Rita Shaheer Shawky

Limitations of Machine Translation of Egyptian Dialect in Social Comedy Series into English: A Comparative Study between AI Subtitling Tools
(Rita Shaheer Shawky, Translator)

Abstract

This comparative study investigates the limitations of Machine Translation (MT) in subtitling Egyptian Arabic comedy series, with a particular focus on the preservation of pragmatic meaning and cultural humor. The selected dataset extracted from Egyptian social comedy shows, subtitled using three AI tools: FreeSubtitles.ai, Veed.io, and Maestra.ai. These machine-generated subtitles are evaluated against the theoretical framework of Mona Baker's (1992) pragmatic equivalence model, an extended version of Pedersen's (2017) FAR model and Gottlieb's (1992) subtitling strategies. Through a qualitative and quantitative comparative approach, the AI subtitles are contrasted with human translation strategies to assess how well meaning, tone, and cultural references are conveyed. The findings indicate that MT tools frequently rely on literal translation, often failing to capture the implicit humor and cultural nuances embedded in Egyptian colloquial Arabic. This study highlights the need to expand AI training datasets to include dialectal Arabic, not just Modern Standard Arabic to improve the quality and contextual accuracy of subtitle translations. Ultimately, the research affirms that human translators remain essential in delivering contextually appropriate and culturally sensitive subtitles in comedic content.

Keywords: Machine Translation, Subtitling, Egyptian Arabic, Pragmatic Equivalence, Dialects, Cultural References

قيود الترجمة الآلية للهجة المصرية في المسلسلات الكوميدية الاجتماعية إلى اللغة الإنجليزية:

دراسة مقارنة بين أدوات الترجمة الآلية المصاحبة للنصوص

المستخلص

تهدف هذه الدراسة إلى تقييم كفاءة الترجمة الآلية في ترجمة نصوص الكوميديا الاجتماعية المصرية إلى اللغة الإنجليزية عبر الترجمة النصية (Subtitling) ، مع التركيز على مدى تحقيق التكافؤ التداولي في النصوص المُترجمة. استخدمت الدراسة أدوات ترجمة آلية مثل FreeSubtitles.ai وVeed.io وMaestra.ai، وقارنت نتائجها بإطار من نماذج تحليل الجودة والترجمة، وهي نموذج التكافؤ التداولي لمنى بيكر (١٩٩٢) ونموذج FAR الموسّع ليان بيدرسن (٢٠١٧) واستراتيجيات جوتليب في الترجمة للأفلام أو الترجمة النصية (١٩٩٢). كشفت النتائج عن قصور واضح في قدرة الأدوات الآلية على نقل المعنى الضمني والفكاهة والسياقات الثقافية المرتبطة باللهجة المصرية العامية، واعتمادها بشكل أساسي على الترجمة الحرفية. وتوصي الدراسة بضرورة تدريب أنظمة الذكاء الاصطناعي على بيانات تشمل اللهجات العربية، لا سيما المصرية، لتعزيز كفاءتها في الترجمة السمعية البصرية، والحفاظ على البعد الثقافي والمعنوي للنص الأصلي.

Limitations of Machine Translation of Egyptian Dialect in Social Comedy Series into English: A Comparative Study between AI Subtitling Tools
(Rita Shaheer Shawky, Translator)

Introduction

In recent years, the field of audiovisual translation (AVT) has witnessed a growing reliance on artificial intelligence (AI) and machine translation (MT) technologies to generate subtitles quickly and at scale. While these advancements have significantly improved translation speed and accessibility, they raise essential concerns regarding the quality, cultural accuracy, and communicative effectiveness of the output, especially when dealing with dialectal and humorous content. Subtitling is not merely a process of rendering words from one language into another; it requires a deep understanding of pragmatics, cultural references, and audience expectations. This is particularly relevant in the context of Egyptian Arabic social comedy, where meaning often depends on tone, idiomatic expressions, and shared cultural knowledge.

Most AI subtitling systems are trained primarily on Modern Standard Arabic (MSA) and fail to account for the richness and variability of Arabic dialects. As a result, when these tools attempt to process colloquial dialogue from Egyptian comedy series, they frequently produce literal, awkward, or even misleading translations. These mistranslations disrupt the viewing experience and, more importantly, strip the content of its humorous and cultural essence. Consequently, the effectiveness of AI-generated subtitles in capturing pragmatic meaning and preserving the source text's intent becomes a critical area of investigation.

To address these challenges, this study evaluates the subtitles produced by three popular AI tools, FreeSubtitles.ai, Veed.io, and Maestra.ai, when applied to selected clips from Egyptian Arabic comedy shows. The analysis draws on Mona Baker's (1992) model of pragmatic equivalence to assess how well the intended meaning is preserved and employs Jan Pedersen's (2017) FAR model (Functionality, Acceptability, Readability) as a scoring system to identify translation errors. Special attention is given to how the tools handle culturally loaded language, humor, implicature, and idiomatic expressions.

By comparing machine-generated subtitles with human translation strategies and analyzing the types and frequency of errors, the study aims to determine whether AI tools are capable of producing pragmatically

equivalent subtitles in dialect-rich, comedic content. Also, this research emphasizes the importance of including Arabic dialects in AI training data to enhance the contextual competence of subtitling systems. Ultimately, the study seeks to contribute to ongoing discussions about the role of AI in translation, the limitations of current MT technology, and the irreplaceable value of human translators in culturally sensitive domains.

Arabic language is a very challenging language to learn. In Arabic the spoken language differs completely from the written form. Despite the advancements in machine translation technology and how it excels in translating formal Arabic, it encounters difficulties in translating dialects like Egyptian dialect, resulting in inaccurate translations particularly in audiovisual media.

1. Literature Review

The FAR model is widely used to evaluate subtitle quality across languages, especially in preserving emotional content. Sanatifar and Ghamsarian (2023) applied the FAR model focusing on functional equivalence, acceptability, and readability to assess English subtitles of five Persian drama films. Using Johnson-Laird's theory of basic emotions, they categorized errors by severity and found an 81% emotional equivalence, highlighting the model's effectiveness in capturing emotional accuracy. Similarly, Koglin et al. (2022) applied the FAR model to machine-translated English-to-Brazilian Portuguese subtitles. Their study showed that minor errors did not greatly impact readability or comprehension, further supporting the FAR model's role in evaluating subtitle quality across emotional and linguistic contexts. This framework can likewise be applied to assess AI-generated subtitles of Egyptian dialect comedy series, especially in terms of emotional expression and humor delivery.

A major challenge in audiovisual translation (AVT) and computational linguistics is the limited evaluation of Automatic Speech Recognition (ASR) systems, especially for Arabic dialects. Akasheh et al. (2024) address this by evaluating ASR performance in generating subtitles for Jordanian Arabic, proposing a new error classification for AI-generated subtitles and identifying both linguistic and technical limitations through quantitative and qualitative methods. Likewise, Rahman et al. (2024) investigate the progress and persistent issues in Arabic ASR systems, emphasizing how Arabic's grammatical and structural complexity affects recognition accuracy. Their survey of speech databases and recognition

techniques reveals that, despite technological advances, Arabic ASR remains imperfect. Both studies highlight the necessity of tailoring ASR systems to Arabic dialects to improve the quality of AVT. These findings provide valuable context for assessing the limitations of current machine translation tools in rendering Egyptian dialects in social comedy series, ultimately contributing to better translation strategies for Arabic-speaking audiences.

Although AI technologies like ChatGPT have rapidly evolved, they still fall short in accurately handling complex multilingual media texts. This limitation necessitates human translators to ensure cultural relevance and linguistic precision. Zayed and Nuirat (2024) explore this gap by analyzing ChatGPT's translations of five media texts sourced from Arab news channels. Using a descriptive-analytical method, their study involves expert translators reviewing the AI-generated outputs and applying post-editing strategies to enhance accuracy and naturalness. The research also incorporates interviews to gather qualitative insights, with the central argument being that tools like ChatGPT should complement, not replace, human translators.

Similarly, Mohamed et al. (2024) investigate the broader landscape of AI-powered translation. Their literature review examines developments and obstacles in the field, along with ethical concerns tied to the use of AI in translation. They review technologies such as machine learning, deep learning, neural machine translation (NMT), and natural language processing (NLP), and base their analysis on established models like statistical machine translation (SMT), deep learning frameworks, and fuzzy logic. Their findings recognize the advances in AI but stress the continued need for human oversight to manage ethical challenges and ensure contextual sensitivity.

Arabic dialects present significant obstacles for machine translation (MT) systems due to their non-standardized nature and structural complexity. These challenges are particularly evident when translating between dialects and Modern Standard Arabic (MSA). Faheem et al. (2024) and Khalati and Al-Romany (2020) investigate these difficulties, focusing on enhancing MT performance, especially for the Egyptian dialect. Faheem et al. explore how neural machine translation (NMT) systems can be optimized through three experimental settings: supervised, unsupervised, and semi-supervised. In the supervised model, they employed 40,000 manually aligned Egyptian-MSA sentence pairs. For the unsupervised model, they used a large monolingual dataset of 20 million sentences from platforms like Wikipedia. The semi-supervised approach combined both datasets, leveraging the advantages of each. Based on NMT

principles, which utilize deep learning to map linguistic patterns, the study found that the hybrid, semi-supervised model yielded the highest BLEU scores, signaling notable progress in translation quality and scalability for low-resource language pairs.

In summary, these studies emphasize that while AI-based solutions are advancing, significant linguistic and technical challenges remain in accurately translating Arabic dialects into MSA. Translating Arabic dialects into Modern Standard Arabic (MSA) remains a complex challenge in the field of machine translation (MT) due to dialectal diversity, lack of standardization, and limited parallel training data.

1.1. Theoretical Framework

This study combines Mona Baker's Pragmatic Equivalence Approach with Pedersen's FAR Model and incorporates Gottlieb's Subtitling Strategies to assess the quality of AI-generated subtitles, especially in translating Egyptian Arabic dialects.

Baker's framework emphasizes conveying the speaker's intended meaning in context, focusing on implicature, coherence, and cultural nuances rather than literal translation. This is particularly crucial in subtitling, where cultural references, tone, and communicative intent must be preserved within time and space constraints.

To systematically assess these issues, the study applies Pedersen's FAR model, which evaluates subtitles based on Functionality (meaning accuracy), Acceptability (grammatical and cultural appropriateness), and Readability (clarity and ease of understanding). Each dimension is scored using a penalty system to pinpoint errors and evaluate the impact on subtitle quality. Baker's pragmatic equivalence criteria are mapped onto the FAR categories, creating a hybrid evaluation tool that highlights where AI falls short in conveying nuanced meaning.

Table 1.1 Alignment of Pragmatic Equivalence Elements with FAR Model

FAR Model Category	Baker's Pragmatic Equivalence Errors	Description	Scoring System (Penalty Deductions)
Functionality (F)	Implicature Loss	AI fails to capture indirect meanings (Ex: sarcasm, humor, tone).	0.5 points (minor distortion) 1.0, 2.0 points (major meaning loss)
	Cultural Mismatch	AI translates literally without adapting cultural references.	0.5, 1.0 points, depending on severity
	Coherence Disruption	AI subtitle makes no sense in context or disrupts the flow of conversation.	1.0, 2.0 points, depending on how much it affects comprehension
Acceptability (A)	Cultural Mismatch	Subtitle is grammatically correct but inappropriate for the cultural context.	0.25 points for minor errors 0.5 for standard errors 1.0 point for serious errors
Readability (R)	Coherence Disruption	Subtitle is unclear, awkward, or difficult to read.	0.25 points for minor errors 0.5 for standard errors 1.0 point for serious errors

Additionally, subtitling strategies adopted from Gottlieb (1992) to assess how translation techniques relate to the types of errors identified in the FAR model. These strategies are crucial in compensating for AI's limitations, particularly in dealing with slang, idioms, and cultural expressions in Egyptian Arabic.

Table 1.2 Gottlieb Subtitling Strategies.

Strategy	Definition	Example
Expansion	Expansion is used when the dialogs in the SL need an explanation to gain the audience comprehension because they cannot retrieve the cultural nuance of the SL.	SL: "ده ذكي وبيفكر بعشر طرق في نفس الوقت" TL: "He's not just smart, he's a genius who can think in ten directions at once!"
Paraphrase	Paraphrase is used when the phrase in the SL cannot be reconstructed in the same syntactic way in the TL. In other words, the translation in the TL is syntactically different from the one of the SL but the meaning is still maintained to be comprehended by the audience.	SL: "الراجل ده ايده طايلة" TL: "This man has powerful connections"
Transfer	Transfer refers to the strategy of translating the SL completely and correctly into the TL.	SL: "هو فاكر نفسه جايب الديب من ديله!" TL: "He thinks he's done something impossible!"
Imitation	Imitation is used to translate the proper noun like names, places, countries, and product brands.	SL: "قابلت أحمد في الأقصر". TL: "I met Ahmed in Luxor."
Transcription	Transcription is used when there is an existence of unusual terms, the third language and nonsense language in the SL.	SL: "Ciao" TL: "Ciao"
Dislocation	Dislocation is adopted when the SL employs some sort of special effect, for example a silly song in a cartoon film, where the translation of the effect is more important than the content.	SL: "يلا نرقص ونغني، دي الدنيا يومين تلاتة" TL: "Come on, let's dance and sing, life's just a fleeting fling!"
Condensation	Condensation is applied to solve the problem of limitation of subtitle lines. This can also create efficiency by eliminating the redundancies.	SL: "بص يا عم، أنا كنت ناوي أقولك على الموضوع ده من زمان، بس ما كانش فيه وقت وكل مرة كنت بتشغلني حاجة". TL: "Look, man, I wanted to tell you earlier, but I was busy."

Decimation	Decimation is omitting important element that are confusing the audience and some taboo words.	SL: " إنت عايزني أعمل إيه يعني؟ "أرقصلك في الشارع زي المجانين؟" TL: "What do you want me to do?"
Deletion	Deletion refers to deal with the total elimination of the parts of a text, such as repetition, filler words and question tags.	SL: " يا سلام! هو ده الكلام اللي يفرّح القلب!" TL: "That's great!"
Resignation	Resignation is applied when the translator does not find the solution in translating the SL subtitle and the meaning is inevitably lost.	SL: "يا نهار أبيض!" TL: "Oh white day!"

Overall, this framework supports a comprehensive analysis of how AI translation tools perform in subtitling dialectal Arabic, and provides insight into how human intervention, guided by pragmatic and linguistic principles, can enhance subtitle quality.

2. Methodology

The data of this study consists of 10 episodes from two Egyptian social comedy series in Egyptian dialects, (Family Matters and Kamel El Adad), 5 episodes of each show. The original scripts will be in Egyptian Arabic. The two series were downloaded from CimaClub without subtitles, to be added to the AI tools (Veed.io, Maestra.ai, FreeSubtitles.ai). The series were chosen to represent the diverse language and culture of Egyptian dialects. They include interactions between different age groups, incorporating a variety of slang, idioms, and cultural references. This diversity is crucial for assessing how well AI translation tools handle Egyptian Arabic, highlighting its complexity and richness.

Social comedy series use realistic, everyday language, making them ideal for studying how AI tools translate natural speech, including slang, humor, and informal expressions.

The age group diversity shows a wide range of language styles. Older characters often use traditional slang, while younger ones use modern or emerging expressions. This helps evaluate how well AI tools adapt to different linguistic styles. Also, Comedy relies on cultural references, idioms, and wordplay, which are difficult for machine translation tools to handle. These series provide rich material for testing how AI tools deal with the unique aspects of humor in Egyptian dialects. Moreover, the fast-paced, overlapping dialogues and shifts in tone in these shows add extra challenges. These features test the AI tools' ability to understand context and produce coherent subtitles.

3. Analysis
Family Matters
Example 1:
Arabic ST

يا اخي بدل مانت عمال تشتم كده قوم صليلك ركعتين

Table 3.1 : _Family Matters:_ Error Analysis Example 1			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	And then, instead of insulting me, pray for me.	And then, instead of insulting me, pray for me.	Bro instead of cursing, go and pray two rak'ahs.
Strategy	**Resignation Strategy**	**Resignation Strategy**	**Transfer Strategy**
FAR Model Scoring	**Functionality:** **Score: 0.5**	**Functionality:** **Score:** 1	**Functionality:** **Score: 0**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 1**
	Readability: **Score: 0.25**	**Readability:** **Score**: 0.25	**Readability:** **Score: 1.25**
Total Score	**0.75**	**1.25**	**1.25**

At the very beginning of Episode 2 of Family Matters, the grandfather, Mounir, was revealing a truth to Ibrahim and then started cursing him and saying Ibrahim was not a respectable man. In response, Ibrahim says: "يا اخي بدل مانت عمال تشتم كده قوم صليلك ركعتين" which means "Oh man, instead of cursing me, go pray to God to forgive you."

Among the three subtitling tools, Maestra.ai offers the most accurate and culturally faithful translation. It uses the transfer strategy and correctly renders the phrase, including the religious reference to "rak'ahs", which refers to the Islamic prayer, but it fails to achieve the pragmatic equivalence as described in Mona Baker's theory as it only transliterates the word "rak'ahs" instead of getting an equivalent in English. In contrast, both Veed.io and FreeSubtitles.ai mistranslate the word "rak'ahs" as "pray for me", which distorts the intended meaning. Ibrahim wasn't asking Mounir to pray for him but rather telling him to pray for Mounir's own forgiveness after hiding the truth for so many years. For this reason, Maestra.ai stands out for preserving both the intended meaning and the cultural context of the original dialogue.

Example (2)
Arabic ST

رمى المسئوليه دي كلها في حجري

Table 3.2: *Family Matters:* Error Analysis Example 2

Tool	1: Veed.io	2: FreeSubtitles.ai	3: Maestra.ai
English MT	And threw all these responsibilities on me now?	Ramel, all my responsibilities are in Hijri.	What does it mean to put all this responsibility on my lap?
Strategy	**Transfer Strategy**	**Resignation Strategy**	**Transfer Strategy**
FAR Model Scoring	**Functionality:** **Score: 0**	**Functionality:** **Score:** 2	**Functionality:** **Score: 1**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 1	**Acceptability:** **Score: 0**
	Readability: **Score: 0.25**	**Readability:** **Score**: 0.5	**Readability:** **Score: 0.5**
Total Score	**0.25**	3.5	1.5

Table 3.3 Example 3

In this scene of Episode 2, Ibrahim struggles to tell the girl that he is not her lawyer, but her father. He is clearly overwhelmed by the emotional weight and responsibility of this revelation. The original statement he makes is metaphorical, conveying a sense of being suddenly and unfairly burdened with responsibility as if everything has been heaved at him without warning.

Veed.io employs the transfer strategy effectively, preserving the metaphorical meaning in a way that sounds both natural and idiomatic. In contrast, FreeSubtitles.ai applies the resignation strategy, which results in a loss of the intended meaning, it seems to have merely transcribed the dialogue rather than translated it. Maestra.ai also uses the transfer strategy, but its version is more literal and lacks fluency, even though the general meaning is retained. Therefore, Veed.io offers the most contextually appropriate translation by achieving functional equivalence and maintaining pragmatic clarity.

Example (3)
Arabic ST

يعني ايه؟ مش هتقولها انك ابوها؟

Table 3.3: *Family Matters:* Error Analysis Example 3			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	You're not gonna tell her you're her father?	What do you mean? Aren't you going to tell her that you're her father?	Aren't you going to tell her that you are her father?
Strategy	**Condensation Strategy**	**Transfer Strategy**	**Condensation Strategy**
FAR Model Scoring	**Functionality:** **Score: 0**	**Functionality:** **Score:** 0	**Functionality:** **Score: 0**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 0**
	Readability: **Score: 0.25**	**Readability:** **Score:** 0	**Readability:** **Score: 0.25**
Total Score	**0.25**	**0**	**0.25**

Still in Episode 2 of Family Matters, Ibrahim continues his conversation with the lawyer. The lawyer questions whether Ibrahim truly intends to keep his identity as her father a secret, clearly expressing disapproval.

Veed.io and Maestra.ai applied the condensation strategy, omitting the initial part of the dialogue. While their translations remain natural and fluent, the omission of the first question weakens the emotional intensity, and the sense of disapproval conveyed in the original. A more accurate and emotional translation would be: "Are you serious? You're not going to tell her that you're her father?" and it should be divided into 2 segments introduced in 2 different time stamps as the lawyer paused between the two sentences. Therefore, the wrong segmentation and timing affects by AI tools affect the readability of the subtitles. In contrast, FreeSubtitles.ai used the transfer strategy and successfully preserved both questions, accurately capturing the original tone and meaning and the timing was accurate.

Example (4):
Arabic ST

صبتب دي اسود يانهار

Table 3.4: *Family Matters:* Error Analysis Example 4			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	Wait, what are you saying?	Oh my god she's looking!	Sweden, look,look
Strategy	**Resignation Strategy**	**Transfer Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 2**	**Functionality:** **Score:** 0	**Functionality:** **Score: 2**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 0**
	Readability: **Score: 0**	**Readability:** **Score**: 0	**Readability:** **Score: 0**
Total Score	**2**	**0**	**2**

In episode 2, same scene, Ibrahim continues speaking with the lawyer, who informs him that he needs to travel to Cairo, while Ibrahim is talking to the lawyer, his daughter is sitting at a nearby café, watching them. When Ibrahim noticed her, he became visibly nervous and said, "صبتب دي اسود يانهار" ("Damn, she's staring at us" or "Oh my God, she's looking at us.")

FreeSubtitles.ai successfully captures the pragmatic meaning rather than translating the phrase literally or word by word, it instead got an equivalent to "اسود يانهار" which is "oh my god" The result is natural and fluent, preserving the emotional nuance of the moment. In contrast, Veed.io misinterprets the phrase entirely, applying the resignation strategy and failing to convey the intended meaning. Similarly, Maestra.ai also misinterprets the line, likely due to a mistranscription of the original Arabic.

Example (5):
Arabic ST

مش فاهم هي شحططة وخلاص ولا ايه!

Table 3.5: *Family Matters:* Error Analysis Example 5			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	I don't understand, did she just charge it and that's it or what?	It doesn't make any sense.	Not in any of them, I mean...
Strategy	**Resignation Strategy**	**Paraphrase Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 1**	**Functionality:** **Score:** 0	**Functionality:** **Score: 2**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 0**
	Readability: **Score: 0**	**Readability:** **Score:** 0	**Readability:** **Score: 0**
Total Score	**1**	**0**	**2**

In Episode 2, after the grandfather's death, the lawyer announces that, according to the will, the grandfather wished to be buried in Hurghada. to which the son replies, "مش فاهم هي شحططة وخلاص ولا ايه!" ("I don't get it! This is just a pointless hassle!"). This line reflects the son's frustration and annoyance at having to travel to Hurghada for the funeral.

Veed.io attempts a literal translation but completely misinterprets the key word "شحططة" as "charge it" which leads to a total loss of meaning. This represents a clear case of cultural mismatch and implicature loss. By contrast, FreeSubtitles.ai uses a paraphrase strategy and successfully conveys the speaker's tone and attitude through a generalized expression of confusion and frustration: "It doesn't make any sense." While the translation omits some cultural specificity, it adheres to Mona Baker's pragmatic equivalence approach, which prioritizes the intended meaning and communicative function over literal accuracy. Maestra.ai, however, produces a translation that is entirely unrelated to the original line, resulting in a severe disruption of coherence. Among the three, FreeSubtitles.ai provides the closest rendering to the original meaning, despite its simplification, as it successfully conveys intent behind the son's statement.

Kamel El Adad

Example (1)

Example 1:

Arabic ST

الله ينور عليك طالما ماتسألتش بشكل مباشر يبقى مش كدب

Table 3.6: *Kamel El Adad*: Error Analysis Example 1			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	Exactly, as long as you don't ask me directly, it's not a lie.	God bless you. As long as you don't ask me directly, it's not a lie.	May God enlighten you. As long as you don't ask directly, it won't be a lie.
Strategy	**Transfer Strategy**	**Transfer Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 0**	**Functionality:** **Score:** 0	**Functionality:** **Score: 0.5**
	Acceptability: **Score: 0**	**Acceptability:** **Score: 0**	**Acceptability:** **Score: 0**
	Readability: **Score: 0**	**Readability:** **Score**: 0	**Readability:** **Score: 0**
Total Score	**0**	**0**	**0.5**

Table 3.12 Example 1

In Episode 2 of *Kamel El Adad*, Ahmed asks Laila not to mention how they met his friend in front of his wife. Laila replies that she doesn't know how to lie, but Ahmed reassures her by saying, "If she doesn't ask you directly, then it's not lying." In response, Laila says:

''الله ينور عليك طالما ماتسألتش بشكل مباشر يبقى مش كدب'' which means ("God bless you! As long as you weren't asked directly, then it's not lying.").

Veed.io replaces the phrase ''الله ينور عليك'' with "Exactly," which, preserves the logical flow and pragmatic implication. It delivers the meaning of the sentence by getting an equivalent that will help viewers to understand. Also, FreeSubtitles.ai offers a more culturally grounded rendering by translating ''الله ينور عليك'' as "God bless you." It is not literal, this choice effectively conveys both the sarcastic tone and the cultural nuance, aligning well with Baker's pragmatic equivalence approach. It maintains the speaker's intent and emotional tone in a way that resonates with English-speaking audiences. Maestra.ai, on the other hand, chooses a literal translation which sounds awkward and lacks the pragmatic tone, making it unnatural.

Example 2:

Arabic ST

اتقابلنا في الأقصر واسوان

Table 3.7: *Kamel El Adad*: Error Analysis Example 2			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	We met in Luxor and Aswan	We met in his family in Aswan.	We met in the palace and in Aswan.
Strategy	**Transfer Strategy**	**Resignation Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 0**	**Functionality:** **Score:** 2	**Functionality:** **Score: 2**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 0**
	Readability: **Score: 0**	**Readability:** **Score**: 0	**Readability:** **Score: 0**
Total Score	**0**	**2**	**2**

In Episode 2, Ahmed takes Laila to his friend's birthday party, where the friend's wife asks Laila how they met. Laila responds that they met in Luxor and Aswan.

Veed.io translates the statement faithfully and accurately. Since Luxor and Aswan are culturally and historically significant cities, preserving their names without distortion is crucial. Veed.io succeeds in maintaining both the geographic reference and cultural relevance. In contrast, FreeSubtitles.ai and Maestra.ai fail to translate Luxor correctly, despite it being one of the most prominent historical cities in Egypt. This failure results in a significant loss of meaning and coherence. The mistranslation may be due to inaccurate transcription or a lack of contextual awareness within the tools. This suggests a broader limitation in their design, potentially a lack of integrated research capabilities that could allow the tool to verify unfamiliar terms via web search or context recognition. Therefore, Veed.io's translation stands out as the most reliable, offering pragmatic equivalence.

Example 3:
Arabic ST

والله زي القمر

Table 3.8: *Kamel El Adad*: Error Analysis Example 3			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	My God! Gorgeous!	You look just like Omar.	You're gorgeous!
Strategy	**Paraphrase Strategy**	**Resignation Strategy**	**Condensation Strategy**
FAR Model Scoring	**Functionality:** Score: 0.5	**Functionality:** Score: 2	**Functionality:** Score: 0
	Acceptability: Score: 0	**Acceptability:** Score: 0	**Acceptability:** Score: 0
	Readability: Score: 0	**Readability:** Score: 0	**Readability:** Score: 0
Total Score	**0**	**2**	**0**

In Episode 2, Ahmed is seen flirting with Laila, telling her she looks gorgeous. In Egyptian Arabic, he uses an idiomatic expression " والله زي القمر" that compares her beauty to that of the moon, a poetic and flirtatious way of complimenting someone's appearance.

Veed.io mistranslates the word "والله" which means "I swear" as "My God". This alters the tone from one of sworn sincerity to an exclamation of surprise, shifting the meaning and emotional nuance. While the use of "gorgeous" captures the general intent of the compliment, the sincerity conveyed by the oath is lost, and the tone leans more toward admiration or shock than honest flattery. FreeSubtitles.ai fails to translate the line by rendering "قمر" as "Omar", likely due to a mistranscription. This recognition error not only distorts the intended meaning but also results in a confusing and inaccurate translation. Maestra.ai, on the other hand, applies a condensation strategy by omitting the oath, the core message of admiration remains intact, preserving the intent of the original line more effectively than the other tools. Overall, Maestra.ai provides the most contextually faithful translation in terms of meaning, though none of the tools fully capture the cultural and emotional richness of the original idiom.

Example 4:
Arabic ST

يبني ماتبقاش طري كده! كله هيبقى كويس!

Table 3.9: *Kamel El Adad*: Error Analysis Example 4			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	Please don't be so dry, it's fine.	It's okay bro!	Don't dry that. It's okay!
Strategy	**Resignation Strategy**	**Paraphrase Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 2**	**Functionality:** **Score:** 1	**Functionality:** **Score: 2**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 0**
	Readability: **Score: 0**	**Readability:** **Score**: 0	**Readability:** **Score: 0**
Total Score	2	1	2

In Episode 2, Ahmed tries to comfort his friend Tarek by saying,
(يبني ماتبقاش طري كده! كله هيبقى كويس!)

"Grow a spine, man! everything's gonna be fine." The line is both teasing and reassuring, Ahmed mocks Tarek's fear in a light-hearted way while also offering comfort.

Veed.io completely misinterprets the word "طري" by translating it as "dry," which is the opposite of its intended meaning "soft" or "fragile." This mistranslation alters the entire tone and intent of the line. Maestra.ai makes the same error, rendering "طري" as "dry" as well. FreeSubtitles.ai, while omitting the sarcastic tone "don't be soft" or "toughen up" part, still conveys the overall intended meaning of reassurance. Although the translation is an under-translation, it manages to produce the most pragmatically appropriate version among the three tools. Overall, none of the tools successfully capture both sarcasm and comfort embedded in the original line. The reason for this mistranslation seems to be that the Arabic word "طري" pronounced "tary" is phonetically identical to the English word "dry" This phonetic overlap likely caused the tools to misrecognize the word during transcription, leading to the flawed output.

Example 5:
Arabic ST

انت أم الذوق كله يا ايه!

Table 3.10: *Kamel El Adad*: Error Analysis Example 5			
Tool	**1: Veed.io**	**2: FreeSubtitles.ai**	**3: Maestra.ai**
English MT	You're ever so polite.	You're the mother of my whole family...	Then you are or the market is very beautiful. Are you supposed to..
Strategy	**Transfer Strategy**	**Resignation Strategy**	**Resignation Strategy**
FAR Model Scoring	**Functionality:** **Score: 0**	**Functionality:** **Score:** 2	**Functionality:** **Score: 2**
	Acceptability: **Score: 0**	**Acceptability:** **Score:** 0	**Acceptability:** **Score: 1**
	Readability: **Score: 0**	**Readability:** **Score**: 0	**Readability:** **Score: 1**
Total Score	**0**	**2**	**4**

In Episode 2, Tarek's wife is eager to know who the woman accompanying Ahmed is. He tries to delay the conversation, so they don't leave their guests and calls her "أم الذوق كله" an Egyptian idiomatic expression that means "the queen of politeness." It's a way to tease her for her persistence.

Veed.io interprets the line correctly, providing an acceptable and idiomatic translation and not a word-by-word translation. In contrast, FreeSubtitles.ai and Maestra.ai completely fail to render the phrase meaningfully. Their translations neither capture the meaning nor the context, likely due to difficulty parsing the colloquial Egyptian expression. This results in a loss of both pragmatic meaning and cultural nuance. This example underscores the challenges of translating culturally embedded sarcasm.

Findings

This study evaluated subtitle outputs from FreeSubtitles.ai, Veed.io, and Maestra.ai using Mona Baker's Pragmatic Equivalence Approach, Jan Pedersen's FAR model, and Gottlieb's subtitling strategies. The analysis focused on how well these AI tools handled Egyptian Arabic's cultural, idiomatic, and pragmatic content in social comedy series.

Across all three tools, resignation emerged as the dominant strategy when processing dialect-rich, idiomatic, or culturally bound expressions. Rather than preserving pragmatic meaning, the tools frequently neutralized or omitted nuance, leading to the loss of tone, intertextuality, and cultural references. Implicature, irony, and indirectness—common in Egyptian Arabic—were often flattened or mistranslated. While Veed.io maintained better sentence structure, it still failed to convey nuanced terms. FreeSubtitles.ai showed inconsistent paraphrasing efforts. Maestra.ai relied heavily on simplification, often stripping content of its cultural depth.

The qualitative analysis showed frequent issues with Implicature loss, such as misrendered sarcasm and humor. Cultural mismatches, where literal translations lacked contextual meaning. Coherence disruption, where omitted or unrelated subtitles broke narrative flow. In terms of subtitling strategies, all tools leaned heavily on resignation. Examples included misinterpretation of metaphors (example: "أسد" rendered literally or as a proper name) and poor handling of colloquial expressions, leading to misleading or awkward translations.

The quantitative analysis, based on the FAR model, revealed that functionality errors dominated, averaging over 70% across tools. These included mistranslations, semantic inaccuracies, and context loss. Acceptability and readability errors were less frequent, with Veed.io showing the lowest acceptability error rate (12.5%) and Maestra.ai performing best in readability (10.75%).

Tool	Error Rate	Most Frequent Issue	Best Performance Area
Veed.io	23.7%	Functionality (70.83%)	Acceptability
FreeSubtitles.ai	39.5%	Functionality (70%)	Some cultural paraphrasing
Maestra.ai	36.7%	Functionality (73.1%)	Readability

The high frequency of functionality errors underscores AI's struggle with fast-paced, dialectal, and humorous content. The tools, largely trained on MSA or formal Arabic, lacked exposure to dialectal variation, leading to literal or irrelevant translations.

Discussion

The findings highlight the current limitations of AI subtitling tools in translating Egyptian Arabic into English, particularly in terms of pragmatic and cultural accuracy. While the tools handle basic syntax and literal meaning, they frequently fail to capture idiomatic expressions, humor, sarcasm, and culturally embedded nuances.

A key factor contributing to these shortcomings is the linguistic and pragmatic complexity of Egyptian Arabic, which differs significantly from Modern Standard Arabic (MSA). Most AI tools are trained on formal Arabic datasets, leaving them ill-equipped to handle dialectal features such as informal speech, cultural metaphors, and fast-paced dialogue. Consequently, automated speech recognition (ASR) components frequently mishear or misrender spoken dialect, disrupting the semantic flow and reducing subtitle quality.

By applying Baker's framework, the study identified specific types of meaning loss (example: implicature, tone, and idioms). Pedersen's FAR model quantified the frequency and severity of these losses, while Gottlieb's strategies revealed a strong tendency toward resignation. Together, these frameworks provided a comprehensive diagnostic lens, underscoring the need for tools that go beyond literal translation to preserve the communicative richness of the source text.

4. Conclusion

This study explored the effectiveness of AI subtitling tools in translating Egyptian Arabic (EA) into English, with a particular focus on pragmatic meaning, cultural nuance, and subtitle quality. By applying Mona Baker's Pragmatic Equivalence framework, Gottlieb's subtitling strategies, and Pedersen's FAR model, the research provided both qualitative and quantitative insights into the performance of three widely used AI subtitling tools: Veed.io, FreeSubtitles.ai, and Maestra.ai.

This study assessed the ability of AI subtitling tools to translate Egyptian Arabic (EA) into English, focusing on pragmatic meaning, cultural nuances, and subtitle quality. The results indicate that all three tools, Veed.io, FreeSubtitles.ai, and Maestra.ai, struggled with dialectal transcription, which emerged as the primary source of errors. The qualitative analysis showed that resignation was the dominant subtitling strategy, leading to frequent loss of idioms, humor, and culture-specific

references. Occasional paraphrasing and transfer strategies appeared but were inconsistent. Quantitatively, the FAR model revealed that functionality errors were the most frequent with an average rate of 71%, reflecting recurring semantic and cultural mismatches.

The findings highlight that AI subtitling tools are still far from replicating the nuance of human translation, especially when dealing with dialect-rich, culturally embedded content like Egyptian Arabic. Cultural specificity often fails to be translated accurately, which can impact viewer reception and lead to misinterpretation of meaning. The study reinforces the importance of dialect-specific corpora, improved ASR, and human post-editing as essential steps toward producing accurate and culturally aware subtitles.

The study was limited by a small dataset, focusing on one social comedy series, and only three AI tools were analyzed. There was also no audience reception testing, and given the rapid pace of AI development, newer models may already show improved performance. Future research should investigate a wider range of genres and AI tools, incorporate audience studies, and continue evaluating tools against frameworks such as Baker's Pragmatic Equivalence and Pedersen's FAR model. AI developers are encouraged to build large dialectal corpora and integrate cultural adaptation modules, while translators should leverage AI tools as drafting aids requiring post-editing rather than complete solutions.

References

Abdelgawad, N. (2022). The Netflix Translation of Humour from the Egyptian Vernacular

Dialect into MSA: the Case of "el-Limby" and "Illy Baly Balak" Films. Journal of Linguistics and Education Research, 5(1). https://doi.org/10.30564/jler.v5i1.4814

Alaa, A. M., & Sawi, I. A. (2023). The analysis and quality assessment of translation strategies in subtitling culturally specific references: Feathers. Heliyon, 9(8), e19095. https://doi.org/10.1016/j.heliyon.2023.e19095

Alnassan, A. (2023, January 9). Automatic standardization of Arabic dialects for machine translation. arXiv.org. https://arxiv.org/abs/2301.03447

Al-Sabbagh, R. (2023). The Negative Transfer Effect on the Neural Machine Translation of Egyptian Arabic Adjuncts into English: The Case of Google Translate. https://doi.org/10.33806/ijaes.v24i1.560

Alzamzami, F., & Saddik, A. E. (2023). OSN-MDAD: Machine Translation Dataset for Arabic Multi-Dialectal Conversations on Online Social Media
https://doi.org/10.48550/arXiv.2309.12137

Abdo, I. B. (2020). Preserving style in translating entertainment of Walt-Disney's tangled: English to Egyptian Arabic dialect. https://doi.org/10.25255/jss.2020.9.4.1559.1574

Baniata, L. H., Park, S., & Park, S. (2018). A neural machine translation model for Arabic dialects that utilizes multitask learning (MTL). Computational Intelligence and Neuroscience, 2018, 1–10. https://doi.org/10.1155/2018/7534712

Derouich, W., Kchaou, S., & Boujelbane, R. (2023). ANLP-RG at NADI 2023 shared task: Machine Translation of Arabic Dialects: A Comparative Study of Transformer Models. 10.18653/v1/2023.arabicnlp-1.75

El-Sayd, R. A. F. & Jan Pedersen. (2024). The assessment of subtitling the Arabic Movie Extracurricular to English on Shahid using the FAR model. In B. M. Mazid & B. A. Hassan, *Bulletin of the Faculty of Arts* (Vol. 73, Issue 2, pp. 129–132) [Journal-article].
https://bfa.journals.ekb.eg/article_379608_7f4be0c78da72fd774edf34718c61ccf.pdf

Faheem, M., Wassif, K., Bayomi, H., & Abdou, S. M. (2024). Improving neural machine translation for low resource languages through non-parallel corpora: a case study of Egyptian dialect to modern standard Arabic translation. Scientific Reports,14(1). https://doi.org/10.1038/s41598-023-51090-4

Gil, S. F. (2023). Quality assessment of the English subtitles in five international Award-Winning Colombian films*.
https://www.redalyc.org/journal/2550/255076378006/html/

Harrat, S., Meftouh, K., & Smaïli, K. (2019). Machine translation for Arabic dialects (survey). Information Processing & Management, 56(2),262273.https://doi.org/10.1016/j.ipm.2017.08.003

Jeblee, S., Feely, W., Bouamor, H., Lavie, A., Habash, N., & Oflazer, K. (2014). Domain and Dialect Adaptation for Machine Translation into Egyptian Arabic. DOI: 10.3115/v1/W14-3627

Kadaoui, K., Magdy, S., Waheed, A., Khondaker, M. T. I., El-Shangiti, A., Nagoudi, E. M. B., & Abdul-Mageed, M. (2023). TARJAMAT: Evaluation of BARD and ChatGPT on machine translation of ten Arabic varieties. https://aclanthology.org/2023.arabicnlp-1.6/

Khered, A. S., Abdelhalim, I. Y. H. A., Abdelhalim, N., Soliman, A., & Batista-Navarro, R. (2023). UNIMANC at NADI 2023 Shared task: A comparison of various T5-based models for translating Arabic dialectical text to modern standard Arabic. https://aclanthology.org/2023.arabicnlp-1.71/

Musleh, M. M. T. (2020). Translatability of Metaphors in the Dubbing of Animation Songs from English into the Egyptian Dialect. https://doi.org/10.34257/GJHSSGVOL20IS12PG1

Meftouh, K., Harrat, S., Jamoussi, S., & Smaïli, K. (2015). Machine Translation Experiments on PADIC: A Parallel Arabic DIalect Corpus. ResearchGate. https://doi.org/10.1016/j.ipm.2017.08.003

Moukafih, Y., Sbihi, N., Ghogho, M., & Smaïli, K. (2021). Improving machine translation of Arabic dialects through multi-task learning. https://www.semanticscholar.org/paper/Improving-MachineTranslationofArabicDialectsMoukafihSbihi/51cbd885d31a785c5720e0150ae5cf9cf29afa7a

Pedersen, J. (2017). The FAR model: assessing quality in interlingual subtitling. The Journal of Specialised Translation, 28, 210–212. https://jostrans.soap2.ch/issue28/art_pedersen.pdf

Sabtan, Y., Hussein, M., Ethelb, H., & Omar, A.,2021). An evaluation of the accuracy of the machine translation systems of social media language. *Journal of Advanced Computer Science and Applications*, *12*(7).https://www.semanticscholar.org/paper/An-Evaluation-of-the-Accuracy-of-the-Machine-of-Sabtan-Hussein/3303a9041eead6b4356bd99160f4a5faa25dfc46

Sadiq, S. (2024). Subtitling against the current: Egyptians' perceptions of English sitcoms subtitled in Standard Arabic and Egyptian Colloquial Arabic. https://doi.org/10.12807/ti.116201.2024.a06

Salloum, W. (2011, January 1). Dialectal to standard Arabic paraphrasing to improve Arabic-English statistical machine translation. https://aclanthology.org/W11-2602/

Sajjad, H. (2014, January 1). Unsupervised word segmentation improves dialectal Arabic to English machine translation. https://doi.org/10.26034/cm.jostrans.2017.239

Sakre, M., Kouta, M. M., & Allam, A. M. (2009). Automated construction of Arabic-English parallel corpus. Journal of the ACS Advances in Computer Science (Print), 3(1), 57–69. https://doi.org/10.21608/asc.2009.158223

Sharaf, A. I., Souliman, M., Gao, I., & Bauer, J. (2023). Unsupervised question answering using custom NLP library built for Egyptian Arabic. https://www.mdpi.com/2624-6511/8/2/62

Shatnawi, M., Yasin, M. B., & Huq, A. A. (2023). Building a framework for identifying Arabic dialects using deep learning techniques. DOI: https://doi.org/10.1145/3630632

Tawfik, A. Y., Emam, M., Essam, K., Nabil, R., & Hassan, H. (2019). Morphology-Aware Word-Segmentation in dialectal Arabic adaptation of neural machine translation. DOI: 10.18653/v1/W19-4602

Tohamy, M. (2021). Cohesion in human and machine translation. doi: 10.21608/opde.2021.226278