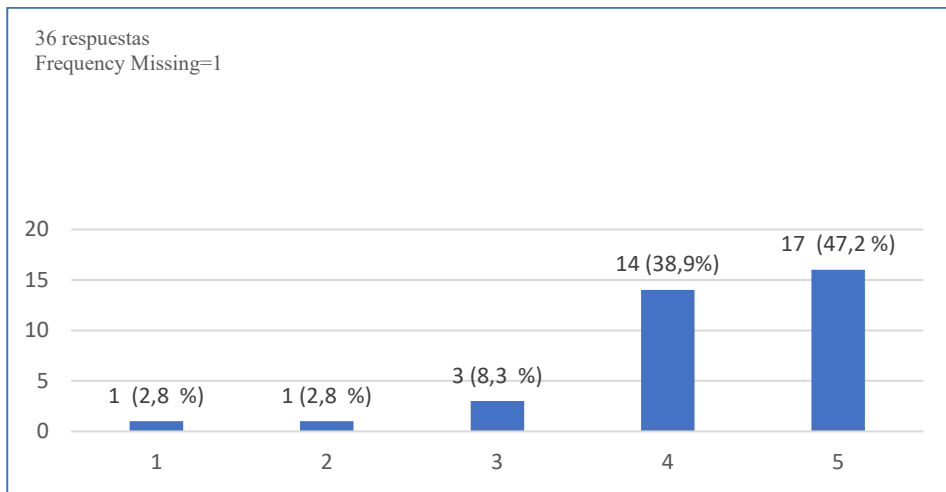


Gráfico 2B. Oración 11: valoración de la naturalidad en el grupo 2.

En este caso, la traducción de *condition* como *condición* es lo que origina la falta de idiomática. La traducción que podríamos considerar como más idiomática sería *estado* (*Si ya no tiene la caja del envío o el material de embalaje, indíquenos cuál era su estado*). En el caso de elegir la primera opción, *estado*, la revisión supondría sustituir el verbo estar por otro como *encontrarse*. Un análisis detallado de las soluciones aportadas por los estudiantes muestra que, en general, en ninguno de los dos grupos se identifica con facilidad el origen del problema. Son muy pocos los estudiantes que aciertan con la solución y existe un porcentaje elevado de estudiantes, claramente superior en el grupo 1, que sobrecorrije.

Las sobrecorrecciones en el grupo 1 tienen que ver con la sustitución del verbo *tener* por otros como *disponer*, *contar* o *poseer*. En otros casos, están relacionadas con la reposición de un sujeto que resulta innecesario en la proposición sustantiva (*en qué condición se encontraba esta/este; en qué condición estaba el pedido, en qué condición estaba el producto*) o la adición de algunas otras palabras que resultan innecesarias y que, en algún caso, dan lugar a problemas de concordancia (*en qué condición estaba el producto al principio, en qué condición estaba cuando la recibió*). En el grupo 2, como se desprende de los números de la tabla anterior, hay mucha menos sobrecorrección y esta consiste en el añadido de la fórmula de cortesía *por favor* o la sustitución del *tener* inicial por *disponer*.

Veamos el siguiente ejemplo. En este caso, la traducción de *condition* como *condición* es lo que origina la falta de idiomática. La traducción que podríamos considerar como más idiomática sería *estado* (*Si ya no tiene la caja del envío o el material de embalaje, indiquenos cuál era su estado*). En el caso de elegir la primera opción, *estado*, la revisión supondría sustituir el verbo estar por otro como *encontrarse*. Un análisis detallado de las soluciones aportadas por los estudiantes muestra que, en general, en ninguno de los dos grupos se identifica con facilidad el origen del problema. Son muy pocos los estudiantes que aciertan con la solución y existe un porcentaje elevado de estudiantes, claramente superior en el grupo 1, que sobrecorrije.

Las sobrecorrecciones en el grupo 1 tienen que ver con la sustitución del verbo *tener* por otros como *disponer*, *contar* o *poseer*. En otros casos, están relacionadas con la reposición de un sujeto que resulta innecesario en la proposición sustantiva (*en qué condición se encontraba esta/este; en qué condición estaba el pedido, en qué condición estaba el producto*) o la adición de algunas otras palabras que resultan innecesarias y que, en algún caso, dan lugar a problemas de concordancia (*en qué condición estaba el producto al principio, en qué condición estaba cuando la recibió*). En el grupo 2, como se desprende de los números de la tabla anterior, hay mucha menos sobrecorrección y esta consiste en el añadido de la fórmula de cortesía *por favor* o la sustitución del *tener* inicial por *disponer*.

Veamos el siguiente ejemplo.

Núm. oración 25	Oración en inglés		Usually, card transfers are instant but in rare cases , like bank holidays or weekends it gets delayed.	
	Traducción automática en español		Normalmente las transferencias con tarjeta son instantáneas, pero, en casos raros , como los días festivos o los fines de semana se retrasan.	
p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,7649	32,35 %	47,6 %	14,71 %	2,7 %

Tabla 4: Análisis de la oración 25: *rare cases*.

En este caso (oración 25), el p-valor es de 0,7649 lo que nos indica que la variación de la naturalidad no es muy diferente en los dos grupos, dato que se confirma en los gráficos 3a y 3b.

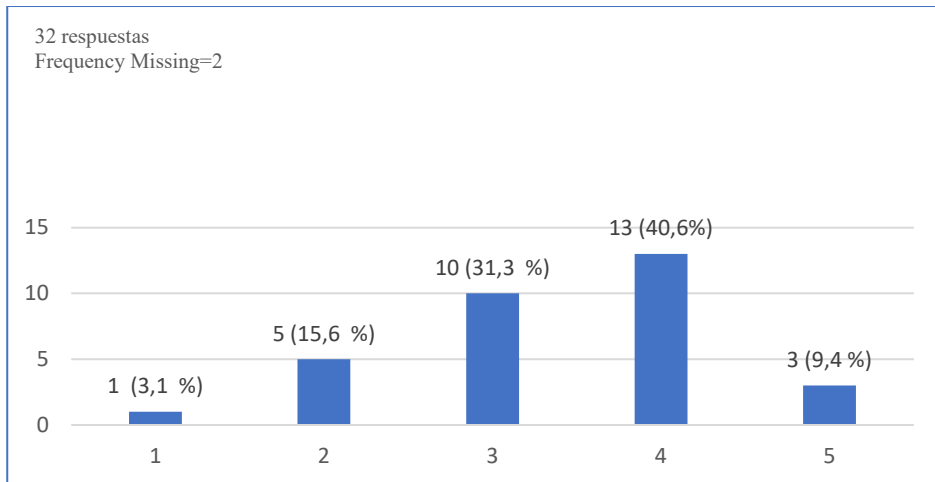


Gráfico 3A. Oración 25: Valoración de la naturalidad en el grupo 1.

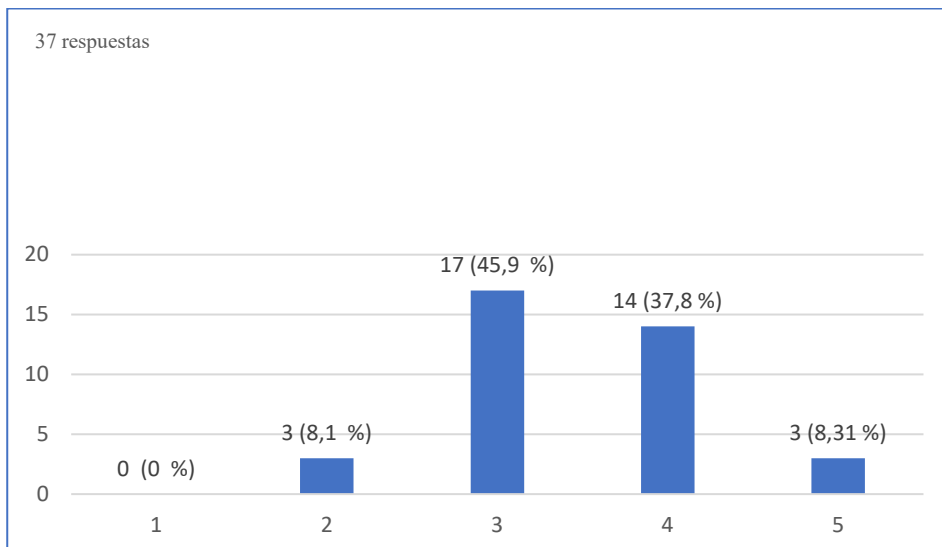


Gráfico 3B. Oración 25: Valoración de la naturalidad en el grupo 2.

En este ejemplo, la traducción de *rare cases* como *raros casos* es lo que origina el problema de falta de idiomática. La traducción que podríamos considerar como más idiomática sería *casos excepcionales* que identifica el adjetivo español *excepcional* como el más adecuado como traducción de *rare* en este contexto. (*Normalmente las transferencias con tarjetas de créditos son instantáneas, pero, en casos excepcionales, como los festivos o los fines de semana se retrasan*).

Un análisis detallado de las soluciones muestra que, en general, se identifica la palabra que origina el problema, incluso en el grupo 1 que no tiene acceso a la oración en inglés.

Las diferencias en las soluciones tienen que ver con el empleo de términos más neutros, que no tienen la connotación de excepcionalidad, como *a veces*, *en algunas ocasiones*, *en ciertas ocasiones*, *en algunos casos* o que no resultan tan adecuados, por ejemplo, *en casos extraordinarios* o *en raras ocasiones*.

Algunos estudiantes optan también por omitir el origen del problema (*Normalmente, las transferencias con tarjeta se realizan de manera instantánea, pero se retrasan los fines de semanas o días festivos*).

Los casos de sobrecorrección más llamativos tienen que ver con soluciones en las que se identifica de manera incorrecta lo que origina la falta de naturalidad. Por ejemplo, se identifica el verbo *retrasar* como el error y se ofrecen soluciones que resultan claramente mucho menos idiomáticas: *Normalmente, las transferencias con tarjetas son instantáneas, pero, a veces, hay retenciones en festivos o fines de semana* (grupo 1).

Compruébese con mayor detalle en la siguiente tabla:

Núm. oración 28	Oración en inglés		Learn how long refunds usually take.	
	Traducción automática en español		Aprende cuánto tiempo suelen tardar los reembolsos.	
p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2

0,0005	35,29 %	32,35 %	14,70 %	0%
--------	---------	---------	---------	----

Tabla 5: Análisis de la oración 28: *learn*.

En este caso, el *p-valor* es de 0,0005 lo que nos indica que la variación de la naturalidad es muy diferente en los dos grupos, dato que se confirma en el gráfico 4A y 4B.

Como se puede observar, en el grupo 1, la mayor parte de las respuestas se sitúan en los valores bajos de la escala de naturalidad; mientras que, en el grupo 2, sucede justamente lo contrario, la mayor parte de los valores están en la escala más alta. Esto podría estar relacionado con el hecho de que el grupo 2 relaciona el *aprender* de la oración en español con *learn* de la versión en inglés.

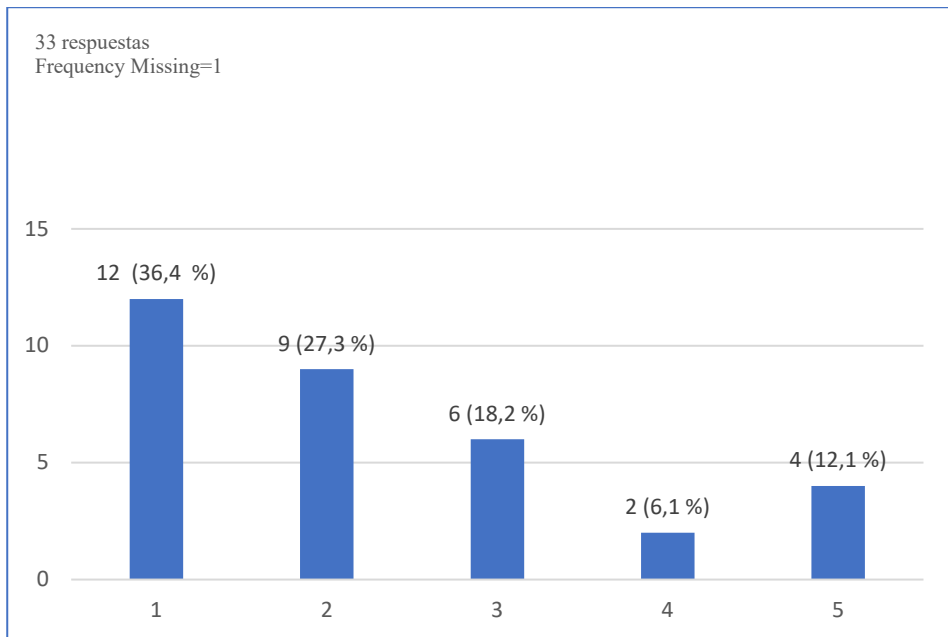


Gráfico 4A. Oración 28: valoración de la naturalidad en el grupo 1.

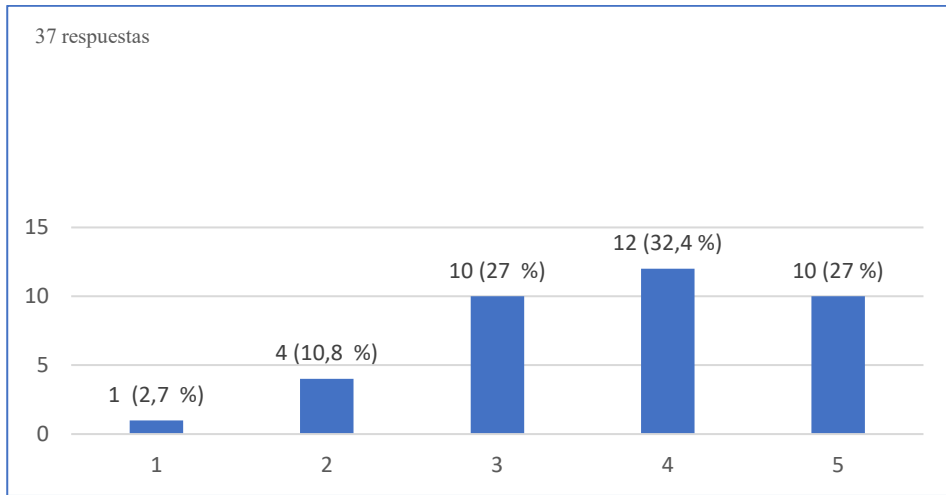


Gráfico 4B. Oración 28: valoración de la naturalidad en el grupo 2.

En esta oración, la traducción de *learn* por *aprende* es la que origina la falta de idiomática. La traducción más idiomática en este caso sería *infórmate*, *conoce* o *averigua*. En general, se identifica de manera correcta el origen del problema. En el grupo 1, se da alguna otra solución más libre poniendo en contexto el mensaje (*El tiempo aproximado de los reembolsos es el siguiente:*).

La sobrecorrección se da solo en el grupo 1 en casos en los que se añade información que no aparece en el original (*Los reembolsos tardan varios días, toda la información se encuentra en la página web/Los reembolsos suelen tardar. Le ofrecemos toda la información que pueda necesitar en nuestra página web/Los reembolsos pueden tardar varios días...*).

3.2.2. Falta de idiomática por interferencias en el nivel morfosintáctico

Núm. oración 43	Oración en inglés	The payment was returned to the sender.
	Traducción automática en español	El pago fue devuelto al remitente.

p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,0151	20,58 %	8,82 %	8,82 %	5,88 %

Tabla 6: Análisis de la oración 43: *pasiva en lugar de pasiva refleja*.

En este caso, el *p-valor* de 0,0151 nos muestra que sí hay ciertas diferencias en la valoración de la naturalidad de los dos grupos, tal y como se puede apreciar en los gráficos 5A y 5B.

En el grupo 2, prácticamente todas las respuestas están situadas entre el 4 y el 5 de idiomática. En el grupo 1, las respuestas están más repartidas entre el 2 y 5 de idiomática, aunque más de la mitad de las puntuaciones están entre el 4 y 5.

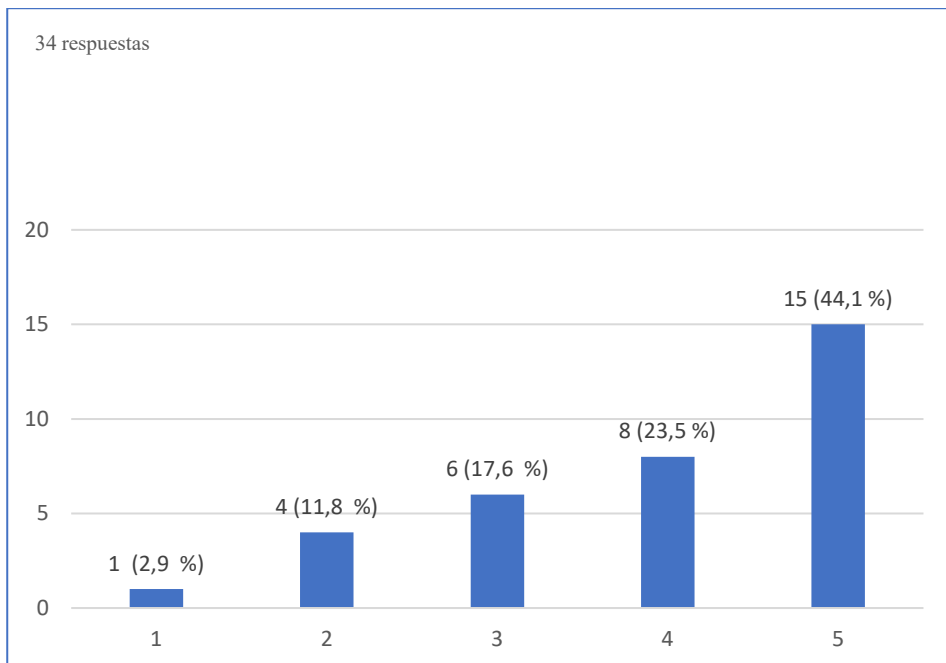


Gráfico 5A. Oración 43: valoración de la naturalidad en el grupo 1.

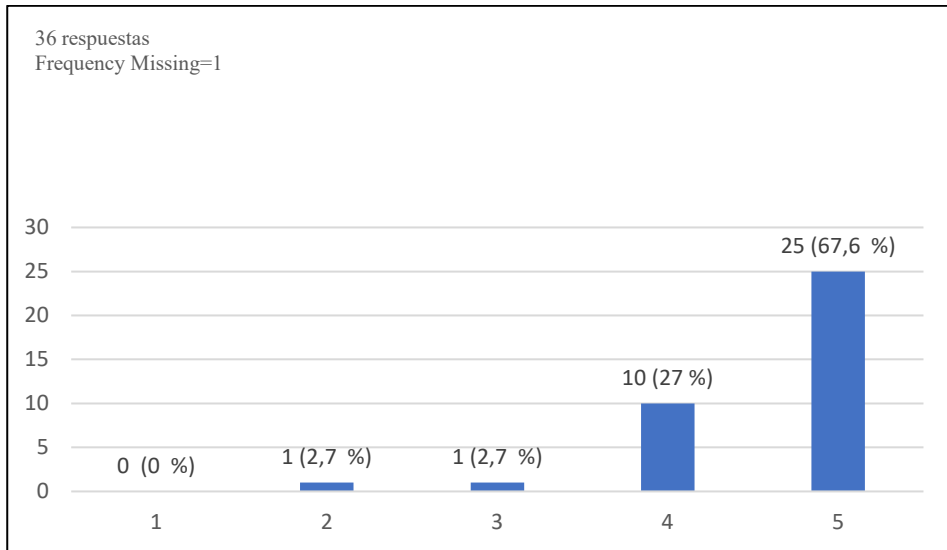


Gráfico 5B. Oración 43: valoración de la naturalidad en el grupo 2.

En esta oración, es la traducción de la pasiva en inglés (*was returned*) por la pasiva en español (*fue devuelto*) lo que causa el problema de idiomática. La solución más idiomática pasaría por la traducción de la pasiva inglesa por una pasiva refleja en español (*El pago se devolvió*). El problema se identifica y corrige de manera significativamente mejor en el grupo 1 (20,58 % de respuestas correctas frente a 8,82 %).

Por otra parte, los casos de sobrecorrección están relacionados con la sustitución del sustantivo *remitente* por el sustantivo *comprador* o *cliente* en el grupo 1 o por la sustitución del pretérito indefinido (*fue devuelto*) por el pretérito perfecto compuesto (*ha sido devuelto*) o por el cambio de *devolver* por *reembolsar* y de nuevo también del pretérito indefinido por el pretérito perfecto compuesto (*El pago ha sido reembolsado*) en el grupo 2.

Núm. oración 4	Oración en inglés	Once we received this information, we will be able to investigate your issue further.
	Traducción automática en español	Una vez que recibamos esa información podemos investigar tu problema más a fondo.

p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,0406	0 %	5,40 %	44,11 %	27,02 %

Tabla 7: Análisis de la oración 4: traducción del adjetivo posesivo del inglés.

En este caso, el p-valor (0,046) nos muestra que sí hay ciertas diferencias, aunque no muy significativas, entre las valoraciones de la idiomática en los dos grupos. Efectivamente, como se puede ver en los gráficos 6A y 6B, la mayor parte de las puntuaciones (el 41,18 % en el grupo 1 y el 54,05 % en el grupo 2) se concentran en el 4. El resto de las puntuaciones en el grupo 2 son de 4 y 5.

El grupo 1 distribuye sus puntuaciones entre todas las comprendidas entre el 1 y el 5 de la escala.

Obsérvese a continuación los gráficos que se han incluido en el texto que nos ocupa:

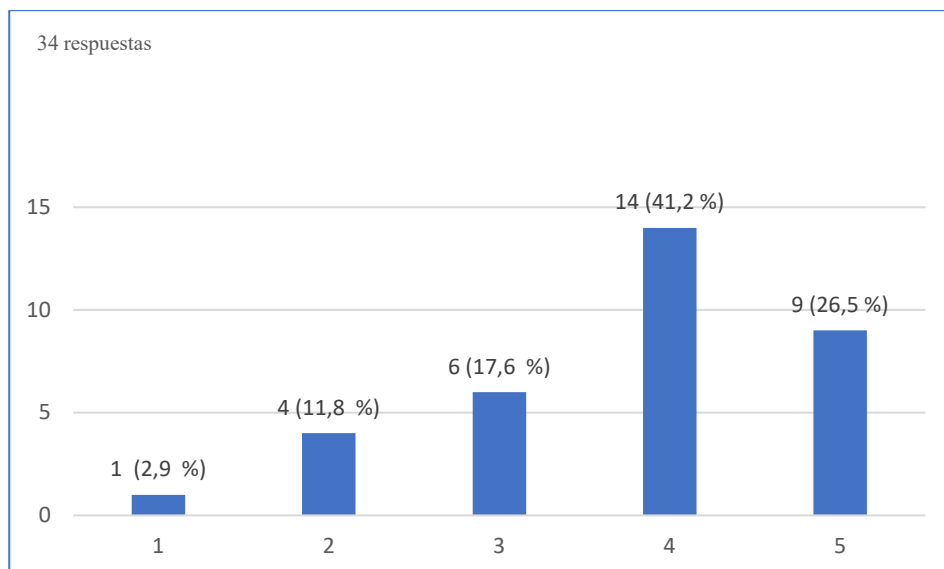


Gráfico 6A. Oración 4: valoración de la naturalidad en el grupo 1.

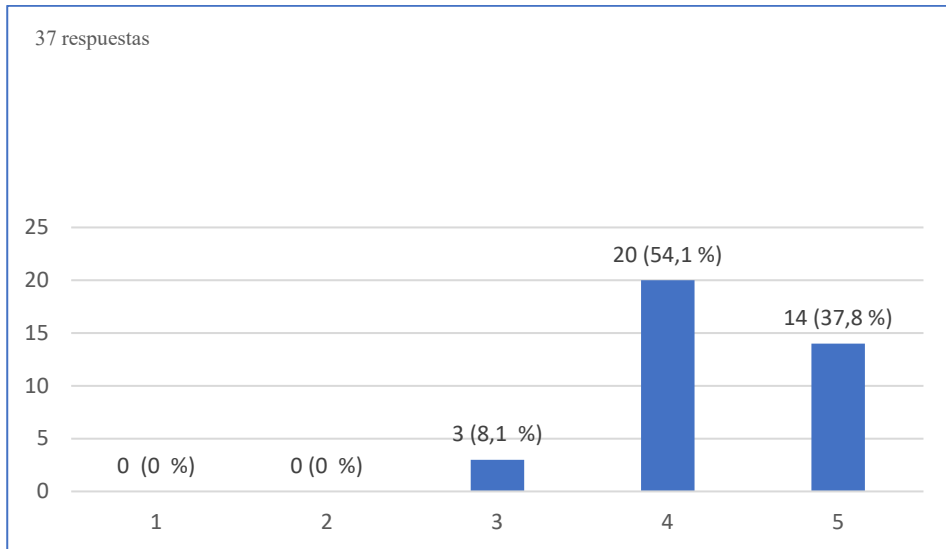


Gráfico 6B. Oración 4: valoración de la naturalidad en el grupo 2.

Si descendemos al análisis de las soluciones aportadas, en el grupo 1, no hay ningún estudiante que haya detectado el posesivo como lo que origina la falta de idiomaticidad.

En el grupo 2, tan solo dos estudiantes detectan y corrigen bien el problema. En los ejemplos de sobrecorrección, se cambia la subordinada temporal en forma personal (*Una vez que recibamos esa información*) por una construcción de participio absoluto (*Una vez recibida la información*) o por otro tipo de complemento circunstancial que podría tener este mismo significado (*Con esta información, Tras recibir la información requerida*). Algún estudiante propone eliminar el “más a fondo” que acompaña al verbo *investigar* o sustituirlo por *más concretamente* (*podremos investigar tu problema más concretamente*).

3.2.3. Falta de idiomaticidad por cuestiones pragmáticas: la cortesía en las peticiones

La tabla 8 nos aporta los necesarios datos de análisis sobre idiomaticidad, cuestiones pragmática y la cortesía en las peticiones:

Núm. oración 3	Oración en inglés		May I ask for the order number?	
	Traducción automática en español		¿Puedo solicitar el número de pedido?	
p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,782	14,70 %	17,64 %	5,88 %	5,88 %

Tabla 8: Análisis de las puntuaciones de la oración 3: cortesía en las peticiones.

En este caso, el p-valor (0,782) nos muestra que la valoración de la idiomaticidad es parecida en los dos grupos.

En otras palabras, como se desprende de los gráficos 7A y 7B, casi la mitad de los alumnos (41,2 %) del grupo 1 y casi un 38 % (37,8 %) dan a esta oración la máxima puntuación en idiomaticidad.

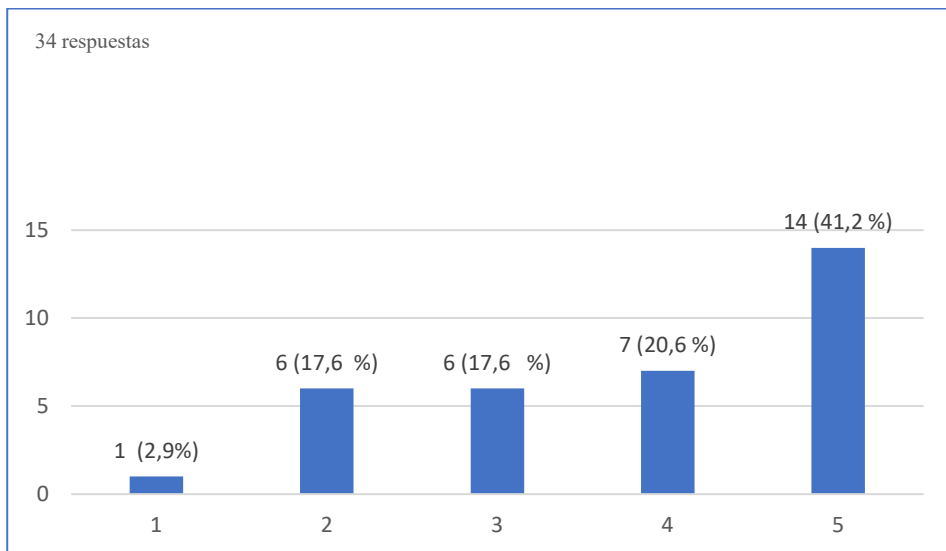


Gráfico 7A. Oración 3: valoración de la naturalidad en el grupo 1.

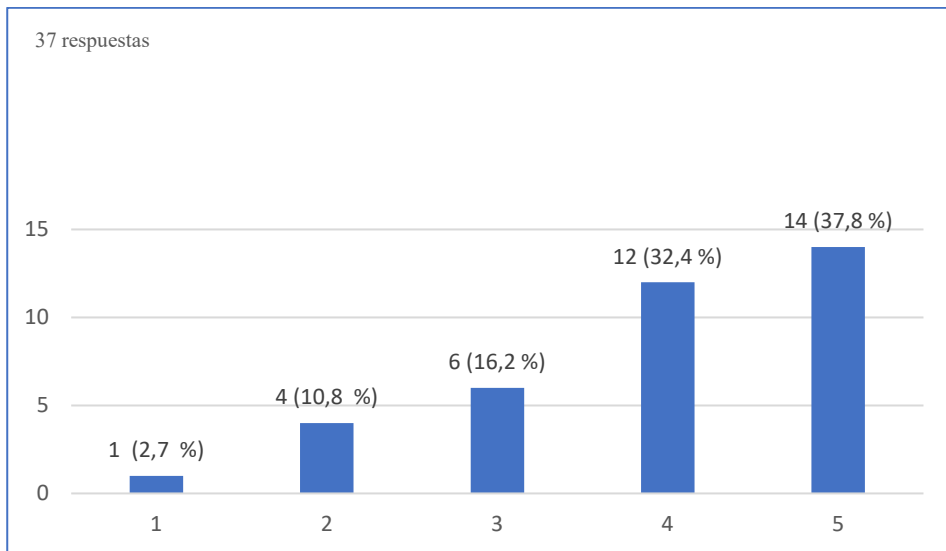


Gráfico 7B. Oración 3: valoración de la naturalidad en el grupo 2.

Como hemos visto en el apartado 3.1 *Corpus de análisis*, en esta oración el problema de la idiomatidad está relacionado con el trasvase al español de las fórmulas indirectas de cortesía del inglés en casos en los que, en español, resultaría más idiomático la utilización de una fórmula directa. En el caso que nos ocupa, ¿Puedo solicitar el número de pedido? en lugar de ¿Puedes darme/proporcionarme el número de pedido?

La sobrecorrección, que es baja en los dos grupos, tiene que ver con el hecho de que, al pasar a utilizar la segunda persona (*puedes*) en lugar de la primera (*puedo*), algunos estudiantes utilizan la variante formal *usted* en lugar de *tú*, que es lo que se les había pedido explícitamente en las instrucciones de la tarea. Este hecho puede explicarse, aunque no justificarse, por la identificación de una petición y la asociación de esta con la necesidad de ser corteses, lo que, en español, lleva a la utilización de la variante de cortesía (*usted*).

3.2.4. Falta de idiomatidad por cuestiones culturales

Recurriremos de nuevo a una tabla, la siguiente:

Núm. oración 49	Oración en inglés		Be safe, stay healthy and take care.	
	Traducción automática en español		Estate seguro, mantente saludable y cuídate.	
p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,0011	23,53 %	10,81 %	35,29 %	32,43 %

Tabla 9: Análisis de las puntuaciones de la oración 49: cuestiones culturales.

El p-valor en ese caso arroja un valor muy bajo, lo que se identifica con diferencias significativas en la percepción de la idiomaticidad. Los gráficos 8A y 8B nos muestran que, efectivamente, en el grupo 1, la mitad de los estudiantes dan a esta oración una puntuación de 1 en idiomaticidad. En el grupo 2, sin embargo, solo el 28 % de los estudiantes da una puntuación baja a esta oración, esto, de nuevo, está relacionado con el hecho de que los estudiantes del grupo 2 tienen acceso al original en inglés.

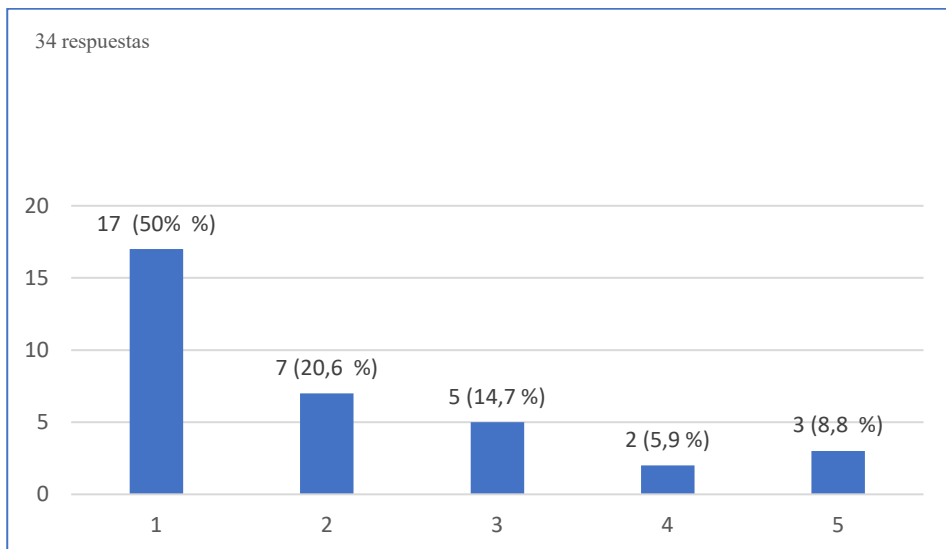


Gráfico 8A. Oración 49: valoración de la naturalidad en el grupo 1.

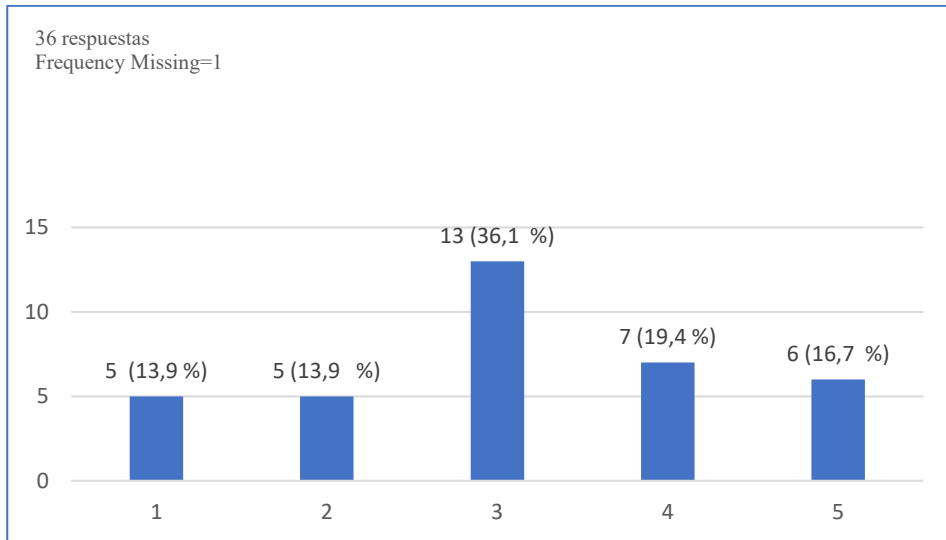


Gráfico 8B. Oración 49: valoración de la naturalidad en el grupo 2.

La falta de idiomática en esta oración está relacionada con la utilización de una fórmula de despedida en inglés cuya traducción literal en español resulta totalmente antinatural. Efectivamente, el español resolvería esta alusión a la salud con fórmulas como *cuídate* o *cuídate mucho* que se hicieron habituales en el contexto de la pandemia.

La valoración del desempeño de nuestros estudiantes muestra un porcentaje más alto de respuestas adecuadas en el grupo 1. La ausencia de la oración en inglés hace que se sientan más libres para proporcionar soluciones que no tienen en cuenta todo lo que se dice en inglés. En este caso, resulta muy llamativa la sobrecorrección que se da en un porcentaje alto tanto en el grupo 1 como en el grupo 2. *Asegúrate, mantén una vida saludable y cuídate, Siéntete seguro, mantente saludables y cuídate, Estate seguro, mantente en forma y cuídate* (en el grupo 1) o *Cuídate y salud, Vete con cuidado, mantente sano y cuídate, Protégete, mantén un estilo de vida saludables y cuídate* (en el grupo 2). Finalmente, entre las soluciones del grupo 1, encontramos oraciones que pierden totalmente la alusión a la salud del original y recurren a las fórmulas de despedida habituales en español (*Un saludo o Muchas gracias, que pases un buen día*).

3.2.5. Falta de idiomática por cuestiones léxicas o sintácticas relacionadas con la variedad latinoamericana (Latam)

Núm. oración 38	Oración en inglés		For any limited warranty claim like a defective product to be replaced, a valid proof of purchase is needed.	
	Traducción automática en español		Para que se reemplace cualquier reclamo de garantía limitada como un producto defectuoso, se necesita un comprobante de compra válido.	
p-valor	Desempeño Grupo 1	Desempeño Grupo 2	Sobrecorrección Grupo 1	Sobrecorrección Grupo 2
0,0319	5,88 %	5,40 %	23,59 %	2,70 %

Tabla 10: Análisis de las puntuaciones de la oración 38: cuestiones léxicas.

En este caso, el p-valor (0,0319) nos muestra que sí hay ciertas diferencias entre las valoraciones de la idiomaticidad en los dos grupos. En los gráficos 9A y 9B, podemos ver que, en el grupo 2, la mayor parte de las respuestas (el 91,64 %) se encuentran en las valoraciones de 3, 4 y 5. Las valoraciones del grupo 1 están más distribuidas a lo largo de la escala del 1 al 5, con cierta concentración (39,4 %) en la valoración de 4.

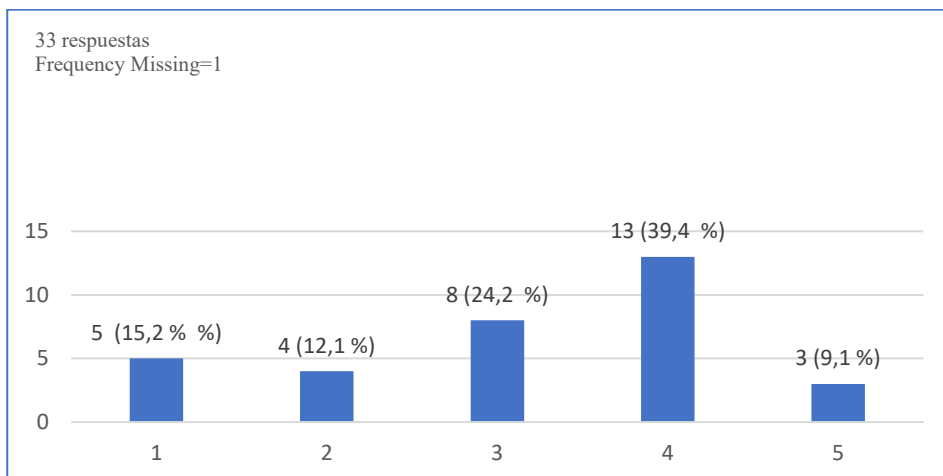


Gráfico 9A. Oración 38: valoración de la naturalidad en el grupo 1.

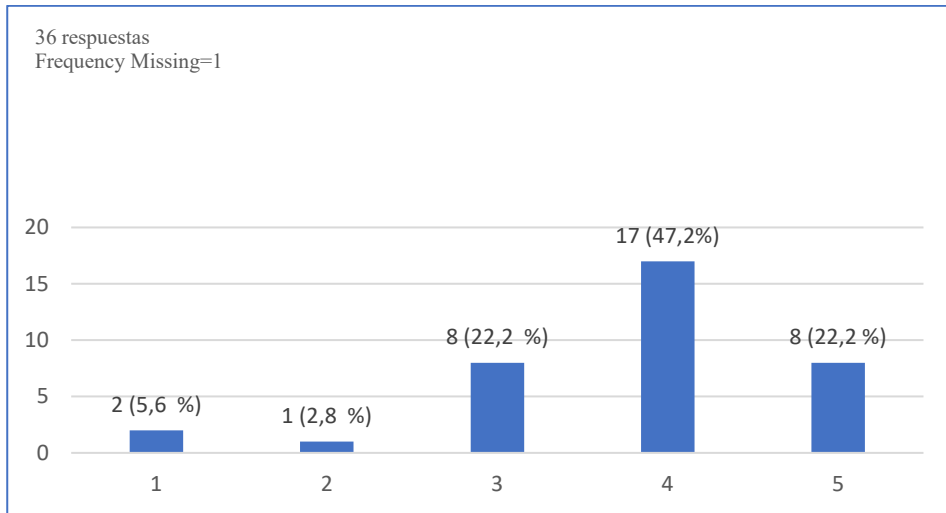


Gráfico 9B. Oración 38: valoración de la naturalidad en el grupo 2.

El problema que origina la falta de idiomática en este caso tiene que ver con un uso léxico, el de la palabra *reclamo*, que se asocia más a la variedad Latam. El *Diccionario de la lengua española* recoge como décima acepción, “Reclamación” (Real Academia Española, *s.f.*, *definición* 10) con la marca de *Am.* (América). El español de España prefiere la forma *reclamación* como sustantivo deverbal correspondiente a la acepción intransitiva del verbo *reclamar* “Clamar contra algo, oponerse a ello de palabra o por escrito” (*Ibid.*, *s. f.* *acepción* 6), que es la que se pone de manifiesto en este ejemplo.

El análisis del desempeño pone de manifiesto que muy pocos estudiantes (5,8 % en el grupo 1 y 5,40 % en el grupo 2) aciertan a identificar *reclamo* como el origen del problema de la falta de idiomática y a sustituirlo por la palabra *reclamación* que es la palabra más usual para este contexto en España. La sobrecorrección, que se da, sobre todo, en el grupo 1, tiene que ver con la sustitución del término *comprobante de compra válido* por *ticket de compra, justificante de compra* o *comprobante de pago*.

4. DISCUSIÓN

A la vista de los resultados obtenidos, nos parece relevante señalar, en primer lugar, el efecto que tiene en los estudiantes el acceso al texto

original en inglés, que los lleva a dar una puntuación significativamente más alta que en aquellos otros casos en los que no tienen el texto en inglés (p -valor $<$,0001). Junto a esta interferencia del texto original, hemos apuntado también como posible causa la excesiva confianza que los estudiantes tienen en los resultados de la traducción automática. A este respecto, es muy interesante comprobar cómo los estudiantes tienen dificultades para marcar la no idiomaticidad en oraciones que contienen calcos léxicos (*existing/existente, investigate your issue / investigar tu problema*) o sintácticos (*So that I can get this resolved for you? / Para que podamos resolver esto por ti*). En el caso de las oraciones con variedad Latam, los resultados de nuestro trabajo apuntan a que los estudiantes, quizá por ser un grupo de edad con mayor exposición a los medios de comunicación masiva, perciben este tipo de oraciones como más naturales.

En segundo lugar, nuestro estudio ha revelado la tendencia de los estudiantes a sobre corregir, un fenómeno que no es extraño en la traducción profesional cuando la persona que traduce acomete la tarea de posesición con su propia representación mental del texto de llegada. Cabe destacar que, a la luz de esta pequeña muestra de análisis cualitativo, la tendencia a la sobre corrección es sensiblemente mayor cuando no tienen acceso al texto en inglés, compárese los resultados de la celda de sobre corrección en el grupo 1 y 2 en las ocho oraciones de muestra del análisis cualitativo. La sobre corrección es un fenómeno ampliamente estudiado por Nitzke y Gros (2021) quienes indican que en el proceso de posesición algunos traductores tienden a mejorar innecesariamente el resultado de la traducción automática con el objetivo de conseguir una traducción perfecta, incluso aunque las instrucciones de posesición indiquen lo contrario. La sobre corrección no es algo que ocurra solo con la traducción automática, sino que se da también al revisar los resultados que se obtienen de las memorias de traducción y que, incluso, pueden llegar a suponer la revisión del 60% de coincidencias exactas que, en realidad, no requieren ninguna modificación (Mellinger y Shreve, 2016). Según estos autores, la causa de este comportamiento está en que los traductores tienen en mente sus propias traducciones cuando se disponen a trabajar con los resultados de la memoria de traducción y, así, cuando comparan el texto original con el traducido, buscan errores de manera activa, aunque estos no existan, con el fin de acomodar el resultado a su propia traducción ideal. Algo similar ocurre, entonces, en el proceso de posesición (Nitzke y Gros, 2021), cuando los traductores se enfrentan a la traducción automática. El resultado automático no tiene por qué ser

necesariamente malo, pero no siempre corresponde a la representación mental que se pueda haber hecho el traductor. Es en este punto donde surge la sobrecorrección y se llevan a cabo cambios que no son esenciales.

Ante esta situación, en la que vemos que la percepción de qué es una oración natural está sujeta a diferentes interpretaciones por parte de los estudiantes en función de una serie de variables, planteamos un conjunto de reflexiones de carácter docente. La primera de ellas se refiere a la definición de umbrales flexibles de aceptabilidad en el resultado de la traducción automática. Esta idea está relacionada con el planteamiento de Way (2018) con respecto a que la traducción automática debe ser evaluada en el contexto del uso específico para el que se genera la traducción (Way 2018, p. 172). A este respecto, podría argumentarse que la formación de los futuros traductores debe aspirar a la producción de textos de la máxima calidad, pero con la creciente incorporación de la traducción automática parece razonable replantear en las aulas la discusión de qué se considera un texto de calidad. El concepto de los umbrales de aceptabilidad en traducción automática ha sido estudiado ampliamente en autores como Bowker (2020), Vashee (2021), Plaza-Lara (2020) o Doherty (2017). De hecho, este último autor subraya precisamente que la evolución y la adopción generalizada de las tecnologías de la traducción –especialmente de la traducción automática– han dado lugar a una plétora de definiciones de la calidad y sus respectivas medidas, típicamente implícitas y operacionalizadas de forma diferente (Doherty 2017, 131). Todo ello afecta, por tanto, a las decisiones sobre cómo evaluamos la calidad de las traducciones y añade, por decirlo de algún modo, una capa adicional de mediación a la traducción y la revisión con «artefactos externos» que colaboran con el traductor humano «hacia el objetivo final de una traducción de calidad» (Mellinger 2018: p. 321). El traductor ya no es el único creador del texto de destino, sino que comparte esta responsabilidad con la máquina.

5. CONCLUSIONES

En este artículo hemos presentado los resultados de un estudio planteado con el objetivo de analizar la percepción de los estudiantes de traducción con respecto a la idiomática en la traducción automática del inglés al español en un corpus de oraciones. El género textual al que hemos aplicado este estudio es la atención al cliente, dada, por una parte, su importancia en el ámbito de la comunicación digital y, por otra, la creciente

tendencia a que estos contenidos se traduzcan automáticamente. El corpus analizado estaba compuesto por cincuenta oraciones, divididas en tres grupos: oraciones idiomáticas, oraciones no idiomáticas y, dentro de estas segundas, oraciones con variedad de lengua Latam.

Las investigaciones futuras apuntan hacia varias líneas. En primer lugar, se podrían profundizar en las diferencias que aportan las cuatro variables señaladas (interferencias léxicas, morfosintácticas, pragmáticas, culturales y las relacionadas con la variedad de lengua) en la percepción de la idiomatidad en los dos grupos. Otra posibilidad sería profundizar en el efecto que puede tener la presencia (o ausencia) del texto en la lengua original. En este artículo hemos contrastado dos grupos de dos cursos diferentes, en el futuro podrían compararse la percepción de la naturalidad y el desempeño al aportar otras soluciones en un mismo grupo de la misma asignatura con un grupo de control, con acceso al texto en inglés, que se tomaría como referencia para comprobar y evaluar los efectos de la ausencia del texto en inglés. Finalmente, otra tercera línea de trabajo sería pasar este mismo cuestionario a un grupo de adultos profesionales de la traducción. La edad podría ser uno de los factores más relevantes en la percepción de lo idiomático.

Los resultados de nuestro trabajo parecen indicar que la llegada de la traducción automática a nuestras aulas implica un cambio disruptivo en la docencia que, entre otras cuestiones, nos obliga a pensar de otro modo para poder acompañar a nuestros estudiantes a la hora de hacerles conscientes de la diferencia entre los textos traducidos automáticamente y los producidos por el traductor humano. Los estudios del efecto que tiene la traducción automática sobre la traducción final de un texto y sus posibles interferencias son incipientes y, precisamente por esta razón, entendemos que es una buena oportunidad para formar una nueva generación de profesionales que sepan tratar con estas variedades de traducción, haciendo que los alumnos sean conscientes de la diferencia entre los textos traducidos automáticamente y los textos directamente producidos en una lengua, esto es, aquellos que siguen el genio de la lengua.

REFERENCIAS BIBLIOGRÁFICAS

Alemán, Vanessa (2018). Traducción automática y fraseología especializada: el caso de la propiedad intelectual e industrial. En: Pino Valero Cuadrado, Analía Cuadrado Rey y Paola Carrión González

(Coords.) *Nuevas tendencias en traducción: Fraseología, Interpretación, TAV y sus didácticas* (pp. 69-87). Peter Lang.

Bhattacharyya, Goury K., Johnson, Richard. A. *Statistical Concepts and Methods* (Ser. Wiley Series in Probability and Mathematical Statistics). John Wiley & Sons.

Baziotis Christos, Mathur, Prashant y Hasler Eva (2022). *Automatic Evaluation and Analysis of Idioms in Neural Machine Translation*. <https://arxiv.org/abs/2210.04545>.

Bowker Lynne (2020). Fit-for-purpose translation. En O'Hagan, M. *The Routledge Handbook of Translation and Technology* (pp. 453-468). Routledge. <https://doi.org/10.4324/9781315311258-27>

Bryman Alan (2016). *Social Research Methods* (5.^a ed.). Oxford University Press.

Candel-Mora, Miguel A. (2022). Fine-tuning machine translation quality-rating scales for new digital genres: The case of user-generated content. En: Chelo Vargas-Sierra y José A. Sánchez Fajardo (Coords.), *Los géneros digitales en la interfaz del trabajo de corpus y el análisis del discurso especializado*. ELUA (pp.117-136). Universidad de Alicante. <https://doi.org/10.14198/ELUA.21900>

Corpas Pastor, Gloria (1996). *Manual de fraseología española*. Gredos.

Coulmas, Florian (1982). Ein Stein des Anstoßes. Ausgewählte Probleme der Idiomatik. *Studium Linguistik*, 13, 17-36.

Daems, Joke, De Clearcq, Orphée, and Macken, Lieve (2017). Translationese and post-editeese: How comparable is comparable quality? *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 16, 89-103. <https://doi.org/10.52034/lanstts.v16i0.434>

Doherty, Stephen (2017). Issues in human and automatic translation quality assessment. En Dorothy Kenny, *Human Issues in Translation Technology* (pp. 131-148). Routledge.

- EMT. (2022). *European Master's in Translation - EMT Competence Framework*. Recuperado el 23/5/2025 de https://commission.europa.eu/news/updated-version-emt-competence-framework-now-available-2022-10-21_en
- Fan, Wu (2018). *Estudio contrastivo de la fraseología china y española desde una perspectiva lingüística*. Comares.
- Fadaee, Marzieh A., Bisazza, Arianna y Monz, Chirstof (2018). Examining the Tip of the Iceberg: A Data Set for Idiom Translation. Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). Recuperado el 23/5/2024 de <https://aclanthology.org/L18-1148.pdf>
- García Rodríguez, Alba (2012) Las perífrasis verbales en el español de América. *Analele Universității Ștefan cel Mare*, tomo XVIII, n.º 1, s. p. https://www.academia.edu/5765203/Las_per%C3%ADfrasis_verbales_en_el_espa%C3%B1ol_de_Am%C3%A9rica
- Gellerstam, Martin (1986). *Translationese in Swedish Novels Translated from English*. The Scandinavian Symposium on Translation Theory.
- Gonçalves, Madalena, Buchicchio, Marianna, Stewart, Craig, Moniz, Helen y Lavie, Alon (2022). Agent and user-generated content and its impact on customer support MT. *Proceedings of the 23rd EAMT*. Recuperado el 23/5/2024 de <https://aclanthology.org/2022.eamt-1.23.pdf>
- Graça, Joao (2022). 7 tips to create machine-friendly, translatable content for customer service. Recuperado el 23/5/2024 de la dirección web <https://resources.unbabel.com/blog/7-tips-to-create-machine-friendly-translatable-content-for-customer-service>
- Jolley, Jason y Maimone, Luciane (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. En Aleidine Konttinen, Kalle, Salmi, Leena y Koponen, Maarit (2021). Revision and post-editing competences in translator education. En: Maarit Koponen, Brian Moosop, Isabelle S. Robert y Giovanna Scocchera,

Translation, Revision and Post-editing. Industry Practices and Cognitive Processes. Routledge.

- Laver, John y Mason, Ian (2018). *A Dictionary of Translation and Interpreting*. Cultural Bridges. Recuperado el 23/5/2024 de <https://www.cultural-bridges.co.uk/1/dictionary-of-translation-and-interpreting/>
- Loock, Rudy y Léchauguette, Sophie (2021). Machine translation literacy and undergraduate students in applied languages: report on an exploratory study. *Revista Tradumàtica. Tecnologies de la Traducció*, 19, 204-225. <https://doi.org/10.5565/rev/tradumatica.281>
- Loock, Rudy, Lechauguette, Sophie y Holt, Benjamin (2022). Dealing with the «elephant in the classroom»: Developing language students' machine translation literacy. *Australian Journal of Applied Linguistics*, 5 (3), 118-134. <https://doi.org/10.29140/ajal.v5n3.53si2>
- Mellado Blanco, Carmen (2016). ¿Qué es la idiomaticidad? En: Suzete Silva (Coord.), *Fraseología & Cía: entabulando diálogos reflexivos* (Vol. 2, pp. 229-256). Pontes.
- Mellinger, Christopher D. y Shreve, Gregory M. (2016). Match evaluation and over-editing in a translation memory environment. En Ricardo Muñoz Martín (Ed.) *Reembedding Translation Process Research.*, (pp. 131-148). John Benjamins. <https://doi.org/10.1075/btl.128.07mel>
- Mellinger, Christopher (2018). Re-thinking translation quality. *Target*, 30:2, 310-331. <https://doi.org/10.1075/target.16104.mel>
- Moeller, Aleidine J. (Ed.), Free Online Machine Translation: Use and perception by Spanish Students and Instructors En Aleidine J. Moeller (Ed.), *Learn Languages, Explore Cultures, Transform Lives* (pp. 181-200). Robert M. Terry.
- Morate Béjar, Esther (2018). *Lengua extranjera profesional (inglés) para la gestión administrativa en la relación con el cliente*. Tutor Formación.

- Moreno Fernández, Francisco y Ueda, Hiroto (2018). Cohesion and particularity in the Spanish dialect continuum. *Open Linguistics*, 4, 722-742. <https://doi.org/10.1515/opli-2018-0035>
- Newmark, Peter (1988). *Approaches to Translation*. Prentice Hall.
- NitkeI, Jean y Gros, Anne-Kathrin (2021). Preferential changes in revision and post-editing. En: Maarit Koponen, Brian Moosop, Isabelle S. Robert y Giovanna Scocchera (Eds.) *Translation, Revision and Post-editing. Industry Practices and Cognitive Processes* (pp. 21-34). Routledge. <https://doi.org/10.4324/9781003096962-3>
- Ortego Antón, María Teresa y Seghiri, Míriam (2019). La traducción automática de locuciones nominales del español al inglés. *A pain in the neck?* En: Cristina Carrasco Bengoa, María Cantarero Muñoz y Coral Díez Carbajo (Coords.) *Localización: traducción y sostenibilidad cultural: sustrato, fundamentos y aplicaciones* (pp. 331-342) Ediciones Universidad de Salamanca. <https://doi.org/10.2307/j.ctvrdflts.57>
- Parmar, Jupinder y Estrada Arias, David (2020): *Idiomatic Language Translation and Transfer Learning*. <https://api.semanticscholar.org/CorpusID:215717832>
- Plaza-Lara, Cristina (2020). How does machine translation and post-editing affect project management? An interdisciplinary approach. *Hikma* 19 (2), 163-182. <https://doi.org/10.21071/hikma.v19i2.12516>
- Ramírez Rodríguez, Pablo (2022). Fraseologismos en el aula de ELE: la problemática de la traducción automática. *Revista Tradumàtica. Tecnologies de la Traducció*, 20, 077-095 <https://doi.org/10.5565/rev/tradumatica.304>
- Real Academia Española / Asociación de Academias de la lengua española. (s. f.) Autenticar. En *Diccionario panhispánico de dudas*. Recuperado el 23/5/2024 de <https://www.rae.es/dpd/autenticar>

Real Academia Española (2009). El verbo (1). Tiempo y aspecto. El aspecto léxico. Los tiempos del modo indicativo. En *Nueva gramática de la lengua española*. (pp. 1161-1267). Espasa Calpe.

Real Academia Española (s. f.). Reclamación. En *Diccionario de la lengua española*. Recuperado el 23/5/2024 de <https://dle.rae.es/reclamaci%C3%B3n>

Real Academia Española. (s. f.) Reclamo. En *Diccionario de la lengua española*. Recuperado el 23/5/2024 de <https://dle.rae.es/reclamo>

Rico Pérez, Celia y Martín de Santa Olalla Sánchez, Aurora (2024). corpus-TA_atencion_cliente [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10953240>

Rico Pérez, Celia y González Pastor, Diana (2022). The role of machine translation in translation education: A thematic analysis of translator educators' beliefs. *Trans-Int*, 14, 1, 177-197. <https://doi.org/10.12807/ti.114201.2022.a010>

Salesforce (2020). *State of the Connected Customer*. Recuperado el 6/3/2023 de <https://www.salesforce.com/resources/articles/customer-expectations/?sfdc-redirect=369>.

Santing, Lukas, Sijstermans, Ryan, Anerdi, Giacomo, Jeuris, Pedro, Ten Thij, Marijn y Batista Navarro, Riz (2022). Food for thought: How can we exploit contextual embeddings in the translation of idiomatic expressions? En *Proceedings of the 3rd Workshop on Figurative Language Processing (FLP)*, (pp. 100-110). Association for Computational Linguistics. <https://aclanthology.org/2022.flp-1.14/> ; <https://doi.org/10.18653/v1/2022.flp-1.14>

TAUS (2020). *Measuring the Impact of Machine Translation Assumption 1: MT will Soon Make TM Redundant*. Recuperado el 23/5/2024 de <https://info.taus.net/taus-business-intelligence-bulletin-september-2020>

Toral, Antonio (2019). Post-editeese: An exacerbated translationese. En Mikel Forcada, Andy Way, Barry Haddow y Rico Sennich (Eds.),

Proceedings of Machine Translation Summit XVII: Research Track, (pp. 273-281), European Association for Machine Translation. <https://aclanthology.org/W19-6627/>

Vargas-Sierra, Chelo y Sánchez Fajardo, José A. (Coords.), (2022). *Los géneros digitales en la interfaz del trabajo de corpus y el análisis del discurso especializado*. ELUA. Universidad de Alicante.

Vashee, Kirti(2021). *The Challenge of Defining Translation Quality*. GALA Knowledge Center. Recuperado el 23/12/2023 de <https://www.gala-global.org/knowledge-center/professional-development/articles/challenge-defining-translation-quality>

Way, Andy (2018). Quality expectations of machine translation. En Joss Moorkens, Sheila Castillo, Federico Gaspari y Stephen Doherty (Eds.). *Translation Quality Assessment. From Principles to Practice*, (pp. 159-178), Springer. https://doi.org/10.1007/978-3-319-91241-7_8

Zuluoaga Ospina, Alberto (1980). *Introducción al estudio de las expresiones fijas*. Peter Lang.