



Expresiones idiomáticas en la era digital: un estudio comparativo de cinco herramientas de traducción automática

Idiomatic expressions in the digital age: a comparative study of five machine translation tools

Mebarka BEDARNIA¹

Universidad Amar Téliidji-Laghouat | Argelia
m.bedarnia@lagh-univ.dz

Resumen: El objetivo de este estudio es evaluar la efectividad de la traducción automática de expresiones idiomáticas del español al árabe. Hemos seleccionado 19 expresiones idiomáticas que se centran en emociones y reacciones humanas. Utilizamos un enfoque cuantitativo para evaluar cinco programas de traducción automática y herramientas de inteligencia artificial: Google Translate, Bing Translator, DeepL, Gemini y ChatGPT 3.5, comparándolos con una traducción humana de referencia. El análisis se realizó mediante la plataforma TILDE. Los resultados muestran que ChatGPT 3.5 y Gemini ofrecieron las traducciones más precisas, aunque todos los motores enfrentaron desafíos al interpretar algunas expresiones idiomáticas.

Palabras clave: Inteligencia artificial; diversidad lingüística; traducción automática; matices culturales, expresiones idiomáticas, español-árabe

Abstract: The objective of this study is to evaluate the accuracy and fidelity of automatic translation of idiomatic expressions from Spanish to Arabic. We selected 19 idiomatic expressions focusing on emotions and human reactions. We used a quantitative approach to assess five automatic translation programs and artificial intelligence tools: Google Translate, Bing Translator, DeepL, Gemini, and ChatGPT 3.5, comparing them with a human reference translation. The analysis was conducted using the TILDE platform. The results show that ChatGPT 3.5 and Gemini provided the most accurate translations, although all the engines faced challenges in interpreting some idiomatic expressions.

Keywords: Artificial intelligence; linguistic diversity; machine translation; cultural matters; idiomatic expressions; Spanish-Arabic.



La globalización ha convertido la comunicación intercultural en una pieza clave para entendernos y colaborar con personas de diferentes culturas (Byram, 1997; House, 2015; Katan, 1999). En este sentido, según Byram (1997), la competencia intercultural va más allá de hablar otro idioma; se trata de interactuar de manera efectiva con personas

¹ Auteur correspondant : MEBARKA BEDARNIA | m.bedarnia@lagh-univ.dz

de distintas culturas, entendiendo y respetando sus valores y perspectivas. De hecho, el idioma no solo comunica ideas, sino que también refleja la cultura de una sociedad, incorporando sus creencias y prácticas (Nida, 1999). Por ejemplo, este vínculo es evidente en la semántica, donde el significado está influenciado por la cultura (Samaniego, 2007). Además, las expresiones idiomáticas son una manifestación clara de cómo el lenguaje refleja el estilo de vida y las costumbres de una comunidad (Luque y Manjón, 1998). Por esta razón, varios estudios han explorado cómo estas expresiones y su carga cultural afectan la comunicación intercultural (Boers y Stengers, 2008; Bueno, 2002; Dobrovól'skij, 1994, 2000; Dobrovól'skij y Piirainen, 2005; González Rey, 2002). En conclusión, comprender estas conexiones es esencial para mejorar la comunicación entre culturas en un mundo cada vez más globalizado.

Por otro lado, la historia de la Traducción Automática (TA) comienza a mediados del siglo XX, cuando Warren Weaver propuso en 1953 que los ordenadores podrían utilizarse para resolver el problema de la TA entre lenguas (Hutchins, 1995). La primera TA se realizó en 1954, con un texto en ruso traducido al inglés. Sin embargo, este primer intento mostró tanto las limitaciones del sistema como su potencial para ofrecer traducciones útiles (Hutchins, 1986). No obstante, en los años 60, la TA enfrentó una crisis de confianza debido a la dificultad de lograr resultados satisfactorios, lo que llevó a la creación del Automatic Language Processing Advisory Committee (ALPAC) en EE. UU. En efecto, el informe del ALPAC concluyó que los resultados cualitativos aceptables no se podían alcanzar en plazos breves, lo que ralentizó la investigación en TA (Hutchins, 1995).

A pesar de ello, la TA revivió en los años 70, impulsada por la creciente necesidad de traducción en contextos multilingües. Por ejemplo, la Comunidad Económica Europea en Luxemburgo adoptó el sistema Systran en 1976, y en Canadá se implementó el sistema MÉTEO (Hutchins, 2002). Estos ejemplos demuestran que estos sistemas desempeñaron un papel clave en la aplicación de la TA en contextos oficiales y multinacionales.

Posteriormente, en los años 80, los sistemas de TA se volvieron más sofisticados, enfocándose en la Traducción por Reglas (TBR), que aplicaba reglas lingüísticas detalladas para mejorar la precisión. Mientras que Harris (1957) contribuyó significativamente a la teoría de la información y la lingüística estructural, sentando las bases para los métodos basados en reglas. Más tarde, la década de 1990, con la expansión de Internet, trajo consigo el desarrollo de sistemas basados en ejemplos y en corpus bilingües, utilizando grandes volúmenes de datos para mejorar las traducciones (Hutchins, 2010).

En los últimos años, la TA ha experimentado una transformación significativa con la llegada de redes neuronales y modelos de aprendizaje profundo. De hecho, los sistemas de Traducción Automática Neuronal (TAN), como los desarrollados por Google y DeepL, han superado en muchos casos a los sistemas estadísticos en términos de precisión y fluidez. Sin embargo, la TAN todavía enfrenta desafíos, especialmente en la traducción de términos específicos de dominio y en la gestión de bases de datos terminológicas. A pesar de estas limitaciones, la TAN ha revolucionado el campo de la traducción, mejorando la calidad y reduciendo tanto los costos como el tiempo de posesición (Monti, 2019).

En la actualidad, los sistemas de TA se dividen en tres categorías principales: los basados en reglas, los basados en corpus y los neuronales. En primer lugar, los sistemas basados en reglas,

como los sistemas directos e indirectos, usan diccionarios y reglas lingüísticas para realizar traducciones. En segundo lugar, los sistemas basados en corpus, como los sistemas estadísticos, utilizan grandes volúmenes de datos para calcular la probabilidad de equivalencias entre términos. Finalmente, los sistemas neuronales, como Gemini y ChatGPT, emplean redes neuronales para ofrecer traducciones más precisas y contextuales.

Con este estudio procuramos evaluar la efectividad de varios motores de traducción automática (MTAs) para determinar cuál ofrece la mejor calidad en las traducciones del español al árabe. Se analizará cómo cada motor maneja la autenticidad lingüística y los matices culturales del texto original, comparándolos con la traducción humana.

En la asignatura de Traducción del Máster 1, especialidad en Didáctica de Lenguas Extranjeras en la Universidad de Laghouat, hemos llevado a cabo un experimento sobre la enseñanza de expresiones idiomáticas y su TA. Hemos observado que cada sistema de TA produce resultados diferentes, con variaciones significativas entre pares de lenguas. Por ejemplo, las traducciones entre español y francés no siempre coinciden con las traducciones entre español y árabe. Este estudio examina cómo estos sistemas no solo afectan la precisión lingüística, sino también la captura de matices culturales y expresivos de las expresiones idiomáticas. La capacidad de los sistemas para reflejar estas sutilezas culturales es esencial para obtener traducciones correctas y culturalmente apropiadas. Este trabajo propone abordar la siguiente cuestión: ¿En qué medida la TA es capaz de captar las connotaciones y asociaciones culturales de las expresiones idiomáticas en español, evitando traducciones literales que puedan resultar extrañas o incomprensibles?

A esta pregunta general, planteamos otras específicas:

- ¿En qué medida los MTAs son capaces de reconocer y manejar las diferencias estructurales entre el español y el árabe al traducir determinadas expresiones idiomáticas?
- ¿Cómo afectan las herramientas de TA a la interpretación de expresiones idiomáticas del español al árabe en términos de precisión y adecuación cultural?

La hipótesis de nuestro estudio es la siguiente:

- Los MTAs evaluados no logran identificar y manejar adecuadamente las diferencias estructurales entre el español y el árabe en cuanto a determinadas expresiones idiomáticas.

El objetivo principal de este estudio radica en investigar la capacidad de los traductores automáticos basados en la inteligencia artificial para reconocer y reflejar adecuadamente las complejas diferencias en la traducción de expresiones idiomáticas entre el español y el árabe. Este análisis se centra específicamente en las expresiones idiomáticas debido a su inherente complejidad y riqueza cultural. Este enfoque ofrece una perspectiva profunda y valiosa sobre el estado actual de la traducción automática, proporcionando una evaluación crítica de su eficacia y precisión en contextos lingüísticos desafiantes.

1. Método y material

En este estudio, adoptamos un enfoque mixto cuantitativo-cualitativo para evaluar el impacto de la TA en la autenticidad lingüística y en la traducción de expresiones idiomáticas del español al árabe. Este enfoque mixto permite un análisis exhaustivo de la eficacia de las herramientas de TA en la preservación de la autenticidad lingüística y cultural en traducciones automáticas.

El diseño de este estudio es no experimental, se enfoca en observar y comparar de manera descriptiva cómo diferentes sistemas de TA manejan la traducción de expresiones idiomáticas del español al árabe sin intervenir en el proceso. Este enfoque permite analizar las capacidades naturales de los sistemas y evaluar la precisión y autenticidad de sus traducciones en condiciones reales (Hernández *et al.*, 2014).

El alcance correlacional de este estudio se centra en examinar la relación entre los sistemas de TA y la precisión en la traducción de expresiones idiomáticas del español al árabe. Los objetivos incluyen identificar cómo distintos MTAs manejan las complejidades culturales y lingüísticas de las expresiones idiomáticas, y evaluar si existen asociaciones significativas entre el tipo de sistema utilizado y la calidad de la traducción obtenida.

Hemos seleccionado un total de 19 expresiones idiomáticas traducidas del español al árabe, todas centradas en describir emociones y reacciones humanas. Estas expresiones idiomáticas son de un nivel avanzado según el Marco Común Europeo de Referencia para las Lenguas (MCER), correspondientes a los niveles C1 o C2, lo que indica que requieren un entendimiento profundo de la lengua y su uso connotativo e implícito. A continuación, presentamos el corpus seleccionado para llevar a cabo este estudio:

- Dar saltos de alegría
- Estar negro
- Romperse el corazón
- Estar hasta las narices
- Estar rebosante de satisfacción
- Estar como un flan
- Helarse la sangre
- Estar como unas pascuas
- Subirse por las paredes
- Estar como un niño con zapatos nuevos
- Sacar el quicio
- Llenar de dicha
- Sentirse helado
- Estar con la moral por los suelos
- Ponerse rojo de ira
- Hervir la sangre
- Estar que muerde
- No levantar la cabeza
- No caber en sí de gozo

Asimismo, para llevar a cabo nuestros objetivos, hemos seleccionado varios motores de TA:

- Google Translate (neural)
- Bing Translator (neural)
- DeepL (neural)

- Gemini (neural)
- ChatGPT 3.5 (neural)

Además, hemos utilizado una referencia de traducción humana como estándar de comparación.

Con el fin de evaluar la efectividad los MTAs, hemos usado la plataforma BLEU proporcionado por TILDE (Para este estudio, hemos utilizado el sistema BLEU proporcionado por TILDE. Tilde Custom Machine Translation) donde, descargamos expresiones idiomáticas en lengua origen, la traducción de referencia y las traducciones de los diferentes motores, establecemos cuál es el motor más eficaz. Los pasos a seguir para la evaluación son los siguientes:

1. *Compilación de datos*: hemos recopilado las 19 expresiones idiomáticas en español y sus traducciones de referencia en árabe.
2. *Traducción automática*: cada expresión idiomática es traducida utilizando los motores de TA seleccionados.
3. *Comparación y análisis*: las traducciones obtenidas de cada motor han sido comparadas con las traducciones de referencia. Hemos evaluado la precisión, naturalidad y coherencia idiomática de cada traducción.
4. *Plataforma TILDE*: hemos usado la plataforma TILDE para analizar las traducciones generadas por los motores. Esta plataforma permite comparar las traducciones en términos de fidelidad y adecuación.

A continuación, presentamos la Tabla 1, que incluye las traducciones aportadas por diferentes motores de traducción automática, así como la traducción humana:

Nota. Elaboración propia.

Tabla 1: Corpus de análisis

	Google Translate	Bing Translator	DeepL	Gemini	ChatGPT 3.5	Traducción humana
1	لقفز من أجل الفرح	القفز من أجل الفرح	القفز من الفرح	رقص فرحا	يقفز من الفرح	رقص فرحا- يقفز من الفرح- يطير فرحا
2	يكون أسود	أن تكون أسود	أن تكون أسود	غاضبا	يكون غاضباً / يكون في حالة من الغضب الشديد	يكون غاضباً - غاضبا جدا - استشاط غضبا
3	يحطم فؤادك	كسر قلبك	أن تكون مكسور القلب	تحطم قلبي - انكسرت روحي	ينكسر القلب	ينكسر القلب - تحطم قلبي
4	سحر	كن على مستوى أنفك	أن يكون المرء حتى أنفه	مللت- ضجرت	يكون متعباً / يكون قد بلغ أقصى درجات الاستياء	مللت - ضجرت- بلغ أقصى درجات الاستياء- طغح الكيل- بلغ السيل الزبي
5	كن مليئاً بالرضا	كن بفيض بالرضا	أن تفيض بالرضا	ملئ بالسعادة - طافح بالبهجة	يكون مملوئاً بالرضا	قلبه بفيض بالسعادة- ملئ بالسعادة- يكون مملوئاً بالرضا
6	كن مثل فلان	أن تكون مثل الكاسترد	أن تكون مثل الفطيرة	لين كالهلام - ضعيف و مهتز	يكون مرتجفاً / يكون خائفاً	يكون مرتجفاً / يكون خائفاً - لين كالهلام - ضعيف و مهتز- ارنعدت فرائضه
7	كن مروغاً للدماء	تجميد الدم	أن تجمد دمك	تجمد دمي- خوفني الامر	يجمد الدم	تجمد دمي- خوفني الامر- جكد الدم في عروقه من الخوف
8	كن مثل عيد الفصح	أن تكون مثل عيد الفصح	أن تكون مثل عيد الفصح	سعيد جدا فرحان للغاية	يكون سعيداً جداً / يكون في قمة السعادة	يكون سعيداً جداً / يكون في قمة السعادة - فرحان للغاية- غمرته السعادة كأنه يوم عيد
9	تسلق الجدران	تسلق الجدران	أن تتسلق الجدران	فقدت صبري ثارت ثائرتي	يكون في قمة الإزعاج / يكون مضطرباً جداً	فقدت صبري- ثارت ثائرتي
10	كن مثل طفل بحذاء جديد	أن تكون مثل طفل يرتدي حذاء	أن تكون مثل طفل بحذاء جديد	فرحان كالاطفال متحمس جدا	يكون سعيداً جداً / يكون في قمة الفرح	يكون سعيداً جداً / يكون في قمة الفرح - يكون متحمسا
11	الحصول على تعليق منه	الحصول على الأعصاب	أن تكون خارج عقلك	اغضبني - جنني	يخرج عن السيطرة / يسبب الغضب	اغضبني- جنني- استفزني
12	املاً بالفرح	املاً بالفرح	أن تمتلئ بالبهجة	ملاني سعادة- اسعدني	يملأ بالسعادة	ملاني سعادة - اسعدني
13	الشعور بالتجميد	الشعور بالبرد	الشعور بالتجمد	يشعر بالخوف يشعر بالبرد	يشعر بالبرودة الشديدة / يشعر بالصدمة	يشعر بالبرد الشديد/ يشعر بالصدمة- يشعر بالخوف
14	أن تكون معنويًا على الأرض	أن تكون في حالة معنوية منخفضة	أن تكون محبطاً	محبط مكتئب	يكون محبطاً جداً / يكون في قمة الإحباط	كانت معنوياته في الحضيض - يكون محبطاً جداً / يكون في قمة الإحباط
15	يتحول إلى اللون الأحمر مع الغضب	تحول إلى اللون الأحمر مع الغضب	التحول إلى اللون الأحمر من الغضب	احمرّ وجهه غضباً اشتعل غضبه	يتحول إلى الأحمر من الغضب	احمرّ وجهه غضباً - اشتعل غضبه
16	يغلي الدم	غلي الدم	أن يغلي دم المرء	ثارت ثائرتي غلي دمي	يغلي الدم من الغضب	ثارت ثائرتي- غلي دمي
17	يكون ذلك لدغات	التعرض للعض	أن تشعر بالعض	كان على وشك الانفجار كان غاضباً جداً	يكون غاضباً جداً / يكون في قمة الغضب	كان على وشك الانفجار كان غاضباً جداً
18	لا ترفع رأسك	لا ترفع رأسك	عدم رفع الرأس	كان مكتئباً	لا يرفع رأسه / يشعر بالإحباط الشديد	طأطأ رأسه / يشعر بالإحباط الشديد- كان مكتئباً
19	لا يصلح داخل نفسه مع الفرح	لا ينبغي أن تشعر بسعادة غامرة	أن تمتلئ بالفرح	طفح فرحه كان سعيداً جداً	لا يبسه الفرح / يكون في قمة السعادة	لا يبسه الفرح / يكون في قمة السعادة

2. Descripción y análisis

Basándonos en los resultados obtenidos de la traducción de 19 expresiones idiomáticas con los MTAs (ChatGPT, Gemini, DeepL, Bing y Google Translate), presentamos un análisis detallado de cada uno. A continuación, describimos y examinamos tanto las traducciones correctas como las incorrectas, y proporcionamos una explicación de los errores cometidos en cada caso. Este análisis revela cómo cada sistema maneja las expresiones idiomáticas y por qué algunas traducciones fallan, ofreciendo una visión sobre las fortalezas y limitaciones de cada motor en la captura de matices y contextos culturales, tal y como se muestra en la Tabla 2:

Tabla 2: Resultados conseguidos de los diferentes MTAs

Google Translate	Gemini	Microsoft Translator	DeepL	ChatGPT
25.54 %	43.67 %	25.31 %	24.31 %	64.75 %

Nota. Elaboración propia basándonos en los resultados de TILDE.

A partir de los datos obtenidos a través de la plataforma TILDE, hemos observado que ChatGPT 3.5 y Gemini son los motores de traducción automática con mayor precisión en la traducción de expresiones idiomáticas del español al árabe. Hemos visto que ChatGPT 3.5 ha logrado la mejor puntuación, con un 64 %, destacándose por su capacidad para interpretar matices idiomáticos. Después, Gemini ha alcanzado un 43.67 % mostrando un desempeño sólido aunque el porcentaje de precisión es menor. En contraste, otros motores como Google Translate, Bing Translate y DeepL mostraron un desempeño inferior, con porcentajes de precisión menores al 26 %. Aseveramos, pues, que estos motores presentan un enfoque de traducción más literal y menos efectivo al manejar matices culturales de las expresiones idiomáticas. Asimismo, resaltamos la eficacia de los modelos basados en redes neuronales avanzadas y su capacidad para manejar contextos culturales complejos, aunque persisten desafíos en la interpretación semántica cultural.

A continuación, presentamos la Tabla 3 con los datos detallados de cada expresión idiomática y su porcentaje de precisión para cada motor de traducción:

Tabla 3: Expresión idiomática y porcentaje de precisión para cada motor de traducción

	ChatGPT	Gemini	DeepL	Bing	Google Translate
1	38.73 %	51.34 %	31.85 %	29.79 %	29.79 %
2	62.38 %	4.43 %	6.43 %	6.43 %	10.33 %
3	54.88 %	27.05 %	22.77 %	21.82 %	21.82 %
4	62.57 %	16.02 %	12.97 %	11.22 %	6.06 %
5	100 %	29.84 %	43.47 %	43.47 %	43.47 %
6	71.65 %	25.57 %	20.96 %	20.96 %	20.12 %
7	26.65 %	59.46 %	27.48 %	26.65 %	25.57 %
8	100 %	21.38 %	21.38 %	21.38 %	20.96 %

9	21.20 %	41.53 %	30.89 %	30.81 %	30.81 %
10	76.59 %	29.75 %	29.64 %	29.64 %	28.77 %
11	17.74 %	100 %	24.44 %	30.21 %	24.44 %
12	26.65 %	100 %	25.57 %	26.65 %	26.65 %
13	51.00 %	39.60 %	4.32 %	4.32 %	4.32 %
14	100 %	38.75 %	40.06 %	41.69 %	40.15 %
15	26.70 %	100 %	26.26 %	26.26 %	26.26 %
16	21.19 %	61.47 %	21.19 %	23.45 %	21.82 %
17	13.95 %	56.37 %	15.51 %	14.62 %	15.51 %
18	100	22.08 %	23.21 %	34.71 %	34.71 %
19	100 %	35.31 %	33.70 %	43.64 %	45.13 %

Nota. Elaboración propia basándonos en los datos de TILDE.

2.1. Resultados de ChatGPT 3.5

Después de realizar el estudio, determinamos que la plataforma ChatGPT es el mejor motor de traducción automática, ya que obtuvo las puntuaciones más altas en la evaluación de la traducción de expresiones idiomáticas. Las traducciones realizadas por este motor son más precisas y se asemejan más al lenguaje natural.

Es importante destacar que, de un total de 19 expresiones idiomáticas traducidas al árabe, ChatGPT logró traducir correctamente 16. De estas, 8 mantuvieron la misma estructura oracional, mientras que las otras 8 presentaron estructuras o terminologías ligeramente diferentes, pero conservaron la precisión idiomática. ChatGPT cometió errores en 3 expresiones idiomáticas, tal y como se refleja en la Tabla 4:

Tabla 4: Errores cometidos por ChatGPT en 3 expresiones idiomáticas

Expresión idiomática en español	Traducción errónea de ChatGPT	Comentario
Estar rebosante de satisfacción	يكون مملوءاً بالرضا	Aunque la traducción es precisa, la estructura puede ser más idiomática con "ملئ بالسعادة".
Estar helado de miedo	يشعر بالبرودة الشديدة / يشعر بالصدمة	La traducción es literal y no captura el matiz de "miedo". Debería ser "يشعر بالخوف الشديد".
Subirse por las paredes	يكون في قمة الإزعاج / يكون مضطرباً جداً	Aunque es comprensible, "تارت ثائرتي" captura mejor la esencia idiomática.

Nota. Elaboración propia.

Los errores de ChatGPT suelen ser mínimos y a menudo relacionados con la elección de palabras que, aunque correctas, pueden no capturar perfectamente el sentido idiomático.

2.2. Resultados de Gemini

El rendimiento de Gemini fue mixto, con algunas traducciones precisas pero una notable variabilidad en la calidad. Aunque logró traducir correctamente algunas expresiones, en general, las traducciones tendieron a ser más literales.

De un total de 19 expresiones idiomáticas, Gemini tradujo correctamente 7. De estas, 3 mantuvieron la misma estructura oracional, mientras que las otras 4 presentaron estructuras o terminologías diferentes. Sin embargo, tal y como se muestra en la Tabla 5, Gemini cometió errores en 12 expresiones idiomáticas.

Tabla 5: Errores cometidos por Gemini en 12 expresiones idiomáticas

Expresión idiomática en español	Traducción errónea de Gemini	Comentario
Dar saltos de alegría	القفز من أجل الفرح	Traducción literal, falta el sentido figurado.
Estar negro	أن تكون أسود	Traducción literal, debería ser "غاضب". او استشاط غضبا
Estar como un flan	أن تكون مثل الكاسترد	Traducción literal, no captura el sentido de "nervioso".
Helarse la sangre	تجميد الدم	Traducción literal, debería ser "تجمد دمي".
Estar como unas pascuas	أن تكون مثل عيد الفصح	Traducción literal, debería ser "سعيد جدا".
Subirse por las paredes	تسلق الجدران	Traducción literal, no captura el sentido idiomático.
Estar como un niño con zapatos nuevos	أن تكون مثل طفل يرتدي حذاء	Traducción literal, falta el sentido de "feliz".
Sacar de quicio	الحصول على الأعصاب	Traducción literal, no captura el sentido de "enfurecer".
Estar con la moral por los suelos	أن تكون في حالة معنوية منخفضة	Traducción literal, debería ser "محبط".
Ponerse rojo de ira	تحول إلى اللون الأحمر مع الغضب	Traducción literal, debería ser "احمر وجهه غضبًا".
Hervir la sangre	غلي الدم	Traducción literal, debería ser "غلي دمي".
Estar que muerde	أن تشعر بالعض	Traducción literal, debería ser "كان غاضبًا جدًا".

Nota. Elaboración propia.

Gemini tiende a proporcionar traducciones más literales, lo que a veces resulta en expresiones menos naturales en árabe. La variabilidad en los resultados sugiere que el motor podría beneficiarse de una mejor capacidad para reconocer y traducir elementos idiomáticos.

2.3. Resultados de DeepL

DeepL mostró un rendimiento razonable, aunque no alcanzó la precisión de ChatGPT. Sus traducciones fueron generalmente comprensibles, aunque a veces carecieron de naturalidad idiomática. De las 19 expresiones idiomáticas, DeepL tradujo correctamente 8. De estas, 4 mantuvieron la misma estructura oracional, mientras que las otras 4 presentaron estructuras

o terminologías diferentes. Sin embargo, DeepL cometió errores en 11 expresiones idiomáticas, tal y como se muestra en la Tabla 6:

Tabla 6: Errores cometidos por DeepL en 11 expresiones idiomáticas

Expresión idiomática en español	Traducción errónea de DeepL	Comentario
Dar saltos de alegría	القفز من الفرح	Traducción literal, falta el sentido figurado.
Estar negro	أن تكون أسود	Traducción literal, debería ser "غاضب".
Estar rebosante de satisfacción	أن تفيض بالرضا	Traducción literal, debería ser "ملى بالسعادة".
Estar como un flan	أن تكون مثل الفطيرة	Traducción literal, no captura el sentido de "nervioso".
Helarse la sangre	أن تجمد دمك	Traducción literal, debería ser "تجمد دمي".
Estar como unas pascuas	أن تكون مثل عيد الفصح	Traducción literal, debería ser "سعيد جدا".
Subirse por las paredes	أن تتسلق الجدران	Traducción literal, no captura el sentido idiomático.
Estar como un niño con zapatos nuevos	أن تكون مثل طفل بحذاء جديد	Traducción literal, falta el sentido de "feliz".
Sacar de quicio	أن تكون خارج عقاك	Traducción literal, no captura el sentido de "enfurecer".
Estar con la moral por los suelos	أن تكون محبطاً	Traducción literal, debería ser "محبط".
Hervir la sangre	أن يغلي دم المرء	Traducción literal, debería ser "غلي دمي".

Nota. Elaboración propia.

DeepL mostró una mejor capacidad para manejar la concordancia y el orden de palabras en comparación con Gemini y Bing; sin embargo, sus traducciones a veces carecieron del sentido figurado necesario para expresar con precisión las expresiones idiomáticas.

2.4. Resultados de Bing

Bing mostró un rendimiento inferior en comparación con otros motores, con traducciones a menudo literales que carecieron de naturalidad y precisión idiomática. De las 19 expresiones idiomáticas, Bing tradujo correctamente solo 6; de estas, 2 mantuvieron la misma estructura oracional, mientras que las otras 4 presentaron estructuras o terminologías diferentes. En total, Bing cometió errores en 13 expresiones idiomáticas, tal y como se muestra en la Tabla 7:

Tabla 7: Errores cometidos por Bing en 13 expresiones idiomáticas

Expresión idiomática en español	Traducción errónea de Bing	Comentario
Dar saltos de alegría	القفز من أجل الفرح	Traducción literal, falta el sentido figurado.
Estar negro	أن تكون أسود	Traducción literal, debería ser "غاضب". او استشاط غضبا
Estar rebosante de satisfacción	أن تفيض بالرضا	Traducción literal, debería ser "ملى بالسعادة".

Estar como un flan	أن تكون مثل الفطيرة	Traducción literal, no captura el sentido de "nervioso".
Helarse la sangre	أن تجمد دمك	Traducción literal, debería ser "تجمد دمي".
Estar como unas pascuas	أن تكون مثل عيد الفصح	Traducción literal, debería ser "سعيد جدا".
Subirse por las paredes	أن تتسلق الجدران	Traducción literal, no captura el sentido idiomático.
Estar como un niño con zapatos nuevos	أن تكون مثل طفل بحذاء جديد	Traducción literal, falta el sentido de "feliz".
Sacar de quicio	أن تكون خارج عقلك	Traducción literal, no captura el sentido de "enfurecer".
Estar con la moral por los suelos	أن تكون محبطاً	Traducción literal, debería ser "محبط".

Nota. Elaboración propia.

2.5. Resultados de Google Translate

Después de usar la plataforma Tilde, hemos encontrado que el motor de traducción Google Translate obtuvo un porcentaje de precisión del 25.54 %, lo que lo sitúa en el último lugar en comparación con los otros motores de traducción. Esto evidencia que este motor presenta una clara tendencia hacia la traducción literal, y muestra una capacidad limitada para captar matices culturales y contextuales inherentes a las expresiones idiomáticas. Según los datos obtenidos, hemos observado que esta plataforma falló en 17 de las 19 expresiones idiomáticas evaluadas, como se muestra en la Tabla 8:

Tabla 8: Errores cometidos por Google Translate en 17 expresiones idiomáticas

Expresión idiomática en español	Traducción errónea de Google Translate	Comentario
Dar saltos de alegría	القفز من أجل الفرح	Traducción literal, no captura el sentido figurado de felicidad extrema.
Estar negro	أن تكون أسود	Traducción literal, debería ser "عاضب" o "استشاط غضباً" para reflejar enojo.
Romperse el corazón	يحطم فؤادك	Traducción literal, falta naturalidad en el idioma árabe.
Estar hasta las narices	كن على مستوى أنفك	Traducción literal, no refleja el significado de "estar harto".
Estar rebosante de satisfacción	كن مليئاً بالرضا	Traducción literal, podría ser más idiomático como "ملى بالسعادة".
Estar como un flan	كن مثل فلان	Traducción literal, no captura el sentido de "estar nervioso".
Helarse la sangre	كن مروغاً للدماء	Traducción literal, no refleja el sentido de "tener miedo".
Estar como unas pascuas	كن مثل عيد الفصح	Traducción literal, no refleja "estar muy feliz".
Subirse por las paredes	تسلق الجدران	Traducción literal, no capta el sentido de "perder la paciencia".
Estar como un niño con zapatos nuevos	كن مثل طفل بحذاء جديد	Traducción literal, falta el sentido figurado de "estar emocionado".

Sacar el quicio	الحصول على تعليق منه	Traducción literal, no refleja "enfurecer".
Llenar de dicha	املاً بالفرح	Traducción literal, falta naturalidad idiomática en árabe.
Sentirse helado	الشعور بالتجميد	Traducción literal, no captura el matiz de miedo como "يشعر بالخوف".
Estar con la moral por los suelos	أن تكون معنويًا على الأرض	Traducción literal, debería ser "محبط" o "في حالة معنوية منخفضة".
Ponerse rojo de ira	يتحول إلى اللون الأحمر مع الغضب	Traducción literal, mejor como "احمرّ وجهه" "غضبًا".
Hervir la sangre	يعلي الدم	Traducción literal, debería ser "علي دمي".
Estar que muerde	يكون ذلك لدغات	Traducción literal, debería ser "كان غاضبًا جدًا".
No levantar la cabeza	لا ترفع رأسك	Traducción literal adecuada, pero con contexto limitado.
No caber en sí de gozo	لا يصلح داخل نفسه مع الفرح	Traducción literal, debería ser لا يسعه الفرح / يكون في قمة السعادة

Nota. Elaboración propia.

3. Discusión de resultados

En nuestro análisis, utilizamos cinco motores de traducción automática: Google Translate, Bing Translator, DeepL, Gemini y ChatGPT 3.5. Los resultados mostraron que ChatGPT 3.5 y Gemini ofrecieron las traducciones más precisas y fieles, con ChatGPT 3.5 alcanzando un 64.75 % de precisión general y Gemini un 43.67 %. Otros motores, como Google Translate, Bing Translator y DeepL, presentaron porcentajes de precisión menores, con 25.31 %, 24.31 % y 25.31 %, respectivamente. Estos resultados se coinciden con el estudio realizado por Alzeebaree (2020) y Olmedo Ruiz (2018).

En el análisis de equivalencia total, encontramos que solo un pequeño porcentaje de las traducciones automáticas logró una equivalencia completa con la traducción de referencia humana. ChatGPT 3.5 y Gemini destacaron nuevamente en este aspecto, aunque aún mostraron un número considerable de traducciones con equivalencia parcial o traducciones literales que no capturaban el sentido idiomático completo.

Un hallazgo notable de nuestro estudio fue la frecuencia con la que los motores de traducción recurren a la paráfrasis y a la traducción literal. Por ejemplo, expresiones como "estar como un flan" fueron traducidas literalmente como "ser como una tarta", lo cual no captura el sentido de "estar nervioso". Este patrón fue más pronunciado en motores como Google Translate y Bing Translator.

También destacamos que las traducciones literales eran comunes y frecuentemente resultaban en frases que, aunque gramaticalmente correctas, eran culturalmente inadecuadas o incluso incomprensibles. Las paráfrasis, aunque a veces más precisas, a menudo no lograban mantener el tono idiomático original. Estos resultados coinciden con el estudio de Alzeebaree (2020).

Observamos que los MTAs presentan dificultades significativas en la traducción de expresiones idiomáticas debido a la complejidad cultural y lingüística inherente. Todos los motores evaluados enfrentaron dificultades notables al traducir expresiones idiomáticas, lo que subraya la complejidad de mantener la autenticidad cultural y lingüística en la traducción automática. Cabe señalar que estos resultados coinciden con los de Olmedo Ruiz (2017).

Conclusión

En este apartado presentamos las conclusiones derivadas de la investigación sobre la traducción automática de expresiones idiomáticas del español al árabe. Inicialmente, retomamos los objetivos fundamentales del estudio, que consistían en evaluar la precisión y fidelidad de los traductores automáticos y herramientas de inteligencia artificial en la traducción de expresiones idiomáticas complejas y culturalmente ricas entre estas dos lenguas.

La aplicación del sistema BLEU, proporcionado por TILDE, permitió una evaluación cuantitativa precisa de las traducciones realizadas por cinco programas de traducción: Google Translate, Bing Translator, DeepL, Gemini y ChatGPT 3.5, comparadas con una traducción humana de referencia. Los datos obtenidos y analizados revelan que, aunque todos los motores enfrentaron ciertos desafíos en la interpretación de algunas expresiones idiomáticas, ChatGPT 3.5 y Gemini ofrecieron las traducciones más precisas y fieles.

Estos resultados confirman que los sistemas de traducción automática, especialmente aquellos basados en inteligencia artificial avanzada, han alcanzado un nivel notable de competencia en la traducción de expresiones idiomáticas. Sin embargo, es importante destacar que las complejas diferencias gramaticales y culturales entre el español y el árabe aún representan un desafío significativo para estos sistemas. Las traducciones automáticas, aunque mejoradas, siguen requiriendo supervisión humana para asegurar la precisión y el contexto adecuado.

En vista de estos resultados, el estudio recomienda una integración híbrida que combine la eficiencia de los traductores automáticos con la revisión y supervisión humana, especialmente en el ámbito de las expresiones idiomáticas y otros elementos lingüísticos culturalmente específicos. Este enfoque no solo mejoraría la calidad de las traducciones automáticas, sino que también contribuiría a un aprendizaje más profundo y contextualizado del idioma.

En conclusión, este estudio establece un marco sólido para entender el estado actual y las capacidades de la TA en el par de lenguas español-árabe. Además, subraya la necesidad de continuar investigando y desarrollando estas tecnologías para superar las limitaciones actuales y optimizar su uso en contextos lingüísticos desafiantes. Para profundizar en este ámbito y evaluar el impacto a largo plazo de las nuevas estrategias propuestas, se recomienda llevar a cabo estudios longitudinales que midan cómo las intervenciones híbridas de traducción influyen en la precisión y comprensión a lo largo del tiempo. Este enfoque brindaría una comprensión más completa de la eficacia de las intervenciones y su contribución al éxito en la traducción automática.

Referencias bibliográficas

- ALZEEBAREE Y. 2020. «Machine translation and issues of multiword units: Idioms and collocations». *Eastern Journal of Languages, Linguistics and Literatures*, 1(2). Consultado el 16 de enero de 2024, file:///C:/Users/win/Downloads/5-Article%20Text-20-1-10-20210215.pdf
- BOERS F. y STENGERS H. 2008. «A quantitative comparison of the English and Spanish repertoires of figurative idioms». *Applications of cognitive linguistics*, 6, 355.
- DOBROVOL'SKIJ D. 1994. «Sobre la equivalencia translingüística de los fraseologismos», en A. Hurtado (ed.) *Estudis sobre la Traducció*. Castellón: Universitat Jaume I, 359-380.
- DOBROVOL'SKIJ D. 2000. «Idioms in contrast: A functional view», en G. Corpas (ed.) *Las Lenguas de Europa. Estudios de Fraseología, Fraseografía y Traducción*. Granada: Editorial Comares, 367-388.
- DOBROVOL'SKIJ D. y PIIRAINEN E. 2005. *Figurative Language. Cross-Cultural and Cross-Linguistic Perspectives*. Amsterdam: Elsevier
- GONZÁLEZ-REY F., MITJÁNS MARTÍNEZ A. 2016. «Una epistemología para el estudio de la subjetividad: Sus implicaciones metodológicas». *Psicoperspectivas*, 15(1), 5-16.
- HARRIS Z. 1957. «Co-occurrence and Transformation in Linguistic Structure», *Language*, vol. 3, no. 33, pp. 283-340. (Citado por su reedición en Harris 1981, pp. 143-210).
- HERNÁNDEZ R., FERNÁNDEZ C. y BAPTISTA P. 2014. *Metodología de la investigación*. McGraw Hill España. House
- HUTCHINS W. J. 1995. *Introducción a la traducción automática*. Madrid: Visor.
- HUTCHINS W. J. 2002. *The state of machine translation in Europe and future prospects*, en HLT Central.
- HUTCHINS W. J. 2010. *Machine translation: a concise history* en Sin-Wai, C. (ed.): *Special issue: The teaching of computer-aided translation*. Hong Kong: Chinese University of Hong Kong, 29-70.
- KATAN D. 1999. *Translating Cultures. An Introduction for Translators, Interpreters and Mediators*. Manchester: St. Jerome Publishing.
- LUQUE J. y MANJÓN R. 1998. «Tipología léxica y tipología fraseológica: universales y particulares», en J.d.D. Luque y A. Pamies (eds.) *Léxico y Fraseología*. Universidad de Granada: Método Ediciones, 139-153.
- MONTI J. 2019. «Statistical vs. neural machine translation: A comparative study». *Journal of Language Technology*, 34(2), 175-192. Consultado el 16 de enero de 2024, https://doi.org/10.1162/jlang_a_00289
- NIDA E. 1999. «Lengua, cultura y traducción», en M.A. Vega y R. Martín Gaitero (eds.) *Lengua y Cultura. Estudios en torno a la Traducción*. Madrid: Editorial Complutense, 1-6
- OLMEDO RUIZ M. 2018. *Los tipos de traducción automática y su evaluación mediante perífrasis verbales y expresiones idiomáticas (alemán-español)*. Consultado el 20 de mayo de 2024, <https://ddd.uab.cat/record/192790>
- SAMANIEGO J. 2007. «El impacto de la lingüística cognitiva en los estudios de traducción», en P.A. Fuertes (coord.) *Problemas Lingüísticos en la Traducción Especializada*. Valladolid: Universidad de Valladolid, 119-154.