



**Universidad de Concepción**

Facultad de Humanidades y Arte

Departamento de Idiomas extranjeros

Traducción/Interpretación en Idiomas Extranjeros

**ANÁLISIS SOBRE LA PERCEPCIÓN DE NATURALIDAD Y  
FIDELIDAD EMOCIONAL EN LA TRADUCCIÓN DE  
EXPRESIONES METAFÓRICAS REALIZADAS POR  
MODELOS EXTENSOS DE LENGUAJE DEL ESPAÑOL  
AL INGLÉS Y ALEMÁN**

Tesina presentada a la Facultad de Humanidades y Arte de la Universidad  
de Concepción para optar al grado de Licenciado en Traductología

POR: Allan Rubilar Rubilar

Diego Villaseca Pereira

Profesor guía: Mario Helm

Concepción, octubre de 2025

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

## AGRADECIMIENTOS

Agradecemos enormemente a todos nuestros profesores, su apoyo y formación durante estos años nos han permitido llegar hasta aquí. Un agradecimiento especial a nuestro profesor guía, Mario Helm, por su ayuda constante, su ánimo, su sabiduría y aliento. No habría sido lo mismo sin él.

De parte de Allan, agradezco a mi madre, Patricia, por el cariño constante y por soportarme en todo y ayudarme cuando lo necesitaba. Si he podido salir adelante en la vida, ha sido por ella. Agradezco a mis abuelos, Alicia y José, por siempre estar conmigo. Estoy feliz de que me vean dando este paso. Agradezco a Josefina, me caí tantas veces y nunca dejaste que me quedara ahí. No habría sido lo mismo sin ti. Gracias por ser parte de mi vida. Agradezco a mis amigos por sacarme incontables sonrisas. Gracias a todos los que me ayudaron en este viaje.

De parte de Diego, gracias a mi madre, María Lorena Pereira Araya, por acompañarme en espíritu cada día y esperarme cada noche durante este largo proceso; y a mis abuelos, Pascual German Pereira Contreras y María Haydee Araya Rojas, por haber allanado el camino hasta este último proyecto.

## RESUMEN

La prominencia de los modelos de inteligencia artificial en la actualidad resulta innegable. Esta presencia se observa en distintos ámbitos de la vida diaria y profesional. Entre los ámbitos donde su impacto se aprecia notoriamente es en la traducción, en la cual los modelos de lenguaje extensos (LLM) han adquirido protagonismo. Si bien es cada vez más común encontrar traducciones generadas por LLM como ChatGPT o DeepSeek, resulta pertinente cuestionar si dichas traducciones son percibidas positivamente por los hablantes nativos. Con este objetivo, la presente investigación analiza la percepción de hablantes nativos del idioma inglés y alemán frente a expresiones metafóricas traducidas por LLM. Para evaluar correctamente esta percepción, este trabajo se apoya en la propuesta de naturalidad planteada por Newmark (1998) y el concepto de fidelidad emocional desarrollado por Feng et al. (2025). Ambos parámetros se midieron por medio de una encuesta en la que los hablantes nativos calificaron las traducciones generadas por tres LLM (DeepSeek V-3, GPT-4o y Llama 4 Maverick). Se observó que el LLM DeepSeek V-3 posee la menor capacidad de traducción de forma natural y fiel a nivel emocional. Por el contrario, GPT-4o obtuvo mejores resultados en naturalidad en el idioma alemán, mientras que Llama 4 Maverick presentó un mejor desempeño en la fidelidad en inglés.

**Palabras claves:** naturalidad, fidelidad emocional, modelos extensos de lenguaje, traducción de metáforas

## ABSTRACT

The prominence of artificial intelligence models is undeniable nowadays. Their presence can be observed in various aspects of daily and professional life. One of the areas where their impact is particularly evident is translation, in which large language models (LLMs) have gained significant prominence. While it is increasingly common to find translations generated by LLMs such as ChatGPT or DeepSeek, it is relevant to question whether such translations are perceived positively by native speakers. To achieve this objective, the present study analyzes the perception of native speakers of English and German toward metaphors translated by LLMs. In order to properly assess this perception, the study draws on Newmark's (1988) proposal of naturalness and the concept of emotional fidelity developed by Feng et al. (2025). Both parameters were measured through a survey in which native speakers evaluated translations generated by three LLMs (DeepSeek V-3, GPT-4o y Llama 4 Maverick). The LLM DeepSeek V-3 performed the worst in terms of natural and emotionally faithful translation. In contrast, GPT-4o performed better in terms of naturalness in German, while Llama 4 Maverick performed best in terms of emotional fidelity in English.

**Keywords:** naturalness, emotional fidelity, large language model, metaphor translation

## TABLA DE CONTENIDOS

RESUMEN.....	4
ABSTRACT .....	5
1. INTRODUCCIÓN .....	8
2. MARCO TEÓRICO .....	10
2.1. Traducción .....	10
2.1.1. Teoría de la equivalencia dinámica .....	11
2.1.2. Naturalidad.....	13
2.2. Metáfora conceptual .....	14
2.2.1. Emociones en las metáforas.....	18
2.2.2. Metáforas en el discurso dramático .....	20
2.3. El reconocimiento y clasificación de emociones .....	22
2.4. Modelos extensos de lenguaje .....	24
2.4.1. Traducción con modelos extensos de lenguaje.....	25
2.4.2. Fidelidad emocional.....	26
3. PREGUNTA DE INVESTIGACIÓN .....	29
4. OBJETIVOS .....	29
4.1. Objetivo general .....	29
4.2. Objetivos específicos .....	29
5. METODOLOGÍA .....	30
6. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS .....	38
6.1. DeepSeek V-3.....	38
6.1.1. Naturalidad.....	38
6.1.2. Fidelidad emocional.....	40
6.2. GPT-4o .....	41
6.2.1. Naturalidad.....	41
6.2.2. Fidelidad emocional.....	43
6.3. Llama 4 Maverick.....	45
6.3.1. Naturalidad.....	45
6.3.2. Fidelidad emocional.....	46
6.4. Observaciones generales.....	47

7. CONCLUSIÓN Y PROYECCIONES .....	58
8. REFERENCIAS BIBLIOGRÁFICAS.....	62
9. ANEXOS .....	69
Anexo 1: Tabla de textos origen de las expresiones metafóricas seleccionadas .....	69
Anexo 2: Tabla de expresiones metafóricas clasificadas según Rueda de Emociones de Plutchik (1980) .....	70
Anexo 3: Tabla de doce expresiones metafóricas traducidas por el LLM DeepSeek V-3 al inglés y alemán .....	71
Anexo 4: Tabla de doce expresiones metafóricas traducidas por el LLM GPT-4o al inglés y alemán .....	73
Anexo 5: Tabla de doce expresiones metafóricas traducidas por el LLM Llama 4 Maverick al inglés y alemán .....	75
Anexo 6: Gráfico de resultados encuesta en naturalidad en inglés de LLM DeepSeek V-3.	77
Anexo 7: Gráfico de resultados encuesta en naturalidad en alemán de LLM DeepSeek V-3 .....	78
Anexo 8: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM DeepSeek V-3 .....	79
Anexo 9: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM DeepSeek V-3 .....	80
Anexo 10: Gráfico de resultados encuesta en naturalidad en inglés de LLM GPT-4o .....	81
Anexo 11: Gráfico de resultados encuesta en naturalidad en alemán de LLM GPT-4o .....	82
Anexo 12: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM GPT-4o .....	83
Anexo 13: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM GPT-4o .....	84
Anexo 14: Gráfico de resultados encuesta en naturalidad en inglés de LLM Llama 4 Maverick .....	85
Anexo 15: Gráfico de resultados encuesta en naturalidad en alemán de LLM Llama 4 Maverick .....	86
Anexo 16: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM Llama 4 Maverick .....	87
Anexo 17: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM Llama 4 Maverick .....	88
Anexo 18: Tabla de expresiones metafóricas según suma total de promedios finales desde el resultado más bajo al más alto.....	89

## 1. INTRODUCCIÓN

La creciente presencia de modelos extensos de lenguaje (LLM) en diversos ámbitos de la vida cotidiana y profesional es innegable. Empresas multinacionales, laboratorios de investigación y comunidades de código abierto han impulsado su desarrollo, y su uso se ha extendido desde la automatización de tareas simples hasta la generación de contenido creativo. Este avance ha alcanzado incluso áreas tradicionalmente consideradas exclusivas del pensamiento humano, como la traducción, donde confluyen lengua, cultura y contexto, factores que dificultan cualquier intento de automatización.

En los últimos años, la investigación académica sobre la traducción automatizada mediante inteligencia artificial ha crecido notablemente. Mohamed et al. (2024) analizan los distintos métodos de traducción automática y evalúan su capacidad para mantener la fidelidad del texto original, a su vez, abordan desafíos éticos asociados a su implementación. Por su parte, Wang (2023) señala deficiencias lingüísticas frecuentes al traducir entre lenguas con estructuras diferentes, mientras que Moneus y Sahari (2024) advierten sobre la pérdida de matices emocionales, incluso cuando el flujo de trabajo de traducción sea más eficiente.

A pesar de estas limitaciones, los estudios coinciden en que los LLM están en constante evolución, con crecientes capacidades para enfrentar retos lingüísticos (Shahmerdanova, 2025). No obstante, persiste una interrogante crucial: ¿pueden

estos modelos captar y transmitir con precisión los aspectos más subjetivos del lenguaje humano, como las emociones?

La traducción de textos emocionalmente cargados, en particular aquellos del ámbito literario, plantea un desafío significativo, puesto que su comprensión y recepción dependen en gran medida de la percepción de los lectores en la lengua meta. En este contexto, la presente investigación propone determinar la capacidad de naturalidad y la fidelidad emocional traducciones generadas por LLM a partir de expresiones metafóricas emocionales extraídas de obras teatrales. Para ello, se aplicará una encuesta a hablantes nativos de inglés y alemán, quienes calificarán dichas traducciones.

El objetivo principal de esta investigación es analizar si los modelos extensos de lenguaje logran reflejar de manera precisa la intención emocional original al traducir expresiones metafóricas presentes en textos teatrales. Particularmente, se indagará en su capacidad para preservar tanto la fidelidad emocional como la naturalidad de las expresiones metafóricas traducidas en contextos narrativos.

## 2. MARCO TEÓRICO

### 2.1. Traducción

Desde sus inicios, la traducción ha sido una disciplina vital. Sin embargo, la reciente explosión de información ha intensificado el enfoque académico en ella, con un número creciente de investigadores dedicados a analizar esta disciplina desde su concepto mismo.

Mayoral (2001, p. 45) declara que existen tantas definiciones de traducción como autores que escriben sobre ella. Esta afirmación se comprueba al observar la variedad de trabajos sobre las bases teóricas de la traducción.

Newmark (1988) define la traducción como el acto de transferir un mensaje entre dos lenguas, enfatizando el respeto por la intención y el texto del autor original. Complementando esta visión, Bassnett (2002) propone una concepción de la traducción como un proceso de codificación y decodificación que busca, en la medida de lo posible, conservar las estructuras originales, sin dejar de considerar las convenciones de la lengua meta. No obstante, también subraya la necesidad de concebir tanto el texto origen como el texto meta como productos creativos equivalentes, capaces incluso de cumplir funciones diferentes.

Desde una perspectiva distinta, Nida y Taber (1982) sitúan al receptor meta en el centro de su concepción traductológica. Para estos autores traducir consiste en reproducir en la lengua receptora el equivalente natural más cercano del mensaje de

la lengua fuente, primero en términos de significado y luego en términos de estilo. En este planteamiento, la percepción del receptor meta es el criterio principal para tomar decisiones traductológicas, ya que la prioridad es que esta audiencia reciba un mensaje equivalente en contenido y carga emocional. Lo anterior conlleva a que, según lo expresado por Nida y Taber, «to preserve the content of the message the form must change» (p. 5) y que otros elementos del texto origen sean desplazados de las prioridades del traductor.

### **2.1.1. Teoría de la equivalencia dinámica**

Entre los autores fundamentales en la evolución de los estudios de traducción, como se mencionó, se encuentra Eugene A. Nida, quien desarrolló la Teoría de la Equivalencia Dinámica. Esta surge en el contexto de la traducción bíblica, en un esfuerzo por hacer que los textos sagrados fuesen comprensibles para culturas muy distintas a la hebrea o grecolatina. Frente a la noción tradicional de equivalencia formal, que se preocupaba por mantener la estructura gramatical y léxica del texto fuente, Nida propuso un enfoque centrado en el efecto comunicativo del mensaje.

En su obra *Toward a Science of Translating* (1964), Nida define la equivalencia dinámica como aquella que busca generar en los receptores del texto meta una respuesta lo más natural y equivalente posible a la que habrían tenido los receptores del texto original. Según el autor, la traducción debe ser natural, comprensible y equivalente en efecto, lo cual implica una adaptación activa a las normas lingüísticas, culturales y contextuales de la audiencia meta, así como también a la interpretación de esta audiencia a la expresión de las emociones.

Una de las ideas centrales de Nida es que el traductor debe interpretar el sentido del mensaje en su totalidad, incluyendo tanto los aspectos semánticos como pragmáticos, para reconstruirlo en la lengua meta con el mismo impacto. Por lo mismo, en lugar de trasladar palabra por palabra, se debe trasladar el sentido completo, incluso si eso implica modificar estructuras o utilizar expresiones culturalmente distintas.

Este enfoque también considera que todo texto está contextualizado dentro de una cultura, y que muchas veces ciertos elementos no son directamente trasladables. Para estos casos, Nida introduce el concepto de transferencia cultural, en el que el traductor reemplaza un referente del texto original por uno culturalmente equivalente en la lengua de llegada, si con ello se logra conservar el efecto original.

Además, Nida propone una metodología basada en tres etapas para el proceso traductológico: análisis, transferencia y reestructuración. En la primera, el traductor analiza el contenido, la estructura y el significado del texto fuente. Luego, en la etapa de transferencia, transfiere el mensaje al sistema conceptual de la lengua meta. Finalmente, en la reestructuración, organiza el mensaje de forma que sea comprensible y aceptable para el público meta, respetando su nivel lingüístico y sus convenciones culturales.

En síntesis, este énfasis en el receptor como centro del acto traslativo y su idea de la equivalencia como efecto, más que como forma, sientan las bases para muchas discusiones actuales sobre adecuación, naturalidad y función comunicativa en la traducción.

### 2.1.2. Naturalidad

La importancia de la naturalidad como concepto dentro de la traducción resulta inescapable si se busca introducir un nuevo texto a la lengua meta. Este concepto ha sido desarrollado a lo largo de los años, reforzando su relevancia dentro del estudio y práctica de la traducción. Es posible encontrar una definición clásica en Newmark (1988) quien define a la naturalidad como la característica de un texto traducido de no poseer interferencias de la lengua origen y que se perciba como un texto propio de la lengua meta. Esta definición ha sido retomada y ampliada por distintos estudios contemporáneos con el fin de limitar de forma clara su alcance a través de perspectivas pragmáticas, funcionales y estéticas.

Por ejemplo, House (2015) plantea del término *covert translation* como «a translation which enjoys the status of an original source text in the target culture» (p. 56) refiriéndose a la cualidad de un texto de apreciarse como un originario del lenguaje meta gracias a la fluidez, a la adecuación pragmática del texto traducido y la capacidad de mantener las convenciones comunicativas y estilísticas propias del género textual en la lengua meta. Este planteamiento se alinea a la definición de Newmark de la naturalidad, otorgándole parámetros que se espera que una traducción natural cumpla.

Desde una perspectiva funcionalista, Nord (2005) enfatiza que la naturalidad debe entenderse en relación con la función comunicativa del texto, lo que implica que un texto natural no solo debe sonar fluido, sino que debe responder

adecuadamente a las expectativas del público meta y al propósito comunicativo definido por el encargo de traducción.

Por otro lado, Gile (2009) señala las adversidades a las que se enfrentan los traductores a la hora de traspasar de forma natural el mensaje de lengua origen a lengua meta, debido a las idiosincrasias de las lenguas, los matices léxicos y las diferencias textuales. Estos factores resultan en un texto que se aprecia como adecuado, pero que conlleva a que se perciba como «falso» por un receptor de la lengua meta. Además, Gile indica que el grado en el que se observa esta problemática se puede entender como un espectro, dado que resulta complejo cuantificar la percepción de los receptores en este ámbito.

Como resultado de la revisión teórica en torno a la naturalidad realizada para la presente investigación, y considerando los planteamientos de Newmark (1988), Nord (2005) y House (2015), se entiende la naturalidad como la cualidad de un texto traducido de ser percibido como si hubiera sido originalmente escrito en la lengua meta, debido a la ausencia de interferencias de la lengua origen y a la adopción de las convenciones y expectativas textuales del receptor meta.

## **2.2. Metáfora conceptual**

La metáfora, como concepto lingüístico, ha sido objeto de estudios desde diversas perspectivas. Entre los enfoques con los que se ha analizado se encuentra la arista clásica en la cual Aristóteles la define como «traslación de un nombre ajeno, o desde el género a la especie, o desde la especie al género, o desde una especie a otra

especie, o según la analogía» (1974, p. 204). La mencionada perspectiva ha sido cuestionada tal como se observa en Laks (2020), debido a que la definición clásica entiende la metáfora como un elemento más cercano al arte, un ornamento del lenguaje y no como parte del pensamiento.

El presente trabajo se sitúa dentro del marco de la lingüística cognitiva, disciplina en la que Lakoff y Johnson (1980) proponen que la metáfora consiste en comprender y experimentar una clase de fenómenos a partir de otra. En esta línea, Kövecses (2002) explica que, en una metáfora conceptual, el dominio conceptual desde el cual se estructuran las expresiones metafóricas se denomina dominio fuente, mientras que el dominio meta es aquel que se busca comprender mediante el primero.

Lakoff y Johnson proponen en su libro *Metaphors We Live By* (1980) que las metáforas se crean inconscientemente en la comunicación cotidiana con la finalidad de reducir a elementos simplificados un concepto de mayor complejidad.

La concepción conceptual de las metáforas (Lakoff y Johnson, 1980; Kövecses, 2002) se encuentra en una oposición casi completa a la definición clásica:

**Tabla 1**

<b>Metáfora según la perspectiva clásica</b>	<b>Metáfora según la perspectiva conceptual</b>
La metáfora es propia de las palabras	La metáfora es propia de los conceptos
Tiene un objetivo estético	Tiene un objetivo práctico para comprender conceptos
Se basa en las similitudes entre dos elementos	No se requiere que exista una similitud
Se usa deliberadamente	Es propia del lenguaje cotidiano y es espontáneo
Es una figura discursiva prescindible	Es inherente al razonamiento humano

*Comparación entre la definición clásica y conceptual de la metáfora*

Es posible observar la perspectiva conceptual de las metáforas en la idea de que «una discusión es una guerra». En esta idea, explicada por Lakoff y Johnson (1980), se observa dos dominios conceptuales: «discusión» y «guerra», el primero corresponde al dominio meta y el segundo al dominio fuente. El dominio meta es el que se busca explicar mediante el dominio fuente. Al asociar estos dos dominios es posible que emerjan nuevas expresiones metafóricas sistemáticas, debido a que su origen es una metáfora conceptual de origen. Por ejemplo, se pueden construir expresiones metafóricas como: «atacar el punto débil de su argumento», «destruir el argumento», «nunca he ganado una discusión contra él» y «su punto es indefendible». Lakoff y Johnson explican con ejemplos como estos que las expresiones metafóricas forman una manera sistemática de comunicación.

La noción de que la metáfora se estructura a partir de dominios conceptuales, y que estos generan múltiples expresiones metafóricas relacionadas entre sí, confirma las características fundamentales atribuidas por Lakoff y Johnson (1980). Según estos autores, la metáfora conceptual no solo se manifiesta de forma recurrente en el lenguaje cotidiano, sino que también opera de manera inconsciente y abarca tanto el pensamiento como el discurso.

También es relevante mencionar que las metáforas conceptuales pueden verse afectadas en su interpretación por elementos externos. A raíz de esto, Kövecses (2021) introduce la importancia del cotexto, el contexto lingüístico inmediato en el que se inserta una metáfora, para su correcta interpretación. Según el autor, aunque las metáforas conceptuales se organizan en torno a *mappings* (relaciones conceptuales entre dominios) relativamente estables, el significado específico de una metáfora en uso depende en gran medida de los elementos lingüísticos que la rodean. Es decir, el cotexto no solo restringe las posibles interpretaciones de una expresión metafórica, sino que también puede activar matices conceptuales particulares que no se desprenden únicamente del esquema general del *mapping*.

Entendiendo este concepto de metáfora, este trabajo hace distinción entre metáfora (la asociación entre dominios, configurada por medio de la relación figurativa entre ellos) de expresión metafórica (oración lingüística que contiene una metáfora). Este último concepto es el más adecuado para el corpus de esta investigación; no obstante, siempre se considera como una extensión de una metáfora.

### 2.2.1. Emociones en las metáforas

El estudio de las metáforas conceptuales ha evolucionado hacia el análisis de su relación con el ámbito emocional, estableciendo que muchas de nuestras experiencias afectivas también son estructuradas mediante *mappings* metafóricos. Tal como sostienen Lakoff y Johnson (1980), estos mapas permiten comprender fenómenos complejos a partir de conceptos más concretos. En el caso de las emociones, se ha observado que se conceptualizan mediante dominios experienciales como el calor, la presión o el movimiento, dichos dominios se proyectan sistemáticamente sobre estados afectivos abstractos.

Kövecses (2002) ha desarrollado ampliamente esta idea, señalando que las metáforas conceptuales vinculadas a las emociones suelen tener un carácter sistemático y recurrente en múltiples lenguas y culturas (p. 177). Por ejemplo, la emoción de la ira suele conceptualizarse como un líquido caliente en un recipiente, lo que da origen a expresiones como «estallar de ira» o «hervir de rabia». Estas expresiones reflejan un patrón cognitivo en el que el cuerpo actúa como un contenedor y la emoción como una sustancia que se acumula hasta desbordarse.

Reforzando la unión de las emociones con su expresión metafórica, Grady (1997) introduce el concepto de *primary metaphors* y fortalece la noción de los mapas como una proyección de estructuras conceptuales de un dominio fuente hacia un dominio meta. En el caso de las emociones, esta proyección permite construir una red coherente de expresiones metafóricas que articulan experiencias internas a partir de elementos perceptibles o físicos.

Mediante una investigación más actual, Gibbs (2017) sostiene que las metáforas emocionales no son simples adornos lingüísticos, sino estructuras fundamentales de nuestra cognición. Afirma que los *mappings* emocionales implican correspondencias sistemáticas entre dominios sensoriales y afectivos, las cuales reflejan y modelan nuestras respuestas emocionales ante el mundo. Así, emociones como la tristeza, el miedo o la vergüenza también se comprenden a través de metáforas espaciales, como en las expresiones «caer en depresión», «estar bajo presión» o «sentirse expuesto».

Diversos estudios han respaldado la idea de que las metáforas no solo representan emociones, sino que también contribuyen a su construcción discursiva y conceptual. Por ejemplo, Charteris-Black (2004) analiza el discurso político y argumenta que las emociones se construyen y persuaden mediante metáforas que apelan tanto a la experiencia corporal como al imaginario colectivo, reforzando así la función cognitiva, cultural y pragmática de estas expresiones figuradas. Por su lado, Barcelona (2000) en Gonzales (2001) profundiza en cómo las metáforas interactúan en la conceptualización emocional, mostrando que muchas expresiones emocionales surgen de esquemas corporales básicos que estructuran nuestras experiencias.

A partir de los estudios consultados, es posible argumentar que las metáforas conceptuales no solo nos permiten estructurar ideas abstractas como el tiempo o la argumentación, sino que también son esenciales para comprender, expresar y modelar nuestras emociones. Esta relación entre emoción y metáfora fortalece la

idea de que el lenguaje figurado es una herramienta fundamental de la cognición humana.

### **2.2.2. Metáforas en el discurso dramático**

Los textos dramáticos son manifestaciones literarias pensadas para ser representadas en un escenario, caracterizadas por su estructura dialógica y su orientación hacia la acción. Como plantea Pavis (1998), el texto dramático no debe entenderse únicamente como literatura escrita. En el género textual no se describen meramente diálogos en una estructura de escenas, sino que también se emplean elementos como el lenguaje, los gestos y la emoción, dando lugar a una experiencia estética multisensorial. Esta condición lo convierte en un terreno fértil para la exploración de fenómenos simbólicos y comunicativos complejos, como es el caso de las metáforas conceptuales.

En sus investigaciones sobre estilo y emoción, Semino (2010) destaca que las metáforas en el teatro contribuyen a una «dramatización cognitiva», en la que los conceptos abstractos se materializan a través del lenguaje figurado, facilitando así la identificación emocional del espectador con los personajes. Esto se observa en el diálogo dramático, ya que este se potencia con el uso de expresiones metafóricas conflictivas o defensivas, como «tus palabras son cuchillos» o «me dejaste en ruinas», debido a que estas expresiones metafóricas sistemáticas configuran relaciones interpersonales de forma automática y con una gran carga emocional. Goatly (2007) ha argumentado que estas metáforas en los textos literarios, especialmente en el

teatro, funcionan como mapas emocionales que codifican valores, tensiones y oposiciones presentes en la experiencia humana.

Esta íntima relación entre metáforas conceptuales y discurso dramático plantea importantes desafíos para la traducción, ya que traducir teatro implica no solo trasladar palabras, sino también conservar la intensidad emocional, la carga conceptual y la función dramática de las metáforas empleadas. Como sostiene Hanić et al. (2016), el traductor debe ser capaz de identificar los mapas conceptuales que estructuran las metáforas en la lengua original y reconstruirlos en la lengua meta, manteniendo su coherencia cognitiva, su fidelidad emocional y su naturalidad.

Es relevante agregar que en contextos escénicos donde la palabra cobra presencia viva, la metáfora despliega su potencial performativo. Tseng (2010) propone que la metáfora no solo transmite significados y pensamientos, sino que también ejerce fuerza ilocutiva y perlocutiva, lo que significa que, en escena, la metáfora puede «actuarse» y producir efectos emocionales en el público, más allá de lo conceptual. Este aspecto se menciona en esta sección, debido a que, a pesar de su importancia a nivel de análisis entorno a los textos dramáticos, este factor ilocutivo se escapa del alcance y enfoque del presente trabajo, por lo cual no se tendrá en consideración.

### **2.3. El reconocimiento y clasificación de emociones**

En palabras de Ekman (1992), las emociones juegan un rol importante en las experiencias humanas, ya sea a la hora de tomar decisiones o de establecer vínculos afectivos. Debido a esto, varias disciplinas se han enfocado en estudiarlas para cantidad diversa de aplicaciones. Trabajos como el de Dhanapala (2024), en el que se explora su influencia en la música, otros, como el de Canal (2022) o el de Lek (2023), revisan la relación de las emociones con las expresiones faciales. Lo que todos estos trabajos tienen en común es uno de los aspectos más importantes y primordiales para el estudio de las emociones: el reconocimiento y clasificación de las emociones.

Esta área del estudio de las emociones ha tenido un crecimiento significativo en los últimos años, incentivado principalmente por la evolución de las redes neuronales y por consiguiente de inteligencias artificiales, a fin de entrenarlas para adecuarlas a los usuarios humanos y sus necesidades (Shaw et al., 2016).

Para clasificar e identificar emociones en textos, es necesario un modelo de emociones, por ejemplo, el modelo de Emociones Básicas de Ekman (1992), basado en aquellas emociones que son universalmente relacionables con expresiones faciales, tal y como se hizo en el trabajo de Ghazi et al. (2016) o el de Bandhakavi et al. (2016). Estas emociones son alegría, tristeza, miedo, enojo, disgusto y sorpresa. En el caso del trabajo de Ghazi, se categorizaron oraciones completas en una de estas seis emociones, mientras que en el de Bandhakavi se hizo con conjuntos de palabras.

Sin embargo, otros modelos agrupan, mezclan o expanden estas emociones de varias maneras y ahondan en mayor profundidad en la vasta gama de emociones humanas y sus matices. Entre esos modelos, existe el modelo de la Rueda de las Emociones de Plutchik (1980), el cual será usado en este trabajo debido a su amplio uso en análisis de textos.

El modelo de análisis de emociones de la Rueda de las Emociones o por su nombre en inglés *Wheel of Emotion* ha sido ampliamente usado en análisis de textos por la versatilidad que ofrece a la hora de identificar la emoción e intensidad específica que se busca transmitir. El modelo consiste en una rueda o rosa cromática en la que se agrupan en torno a un eje un conjunto de emociones, de tal manera que las emociones opuestas quedan en los extremos del eje. Las emociones recopiladas por Plutchik son: éxtasis, admiración, terror, asombro, pena, odio, furia y vigilancia. Dicha organización de las emociones fue utilizada por Mondal (2020) en su trabajo y las agrupó en cuatro dinámicas binarias: amor vs. odio, alegría vs. tristeza, confianza vs. disgusto, y anticipación vs. sorpresa. Así mismo, el modelo de Plutchik mide la intensidad de las emociones en función de su distancia con eje: las emociones más intensas están más cerca al eje mientras que las menos intensas se encuentran más lejos.

Entre los trabajos que ocuparon este modelo para analizar textos se encuentran el de Arroyo-Barrigüete (2023) en el que se analizaron obras de ficción narrativa de H.P. Lovecraft, el de Mondal (2020) en el que se trabajó con tweets, el de Ganganwar (2021) en el que se identificaron emociones en correos electrónicos

legales, e incluso para la enseñanza de lenguas extranjeras como el trabajo de Brkić (2024). Varios de estos trabajos eligen este modelo no solo por su simpleza y precisión, sino también por la posibilidad de modificar las emociones dentro de la rueda para adaptarla a las necesidades del trabajo. Estos beneficios se observan, por ejemplo, en los resultados de Arroyo-Barrigüete, en los cuales, gracias al modelo de Plutchik, se observa la presencia dominante de las emociones terror, tristeza y admiración en las obras de Lovecraft. Además, consigue determinar una relación entre el terror y la tristeza debido a su cantidad heterogénea de coincidencias.

Por estas características, en este trabajo se utilizará el modelo de Plutchik para identificar las emociones en las metáforas encontradas en el conjunto de textos dramáticos escogidos antes de ser traducidas por los LLM y para comprobar si esas mismas emociones se conservan después de la traducción.

#### **2.4. Modelos extensos de lenguaje**

Los modelos extensos de lenguaje, también llamados grandes modelos de lenguaje, o por sus siglas en inglés LLM, son modelos complejos de inteligencia artificial diseñados específicamente para trabajar con el lenguaje humano, ya sea para analizar textos o crearlos. Como explica Shanahan et al. (2023), los LLM funcionan con un sistema de *tokens*: dada cierta secuencia de palabras, signos o símbolos a la que se le asignó determinados *tokens*, el modelo trata de predecir cuáles deberían ser los siguientes.

Si bien su uso generalizado ha tenido un amplio recibimiento por parte del público, ya sea por estudiantes, docentes, académicos, redactores, editores, entre otros grupos relacionados, también ha generado olas de críticas por su uso indiscriminado y denominado «poco ético», el cual se acentúa por la falta de reflexión ante esta problemática en su uso (Heilinger,2022). Sin embargo, como bien explica Zhao et al. (2023), comprender y explicar estos modelos es crucial para dilucidar sus comportamientos, limitaciones e impactos en la sociedad. Relevante para este tema es el trabajo de Chang et al. (2023), donde explora los métodos y las áreas que deben ser evaluadas, entre ellas el procesamiento general del lenguaje, su razonamiento, el uso en el ámbito médico y su uso ético.

Según Naveed et al. (2024), la necesidad de modelos especializados en tareas relacionadas con el lenguaje surge de la creciente demanda por modelos capaces de resumir, recopilar información, mantener conversaciones y traducir.

#### **2.4.1. Traducción con modelos extensos de lenguaje**

Los LLM han mostrado capacidades destacables a la hora de traducir textos. Como bien se explica en el trabajo de Moslem (2023), estos modelos pueden adaptarse a diferentes pares de oraciones y terminología en su contexto a diferencia de la mayoría de los modelos tradicionales de traducción automática, tal y como se muestra en el trabajo de Wang et al (2023), donde los modelos de ChatGPT, GPT-3.5 y GPT 4, son capaces de superar los modelos tradicionales de traducción automática a la hora de traducir documentos, y que, gracias al constante desarrollo y evolución, muestran un potencial prometedor para el futuro.

Sin embargo, como se explica en el trabajo de Xu (2024), solo los modelos de lenguaje más grandes y con mayores bases de datos muestran este tipo de resultados, y solo si son idiomas con vastos recursos en sus bases de datos, ya que, en lo que respecta a idiomas que en comparación tienen menos recursos, el bajo rendimiento de los modelos se hace evidente frente a modelos de traducción tradicionales. Así mismo, en el trabajo de Zhang et al. (2023), se encontró que estos modelos pueden confundir o malinterpretar los *prompts*, mezclar parte de este con el texto al traducir y generar texto no relacionado con la petición hecha, también conocido como «alucinaciones». Además, se observó que la calidad de las traducciones puede variar en función de la combinación de idiomas, así como de la cantidad del texto.

Mientras que los trabajos enfocados en las traducciones realizadas por los LLM han hecho hincapié en la calidad de las traducciones, los análisis y evaluaciones no han prestado particular atención a al análisis literario de las mismas, y más específicamente, a si estas traducciones pueden ser consideradas naturales, si mantienen la carga emocional o si producen el mismo impacto al lector en la lengua meta que en la lengua origen.

#### **2.4.2. Fidelidad emocional**

Con el considerable desarrollo tecnológico que han experimentado los LLM en los últimos años, resulta evidente que su capacidad para procesar y emular emociones expresadas en textos se ha visto potenciada. Esta evolución ha generado

nuevas posibilidades en el campo de la traducción automatizada, especialmente en lo que respecta a la preservación de la carga afectiva de los textos originales.

Si bien las primeras generaciones de sistemas de traducción automática mostraban limitaciones en el reconocimiento de matices emocionales o metáforas figuradas, los modelos más recientes, basados en arquitecturas *transformer* (como GPT o BERT), han demostrado una sensibilidad creciente hacia la estructura emocional del discurso (Brown et al., 2020, p. 12).

La importancia de esta capacidad radica en que la emoción no solo contribuye al significado del mensaje, sino que también influye en la manera en que los receptores interpretan y reaccionan a la información. La incapacidad de reproducir correctamente la carga afectiva puede llevar a malentendidos, pérdida de intención comunicativa o a una experiencia de usuario insatisfactoria, especialmente en ámbitos sensibles como la literatura, la publicidad o la comunicación intercultural (Balahr & Turchi, 2014). No obstante, el reconocimiento y la reproducción precisa de emociones en la traducción automática siguen siendo un desafío debido a la complejidad y subjetividad inherentes a las expresiones emocionales, así como a las variaciones culturales en su manifestación.

Investigaciones recientes indican que la incorporación de mecanismos de atención y aprendizaje profundo ha permitido a estos modelos captar no solo la semántica literal, sino también las sutilezas emocionales presentes en los textos (Vaswani et al., 2017; Devlin et al., 2019). Esto es fundamental para lograr una traducción emocionalmente precisa, dado que la emoción juega un papel crucial en

la comunicación efectiva y en la representación adecuada de sentimientos entre lenguas.

Según Feng et al. (2025) denominan esta capacidad como fidelidad emocional, por su término en inglés *emotional fidelity*. En palabras de estos autores, la fidelidad emocional se define como el grado en que un sistema de inteligencia artificial es capaz de simular una emoción conforme a las directrices que se le proporcionan. Esta definición resulta significativa dado que determina que esta es una capacidad con una escala no estática, es decir, en virtud de la base de datos de cada LLM estos pueden tener una mayor o menor fidelidad emocional.

Debido a lo anterior, esta definición es la que se utiliza en el presente trabajo. Sumando a esta justificación, el término definido por Feng et al. (2025) es el considerado en esta investigación, ya que provee una formulación que permite ser simplificada para ser presentada a individuos sin previos conocimientos en esta materia, lo cual es de vital importancia para los propósitos de este trabajo.

### **3. PREGUNTA DE INVESTIGACIÓN**

¿Las traducciones de expresiones metafóricas realizadas por modelos extensos de lenguaje logran aplicar los conceptos de naturalidad y fidelidad emocional?

### **4. OBJETIVOS**

#### **4.1. Objetivo general**

Describir la fidelidad emocional y la naturalidad de las traducciones de expresiones metafóricas realizadas por modelos extensos de lenguaje en un corpus compuesto por textos dramáticos.

#### **4.2. Objetivos específicos**

1. Identificar la emoción expresada en las expresiones metafóricas del corpus según el modelo de Plutchik.
2. Determinar si se traspasa adecuadamente la fidelidad emocional y la naturalidad mediante los resultados de la encuesta a hablantes nativos.
3. Calificar el desempeño de los modelos extensos de lenguaje a partir de los resultados de la encuesta según los criterios de fidelidad emocional y naturalidad.

## 5. METODOLOGÍA

El diseño de esta investigación es mixto, puesto que busca describir la naturalidad y fidelidad emocional de expresiones metafóricas relacionadas con la exposición de emociones traducidas con LLM mediante la recolección de datos a partir de una encuesta, para luego ser analizados en función de su desempeño. A su vez, se pretende analizar los resultados obtenidos mediante la comparación con el marco teórico desarrollado previamente y entregar razones tras estos resultados. Además, cuenta con un alcance descriptivo, dado que pretende organizar y presentar los resultados de la encuesta.

El corpus consistió en textos dramáticos, los cuales fueron extraídos de la página oficial del Centro Latinoamericano de Creación e Investigación Teatral (CELCIT). Estas obras fueron escogidas en primera instancia con base en dos criterios: 1) que el texto haya sido publicado en los últimos treinta años y 2) que haya sido descargado al menos 20.000 veces. El criterio de los treinta años se estableció con el objetivo de garantizar que el lenguaje y expresiones presentes en las obras resultasen contemporáneos, evitando una obstrucción en su comprensión por parte de los LLM o los encuestados y de establecer que las obras seleccionadas fueran seleccionadas. Por su parte, las descargas mínimas se utilizaron como margen para emplear obras que generen el mayor interés por parte del público general, lo que, adicionalmente, podría conllevar a que estas obras sean traducidas con un LLM.

Tras ser seleccionadas, las seis obras que cumplieron estos criterios fueron revisadas manualmente, con el fin de encontrar obras que contuvieran, como mínimo, una expresión metafórica que cumpliera los siguientes requisitos: ser parte de un diálogo en el que un personaje exprese su estado anímico y ser explicable mediante el concepto de dominios propuesto por Lakoff y Johnson (1980).

Como resultado de este segundo proceso de selección, se obtuvieron seis textos, los cuales formaron el corpus de esta investigación. Los textos dramáticos seleccionados fueron: Volpone de David Amitín y Mauricio Kartun (2000), La Cena de Roberto Perinelli (2000), La Caja Negra de César De María (2000), La Casa de los Viejos de Mauricio Kartun (1999), Sumario de la Muerte de Kleist de Alejandro Tantanian (1998) y Parejas de Susana Lastrero (1995).

Una vez establecido el corpus, se analizaron los textos en búsqueda de las expresiones metafóricas que se usarían como muestra. Estas expresiones metafóricas fueron escogidas según los siguientes criterios: ser expresiones que puedan entenderse su función metafórica en concordancia a los parámetros de Lakoff y Johnson (1980) y ser clasificables por los investigadores con una de las emociones presente en el modelo de Plutchik (1980) (ver anexo 1).

Con los resultados obtenidos se eliminaron las dos emociones con menos cantidad de expresiones metafóricas, otorgando un resultado final de doce expresiones metafóricas (anexo 2) con dos expresiones metafóricas representado cada emoción.

**Tabla 3**

Expresión metafórica	Emoción propuesta por Plutchik

La siguiente fase de la investigación fue utilizar tres LLM para que estos tradujeran las expresiones metafóricas a inglés y alemán. En la presente investigación se emplearon los siguientes LLM: DeepSeek V-3, GPT-4o y Llama 4 Maverick. Estos LLM fueron seleccionados por representar distintos enfoques en el desarrollo de modelos de lenguaje de última generación y por su presencia contemporánea.

Para realizar el *prompt*, se tuvo en consideración, como explica Ekin, 2023, la eficiencia de dar la mayor cantidad de contexto posible, de la manera más clara que se pueda, en una cantidad controlada de texto, enfocada en la precisión. Esto se basa en el principio de que una cantidad excesiva de texto puede llevar a que el modelo de lenguaje produzca resultados que no sean los esperados. El *prompt* utilizado en esta investigación solicitó que cada expresión metafórica fuera traducida a la vez al inglés y alemán, que tuviera en consideración la naturalidad, la fidelidad emocional; a su vez, se le entregó el contexto de las expresiones metafóricas y un contexto general sobre la situación en la que es pronunciada la expresión metafórica. El *prompt* utilizado es el siguiente:

Traduce solo la frase entre comillas al inglés y alemán. Procura dar una sola traducción, lo más natural posible y manteniendo la intencionalidad y emoción de la expresión metafórica original:

—

— “”

—

Ten en cuenta el siguiente contexto:

*Muestra de prompt utilizado*

Las traducciones proporcionadas por los LLM fueron recopiladas en tres tablas, una designada para cada LLM, con el fin de organizarlas para que su información fuera plasmada en la encuesta.

**Tabla 3**

Expresión metafórica original	Traducción al inglés (LLM1)	Traducción al alemán (LLM1)

Posteriormente, se elaboró la encuesta. Para ello se estableció que la encuesta tendría como grupo de muestra a estudiantes de intercambio que estuvieran realizando sus estudios, sin importar su carrera, en la Universidad de Concepción al momento de la realización de este trabajo. Como segundo parámetro, estos estudiantes debían poseer al inglés o al alemán como primera lengua.

Resulta notable destacar que al no controlar la carrera que estudiaban los encuestados, se puede dar la situación en que alguno de los sujetos cuente con un trasfondo académico que les permita tener una comprensión más profunda de las expresiones metafóricas, como lo podría ser un estudiante de idiomas. Debido a limitaciones de tiempo y de alcance de este trabajo, se optó por no considerar este factor en los criterios de esta investigación.

La mencionada encuesta fue revisada y aprobada por el Comité de Ética, Bioética y Bioseguridad de la Facultad de Humanidad y Artes de la Universidad de Concepción. Además, la encuesta en línea contó con un consentimiento informado para los encuestados.

Con el fin de facilitar el proceso de evaluación para los encuestados se les envió una encuesta con todos los textos (salvo las expresiones metafóricas antes de ser traducidas) en su primer idioma; es decir, se elaboraron dos encuestas, una en inglés y otra en alemán.

La estructura de la encuesta fue la siguiente: en primer lugar, se explicó el propósito de la investigación y la actividad que se les solicitaba, siendo esta que evaluaran, bajo su criterio, en una escala del 1, como nota mínima, al 5, como nota máxima, la naturalidad y fidelidad emocional de expresiones metafóricas traducidas que se les presentaría más adelante. Para facilitar este proceso, después de la explicación inicial se describieron ambos conceptos siguiendo la definición de naturalidad propuesta por Newmark (1988) y de fidelidad emocional de Feng et al. (2025).

A continuación, se presentaba el consentimiento informado. Una vez aceptado, se les pedía a los encuestados que confirmaran que tuvieran al alemán o el inglés como primera lengua, así como que indicaran su país de origen.

En la sección respectiva de cada expresión metafórica, se presentó la expresión metafórica en español, luego un párrafo que explicara el contexto de la escena, y posteriormente se indicó la emoción a la que estaba asociada la expresión metafórica.

**"Todas las aves de rapiña revolotean creyendo que pronto seré carroña"**

The character who utters this line is an old man, lying in bed, watching as his lawyer, an aging moneylender, and a merchant, visit him bearing gifts—competing for his alleged fortune and quietly awaiting his death.

Emotion: loathing

*Ejemplo de estructura utilizada en encuesta para presentar expresión metafórica*

A continuación, se les enseñó a los participantes tres traducciones realizadas por los LLM y que calificaran estas traducciones entre 1 y 5 en términos de naturalidad y fidelidad emocional, en consideración de la definición entrega previamente.

Evaluate the **naturalness** of the following expressions:

1. All the birds of prey are circling, thinking I'll soon be carrion
2. All the vultures are circling, thinking I'm about to kick the bucket
3. All the birds of prey are circling, believing I'll soon be carrion

	1	2	3	4	5
Translation 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Translation 2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Translation 3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*Ejemplo de estructura utilizada en encuesta para evaluar naturalidad*

Evaluate the **emotional fidelity** of the mentioned expressions:

	1	2	3	4	5
Translation 1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Translation 2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Translation 3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*Ejemplo de estructura utilizada en encuesta para evaluar fidelidad emocional*

Se evaluaron las doce expresiones metafóricas que formaron la muestra. Cada expresión metafórica tuvo dos evaluaciones asociadas, una de naturalidad y otra fidelidad emocional.

Los resultados de esta encuesta fueron recopilados para su posterior análisis, comparación y ponderación. De esta forma se pudo observar el desempeño de cada LLM en promedio. Las calificaciones del desempeño fueron evaluadas utilizando una escala con un nivel de exigencia del 50 %, basada en un rango de 1 a 5. A partir de ello, se establecieron los siguientes parámetros de interpretación:

- Promedio menor a 3,0 = reprobatorio
- Promedio igual a 3,0 = neutro
- Promedio sobre 3,0 = aprobatorio

La encuesta fue respondida finalmente por diez participantes, cinco encuestados hablantes de inglés y cinco encuestados hablantes de alemán en un plazo de un mes.

## 6. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

En esta sección se presentan y se discuten los resultados obtenidos a partir de la encuesta aplicada a hablantes nativos de ambos idiomas, tanto en términos de naturalidad como de fidelidad emocional.

En primera instancia, se expondrán algunos de los resultados obtenidos por los LLM en sus traducciones según las dos métricas mencionadas. Posteriormente, se mostrarán los resultados ponderados de dichas métricas, con el fin de comparar el desempeño traductológico de los modelos y la percepción de los hablantes nativos ante sus traducciones. Finalmente, se llevará a cabo un análisis general de los resultados, con el objetivo de identificar posibles patrones que indiquen en qué tipos de expresiones metafóricas los LLM tuvieron un mejor o peor desempeño.

La encuesta realizada de forma *online* fue respondida por diez hablantes nativos (cinco hablantes por cada idioma) en el plazo de un mes. En dicha encuesta calificaron en una escala del 1, siendo esta la calificación menor, al 5, la calificación más alta, la naturalidad y fidelidad emocional de doce expresiones metafóricas traducidos por LLM.

### 6.1. DeepSeek V-3

#### 6.1.1. Naturalidad

Las calificaciones realizadas a la naturalidad de las expresiones metafóricas en inglés de este LLM resultaron en un promedio de 2,93 en una escala de 5, el

promedio más bajo de esta investigación. Este valor indica que DeepSeek presentó problemas para mantener este elemento, lo cual puede interpretarse como una interferencia para el entendimiento del hablante nativo.

Además, se puede observar que la moda de las calificaciones en este apartado fue 3, lo cual indicaría una capacidad mediana de producir una traducción de expresiones metafóricas de forma natural.

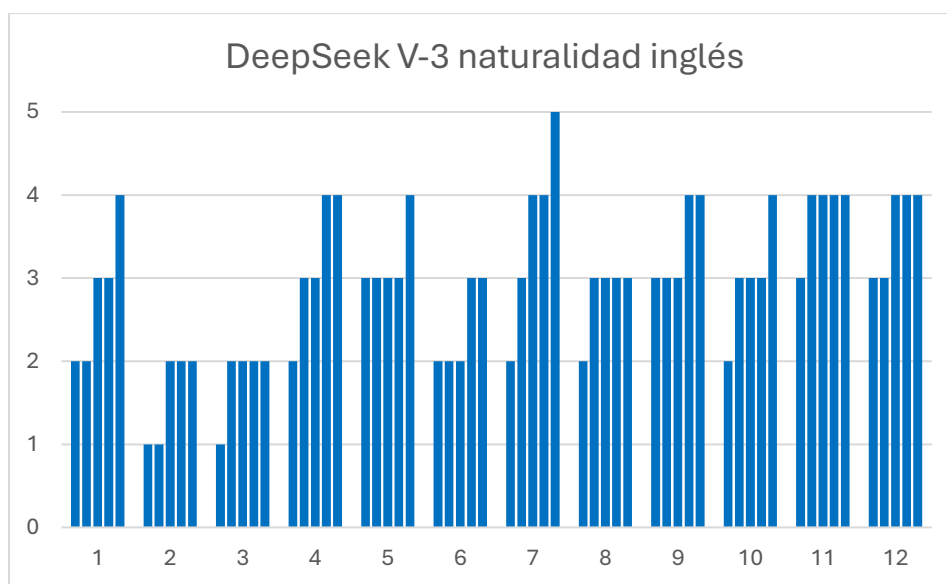


Gráfico 1: Resultados doce expresiones metafóricas por parte de hablantes nativos de inglés en términos de naturalidad. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.

El desempeño del LLM DeepSeek V-3 presenta una variación al momento de traducir al idioma alemán (ver anexo 7). Se observó un cambio en la moda de las calificaciones, siendo 2 el valor más repetido en la evaluación de los hablantes nativos. Este hecho supondría un indicativo de una peor ejecución de la traducción en términos de naturalidad; no obstante, a su vez, hubo un incremento de las evaluaciones con calificación de 4 y 5. Lo anterior indica una mayor polarización en la

percepción de los hablantes nativos. Este fenómeno podría responder a elementos como el entendimiento lingüístico de naturalidad según el lugar geográfico de los encuestados, esta posibilidad no es posible detallar con certeza.

El desempeño promedio en la categoría de naturalidad en inglés obtuvo un resultado de 2,93. Por su lado, el desempeño de DeepSeek V-3 en este concepto en el idioma alemán aumentó a 2,98.

### **6.1.2. Fidelidad emocional**

El escenario anterior se contrapone al rendimiento mostrado por este LLM en cuanto a fidelidad emocional, dado que al traducir al idioma alemán presentó un resultado de 3,01. Este promedio es el más alto obtenido por DeepSeek V-3.

Dicho promedio podría indicar que, aún con ciertas dificultades, tal como lo ejemplifican las calificaciones con valor de uno que obtuvo en algunos casos, DeepSeek V-3 tendría como mayor fortaleza el realizar traducciones al alemán fieles a nivel emocional.

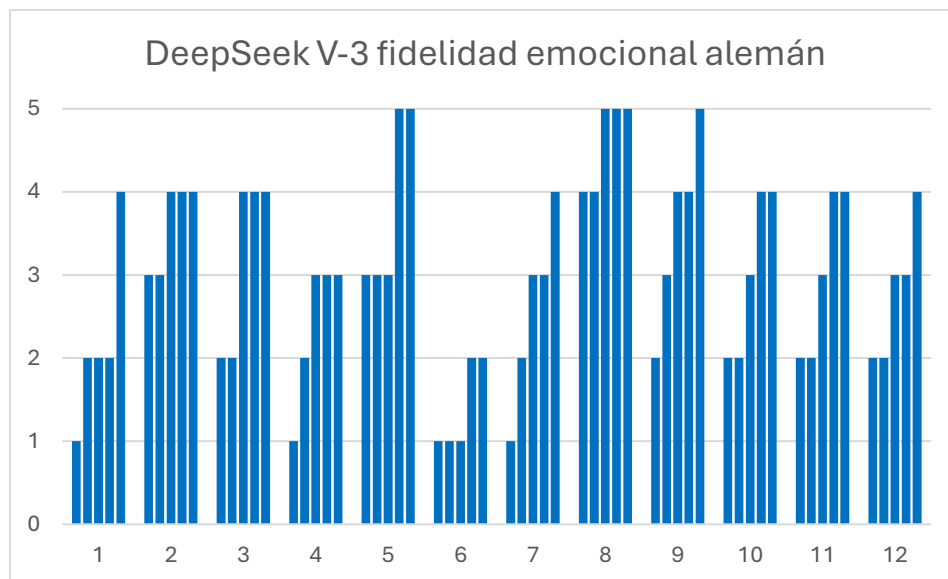


Gráfico 2: Resultados doce expresiones metafóricas por parte de hablantes nativos de inglés en términos de fidelidad emocional. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.

Sin embargo, resulta importante denotar que este desempeño no se observa en el caso de sus traducciones al inglés (ver anexo 8), en las cuales obtuvo un promedio de 2,96. Esto podría ser un indicativo de una dificultad en el traspaso de las emociones que se acentúa en este idioma.

El promedio final de fidelidad emocional mostrado por DeepSeek V-3 en inglés dio un resultado de 2,96. El promedio final de esta categoría en alemán mostró un aumento con un resultado de 3,01.

## 6.2. GPT-4o

### 6.2.1. Naturalidad

De forma similar a lo visto en el LLM anterior, GPT-4o presenta una capacidad de naturalidad más notable en el idioma alemán que en el inglés. En el caso del primer idioma, obtuvo un promedio de 3,81, uno de los más altos conseguidos por

este LLM. Si bien estos resultados tienen un gran elemento de subjetividad por parte de los encuestados, es posible teorizar que este resultado notablemente positivo podría deberse a que GPT-4o logró adaptar de mejor manera sus traducciones al alemán debido a la estructura gramatical más estricta de este idioma, lo que podría facilitar su aceptación como natural por parte de los hablantes nativos.

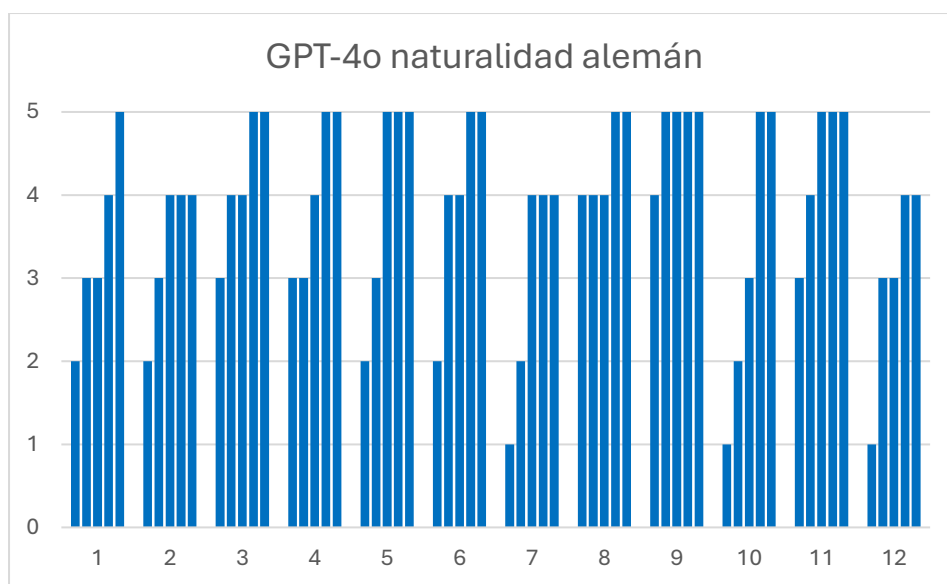


Gráfico 3: Resultados doce expresiones metafóricas por parte de hablantes nativos de alemán en términos de naturalidad. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.

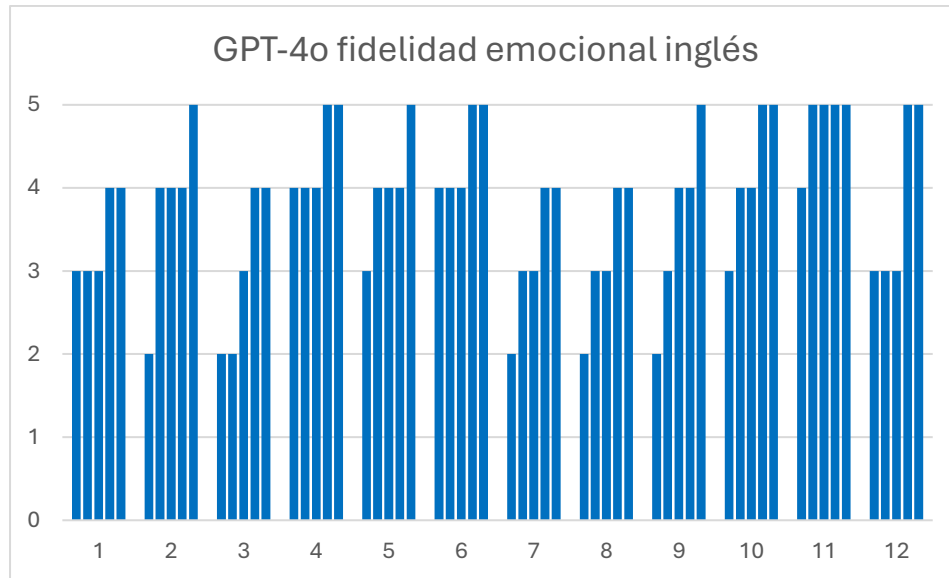
Este rendimiento considerablemente positivo se aprecia, a su vez, en la repetición de la calificación 5 en sus resultados. No obstante, es destacable que incluso en las expresiones metafóricas con mejores resultados totales algunos hablantes le otorgaron la calificación mínima, dando cuenta así de que la traducción realizada por GPT-4o no alcanzó una naturalidad generalizada.

Con respecto al rendimiento de naturalidad en el idioma inglés, si bien obtuvo un promedio final menor al visto en alemán, es relevante notar que obtuvo un menor número de calificaciones mínimas (ver anexo 10). Esto deriva en la conclusión de que su promedio es resultado de una mayoría de calificaciones entre el 3 y el 4, dando a entender que realizó un desempeño aceptable.

Finalmente, GPT-4o consiguió un promedio de 3,48 en su desempeño de naturalidad en inglés, mientras que en el caso del idioma alemán el resultado fue 3,81.

### **6.2.2. Fidelidad emocional**

Respecto a la capacidad de fidelidad emocional, GPT-4o presenta un fenómeno único para esta investigación. A raíz de los resultados obtenidos fue posible notar que el idioma en que consiguió un mayor puntaje en su promedio final de fidelidad emocional fue distinto al visto en el apartado de naturalidad. Esto quiere decir que este LLM presentó una mayor capacidad de fidelidad emocional en el idioma inglés que en el alemán (ver anexo 13).



*Gráfico 4: Resultados doce expresiones metafóricas por parte de hablantes nativos de inglés en términos de fidelidad emocional. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.*

Si bien la diferencia entre promedios no resulta drástica, es remarcable que la perspectiva de los hablantes nativos respecto a naturalidad y fidelidad emocional no coincida en el mismo idioma.

Este resultado podría ser consecuencia de la distribución subjetiva de las respuestas de los hablantes encuestados o podría deberse a que el LLM GPT-4o posee una capacidad con mayor desarrollo de traspaso de emociones en el idioma inglés que en alemán y que esta capacidad no sea condicionada por su capacidad de producción de expresiones metafóricas naturales.

Con un promedio idéntico a lo visto anteriormente, GPT-4o consiguió un resultado de 3,81 en la categoría de fidelidad emocional en inglés. Por el contrario, a

diferencia de lo observado con anterioridad, el idioma alemán sufrió una disminución en su resultado final con un 3,73.

### **6.3. Llama 4 Maverick**

#### **6.3.1. Naturalidad**

En términos de naturalidad, el LLM Llama 4 Maverick presentó un rendimiento significativamente positivo en el idioma inglés, con un promedio final de 4,15. Este promedio es el más alto obtenido en el desarrollo de esta investigación.

Este resultado encuentra su causa en la mayoría de los encuestados calificando la naturalidad de las traducciones con el número máximo. Este hecho solo se observó con este LLM, lo cual sería un indicio de una capacidad superior por parte de Llama 4 Maverick para cumplir las expectativas de un texto en inglés en términos de naturalidad, según lo observado en esta investigación.

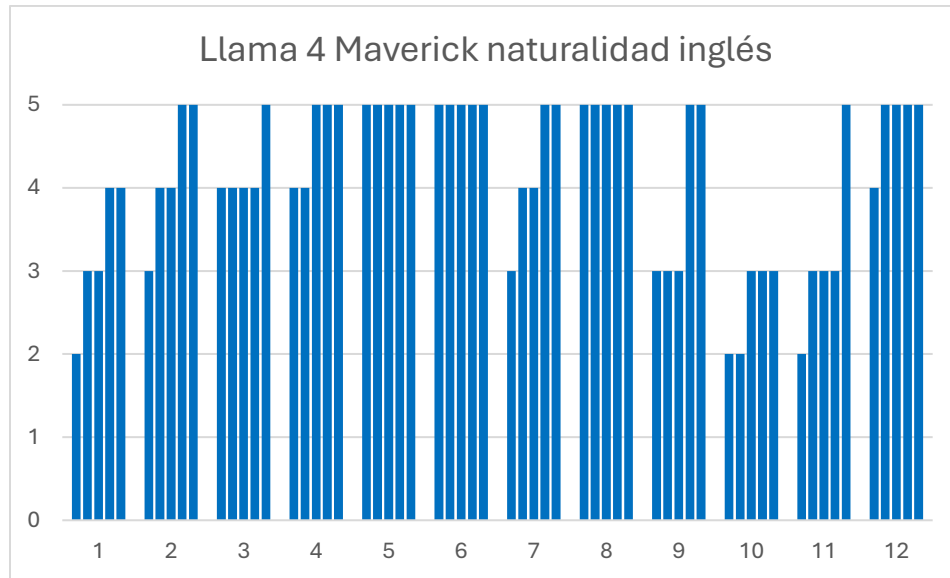


Gráfico 5: Resultados doce expresiones metafóricas por parte de hablantes nativos de inglés en términos de naturalidad. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.

Esto resulta un contraste considerable al comparar con los resultados de naturalidad en alemán (ver anexo 15). Este apartado otorgó un resultado que evidenciaría que este LLM no poseería la misma capacidad de naturalidad en este idioma.

En términos de naturalidad en el idioma inglés, el LLM Llama 4 Maverick fue calificado con 4,15 en su promedio final. Este resultado crea una brecha significativa al compararlo con su promedio final 3,26 de naturalidad en alemán.

### 6.3.2. Fidelidad emocional

El patrón observado en el apartado anterior se presenta nuevamente, conllevando a que el LLM ofrezca mejores resultados, en este caso fidelidad emocional, en el idioma inglés.

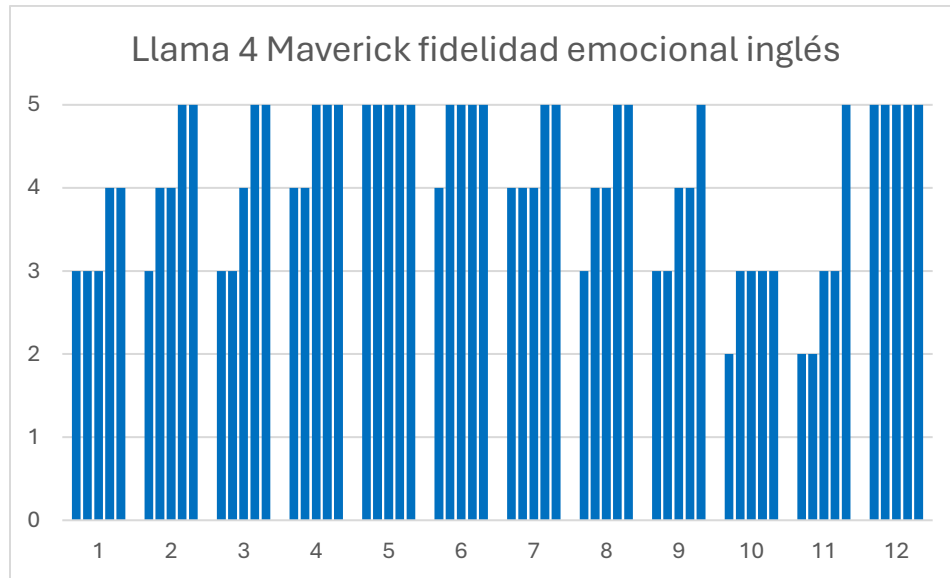


Gráfico 6: Resultados doce expresiones metafóricas por parte de hablantes nativos de inglés en términos de fidelidad emocional. Eje X: expresión metafórica evaluada, eje Y: calificación otorgada por cada encuestado.

Esta diferencia notable se produce a pesar de que la perspectiva de los hablantes nativos le otorgó una moda de 4 a las traducciones en alemán (ver anexo 17) y un 5 a las traducciones inglés. El resultado final parece haber sido afectado por la cantidad de calificaciones de 1 y 2, lo cual conllevó a que el promedio de fidelidad emocional en alemán disminuyera en mayor cantidad que la vista en inglés.

Los resultados finales en la categoría de fidelidad emocional conseguimos por este LLM fueron 4,1 para el idioma inglés y 3,43 para el idioma alemán.

#### 6.4. Observaciones generales

A partir de los resultados recopilados es posible determinar y clasificar el rendimiento de los LLM percibido por los hablantes nativos con el propósito de

establecer una propuesta de jerarquía para su uso en el proceso de traducción de expresiones metafóricas. Con esta determinación, ofrecemos el siguiente análisis:

Tanto en el aspecto de naturalidad como de fidelidad emocional, DeepSeek V-3 demostró ser el LLM con peor desempeño. Este hecho se evidencia en sus bajos promedios finales, tres de ellos obteniendo una calificación menor a 3,0. Lo anterior también indica que este desempeño deficiente se aplica a ambos idiomas estudiados. Además, como fue mencionado con anterioridad, este LLM posee el promedio final más bajo registrado en esta investigación.

Los datos recopilados son un indicativo de que DeepSeek V-3 no estaría cumpliendo los aspectos de naturalidad y fidelidad emocional, a pesar de haber sido solicitados en el *prompt* de traducción.

Sobre GPT-4o, este LLM obtuvo calificaciones destacables en el aspecto de naturalidad y fidelidad emocional en sus traducciones al idioma alemán, siendo las más altas observadas en este idioma. Dichos resultados indicarían que GPT-4o posee una capacidad aceptable de aplicar estos elementos al traducir, al menos al idioma alemán. Sin embargo, también resultaría relevante mencionar que su promedio final respecto a la fidelidad emocional en el idioma inglés fue igualmente positivo, siendo una señal de un desarrollo de esta capacidad en ambos idiomas.

Con respecto al último LLM, Llama 4 Maverick, se pudo demostrar su eficacia al aplicar el concepto de naturalidad y fidelidad emocional en sus traducciones al inglés gracias a los puntajes establecidos por los encuestados. Lo anterior resultó en

que este LLM consiguiera los promedios más altos para el idioma inglés y, a su vez, el promedio más alto de esta investigación. Por el contrario, sus resultados en las traducciones al alemán no lograron las mismas calificaciones, indicando una posible facilidad en la labor de traducción al inglés por sobre el alemán.

Lo analizado indicaría que DeepSeek V-3 no tendría la capacidad de generar texto natural, según la definición de Newmark (1988) ni de que estos sean fieles a nivel emocional, siguiendo el concepto de Feng et al. (2025). Por el contrario, GPT-4o y Llama 4 Maverick sí cumplen los planteamientos de estos autores.

Resulta importante destacar la diferencia matemática entre el promedio más bajo (2,93 – DeepSeek V-3) y el más alto (4,15 – Llama 4 Maverick) de 1,22. Esto tiene el fin de establecer la disparidad de resultados otorgados por los hablantes nativos, lo que podría ser un indicio de la capacidad de estos modelos para realizar traducción de expresiones metafóricas considerando los aspectos de naturalidad y fidelidad emocional.

Además del análisis general que se puede realizar respecto al desempeño de los LLM, resulta relevante para esta discusión el observar el recibimiento de las traducciones. Con este fin, se han seleccionado algunas de las expresiones metafóricas como ejemplos de las distintas posturas reflejadas en las respuestas de los encuestados.

El primer caso notable se encuentra en las expresiones metafóricas 6 y 8, las cuales fueron las únicas expresiones metafóricas que obtuvieron una calificación

máxima de forma unánime en términos de naturalidad. En particular, los mencionados resultados se presentaron en la evaluación de naturalidad en inglés y fueron traducidos por el LLM Llama 4 Maverick.

**Tabla 4**

<b>Expresión metafórica en español</b>	<b>Traducción al inglés por Llama 4 Maverick</b>
Horizonte negro para usted, eso es lo que veo por delante	Black horizon ahead for you, that's what I see coming
Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	From then on, Mom was like a sleeping swan on the shore of my life

La aprobación por parte de los hablantes nativos indicaría que estas traducciones responden adecuadamente a los aspectos que autores como House (2015) enfatizaban para que un texto se percibiera como propio de una lengua. Entre los principales aspectos de los que hablaba House, se encuentran la fluidez y la capacidad de mantener las convenciones en la lengua meta, lo que se puede observar en las traducciones. Dado que estas no presentan elementos ajenos a la lengua inglesa, se utiliza una estructura gramatical estándar para este idioma, además de emplear un léxico accesible para los hablantes nativos.

Sumado a lo anterior, es posible aseverar que también se cumplen las definiciones de Newmark (1988) y Nord (2005) en estas traducciones, puesto que estos textos no presentan interferencias que denoten la procedencia original de las traducciones, siendo un elemento clave para dichos autores al momento de

determinar la naturalidad. Bajo estas consideraciones, se puede declarar que estas expresiones metafóricas cumplen con los estándares esperables en este concepto.

Por su parte, el aspecto de la fidelidad emocional también encontró su mejor representación en las traducciones al inglés de Llama 4 Maverick, siendo estas las expresiones metafóricas 5 y 12.

**Tabla 5**

<b>Expresión metafórica en español</b>	<b>Traducción al inglés por Llama 4 Maverick</b>
El cosmos está a punto de reventar	The cosmos is about to burst
Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire	I feel the growing of wings, my soul rushes ahead into the quiet expanses of the air

La calificación positiva por parte de los encuestados ante estas traducciones indica que, bajo su perspectiva, la emoción indicada dentro de la encuesta se replica adecuadamente en las traducciones. Siendo este el caso, se entiende que el LLM Llama 4 Maverick logra solventar las problemáticas propuestas por Balahur & Turchi (2014) ante la capacidad de reproducir la carga afectiva dentro de un texto. Este mérito no resulta menor, dado que esto sustenta el trabajo de Vaswani et al. (2017) ante la mejora continua presente en estos modelos respecto del entendimiento de sutilezas emocionales en textos.

Además de lo ya dicho, el resultado positivo unánime indica que expresiones metafóricas como esta logran cumplir la definición misma de la fidelidad emocional de Feng et al. (2025), puesto que la capacidad de estos modelos de traspasar carga

emocional en esta instancia y en varias más observadas en la investigación queda puesta en evidencia.

No obstante, no resulta simple establecer si la naturalidad y la fidelidad emocional fue alcanzada por los LLM. Como se ha mencionado, los mejores resultados se obtuvieron de Llama 4 Maverick y en el idioma inglés. Esto podría deberse a que los modelos poseen una mayor cantidad de datos de este idioma, algo notado también en el trabajo de Shanahan et al. (2023). Dicha situación se entendería como una mayor cantidad de *tokens* obtenidos de datos o textos producidos en inglés. Lo anterior produce escenarios como un LLM consiguiendo mejores resultados al traducir el mismo texto en un idioma que en otro.

**Tabla 6**

<b>Expresión metafórica español</b>	<b>Traducción al inglés GPT-4o</b>	<b>Promedio final naturalidad</b>
Tu inspiración anda tan afilada como mi navaja	Your muse is as sharp as my blade	4,6

**Tabla 7**

<b>Expresión metafórica español</b>	<b>Traducción al alemán GPT-4o</b>	<b>Promedio final naturalidad</b>
Tu inspiración anda tan afilada como mi navaja	Deine Inspiration ist so scharf wie mein Klappmesser	3,2

Esta disparidad de desempeños conlleva que la percepción, tanto de naturalidad como de fidelidad emocional, se vean afectadas de forma significativa.

La mencionada situación ya ha sido recogida por Xu (2024), lo cual indicaría que es una problemática con una existencia documentada y que resulta vigente a la fecha de esta investigación. Si bien anteriormente ya se ha mencionado en este trabajo las diferencias de los LLM utilizados para la investigación, este hecho acentúa la responsabilidad de los usuarios al momento de elegir los modelos más adecuados para sus necesidades, puesto que se ha observado en este como otros trabajos (Xu, 2024; Zhang et al. 2023) que dependiendo del LLM y el idioma se pueden obtener resultados aprobatorios, conclusiones «polarizadas» o traducciones que no obtengan una respuesta favorable por parte de los hablantes nativos.

Este último caso se pudo observar en numerosos casos; sin embargo, fue predominante en las traducciones realizadas por el LLM DeepSeek V-3.

Aunque sea difícil otorgar una razón conclusiva para estos resultados, dado que el presente trabajo no se enfoca en determinar los procesos psicológicos, sociales y lingüísticos que conllevaron a los encuestados para dar calificaciones reprobatorias, es posible realizar ciertas hipótesis, siempre considerando las limitaciones ya mencionadas.

Primeramente, es vital destacar que los peores resultados se observaron en las evaluaciones respecto a la naturalidad y en las traducciones al alemán. Esto podría ser un sustento para el planteamiento ya postulado de que la menor base de datos de un idioma (en este caso, alemán) afecta a la calidad de las traducciones a ese idioma. Adicionalmente, el que en materia de naturalidad se obtuvieran los peores resultados sería un indicio de un problema transversal para los LLM. Es decir,

los LLM tendrían dificultad para elaborar texto en consideración de las características propias de la naturalidad, como las propuestas por Newmark (1988) o House (2015).

A su vez, dicha dificultad podría asociarse a las conclusiones de Charteris-Black (2004). Lo anterior sería posible debido a que este autor logró establecer el vínculo metáfora e imaginario colectivo o conciencia social; a saber, que las metáforas se entienden en tanto sea posible asociarlas a un concepto que pueda ser entendido incluso si no ha sido vivido personalmente. Esta propuesta podría evidenciar que actualmente los LLM no poseen una capacidad plena para emplear experiencias o ideas del imaginario colectivo para establecer conexiones, ergo, metáforas.

**Tabla 8**

<b>Expresión metafórica español</b>	<b>Traducción al inglés DeepSeek V-3</b>	<b>Promedio final fidelidad emocional</b>
¡Vaya a su casa, húndase en la melancolía o la locura!	Go home! Sink into melancholy or madness, for all I care!	2

**Tabla 9**

<b>Expresión metafórica español</b>	<b>Traducción al alemán DeepSeek V-3</b>	<b>Promedio final fidelidad emocional</b>
¡Vaya a su casa, húndase en la melancolía o la locura!	Gehen Sie nach Hause, versinken Sie in Schwermut oder im Wahnsinn	3,6

Adicionalmente, frente a los problemas detectados referentes a la aplicación de los conceptos de naturalidad y fidelidad emocional, resulta importante destacar que, si bien, el *prompt* utilizado solicita directamente la consideración de estos conceptos, es posible que los LLM pudieran haberlo malinterpretado o ignorado, situaciones recogidas por Zhang et al. (2023).

Otro fenómeno que se observó en los resultados fue la dificultad que presentaron algunas expresiones metafóricas para todos los LLM. Un ejemplo de lo mencionado es el caso de las expresiones metafóricas 1, 2 y 3 en la traducción al inglés: tanto en la evaluación de naturalidad como en la de fidelidad emocional se presentaron resultados significativamente más bajos a los de otras expresiones metafóricas. Lo dicho se observa al sumar los promedios finales de todos los LLM, estas expresiones metafóricas presentaron los valores más bajos (ver anexo 18). Analizando los datos obtenidos de sumar los promedios finales, es posible notar que la peor expresión metafórica en la traducción al inglés (expresión metafórica 2: calificación 8,0) se encuentra a una distancia de 4,2 puntos de la expresión mejor evaluada (expresión metafórica 12: calificación 12,2).

El ejercicio de sumar los promedios finales de cada expresión metafórica se realiza con el fin de acentuar el desempeño generalizado de dichas expresiones metafóricas entre todos los LLM.

Dado que las expresiones metafóricas peor calificadas están asociadas, según la clasificación de Plutchik, a las emociones de furia y odio, podría existir una relación entre estas emociones y la dificultad de los LLM para traducirlas.

**Tabla 10**

<b>Emoción según Plutchik</b>	<b>Expresión metafórica original</b>	<b>Suma de promedios en naturalidad inglés</b>	<b>Sumada de promedios en fidelidad emocional inglés</b>
Odio	Todas las aves de rapiña revolotean creyendo que pronto seré carroña	8,0	9,8
Furia	Me trago el veneno, no como otros	9,6	9,8

*Expresiones metafóricas peor resultado de promedios sumados. Extracto anexo 18*

No obstante, la relación mencionada previamente podría ser meramente coincidencia, dado que también se observó que en algunos casos los LLM tuvieron una mejor capacidad de implementar la naturalidad y la fidelidad emocional en expresiones metafóricas de forma generalizada sin que estas estuvieran asociadas a emociones relacionables. Tal caso se puede observar en las expresiones metafóricas 8 y 9 en la traducción al alemán. Estas expresiones metafóricas obtuvieron calificaciones positivas con todos los LLM (ver anexo 18); sin embargo, la expresión metafórica 8 está asociada a la emoción de tristeza y la expresión metafórica 9 a la emoción de admiración.

**Tabla 11**

<b>Emoción según Plutchik</b>	<b>Expresión metafórica original</b>	<b>Suma de promedios en naturalidad alemán</b>	<b>Suma de promedios en fidelidad emocional alemán</b>
Tristeza	Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	12,0	12,6
Admiración	La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola	12,8	12,8

*Expresiones metafóricas mejor resultado de promedios sumados. Extracto anexo 18*

Los ejemplos mencionados previamente, tanto al proveer indicios de una posible relación causal entre emoción y dificultad de traducción, como de indicativos que refutan dicha interpretación, llevan a la necesidad en un trabajo que profundice esta temática para obtener un resultado conclusivo.

Como finalización de la discusión y respondiendo a la pregunta de esta investigación, los LLM poseen la capacidad de aplicar los conceptos de naturalidad y fidelidad emocional en sus traducciones de expresiones metafóricas; no obstante, se observó que esta capacidad no es aplicada efectivamente en todas sus traducciones. La percepción variable de los hablantes nativos ante estos textos y los problemas analizados debido a sus respuestas permiten determinar que los LLM, a fecha de esta investigación, pueden omitir o fallar en la aplicación de estos conceptos, tanto por el LLM utilizado, como el idioma al que es traducido el texto, entre otros factores.

## 7. CONCLUSIÓN Y PROYECCIONES

Tras el análisis en profundidad de la traducción de expresiones metafóricas, el uso de los modelos de lenguaje extenso (LLM) en este ámbito, según lo documentado por diversos autores, y de la percepción de los hablantes nativos en términos de naturalidad y fidelidad emocional, se pueden establecer las siguientes conclusiones.

El modelo LLM DeepSeek V-3 obtuvo los puntajes más bajos en los aspectos de naturalidad y fidelidad emocional en ambos idiomas en comparación con los otros modelos, siendo clasificados como reprobatorios al aplicar una escala de exigencia del 50 %. Por estas razones, según los parámetros del presente trabajo, este modelo no se considera fiable para la traducción de textos con alta carga emocional.

El análisis del modelo GPT-4o evidenció un mejor desempeño en las traducciones al alemán en ambos aspectos evaluados. En el caso del inglés, si bien las calificaciones fueron aprobatorias, no alcanzaron los mejores resultados dentro del estudio. Cabe destacar que el modelo obtuvo el mismo puntaje en naturalidad para el alemán y en fidelidad emocional para el inglés. Estos resultados sugieren que el modelo ofrece traducciones apropiadas a nivel de naturalidad y fidelidad emocional en ambos idiomas, aunque demuestra una mayor competencia en el alemán.

El tercer LLM, Llama 4 Maverick, obtuvo puntajes aprobatorios en todas las categorías, destacando por su mejor desempeño en el idioma inglés, donde alcanzó las calificaciones más altas del estudio. Estos resultados sugieren que este modelo sería el más adecuado para la traducción de textos con abundante presencia de metáforas o con una significativa carga emocional.

Una vez presentados estos resultados, es posible determinar que si un LLM posee la capacidad de traducir un texto de manera natural y emocionalmente fiel depende en gran medida del modelo utilizado. Tal como se ha señalado, modelos como DeepSeek V-3 demostraron una menor capacidad de naturalidad en comparación con modelos como Llama 4 Maverick. En consecuencia, no es posible ofrecer una conclusión general que abarque a todos los LLM, ya que la fidelidad emocional y la naturalidad son características variables entre los distintos modelos.

Si bien los motivos detrás de esta diversidad de desempeños entre LLM se encuentran fuera del alcance de este estudio, resulta evidente que dicha variación existe. Esto indica que la elección del modelo que se utilice para traducir textos marcadamente emocionales, como lo son las metáforas, es un aspecto que no se puede obviar si se desea una traducción que cumpla los estándares de naturalidad y fidelidad emocional.

Es de vital relevancia destacar que las conclusiones que se pudieron extraer de esta investigación no pueden considerarse representativas fuera del alcance de este trabajo; por lo que lo anterior debe entenderse como el análisis y conclusión de una observación limitada, debido a razones que serán explicadas más adelante.

Pese a que podamos extraer ciertas conclusiones, aquellas relacionadas con el desempeño de los LLM no pueden ser consideradas concluyentes. Este elemento de variabilidad se debe principalmente a la muestra de encuestados abarcados en esta investigación. En este estudio se encuestó a diez personas (cinco nativas del idioma inglés y cinco del idioma alemán), las cuales permitieron realizar un análisis inicial de la percepción de hablantes nativos respecto a las traducciones realizadas con LLM. No obstante, resulta evidente que esta muestra de encuestados no es un número considerable que permita representar a la generalidad de hablantes. Además, los encuestados fueron estudiantes extranjeros que se encuentran en su proceso de intercambio en la Universidad de Concepción, lo cual podría incluir variables en su percepción de las traducciones, tales como la exposición de los hablantes al español o su nivel de estudio.

Otro factor que limita el alcance de este trabajo es la posible obsolescencia que pueda producirse, debido a las actualizaciones que sufren constantemente estos modelos de lenguaje. Lo dicho delata que, si bien estas conclusiones pueden realizarse con base a los resultados obtenidos en este estudio, estos podrían verse desfasados si el desempeño de los LLM se modifica con el tiempo.

A raíz de lo mencionado, consideramos que una investigación con un mayor número de encuestados constituiría una fructífera línea de estudio para este tema en el futuro. Además, un continuo análisis en trabajos académicos en torno a estos conceptos resulta relevante dada la naturaleza cambiante que tienen los LLM, debido a su constante actualización. Adicionalmente, aunque no es una respuesta a una

limitación de este trabajo, una expansión de esta metodología aplicada a otras tipologías de textos que cuenten con una significativa carga emocional, tales como la poesía, resultaría beneficiosa para la investigación traductológica.

## 8. REFERENCIAS BIBLIOGRÁFICAS

- Amitín, D. & Kartun, M. (2000). Volpone. CELCIT.
- Aristóteles (1974). *Póetica de Aristóteles*. Gredos. Obra original publicada ca. 350 B. C. E.
- Arroyo-Barrigüete, J. (2023). Sentiment analysis of Lovecraft's fiction writings. *Heliyon*, 9(1) [10.1016/j.heliyon.2022.e12673](https://doi.org/10.1016/j.heliyon.2022.e12673)
- Balahur, A. & Turchi, M. (2014). Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. *Computer Speech & Language*, 28 (1), 56-75. <https://doi.org/10.1016/j.csl.2013.03.004>
- Bandhakavi, A., Wiratunga, N., Massie, S., & Deepak, P. (2016). *Emotion-Corpus Guided Lexicons for Sentiment Analysis on Twitter*. En Bramer, M., Petridis, M. (eds) Research and Development in Intelligent Systems XXXIII. SGAI 2016. Springer. [https://doi.org/10.1007/978-3-319-47175-4\\_5](https://doi.org/10.1007/978-3-319-47175-4_5)
- Bassnett, S. (2002). *Translation Studies* (3ra. ed.). Routledge
- Brkić, G. (2024). Vocabulary Building in EFL Using Plutchik's Wheel of Emotions. *European Journal Of English Language Teaching*, 9(2), 48-62. <https://doi.org/10.46827/ejel.v9i2.5387>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S.,

- Radford, A., Sutskever, I., & Amodei, D. (2020). Language Models are Few-Shot Learners. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2005.14165>
- Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., De Sa, A. R., Junior, Pozzebon, E., & Sobieranski, A. C. (2021). A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Information Sciences*, 582, 593-617. <https://doi.org/10.1016/j.ins.2021.10.005>
- Chang, Y., Wang, X., Wang, J., Wu, Y., Zhu, K., Chen, H., Yang, L., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2023). A Survey on Evaluation of Large Language Models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2307.03109>
- Charteris-Black, J. (2004). *Corpus Approaches to Critical Metaphor Analysis*. Palgrave Macmillan.
- Dhanapala, S., & Samarasinghe, K. (2024). Music Emotion Classification: A Literature review. *Journal Of Research In Music*, 2(2), 14-28. <https://doi.org/10.4038/jrm.v2i2.27>
- De María, C. (2000). *La Caja Negra*. CELCIT.
- Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018a). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1810.04805>
- Ekin, S. (2023). Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices. *TechRxiv*. [10.36227/techrxiv.22683919.v1](https://doi.org/10.36227/techrxiv.22683919.v1)

- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6(3). 169-200.
- Feng, Q., Xie, Q., Wang, X., Li, Q., Zhang, Y., Feng, R., Zhang, T., & Gao, S. (2025). EmoCharacter: Evaluating the Emotional Fidelity of Role-Playing Agents in Dialogues. *Proceedings Of The 2025 Conference Of The Nations Of The Americas Chapter Of The Association For Computational Linguistics: Human Language Technologies, 1: Long Papers*, 6218-6240. <https://doi.org/10.18653/v1/2025.naacl-long.316>
- Ganganwar, V., Babu, N., Kudak, P., Singh, R. & Tanwar, S. (2021). Sentiment analysis of legal emails using Plutchik's Wheel of Emotions in quantified format. *Turkish Journal of Computer and Mathematics Education*, 12 (6), 4979-4987.
- Ghazi, D., Inkpen, D. & Szpakowicz, S. (2010). Hierarchical Approach to Emotion Recognition and Classification in Texts. En *Advances in Artificial Intelligence: Canadian AI 2010 - Lecture Notes in Computer Science*, 40-50. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-13059-5\\_7](https://doi.org/10.1007/978-3-642-13059-5_7)
- Gibbs, R. W., Jr. (2017). *Metaphor wars: Conceptual metaphors in human life*. Cambridge University Press.
- Gile, D. (2009). *Basic Concepts and Models for Interpreter and Translator Training*. John Benjamins.
- Goatly, A. (2007). *Washing the Brain – Metaphor and Hidden Ideology*. John Benjamins.

- Gonzales, M. A. (2001). *Reseña de Barcelona, A.(ed.)(2000), Metaphor and metonymy at the crossroads A cognitive perspective, New York, Mouton de Gruyter*. Universidad de Cádiz, Cádiz.
- Grady, J. E. (1997). *Foundations of Meaning: Primary Metaphors and Primary Scenes*. Tesis doctoral de filosofía en lingüística. Universidad de California, Berkeley.
- Hanić, J., Pavlović, T., & Jahić, A. (2016). Translating emotion-related metaphors: A cognitive approach. *ExELL*, 4(2), 87-101. <https://doi.org/10.1515/exell-2017-0008>
- Heilinger, J. C. (2022) *The Ethics of AI Ethics. A Constructive Critique*. *Philosophy & Technology*. 35, 61. <https://doi.org/10.1007/s13347-022-00557-9>
- House, J. (2015). *Translation Quality Assessment: Past and Present*. Routledge.
- Kartun, M. (1999). *La Casa de los Viejos*. CELCIT.
- Kövecses, Z. (2000). *Metaphor and Emotion*. Cambridge University Press.
- Kövecses, Z. (2002). *Metaphor: A Practical Introduction*. Oxford University Press.
- Kövecses, Z. (2021). A multilevel and contextualist view of conceptual metaphor theory. *Journal of Language and Communication*, 8 (2), 133-134.
- Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. The University of Chicago Press.
- Laks, A. (2020). ¿Qué es cognitivo en la metáfora según Aristóteles? *Revista Archai*, 30. [https://doi.org/10.14195/1984-249x\\_30\\_32](https://doi.org/10.14195/1984-249x_30_32)
- Lastro, S. (1995). *Parejas*. CELCIT.

- Lek, J. X., & Teo, J. (2023). Academic Emotion Classification Using FER: A Systematic Review. *Human Behavior And Emerging Technologies*, 2023, 1-27.  
<https://doi.org/10.1155/2023/9790005>
- Mayoral, R. (2010). *Aspectos Epistemológicos de la Traducción*. Publicacions de la Universitat Jaume I.
- Mohamed, Y. A., Khanan, A., Bashir, M., Hakim, A., Adiel, M. & Elsadig, M. (2024) The Impact of Artificial Intelligence on Language Translation: A Review. *IEEE Access*, 12. <https://ieeexplore.ieee.org/document/10438431>
- Mondal, A. & Gokhale, S. S. (2020). Mining emotions on Plutchik's Wheel. En 2020 *Seventh International Conference on Social Networks Analysis, Management and Security*, 1–6. <https://doi.org/10.1109/SNAMS52053.2020.9336534>.
- Moneus, A. M., & Sahari, Y. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10(6).  
<https://doi.org/10.1016/j.heliyon.2024.e28106>
- Moslem, Y., Haque, R., Kelleher, J. D. & Way, A. (2023). Adaptive Machine Translation with Large Language Models. *arXiv (Cornell University)*.  
<https://doi.org/10.48550/arxiv.2301.13294>
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Barnes, N., & Mian, A. (2023). A Comprehensive Overview of Large Language Models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2307.06435>
- Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall.

- Nida, E. A. (1964). *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. Brill.
- Nida, E. A., & Taber, C. (1982). *Helps for Translators prepared under the auspices of the United Bible Societies – Volume VIII: The Theory and Practice of Translation*. Brill
- Nord, C. (2005). *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis* (2da. ed.). Rodopi.
- Pavis, P. (1998) *Diccionario del Teatro*. Paidós Ibérica.
- Perinelli, R. (2000). La Cena. CELCIT.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. En R. Plutchik, & H. Kellerman (Eds.), *Emotion: Theory, Research, and Experience* (pp. 3-33). Academic Press.
- Semino, E. (2010). Descriptions of Pain, Metaphor, and Embodied Simulation. *Metaphor And Symbol*, 25(4), 205-226. <https://doi.org/10.1080/10926488.2010.510926>
- Shahmerdanova, R. (2025). Artificial Intelligence in Translation: Challenges and Opportunities. *Acta Globalis Humanitatis Et Linguarum*, 2(1), 62-70. <https://doi.org/10.69760/aghel.02500108>
- Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623, 493-498. <https://doi.org/10.1038/s41586-023-06647-8>
- Shaw, A., Vardhan, R. K. & Saxena, S. (2016). Emotion Recognition and Classification in Speech using Artificial Neural Networks. *International Journal of Computer Applications*, 145 (8), 5-9.

- Tantanian, A. (1998). Sumario de la Muerte de Kleist. CELCIT.
- Tseng, M. (2010). The performative potential of metaphor. *Semiotica*, 2010(180), 115-145. <https://doi.org/10.1515/semi.2010.033>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1706.03762>
- Wang, L. (2023). The Impacts and Challenges of Artificial Intelligence Translation Tool on Translation Professionals. *SHS Web Of Conferences*, 163. <https://doi.org/10.1051/shsconf/202316302021>
- Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D., Shi, S., & Tu, Z. (2023). Document-Level Machine Translation with Large Language Models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2304.02210>
- Xu, H., Kim, Y. J., Sharaf, A., & Awadalla, H. H. (2023). A Paradigm Shift in Machine Translation: Boosting Translation Performance of Large Language Models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2309.11674>
- Zhang, B., Haddow, B., & Birch, A. (2023). Prompting Large Language Model for Machine Translation: A Case Study. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2301.07069>
- Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., Wang, S., Yin, D., & Du, M. (2023). Explainability for Large Language Models: A Survey. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2309.01029>

## 9. ANEXOS

### Anexo 1: Tabla de textos origen de las expresiones metafóricas seleccionadas

<b>Obra original</b>	<b>Expresiones metafóricas seleccionadas</b>
Volpone	Todas las aves de rapiña revolotean creyendo que pronto seré carroña
Volpone	¡Vaya a su casa, húndase en la melancolía o la locura!
La Casa de los Viejitos	Me trago el veneno, no como otros
La Casa de los Viejitos	No soporto que me moje el felpudo con las lágrimas
La Cena	El cosmos está a punto de reventar
La Cena	Horizonte negro para usted, eso es lo que veo por delante
Parejas	Y ahora todo está gastado, leproso, muerto. Mi corazón tiene gangrena y se cae a pedazos
Parejas	Desde entonces mamá fue como un cisne dormido en la orilla de mi vida
Parejas	La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola
La Caja Negra	Tu inspiración anda tan afilada como mi navaja
Parejas	Sentí que me brotaban alas y que me iba volando lejos, rozando la cima de los rascacielos
Sumario de la Muerte de Kleist	Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire

**Anexo 2: Tabla de expresiones metafóricas clasificadas según Rueda de Emociones de Plutchik (1980)**

<b>Expresión metafórica</b>	<b>Emoción según Rueda de Emociones</b>
Todas las aves de rapiña revolotean creyendo que pronto seré carroña	Odio
¡Vaya a su casa, húndase en la melancolía o la locura!	Odio
Me trago el veneno, no como otros	Furia
No soporto que me moje el felpudo con las lágrimas	Furia
El cosmos está a punto de reventar	Vigilancia
Horizonte negro para usted, eso es lo que veo por delante	Vigilancia
Y ahora todo está gastado, leproso, muerto. Mi corazón tiene gangrena y se cae a pedazos	Tristeza
Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	Tristeza
La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola	Admiración
Tu inspiración anda tan afilada como mi navaja	Admiración
Sentí que me brotaban alas y que me iba volando lejos, rozando la cima de los rascacielos	Éxtasis
Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire	Éxtasis

**Anexo 3: Tabla de doce expresiones metafóricas traducidas por el LLM DeepSeek V-3 al inglés y alemán**

	<b>Expresión metafórica</b>	<b>DeepSeek V-3 - inglés</b>	<b>DeepSeek V-3 - alemán</b>
1	Todas las aves de rapiña revolotean creyendo que pronto seré carroña	All the birds of prey are circling, thinking I'll soon be carrion	Alle Raubvögel kreisen schon, in der Hoffnung, ich werde bald Aas sein
2	¡Vaya a su casa, húndase en la melancolía o la locura!	Go home! Sink into melancholy or madness, for all I care!	Gehen Sie nach Hause, versinken Sie in Schwermut oder im Wahnsinn
3	Me trago el veneno, no como otros	I swallow the poison—unlike others!	Ich schluck das Gift runter – im Gegensatz zu anderen!
4	No soporto que me moje el felpudo con las lágrimas	I can't stand him soaking the doormat with his tears anymore...!	Ich ertrag's nicht mehr, wie er mir mit seinen Tränen den Fußabtreter durchnässt...!
5	El cosmos está a punto de reventar	The cosmos is about to explode	Das Universum steht kurz vor dem Explodieren
6	Horizonte negro para usted, eso es lo que veo por delante	A dark horizon awaits you, that's what I see ahead	Ein schwarzer Horizont liegt vor dir, das ist es, was ich sehe
7	Y ahora todo está gastado, leproso, muerto. Mi corazón tiene gangrena y se cae a pedazos	And now everything is worn out, leprous, dead. My heart has gangrene and it's falling to pieces	Und jetzt ist alles verbraucht, aussätzig, tot. Mein Herz ist voller Brandwunden, es fällt in Stücken ab
8	Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	And from that moment on, Mama was like a sleeping swan on the shore of my life	Und seitdem war Mama wie ein schlafender Schwan am Ufer meines Lebens

9	La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola	The City of Light, she'd say, with an accent so charming it melted my heart just to hear her	Die Stadt des Lichts, sagte sie – mit einem so bezaubernden Akzent, dass mir das Herz schmolz, wenn ich ihr zuhörte
10	Tu inspiración anda tan afilada como mi navaja	Your inspiration is as sharp as my razor	Deine Inspiration ist so scharf wie mein Rasiermesser
11	Sentí que me brotaban alas y que me iba volando lejos, rozando la cima de los rascacielos	I felt wings sprouting from my back, lifting me up, flying far away, barely grazing the tops of the skyscrapers	Ich spürte, wie mir Flügel wuchsen und ich davonflog, hoch über den Wolkenkratzern
12	Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire	I feel wings growing, my soul soars ahead through the still spaces of the air	Ich spüre, wie Flügel wachsen, meine Seele eilt voraus in die stillen Räume der Luft

**Anexo 4: Tabla de doce expresiones metafóricas traducidas por el LLM GPT-4o al inglés y alemán**

	<b>Expresiones metafóricas</b>	<b>GPT-4o - inglés</b>	<b>GPT-4o - alemán</b>
1	Todas las aves de rapiña revolotean creyendo que pronto seré carroña	All the vultures are circling, thinking I'm about to kick the bucket	Alle Aasgeier kreisen schon, weil sie glauben, ich werde bald zur Beute
2	¡Vaya a su casa, húndase en la melancolía o la locura!	Get out of my sight, go home and wallow in your misery!	Scheren Sie sich nach Hause, vergraben Sie sich in Ihrer Melancholie oder Ihrem Wahnsinn!
3	Me trago el veneno, no como otros	I swallow my pride and my problems, not like some people who can't keep their mouths shut!	Ich schlucke das Gift, nicht wie andere...!
4	No soporto que me moje el felpudo con las lágrimas	I'm sick of him coming back here in tears all the time!	Ich kann nicht mehr, dass er mir die Fußmatte mit seinen Tränen vollheult
5	El cosmos está a punto de reventar	Things are about to blow up	Der Kosmos steht kurz vorm Explodieren
6	Horizonte negro para usted, eso es lo que veo por delante	Dark days ahead for you, that's what I see	Schwarzes Loch vor Ihnen, das sehe ich in Ihrer Zukunft
7	Y ahora todo está gastado, leproso, muerto. Mi corazón tiene gangrena y se cae a pedazos	And now everything is worn out, rotten, dead. My heart is gangrenous and falling apart	Und jetzt ist alles abgenutzt, verfault, tot. Mein Herz hat Gangrän und fällt auseinander
8	Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	From that day on, my mom was like a gentle memory, always with me.	Und seitdem war Mama wie ein schlafender Schwan am Ufer des Lebens

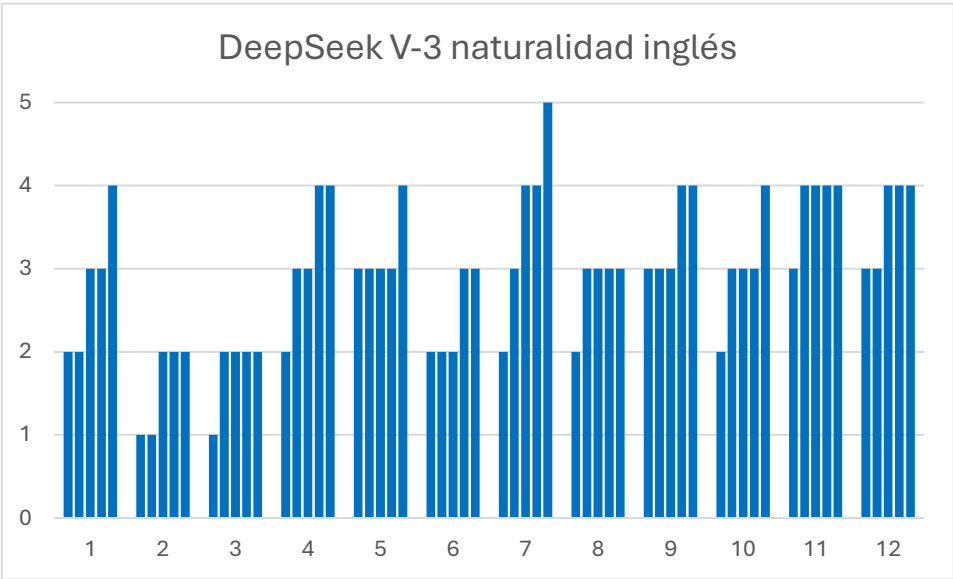
9	La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola	The City of Light, with this adorable accent that would melt my heart every time	Die Stadt des Lichts, sagte sie mit einem so charmanten Akzent, dass mir das Herz schmolz, wenn ich ihr zuhörte
10	Tu inspiración anda tan afilada como mi navaja	Your muse is as sharp as my blade	Deine Inspiration ist so scharf wie mein Klappmesser
11	Sentí que me brotaban alas y que me iba volando lejos, rozando la cima de los rascacielos	I felt wings sprouting from me and I was flying away, skimming the tops of skyscrapers	Mir war, als würden mir Flügel wachsen und ich würde davonschweben, die Dächer der Wolkenkratzer berühren
12	Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire	I feel my wings unfolding, my soul soars through the stillness of the air	Ich spüre das Wachsen von Flügeln, meine Seele eilt voraus in die stillen Räume der Luft

**Anexo 5: Tabla de doce expresiones metafóricas traducidas por el LLM Llama 4 Maverick al inglés y alemán**

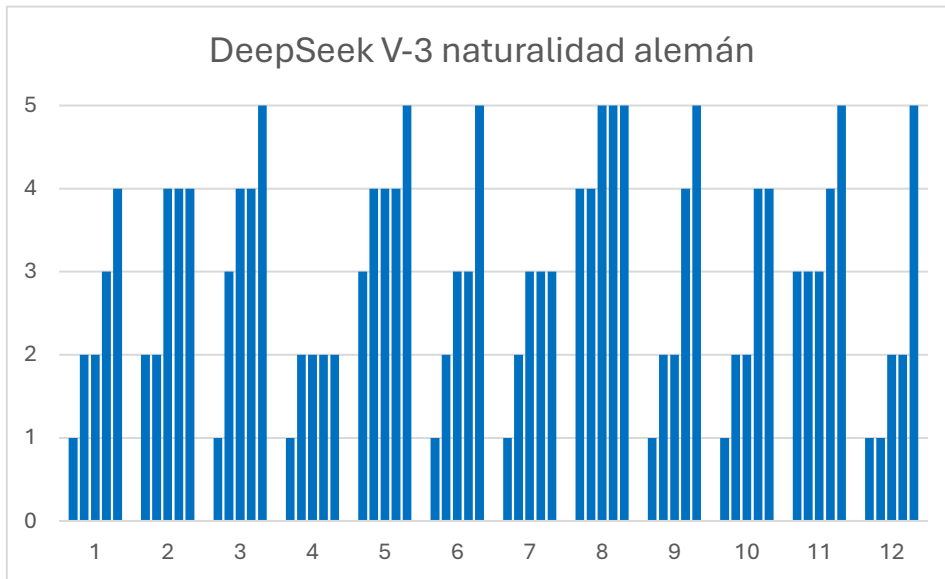
	<b>Expresiones metafóricas</b>	<b>Llama 4 Maverick – inglés</b>	<b>Llama 4 Maverick - alemán</b>
1	Todas las aves de rapiña revolotean creyendo que pronto seré carroña	All the birds of prey are circling, believing I'll soon be carrion	Alle Raubvögel kreisen, in dem Glauben, ich würde bald Aas sein
2	¡Vaya a su casa, húndase en la melancolía o la locura!	Go home, sink into melancholy or madness!	Geh nach Hause, versinke in Melancholie oder Wahnsinn!
3	Me trago el veneno, no como otros	I swallow the poison, unlike others	Ich schlucke das Gift, anders als andere
4	No soporto que me moje el felpudo con las lágrimas	I can't stand you getting the doormat wet with your tears	Ich ertrage es nicht, dass du den Fußabtreter mit deinen Tränen nass machst
5	El cosmos está a punto de reventar	The cosmos is about to burst	Der Kosmos steht kurz davor zu platzen
6	Horizonte negro para usted, eso es lo que veo por delante	Black horizon ahead for you, that's what I see coming	Schwarzer Horizont für Sie, das ist das, was ich vor mir sehe
7	Y ahora todo está gastado, leproso, muerto. Mi corazón tiene gangrena y se cae a pedazos	And now everything is worn out, leprous, dead. My heart has gangrene and is falling to pieces	Und jetzt ist alles abgenutzt, leprös, tot. Mein Herz hat Gangrän und fällt in Stücke
8	Desde entonces mamá fue como un cisne dormido en la orilla de mi vida	From then on, Mom was like a sleeping swan on the shore of my life	Seitdem war Mama wie ein schlafender Schwan am Ufer meines Lebens

9	La Ciudad Luz, decía, con un acento tan encantador que se me derretía el corazón escuchándola	The City of Light, she'd say, with such a charming accent it melted my heart to hear her	Die Stadt der Lichter, sagte sie mit einem so bezaubernden Akzent, dass es mir das Herz zerfließen ließ, sie zu hören
10	Tu inspiración anda tan afilada como mi navaja	Your inspiration is as deadly as my razor	Deine Inspiration ist so scharf wie mein Rasiermesser
11	Sentí que me brotaban alas y que me iba volando lejos, rozando la cima de los rascacielos	I felt wings sprouting from me, carrying me far away, skimming the tops of skyscrapers	Ich spürte, wie mir Flügel wuchsen und ich davonflog, an den Spitzen der Wolkenkratzer entlang
12	Siento un crecer de alas, se adelanta mi alma en los quietos espacios del aire	I feel the growing of wings, my soul rushes ahead into the quiet expanses of the air	Ich spüre ein Flügelwachsen, meine Seele eilt voraus durch die stillen Räume der Luft

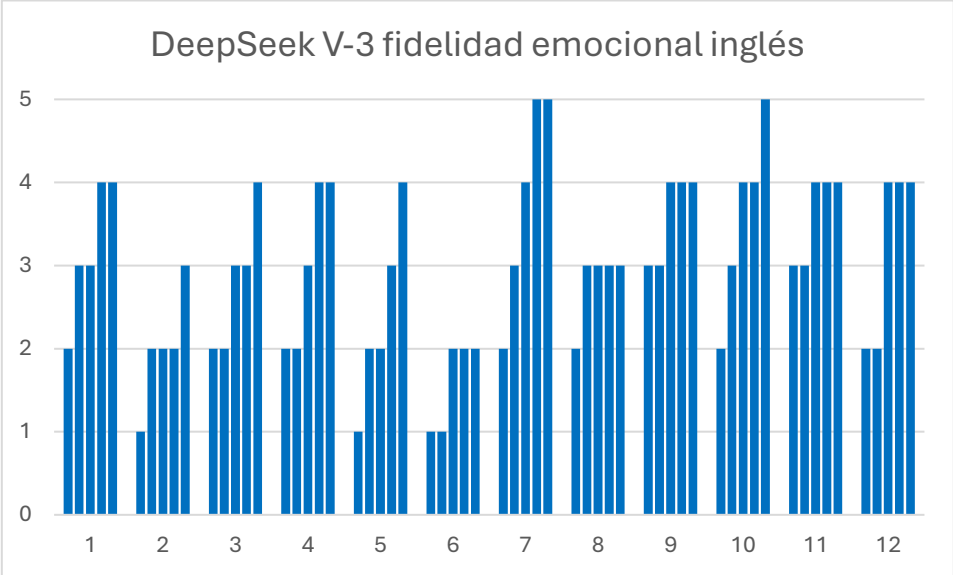
**Anexo 6: Gráfico de resultados encuesta en naturalidad en inglés de LLM DeepSeek V-3**



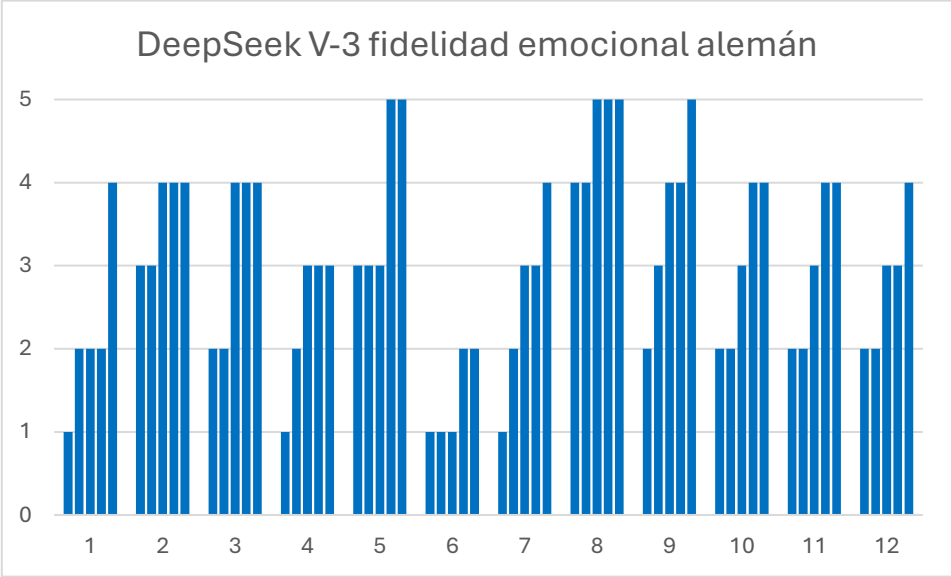
**Anexo 7: Gráfico de resultados encuesta en naturalidad en alemán de LLM DeepSeek V-3**



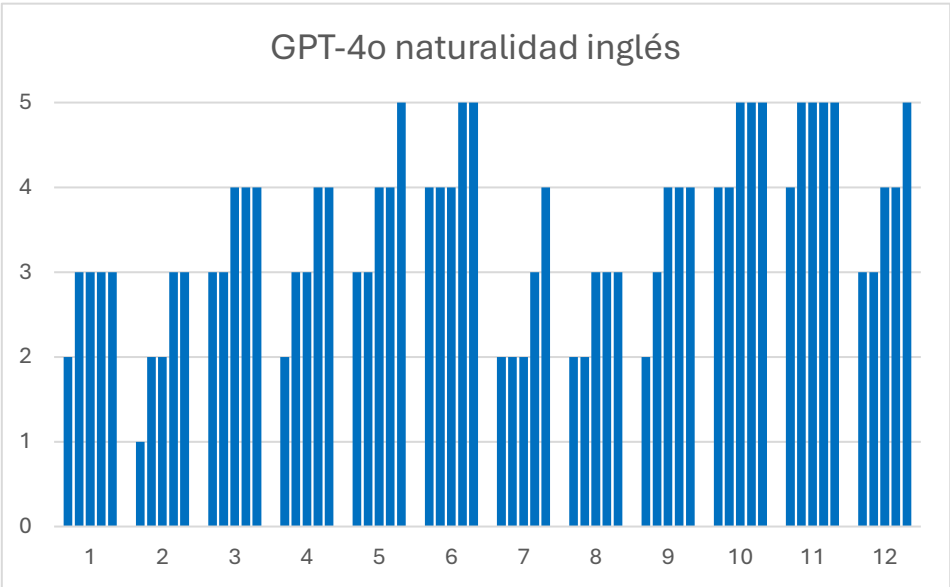
**Anexo 8: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM DeepSeek V-3**



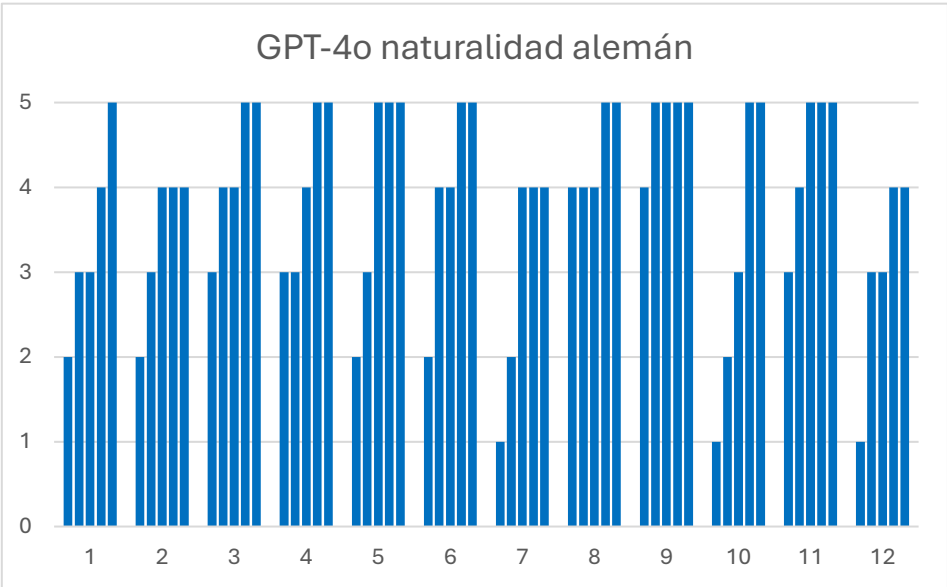
**Anexo 9: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM DeepSeek V-3**



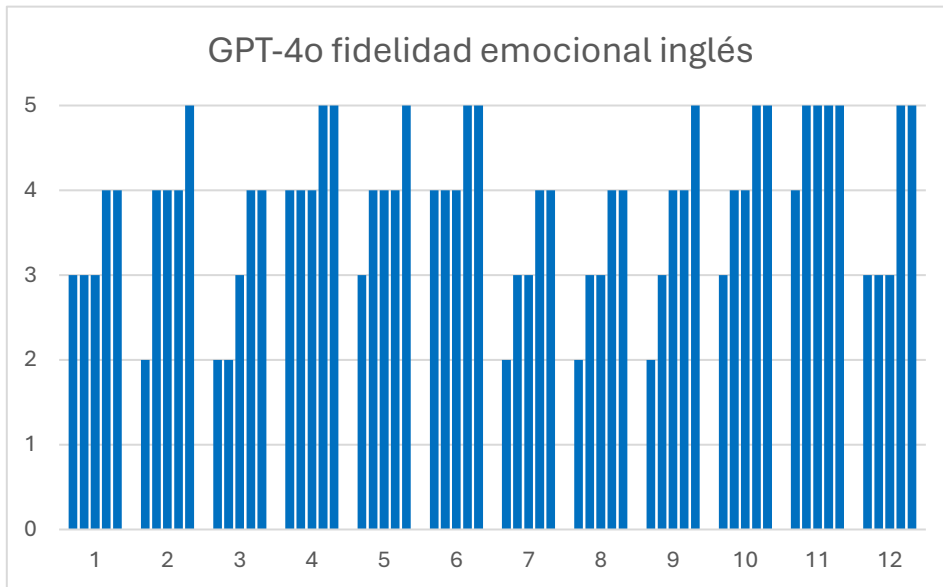
**Anexo 10: Gráfico de resultados encuesta en naturalidad en inglés de LLM GPT-4o**



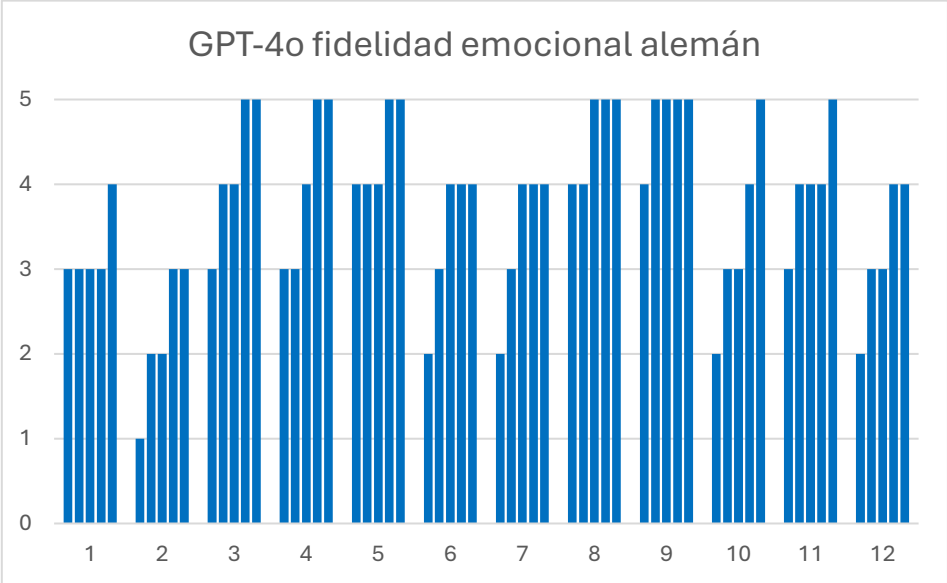
**Anexo 11: Gráfico de resultados encuesta en naturalidad en alemán de LLM GPT-4o**



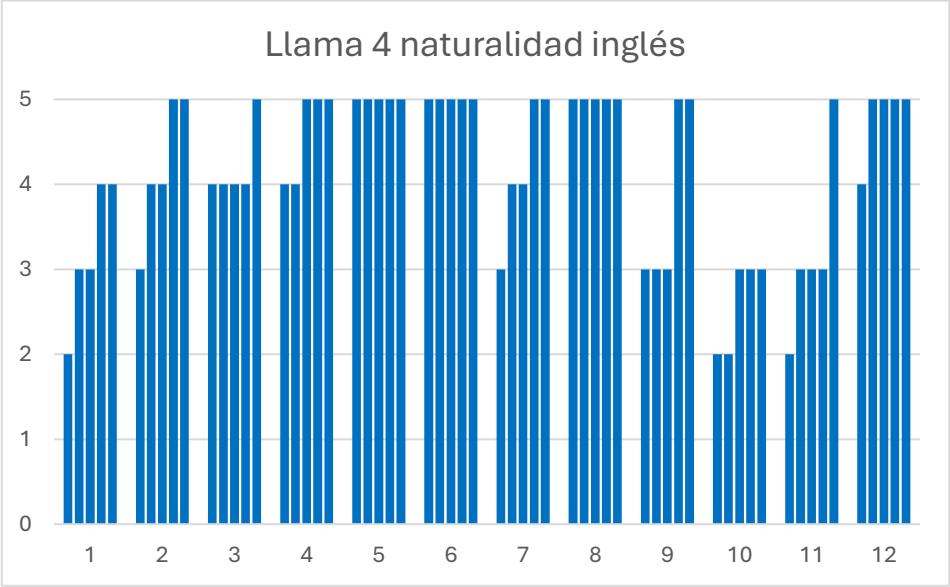
**Anexo 12: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM GPT-4o**



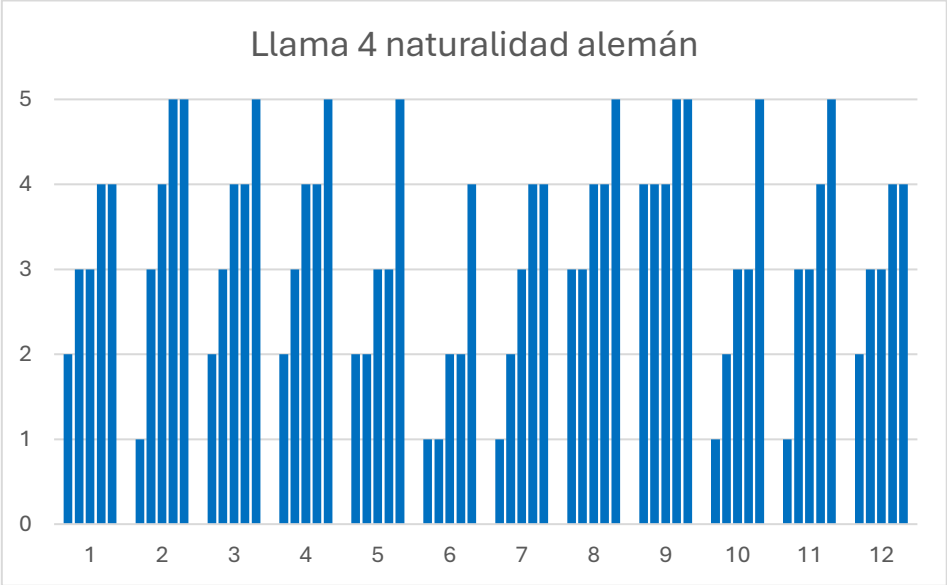
**Anexo 13: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM GPT-4o**



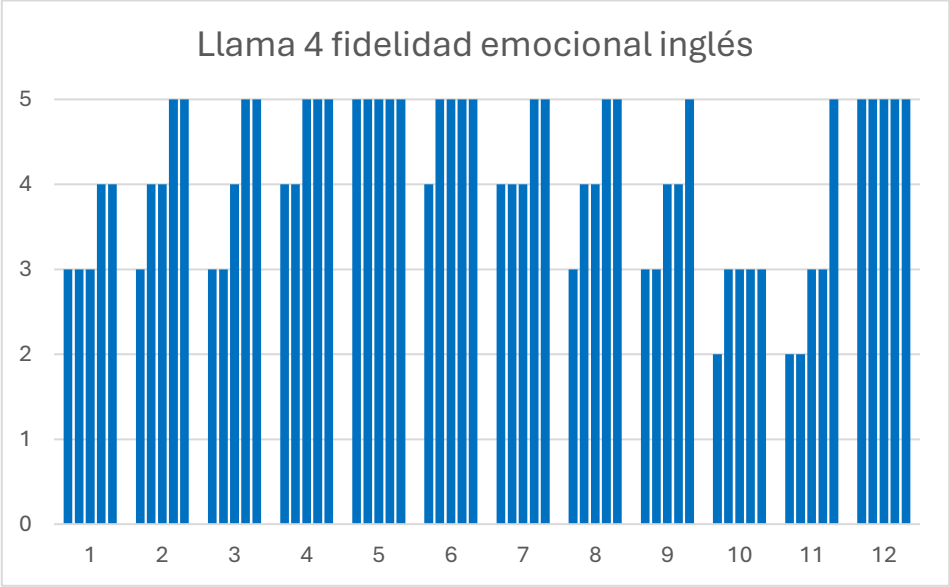
**Anexo 14: Gráfico de resultados encuesta en naturalidad en inglés de LLM Llama 4 Maverick**



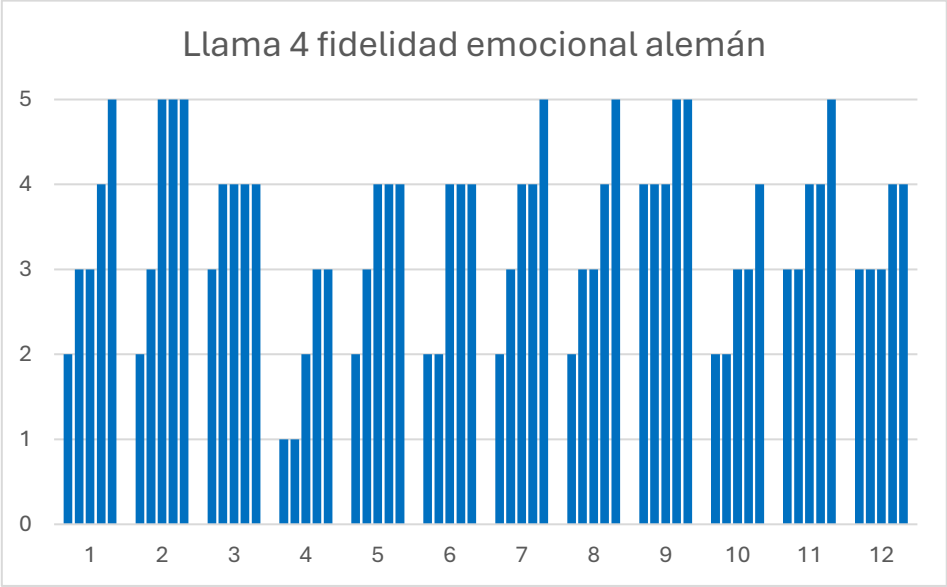
**Anexo 15: Gráfico de resultados encuesta en naturalidad en alemán de LLM Llama 4 Maverick**



**Anexo 16: Gráfico de resultados encuesta en fidelidad emocional en inglés de LLM Llama 4 Maverick**



**Anexo 17: Gráfico de resultados encuesta en fidelidad emocional en alemán de LLM Llama 4 Maverick**



**Anexo 18: Tabla de expresiones metafóricas según suma total de promedios finales desde el resultado más bajo al más alto**

Naturalidad inglés		Fidelidad inglés		Naturalidad alemán		Fidelidad alemán	
Expresión 2	8	Expresión 3	9,8	Expresión 7	8,2	Expresión 6	8
Expresión 1	8,8	Expresión 1	10	Expresión 12	8,4	Expresión 4	8,4
Expresión 3	9,6	Expresión 2	10	Expresión 10	8,6	Expresión 1	8,8
Expresión 10	10,2	Expresión 8	10,2	Expresión 6	8,8	Expresión 10	9,2
Expresión 7	10,4	Expresión 10	10,6	Expresión 1	9	Expresión 12	9,4
Expresión 8	10,4	Expresión 6	10,8	Expresión 4	9,4	Expresión 7	9,6
Expresión 9	10,6	Expresión 9	11	Expresión 2	10,2	Expresión 2	9,8
Expresión 4	11	Expresión 5	11,4	Expresión 5	11	Expresión 11	10,8
Expresión 6	11,8	Expresión 7	11,4	Expresión 3	11,2	Expresión 3	11,2
Expresión 11	11,8	Expresión 11	11,4	Expresión 11	11,2	Expresión 5	11,6
Expresión 5	12	Expresión 4	12	Expresión 9	12	Expresión 8	12,6
Expresión 12	12,2	Expresión 12	12	Expresión 8	12,8	Expresión 9	12,8