BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

YouTube's Automated Subtitling from English into Arabic: A Case Study of Harry Potter and the Prisoner of Azkaban

Ahmad Mohammad Al-Harahsheh * **, *Ro'a Hasan Rababah***

Translation Department, Faculty of Arts, Yarmouk University, Irbid , Jordan

ABSTRACT

Recently, the development of speech-to-text technology, together with machine translation, has led to the development of simultaneously translating the captions of videos into other languages. YouTube, a video-sharing platform, offers multilingual subtitles using this feature. The current automated caption system captures audio data during video uploads and generates a subtitle file in text format. The current study aims at examining whether YouTube machine translation from English into Arabic is reliable in rendering the intended meaning on subtitling, depending on the FAR model (functional equivalence, readability, and acceptability). The data of this study consisted of 30 examples that were taken from the YouTube platform and their translated versions into Arabic using YouTube's machine translation. The study is both descriptive and comparative. The results of the study indicate that YouTube machine translation represents varying levels of inadequate translation according to its system and database, revealing many deficiencies. The total approval rate is 68.5%, which gives the impression that the translation is very poor. Therefore, the machine requires the development of its system and the enrichment of its databases, specifically the Arabic ones.

Keywords: Machine Translation; Subtitling; Far Model; YouTube; Harry Potter

*CORRESPONDING AUTHOR:

Ahmad Mohammad Al-Harahsheh, Translation Department, Faculty of Arts, Yarmouk University, Irbid, Jordan; Email: harahsheh77@yu.edu.jo

ARTICLE INFO

Received: 1 December 2024 | Revised: 10 January 2025 | Accepted: 13 January 2025 | Published Online: 18 February 2025
DOI: https://doi.org/10.30564/fls.v7i2.8163

CITATION

Al-Harahsheh, A.M., Rababah, R.H., 2025. YouTube's Automated Subtitling from English into Arabic: A Case Study of Harry Potter and the Prisoner of Azkaban. Forum for Linguistic Studies. 7(2): 583–602. DOI: https://doi.org/10.30564/fls.v7i2.8163

1. Introduction

The field of Machine Translation (MT) explores several approaches for the translation of written or spoken content from one human language to another. It is a subdomain within the discipline of computational linguistics, which incorporates concepts and principles from several disciplines, including linguistics, information theory, artificial intelligence, computer science, and statistics. According to Sepesy and Donaj [1], historically, it was subject to negative perceptions due to its apparent lack of quality. Significant advancements in machine translation quality have been seen, particularly over the last two decades. These advancements have generated considerable interest in using MT within the translation sector. Although the quality of machine translation remains inferior to that of human translation, it should not be dismissed as lacking practical use. Throughout history, the translation market has been mostly dominated by translation firms and expert translators. However, in recent times, there has been a notable surge in the availability and practicality of machine translation solutions, leading to a significant expansion of options in the market. The translation market is now facing growing challenges in terms of pricing, volume, and turnaround time. The advent of commercial applications for MT is a positive development in the field of translation methodologies. In professional or formal contexts, the use of human translation is essential since human involvement is crucial for ensuring grammatical accuracy and preserving the intended meaning of the original text [1]. Machine translation systems are used by a substantial number of people on a daily basis. Specialized machine translation systems have been widely utilized and have shown effective performance in certain domains or for specific companies. Currently, MT has further extended its application into other sectors, including technical and audiovisual domains, among others.

Audiovisual Translation (AVT) often involves the translation of the spoken elements included in a video. The primary distinguishing characteristic of this phenomenon is the coordination of verbal and nonverbal elements. When working with an audiovisual output, translators are not just concerned with textual elements but also deal with the multifaceted facets of media art. Consequently, the creators of the video use many elements, such as conversations, remarks, sound effects, images, and ambiance, to enhance its overall impact. The concept of audiovisual translation encompasses more than just grammatical equivalency between two languages. It also involves establishing a suitable connection between the verbal and nonverbal components included in both the original work and its translation [2].

In light of contemporary advancements, scholars have devised two primary procedures for AVT that facilitate effective comprehension of the conveyed material by the audience. Also, Matkivska [2] points out in her research that there exists a minimum of 10 distinct forms of audiovisual translation, with subtitling and dubbing being identified as the two most significant methodologies.

The current study focuses on the use of machine translation for subtitling purposes on the YouTube platform. The analysis focuses on identifying stylistic issues and differences in lexical choices between the original text in the source language and the machine-generated translation. This study also aims to analyze the readability and acceptability of machine translations in Arabic, in terms of their ability to accurately represent the intended meaning of the source text.

The current study aims to investigate the manner in which social media machine translation, particularly on the platform of YouTube, effectively preserves the intended meaning within the domain of subtitling. Furthermore, it is important to assess the acceptability and readability of the content for the intended target audience. Additionally, the objective of this study is to emphasize the disparities between the original text in the source language and the translated text generated by machine translation systems. Therefore, the study attempts to answer the following questions:

(1) To what extent can machine translation effectively convey the adequate meaning in subtitling based on the terminology choices?
(2) To what extent can machine translation be a viable solution for subtitling in the future?

The significance of the study arises from the limited number of research endeavors that have explored YouTube automated subtitling from English into Arabic. Moreover, people throughout the globe engage in the habitual practice of consuming content on the popular online platform, YouTube, many times throughout the day. This research aims to assess the acceptability and readability of machine translations on YouTube by conducting a comparative analysis between the original source language text and the translated output generated by the machine translation system. Therefore, the

present study may have the potential to assist researchers engaged in the use of artificial intelligence within the domain of translation.

2. Literature Review

The translation industry is influenced by the integration of Artificial Intelligence (AI), leading to the development of various software applications, databases, corpora, and machine translation systems. However, the quality of translation or subtitle provided by machine translation and AI is still doubtful because this translation or subtitle has some linguistic, and cultural mistakes. Yao [3] investigates the quality of the automated subtitles created by the NetEaseSight platform. The study used a selection of the top 20 TED speeches that have been posted on the NetEaseSight platform in order to develop machine-translated subtitles. He investigates the accuracy of voice recognition and cut scores, using the FAR approach to assess the efficacy of machine-generated Chinese subtitles. The findings indicate that there is a need to enhance the accuracy of machine translation engines. Certain terms occasionally do not appear in standard dictionaries but are used to name things such as people, locations, establishments, and registered brands, as well as denoting temporal references, numerical values, and new words.

In addition, it is worth noting that although the translation exhibits a certain level of coherence, the overall readability remains a significant concern. During the training phase of the machine translation engine, low-frequency and uncommon phrases are removed in order to reduce the complexity of the module and save storage space. There is also a need to enhance the accuracy of word translations. Certain terms occasionally do not appear in standard dictionaries but are used to name things such as people, locations, establishments, and registered brands, as well as denoting temporal references, numerical values, and new words. In addition, it is worth noting that although the translation exhibits a certain level of coherence, the overall readability remains a significant concern [3].

Besides, there is a need for enhancing the accuracy of speech recognition systems and optimizing the segmentation process. Despite significant advancements in the technological maturity of speech recognition, achieving one hundred percent accuracy remains unachievable. Speech interaction is impacted by several factors, including background noise and speech pace, leading to significant variations in recognition rates across different scenarios. One limitation of speech recognition is its inability to modify text based on contextual cues. Moreover, inadequate semantic understanding is a major obstacle in this field. Yao [3] suggests that in order to address this issue, it is essential to enhance the algorithm and acquire a substantial volume of dependable data for algorithmic training. These measures are necessary to facilitate the algorithm's progress toward a certain degree of complexity. Nevertheless, it is important to recognize that machine-translated subtitles, although not yet meeting the necessary standards for direct market use, do exhibit a certain level of accuracy, devoid of grammatical errors or omissions. Furthermore, the automated generation of timetables does not only save time for subtitlers but also improves overall efficiency.

In their research, Hagström and Pedersen [4] undertook a diachronic analysis of subtitles, examining the changes that occurred both before and subsequent to the integration of machine translation into the translation process. They conduct a comparative analysis of a corpus of Swedish subtitles derived from Anglophone TV programs made after the implementation of machine translation and a corpus of subtitles from the pre-machine translation era. This study aimed to examine if there were differences in the quality of subtitles generated in the 2020s compared to those produced in the 2010s. They adapted the FAR approach, which encompasses an analysis of three distinct dimensions of quality from the viewers' perspectives, namely functional equivalence, acceptability, and readability. The findings indicated that the post-edited subtitles generated in the 2020s had certain characteristics when evaluated based on established standards and the FAR model. Specifically, these subtitles were observed to be faster, less cohesive, more oral, and less complete with less meticulous punctuation, and line breaks compared to the subtitles created in the 2010s. The items examined exhibited notably lower quality across all assessed areas.

In the same line, Karakanta [5] concentrates on automated and PE-based assessments of automatic subtitling. Initially, she evaluated automatic subtitling in terms of technological advancements, assessment methods, and empirical studies. Secondly, she emphasized existing shortcomings and aspects that require more attention to fully understand and enhance automation in subtitling through the application

of effective approaches utilizing advancements in both MT and AVT. She analyzed publications that provided at least one form of experimental design, using automated and/or human evaluation. She also noted that, while there had been studies undertaking experimental research on automated processes for interlingual subtitling, the transition from source to target language in interlingual subtitling introduces a further level of complexity to the approach and evaluation. Therefore, she only examined works utilizing MT/ST in interlingual subtitling, including both automated and human assessments. Furthermore, she addressed the emerging paradigm of automatic subtitling, which presents additional obstacles manifested in a variable number of segments, necessitating auto-spotting and segmentation, as well as the disentanglement of variables.

Karakanta comes up with a set of suggestions that should help keep experimental designs for studying automatic subtitling from running into problems. The main recommendations were that research in automatic subtitling should encompass all aspects of subtitling, enhanced interfaces, adherence to reporting standards, provision of test data and benchmarks, and assessment should be independent of generation. In conclusion, Karakanta [5] asserts that her selection does not diminish the necessity for perception studies, which will enhance the understanding of experimental research and gain significance as technological quality advances. Consequently, standardization and harmonization are deemed essential for the prosperous future of the AVT and MT industries.

In Varga's [6] study, the primary focus was on the fundamental framework of automatic subtitle systems. The study primarily centers on the examination of nine online subtitling platforms, with a particular emphasis on the analysis of their features. It is worth noting that out of the nine services, only five provide free automated subtitles, each varying in terms of their quality. The same video clip was used to evaluate these internet platforms, and their results were examined using both quantitative and qualitative analyses to emphasize the main characteristics of each site. To comprehensively evaluate their competencies, the researcher chose the opening sequence of Quentin Tarantino's film Reservoir Dogs as the chosen video clip.

The empirical data highlights many types of errors, including missing text, text coherence, speaker recognition, text layout, spelling problems, and punctuation issues. These categories of errors provide a thorough understanding of the current capabilities of automated transcription technology. The study suggests that these systems lack autonomy and rely on expert intervention to achieve optimal transcription quality. The study findings indicate that the online apps for MT need adequate training and calibration. The text segmenters face challenges related to spatial and temporal constraints that are unique to the field of subtitling.

Matusov et al. [7] provide a comprehensive description of the process by which a state-of-the-art Neural Machine Translation (NMT) system may be successfully tailored for the purpose of subtitling. They put forward a straightforward approach to include inter-sentence context in the translation of brief utterances and dialog turns. They also modified the NMT system to accommodate linguistic diversity, namely Latin American Spanish, as well as subtitling style and domain. They present a unique approach for the segmentation of subtitles that integrates a recurrent neural network model with both hard and soft restrictions on subtitle length and duration inside a beam search framework. A comprehensive assessment, both automated and human-based, was conducted to assess the quality of the modified machine translation output when segmented into subtitles using the suggested method. The results of this evaluation show significant improvements compared to the baseline MT system output, which used line breaks based on heuristics. The implementation of this quality enhancement resulted in significant improvements in productivity and time efficiency when the modified machine translation output was post-edited by impartial professional translators. These improvements were seen in comparison to the processes of translating from scratch and post-editing the translations generated by the original MT system.

Song et al. [8] put up an innovative method for sentence segmentation that involves the use of deep neural networks to automatically create period marks. The primary objective of this technique is to enhance the precision of the automated translation of YouTube subtitles. The study introduces a new method for phrase segmentation that utilizes neural networks and YouTube scripts, and is less dependent on word order and sentence structure. The performance of this strategy was measured. They constructed the input in a manner that closely resembles YouTube scripts and tried to identify punc-

tuation marks only based on textual characteristics. For this investigation, they used a total of 27,826 subtitles extracted from the online courses offered by Stanford University. They use Long-Short Term Memory (LSTM) of Recurrent Neural Network (RNN), a very effective technique in the field of natural language processing, to construct a model using available data. This model is then utilized to make predictions about the placement of punctuation marks. The LSTM model has shown promise for its applicability in the restoration of punctuation in voice transcripts. This approach involves the integration of textual elements and the length of pauses. Despite the fact that RNNs have shown commendable performance over a range of input durations, they have compromised some of these advantages by aligning the data length to that of YouTube subtitles. An attempt was made to forecast the occurrence of periods between consecutive words. The experiment included measuring the accuracy of the approach, which was found to be 70.84%.

In their study on automatic translations, Gupta et al.[9] identify and provide explanations for the challenges encountered. The researchers categorize each difficulty into three distinct categories. First, "the problems directly related to textual translation". Secondly, "problems related to subtitle creation guidelines". Lastly, "problems due to adaptability of MT engines"[9]. The researchers determine the frequency occurrence of 16 significant issues in the automatic translation of subtitles from English to six specific target languages, namely German, Chinese (simplified), French, Castilian Spanish, Arabic, and Brazilian Portuguese. The experiment was conducted using a dataset consisting of 56 movie subtitle files, with a cumulative count of 17,977 subtitle blocks. The English subtitles were produced by humans, while the target subtitles were created using a machine translation system that was trained using a specific methodology.

The findings indicate that the researchers have seen the presence of certain difficulties across a majority of languages. The primary issue in all languages, with the exception of Chinese (simplified), is the high level of paraphrasing errors. However, there are some issues that are peculiar to certain languages and hence need specialist solutions. In comparison to other languages, German translation has a greater prevalence of issues concerning structure errors and word order errors. The occurrence of non-text character translation is often seen in the context of Chinese and Arabic translations. The issue of word structure errors ranked as the second most significant challenge in the French language. The issue of lexical translation posed a considerable challenge for the languages of German, Spanish, and Arabic.

Hiraoka[10] conducted a study on the effective pre-editing rules for subtitling TED Talks using neural machine translation. The study seeks to formulate and evaluate a set of straightforward, efficient pre-editing rules for audiovisual materials, including TED Talk subtitling, to translate Japanese source text into English, utilizing an NNT engine created by the National Institute of Information and Communications Technology (NICT) in Japan.

Pre-editing is classified into two methods: bilingual pre-editing and monolingual pre-editing. Bilingual pre-editing enables the pre-editor to modify the source text while referencing the MT outcome, in contrast to monolingual pre-editing, which does not permit this. Thus, monolingual pre-editing needs no proficiency in the target language. This study concentrates on monolingual pre-editing, since Hiraoka aims to empower content producers or those with limited proficiency in the target language to pre-edit source texts in their native language for content dissemination. It focuses on monolingual pre-editing, since Hiraoka seeks to enable content creators or individuals with limited proficiency in the target language to pre-edit the source text in their SL for content dissemination. The efficacy of the pre-editing rules was assessed based on the enhancement of MT output quality, considering the 21-character-per-second (CPS) constraint. Given that the translation aim is TED subtitling, it is essential to consider character limitations.

The assessment results indicated that, in comparison to the MT output of the raw source text, the MT output of the pre-edited source text showed a quality enhancement in the average scores of both human evaluation and BLEU. The overall percentage of subtitle segments that contributed to a score gain is 41%. Despite the observed score declines in the pre-edited MT, the majority of parts remained over the 'Acceptable' level on the human evaluation scale. Besides translation quality, the study has also investigated the character limitations of subtitling and confirmed that the instances of segments in both raw MT and pre-edited MT outputs that violate the 21-character per segment guideline established by TED were almost nonexistent. Therefore, it is determined that pre-editing according to the prior rules does not hinder

compliance with the 21-CPS requirement.

Athanasiadi[11] investigates the potential of MT and other linguistic assisting technologies in subtitling. This study looks at the lack of commercial subtitling software that includes linguistic assistive tools. The goals are to find out what programs are already on the market, what their limitations are, and whether customers want these tools to be added. Quantitative research was done through an online questionnaire using Google Forms, incorporating both structured (multiple choice), and unstructured (open-ended) questions to get robust results.

The study developed a model of a fully automated MT engine for subtitling, intending to illustrate the optimal functioning of such a system. This concept is predicated on an SMT engine rather than a rule-based or hybrid engine. The model segments the engine's processing into three phases. The initial phase entails preparing the corpus and integrating it into the system to build the engine. The second step, ST editing, has a voice recognition component together with text condensation and segmentation elements. The third process, referred to as ST refinement, entails the automatic modification of the script to reduce post-editing effort. The voice recognition system automatically initiates the translation of previously detected and transcribed subtitles upon completion of all steps. The translated screenplay incorporates the timecodes from the transcribed subtitles, providing the subtitle with two .srt files and one.txt file. One .srt file has timecoded SL subtitles, whereas the second .srt file contains timecoded TL subtitles. The .txt file contains the SL script devoid of timecodes, serving as a reference for post-editing.

The previous MT model was developed as a fully automated machine translation system to emphasize the advantages that a machine translation engine for subtitling offers to subtitlers, particularly regarding time efficiency. Nonetheless, the efficacy of such an engine can only be assessed through deployment. The questionnaire findings indicated a strong preference for TM components in subtitling software over all other alternatives. This indicates that TM tools are desired by subtitlers, and may be seen as a significant oversight in the evolution of subtitling software. The questionnaire's findings indicated that TBs are the respondents' second preference for integration into a subtitling system with a TM component. The primary conclusion is that traditional subtitling software is gradually evolving into online, accessible, and adaptable applications, ushering in a new era of subtitling.

3. Methods and Procedures

3.1. Data Collection

The study focuses on the comparison of the source text and the machine translation output for subtitling on YouTube platform. The data of the study were selected from *Harry Potter and the Prisoner of Azkaban*, it consisted of 30 examples that were taken from the YouTube platform and their translated versions into Arabic using YouTube's machine translation. These examples were purposefully used because they contained certain errors that affect the fluency of subtitling and therefore, affect the process of understanding. The Harry Potter series comprises a collection of seven fantasy books written by the renowned British novelist J. K. Rowling. The literary works document the experiences of a young wizard named Harry Potter, together with his friends Hermione Granger and Ron Weasley, who are enrolled as students at the esteemed school known as Hogwarts School of Witchcraft and Wizardry. The primary narrative trajectory centers on the protagonist Harry's confrontation with Lord Voldemort, a malevolent wizard who seeks immortality, aims to topple the ruling institution of wizards called the Ministry of Magic, and seeks to dominate both wizards and Muggles (people without magical abilities).

3.2. Data Analysis

The data were analyzed based on the FAR model (Functional Equivalence, Acceptability, and Readability) suggested by Pedersen[12]. Moreover, the data were analyzed to explore if YouTube machine translation is reliable in rendering the intended meaning depending on film subtitling. The concept of quality in translation is a multifaceted issue, and it becomes much more complex when applied to subtitling. The assessment of subtitling quality is often evaluated based on internal criteria[12]. The proposed model serves as a comprehensive framework for evaluating the overall quality of pre-existing interlingual subtitles. Its applicability has been seen in the assessment of quality in both fansubs and professional subtitles. Furthermore, it has been seamlessly incorporated into the quality assessment process of the Trados subtitling unit.

The FAR model incorporates the correlation between interlingual subtitles and the ultimate consumer (the viewer). The fundamental unit of evaluation under the FAR model is the subtitle itself. The subtitle is used as the fundamental unit of evaluation, for instance, the word, phrase, or minute of airtime. This is due to the fact that subtitling entails the linguistic compression of information, and the level of conversation intensity may significantly range across various shows. The FAR approach examines three categories of quality as perceived by viewers: functional equivalence, acceptability, and readability. The concept of functional equivalence centers on the communication or significance conveyed in the original text and the extent to which it has been accurately conveyed in the target language. The concept of acceptability involves evaluating whether the linguistic standards of the specific language being targeted have been followed. Finally, readability refers to the capability of the audience to effectively read the subtitles and comprehend the conveyed material.

A suggested penalty point system is included, along with methods for identifying faults and categorizing their severity as intersubjectively as feasible for each of the FAR categories. This gives the users the ability to evaluate each subtitled text from these three angles. The utilization of the penalty point system facilitates the identification of problematic areas in a subtitle's text. Consequently, it may be employed to offer subtitlers constructive feedback, which could be beneficial in an educational setting. The mistake classifications and scores are imported from the NER model, which are "minor," "standard," and "serious".

The rationale behind using FAR model is that it is functional and easy to apply. In addition, it allows the assessor of subtitling to easily recognize the mistakes in translation such as providing the functional or equivalent term of the SL in the TL, and to decide whether this term is readable and acceptable for the target readers or not. Therefore, it is easy to spot and evaluate the error on screen and to suggest an alternative translation.

4. Findings and Discussion

4.1. Quantitative Analysis

Thirty examples were analyzed to investigate the machine subtitling translation in three categories: functional equivalence, acceptability, and readability. We should mention that the researcher analyzed the scenes of the movie from the videos that are available on the YouTube platform, which are estimated to be an hour and a half from the movie.

Table 1 shows that the highest approval rate was in functional equivalence with 111.8%, which means that the machine has a serious problem at the semantics level; the score of serious errors is considered very high, and the viewer certainly won't understand the scenes, and the standard errors are also high, which confuse the viewer's comprehension. Nevertheless, the high percent of untranslated utterances, which is estimated at 34.3% of standard errors, then the acceptability with a 34% approval rate, and that seems like a bad percentage, means that the YMT has an issue with Arabic norms and made the text foreign to the Arab people. After that, the readability had a 30.9% approval rate. The total approval rate is 68.5%, which gives the impression that the translation is very poor. The study used the FAR model, and the researcher looked at how good the examples were by giving each one a penalty point value: minor: 0.25 points, standard: 0.5 points, serious: 1 point for all categories except semantic errors, which are: minor: 0.5 points, standard: 1, serious: 2, and that according to Pedersen[12]. The findings reveal that the YMT was not successful in most examples.

Table 1. Number of Errors, Error Scores, and Approval Rates of the data.

	Number of Errors	Error Score
Functional equivalence	201	224.75
Semantics errors	196	222.5
- Serious errors	40	80
- Standard errors	129	129
- Minor errors	27	13.5
Stylistic errors	5	2.25
- Standard errors	4	2
- Minor errors	1	0.25
Approval rate: 111.8%		
Acceptability	72	24.5
Grammar errors	67	22.75
- Standard errors	24	12
- Minor errors	43	10.75
Spelling errors	3	0.75
- Minor errors	3	0.75
Idiomaticity errors	2	1
- Standard errors	2	1
Approval rate: 34%		
Readability	166	51.25
Segmentation errors	29	12.75
- Serious errors	4	4
- Standard errors	10	5
- Minor errors	15	3.75

Table 1. *Cont.*

	Number of Errors	Error Score
Punctuation & graphics	137	38.5
- Serious errors	1	1
- Standard errors	14	7
- Minor errors	122	30.5
Approval rate: 30.9%		
Total:	439	300.5
Total approval rate:	68.5%	

4.2. Qualitative Analysis

The study explores the subtitling techniques of *Harry Potter and the Prisoner of Azkaban* used by the YouTube machine translation. These examples were analyzed according to the FAR model (functional equivalence, acceptability, and readability). The examples were classified and analyzed depending on the type of error.

4.2.1. Functional Equivalence

According to Pedersen [12], functional equivalence, also known as dynamic equivalence or meaning-based translation, is a translation technique whereby the translator attempts to convey the intended meaning and thought of the reader in the source language, rather than focusing only on the literal words and structures used. Ideally, a subtitle should effectively communicate both the explicit content and the underlying intention. If there is a failure to accurately convey both the literal content and the intended meaning, it would result in a clear mistake. If the intended meaning is accurately delivered without any additional information, this should not be considered a mistake. Instead, it is a common technique in subtitling and may be preferable to providing a word-for-word translation. If just the literal words said or written are considered without taking into account the intended meaning behind them, this would also be considered a mistake since it might lead to misinterpretation or confusion. However, since YMT is a speech-to-text engine, we will focus on whether it conveys the literal meaning in a way that the viewers will get the idea. There are two types of equivalence errors, namely semantic and stylistic.

(1) Semantic Errors

In consideration of the significance of semantic equivalence in interlingual subtitling and the users' presumed reduced error tolerance, the penalty points for semantic equivalence are as follows: minor: 0.5, standard: 1, and serious: 2.

Serious Errors

A serious semantic equivalence error refers to a subtitle that contains such substantial inaccuracies. It makes the viewers' comprehension of the subtitle completely ineffective. This error not only hinders the viewers' understanding of subsequent subtitles but also has the potential to cause misunderstandings in the plot or disturb the overall illusionary experience for more than a single subtitle.

Example (1)

ST: Give me the cup. Oh, my dear boy. My dear...you have the Grim.

TT: أوه أوه أنا هل تجرؤ على الحصول على خبز كريمي،

(**Back translation:** Oh, Oh, I, do you dare to get a creamy bread,)

In example (1), the dialogue was between Professor Trelawney and Harry Potter in the class of divination. The students learned ***tasseomancy,*** the art of reading tea leaves, to have a sight of the future. Therefore, the professor takes Harry's cup to tell him what the sight in it. She is shocked about what he has, and with all sadness, she tells him that he has the *"Grim,"* which is a form of giant spectral dog. It's among the darkest omens in the world, and it's an omen of death. In YMT, it translates the word *"Grim"* as "الكريمي الخبز"(creamy bread) which is far cry from the intended meaning in the ST(See **Figure 1**). Thus, the absence of a connection in the translation causes an inadequate understanding of the meaning by the viewers. Furthermore, the machine omitted the first sentence when the professor asked Harry to give her the cup. Also, it adds the question tool "هل" (do), and translates the words "My dear boy. My dear" into "تجرؤ" (dare) which is a mistake that transfers very different information. This error is serious since it hampers comprehension of both the individual sentence and the broader context. The occurrence of this error might perhaps be attributed to the fact that the machine did not comprehend the British accent. **Suggested translation**: طفلي العزيز! عزيزي، لديك الغريم. أعطني الكأس. أوه! يا (**BT**: Give me the cup. Oh! My dear child! My dear, you have a Grim.).

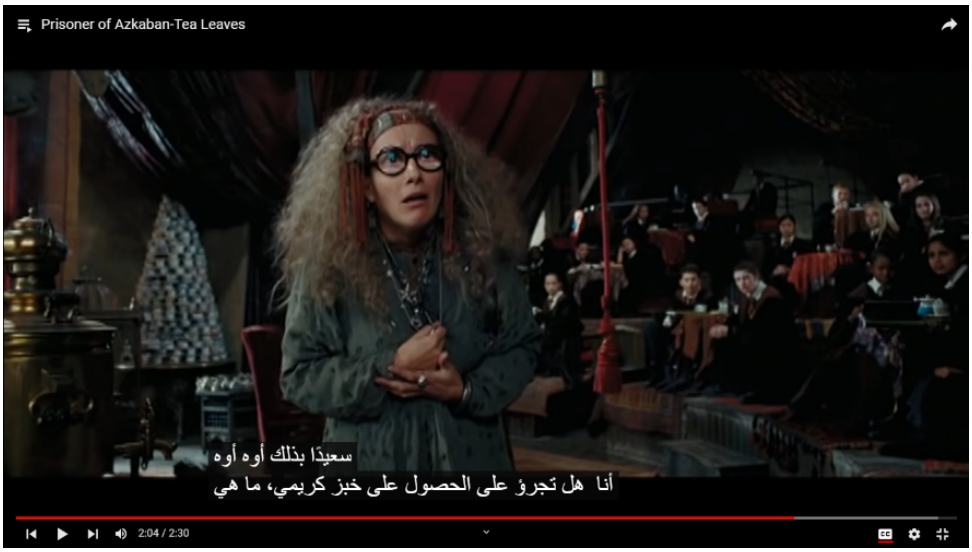

Figure 1. An example on serious error.

Standard Errors

A standard semantic equivalence error may be defined as a subtitle that includes errors but remains relevant to the intended meaning and does not significantly hamper the comprehension of the viewers beyond that particular subtitle.

Example (2)

ST: Who is that? Who is? That is **Sirius Black**, that is. Don't tell me you've never been **hearing of Sirius Black**. He's a murderer. Got himself locked up in Azkaban for it.

TT: ذلك الرجل الذي هو ذلك **الاسود الجاد** الذي لا تخبرني أنك لم تكن هنا أبدًا **بظهر أسود خطير**، إنه قاتل حصل على نفسه تم حبسه في أسكابان بسبب ذلك،

(**BT:** This man who is the serious black that you didn't tell me that you were here at all in a black dangerous back, he is a killer got himself, he was prisoned in Azbakan because of this)

In example (2), Harry sees a picture of a man in the newspaper, so he asks Stan Shunpike, who holds the newspaper, about him. Stan, in a surprising way because Harry doesn't know him, tells him that his name is Sirius Black, a murderer who was locked up in Azkaban prison. In this scene, YMT uses word-for-word translation. It translates "Sirius Black" in two terms: " الاسود الجاد " (serious black) and " أسود خطير " (dangerous black) and both are clearly wrong that it may cause misunderstanding. Moreover, the machine made a mistake in translating the word "hearing" into " بظهر ," (back) which made the sentence have a different meaning from that in the ST(See **Figure 2**). Furthermore, the translation doesn't use the punctuation correctly, especially question marks, which will improve the quality of the translation and make it more acceptable. This may be considered a standard error since it changes the sentence's semantic meaning without changing the viewer's comprehension of the preceding information. Suggested translation: من هو؟ هذا سيريوس بلاك. لا تخبرني أنك لم تسمع عن سيريوس بلاك! أنه قاتل، تم حبسه في أزكابان بسبب من هذا؟ ذلك.

(**BT:** Who is this? Who is he? This is Sirius Black. Don't tell me you haven't heard of Sirius Black! He's a murderer, he was prisoned in Azkaban because of it.)

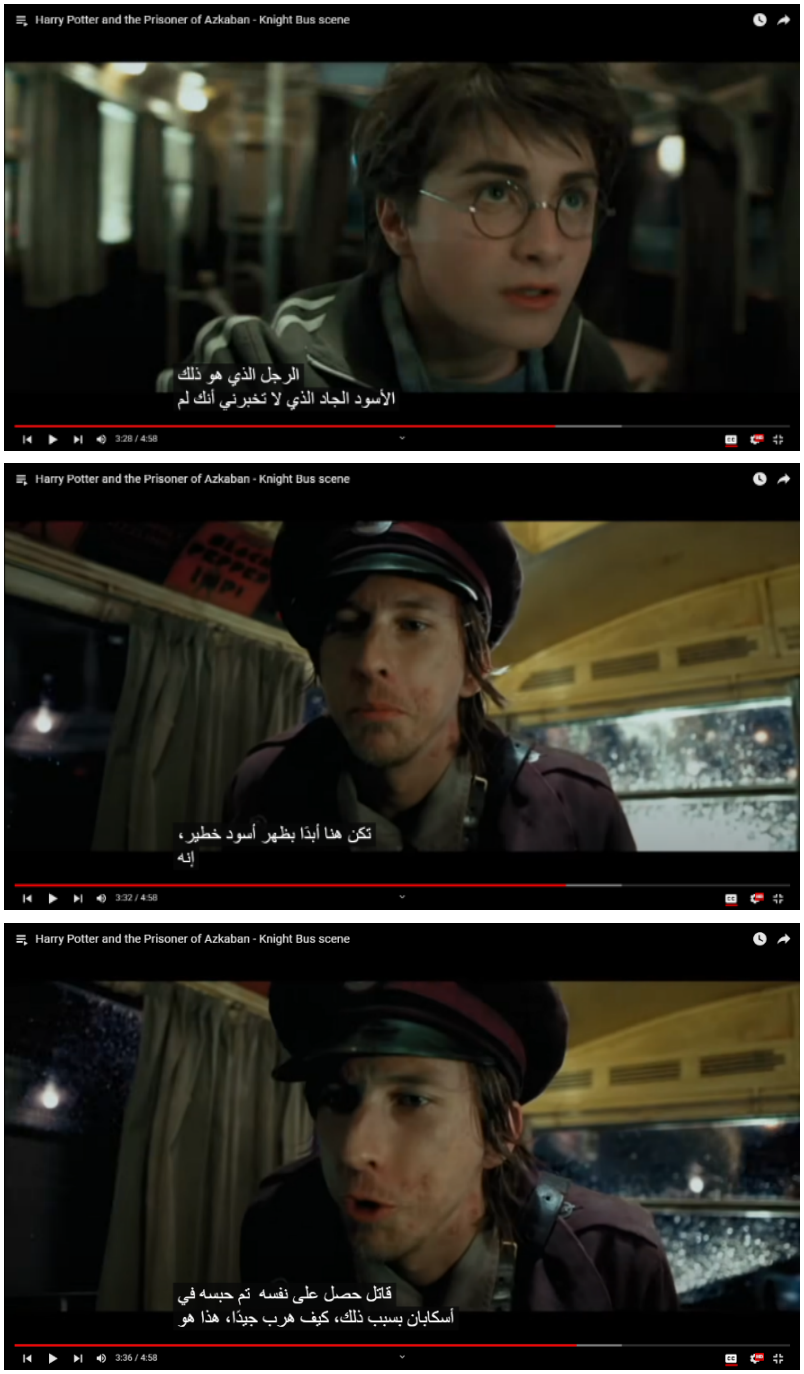

Figure 2. An example of standard error.

(2) Untranslated

According to Pedersen[12], when important statements within the narrative are left uninterpreted, the examples will be classified as standard semantic errors.

Example (3)

ST: A cat? Is that what they told you? Looks like a pig with hair.

TT: أخبرتك حقًا أنك تبدو مثل خنزير بشعر إذا سألتني،

(**BT:** I told you that you look like a hairy pig if you ask me."

In example (3), Ron is talking with Hermione about their pets. He hates her cat because it is always chasing his rat. YMT didn't translate the word "a cat?" the main word in the dialogue and translates the whole other words incorrectly, which means that the viewers will conclude that Ron is directing all this to Hermione, so the whole translation is inaccurate(See **Figure 3**). Far from that, the machine didn't use the correct punctuation, and the untranslated word caused the wrong segmentation. The standard error in untranslated word also prompted a serious error in translating the rest of the sentence. Suggested translation: قطة؟ هل هذا ما أخبروكِ به؟ تبدو مثل خنزير مغطى بالشعر. (**BT**: A cat? Is this what they told you? It looks like a pig covered in hair).

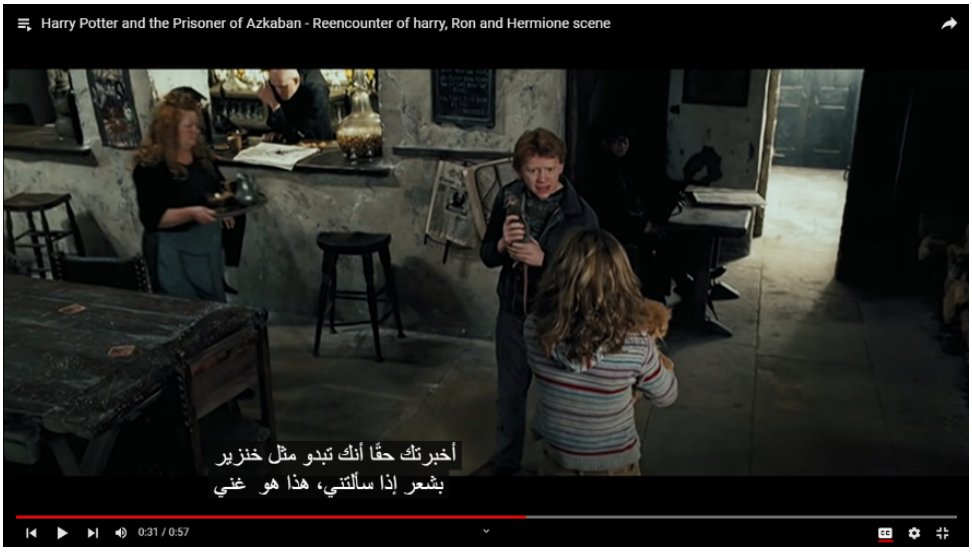

Figure 3. An example on untranslated material.

Minor Errors

Minor functional equivalence errors refer to lexical flaws that primarily involve terminology and do not have a significant impact on the overall narrative of the film.

Example (4)

ST: But where is it? I saw the **beast**, just now. Not a moment ago!

TT: فأين هي، لقد رأيت **البنجر** الآن وليس للحظة قبل ذلك،

(**BT**: Where is it, I saw the beet now and not before a moment.)

In example (4), after the ministry sentenced the death of Buckbeak, a kind of bird that belonged to Hagrid, a friend of Harry, the minister went to Hagrid's house to implement the ruling, but at the same moment, when the minister was busy talking with Professor Dumbledore. Harry and Hermione succeeded in smuggling the bird. The YMT translates the word "beast" as " البنجر ," which means in Arabic "beets," and that may cause a little misunderstanding for the viewers(See **Figure 4**). Moreover, the machine uses the wrong punctuation in this segment; it changes the question mark into a comma and deletes the point that affected the segment and the translation. The error can be a minor one because It's a lexical error that has no bearing on the broader plot of the film. Suggested translation: رأيت الوحش للتو، قبل لحظة! ولكن أين هو؟ لقد (But where is he? I saw the beast just now, not a moment ago).

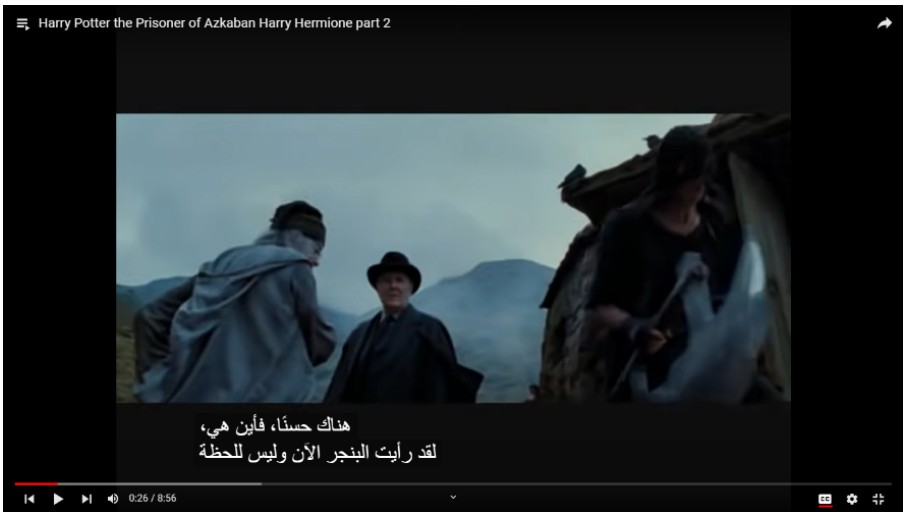

Figure 4. And example of minor error.

Unnecessary Addition

Example (5)

ST: We have a killer on the loose.
TT: لدينا قاتل طليق **طليق،**
(**BT:** We have a free free killer)

In example (5), the minister is telling Harry to take caution, because Sirius has already escaped, and he must not move alone. YMT conveys the meaning correctly; however, it repeats the word " طليق" twice, which doesn't help the viewers and doesn't add any information or explanation(See **Figure 5**). The error is classified as a minor error.

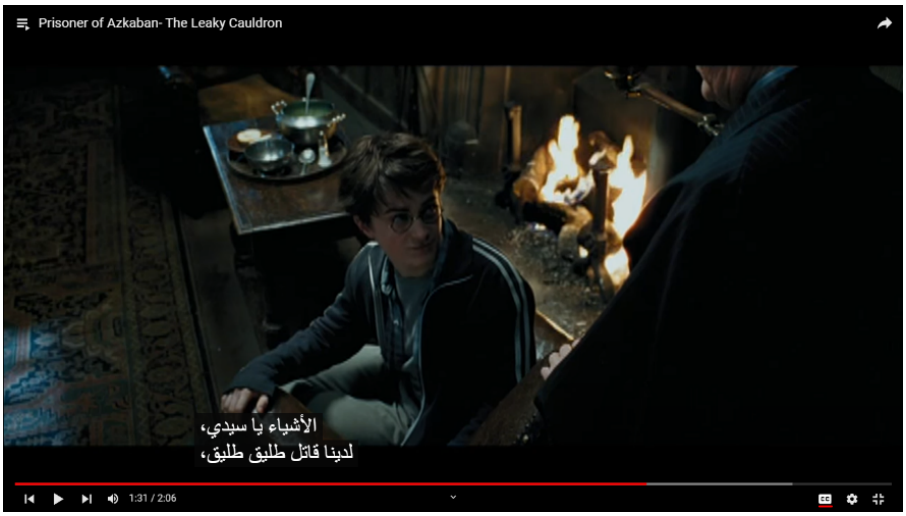

Figure 5. An example of unnecessary addition.

(3) Stylistic Errors

Stylistic errors are comparatively less consequential than semantic errors since they mostly result in inconveniences rather than misunderstandings. For example, using the wrong words for address, speaking in the wrong register (either too high or too low), or using language in a way that doesn't follow the rules set by the original context—for example, using modern language in historical films [12].

Example (6)

ST: Ernie! They're right on top of us! Mind your head.
TT: **إيه** الحق فوقنا، اهتم برأسك

(**BT:** Yeah, the right is above us, take care of your head)

In example (6), Harry is on a wizard bus to take him to the Leaky Cauldron in London. In the meantime, they face the double bus. So, the driver assistant alerts the driver that they're directly in front of them and says, "Ernie, they're right on top of us!". YMT translates it in a word-for-word way as " فوقنا إيه الحق " that doesn't convey neither what it said nor the intended meaning(See **Figure 6**). Also, as we noticed, the machine didn't catch the name of the driver clearly; thus, it translates as "إيه" and this is a word in colloquial Arabic, especially Egyptian, used in wonder or the question; therefore, the viewers may read the subtitle in that way. The machine changed from normal Arabic into colloquial, which caused a misunderstanding in the style and changed the whole meaning of the sentence. The error is classified as a standard error because it affects the sentence's meaning, not the whole scene. Suggested translation: آيرني! أنهن أمامنا مباشرة! انتبه لرأسك. (BT: Ernie! They are right in front of us! Watch your head!)

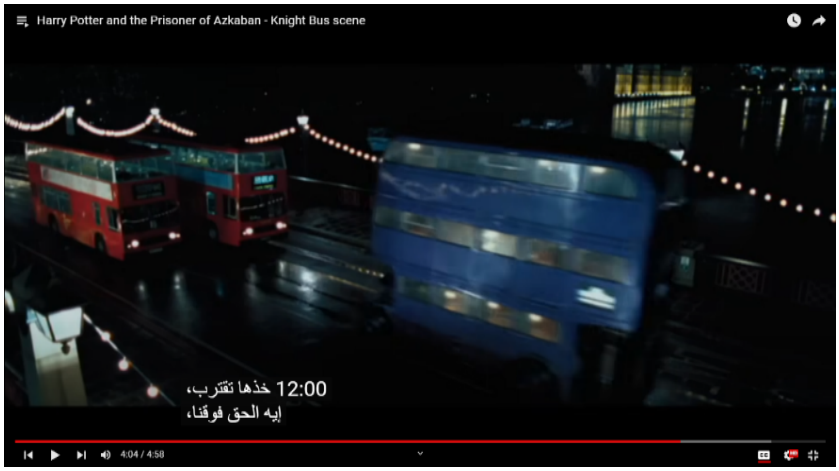

Figure 6. An example of stylistic error.

4.2.2. Acceptability

Acceptability refers to the degree to which the target text adheres to the norms of the target language. The mistakes in this category are those that make the subtitles seem foreign or otherwise odd. These errors also break the contract of illusion as they direct attention to the subtitles. These mistakes are of three kinds: Grammar errors, spelling errors, and errors of idiomaticity[12].

(4) Grammar Errors

These cases are just instances of grammatical errors in the target language, appearing in different forms. Nevertheless, it is important to note that the grammar used in subtitling is specifically tailored to the target language. Pedersen[12] points out that subtitling might be seen as a hybrid manifestation of both spoken and written language, suggesting that a rigid adherence to the grammatical norms of written language may be inappropriate. The presence of a serious grammatical mistake in the subtitle hinders its readability and/or comprehension. Minor errors, such as the misuse of 'whom' in the English language, are considered the pet peeves that cause discomfort to purists. Standard errors are located in between.

Example (7)

ST: Before **I fainted**, I heard something. A woman screaming.

TT: قبل أن **أغمي عليه**، سمعت شيئًا امرأة تصرخ

(**BT**: Before he fainted, I heard something, a woman screaming)

In example (7), Harry was talking with Professor Lupin describing to him how he felt before he fainted because of the Dementors. YMT translates it in a good way that conveys the intended meaning. However, the machine translates "I fainted" as "أغمي عليه", which is a grammatical mistake that changes the agent; like, it's not Harry who fainted but someone else (See **Figure 7**). The error is classified as a standard one. Suggested translation: عليّ، سمعت شيئًا ما. أمرأة تصرخ. قبل أن يُغمى (**BT**: Before I fainted, I heard something, a woman screaming.)

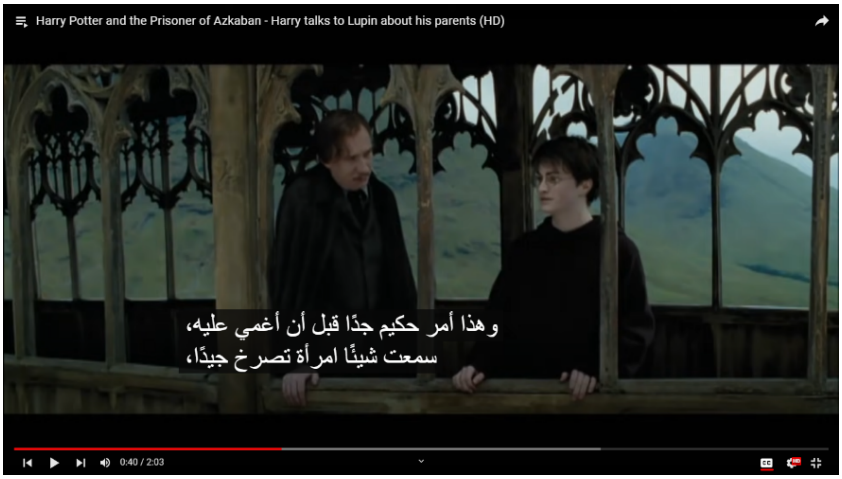

Figure 7. An example of grammatical error.

Problems with Pronouns

Example (8)

ST: Why don't you run along and play with your chemistry set!

TT: لماذا لا تركضان وتلعبان في مجموعة الكيمياء الخاصة بك،

(BT: Why don't you (dual/two people) run and play in your chemistry set!**)**

In example (8), Sirius tells Professor Snape to leave the place. YMT conveys the intended meaning correctly; however, it translates it as if he were talking with two people. It's a minor error that doesn't affect the overall meaning (See **Figure 8**). The error may be considered standard. Suggested translation: في مجموعة الكيمياء الخاصة بك! لماذا لا تركض وتلعب (BT: Why don't you run and play in your chemistry set!).

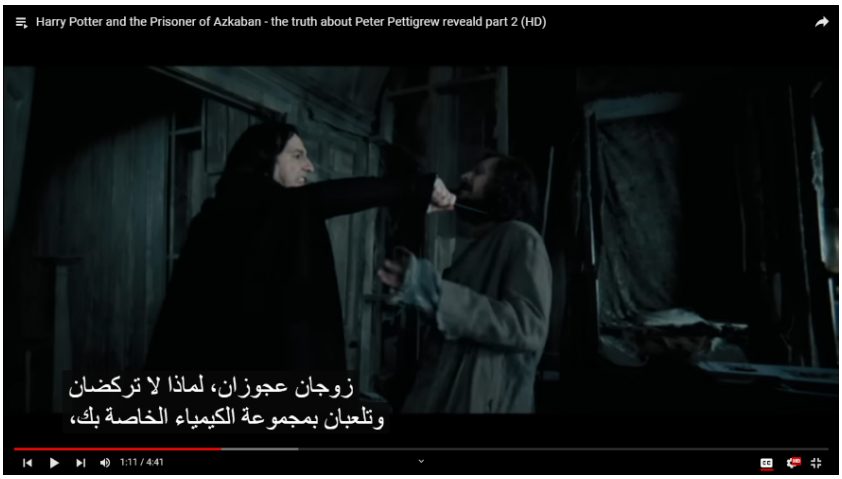

Figure 8. An example of problems with pronouns.

(5) Spelling Errors

Spelling refers to the accurate arrangement of letters in a given word. The assessment of spelling errors may be classified based on their severity as follows: minor errors include all spelling mistakes, standard errors alter the intended meaning of a word, and serious errors render a word illegible.

Example (9)

ST: This term, we'll focus on **Tasseomancy**, the art of reading tea leaves.

TT: هذا الفصل الدراسي الذي سنركز فيه على **tezo MEMC** وهو فن قراءة أوراق الشاي

(**BT**: In this semester, we we'll focus on **tezo MEMC**, the art of reading tea leaves.)

In example (9), the word "tasseomancy," which is the subject of the course that Professor Trelawney started teaching to her students. YMT didn't translate it, but it conveys it as loan word with incorrect spelling. These kinds of errors don't affect the meaning of the scene, especially if it's the name of things, places, etc. Although it may occasionally bother the viewers(See **Figure 9**). Thus, the error is considered minor error. Suggested translation**:** فن قراءة أوراق الشاي. هذا الفصل، سنركز على التاسيومانسي؛ وهو . (BT: In this term, we'll focus on **Tasseomancy**, the art of reading tea leaves).

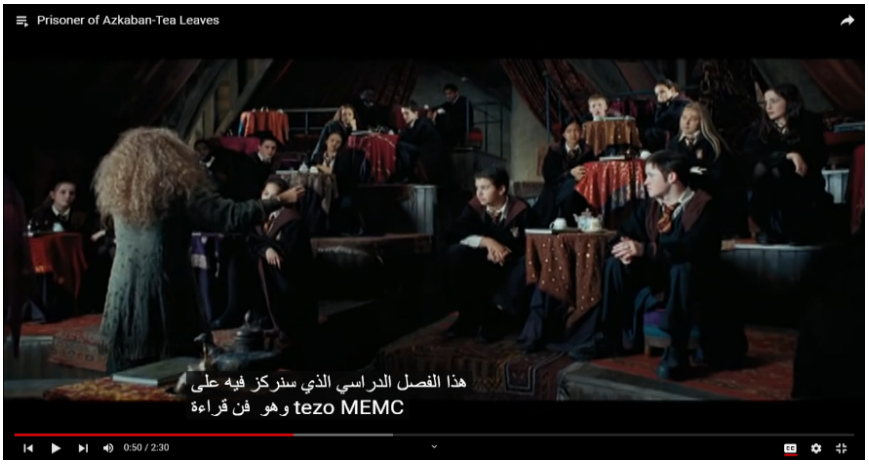

Figure 9. An example of spelling errors.

(6) Idiomaticity Errors

Within this model, the concept of idiomaticity encompasses more than just the use of idioms. It also entails the natural use of language, specifically referring to expressions and phrases that would be considered natural by a native speaker of that particular language. Problems falling under this category pertain not to grammatical flaws but rather to errors that result in unnatural-sounding language in the target language since they give rise to regressions, hinder comprehension, and thus impact reading speed.

Example (10)

ST: I'm sure Madam Pomfrey will fix it in a **heartbeat.**

TT: أنا متأكد من أن مدام بومبس ستصلح **نبضات قلبهم،**

(**BT:** I am sure that Madam Bombos will fix their heartbeat).

In example (10), the dialogue was between Hermione and Ron about Ron's leg, which was almost torn off by the dog after it bit him. For this reason, Hermione gives him the comforting assurance that Madam Pomfrey, who was a skilled healer at Hogwarts, would undoubtedly cure him. The YMT uses the word " نبضات القلب " when translating the word "heartbeat," which is a literal translation in the wrong context. This is an idiom, which means that Madam Pomfrey will fix it in no time or in brief. Therefore, the subtitles don't convey the intended meaning correctly, which may confuse

the viewers. Also, the machine transliterates the name of Madam Pomfrey erroneously as "مدام بومبس" (See **Figure 10**). This may be seen as a standard error since it alters the natural-sounding language and hampers the understanding of viewers, but it doesn't affect the whole scene. Suggested translation**:** أنا متأكدة أن مدام بومفري ستعالجها في طرفة عين. (**BT**: I am sure that Madam Pomfrey will fix it in a twinkle of an eye).

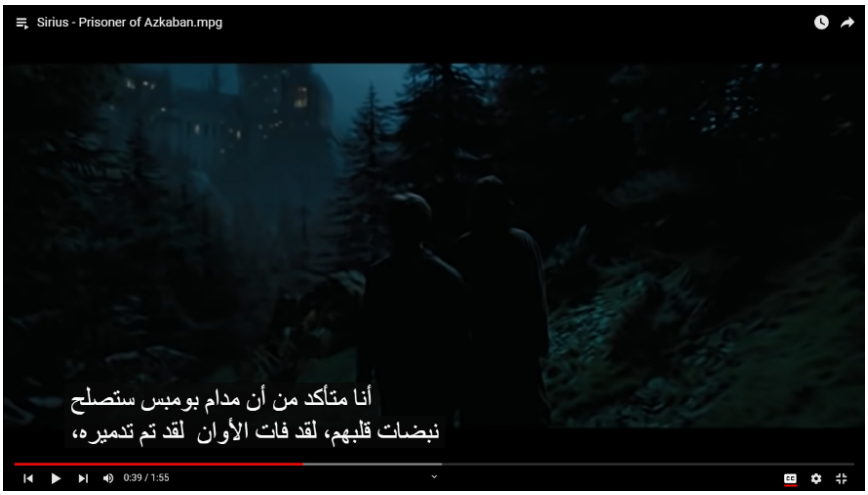

Figure 10. An example of idiomaticity error.

4.2.3. Readability

Within this particular area, one encounters situations that are often referred to as technical norms or problems in other contexts. The rationale for categorizing them as readable in this context comes from the viewer-centric approach of the FAR model. It is assumed that viewers are primarily concerned with the simple nature of reading subtitles rather than the technical aspects involved. The readability issues include mistakes related to segmentation and spotting, punctuation, reading speed, and line length.

(7) Segmentation and Spotting

Literature on subtitling emphasizes the significance of accurate segmentation and spotting. It is worth noting that the presence of defective segmentation has the potential to distract the attention of the viewer. This assertion is supported by empirical evidence from eye-tracking experiments, which have shown that atypical segmentation "increases considerably the time in the subtitled area"[13]. The occurrence of spotting mistakes may be attributed to poor synchronization between speech and other visual elements, such as subtitles appearing prematurely or remaining beyond the acceptable lag time for out-times. In YMT, the translation begins as soon as the scene opens and consists of two lines that are followed by the disappearance of one and the appearance of another. As a result, viewers will focus on the translation more than the scenes, which is seen as a standard error. So, we won't be discussing the spotting here.

Example (11)

ST: Very well. Kill him. But wait one more minute. Harry has the right to know why.

TT: القتل جيدًا ولكن انتظر دقيقة واحدة لاري لديه الحق في معرفة

(**Back translation:** Good killing, but wait for one minute **Larry** has the right to know why).

ST: I know why. You betrayed my parents. You're the reason they're dead!

TT: سبب خيانتك لوالداي، فأنت السبب في موتهما،

(**BT)**: The reason for your betrayal to my parents, you are the reason for their death.

In example (11), Professor Lupin is talking with Sirius Black and reminding him that Harry has to know the truth before killing Peter Pettigrew, and in the meantime, Harry interrupts them by saying that he knows why he did

that, because he betrayed his parents. YMT makes two mistakes at the segmentation level. Firstly, it translates the words "Very well. Kill him." as " القتل جيدًا ", which combines the two sentences together and makes it as one sentence. Also, it omits the pronoun "him", and that clearly affects conveying the intended meaning. Second, the machine omits the word "why" at the end of Lupin's speaking and the first sentence, "I know why." by Harry's speaking, and combines the dialogue to make the translation of theirs as a single sentence. There is also a spelling mistake in translating the name "Harry" as "لاري" (See **Figure 11**). This error may be considered a standard error that causes a misunderstanding and an inconvenience to the viewers. Suggested translation:

- حسنًا إذا! اقتله، لكن انتظر دقيقة واحدة، هاري لديه الحق في معرفة السبب.
- أعلم لماذا. لقد خنت والديّ، أنت السبب في موتهما.

BT: - Alright! Kill him, but wait a minute, Harry has the right to know the reason.
- I know why, you betrayed my parents, you caused their death

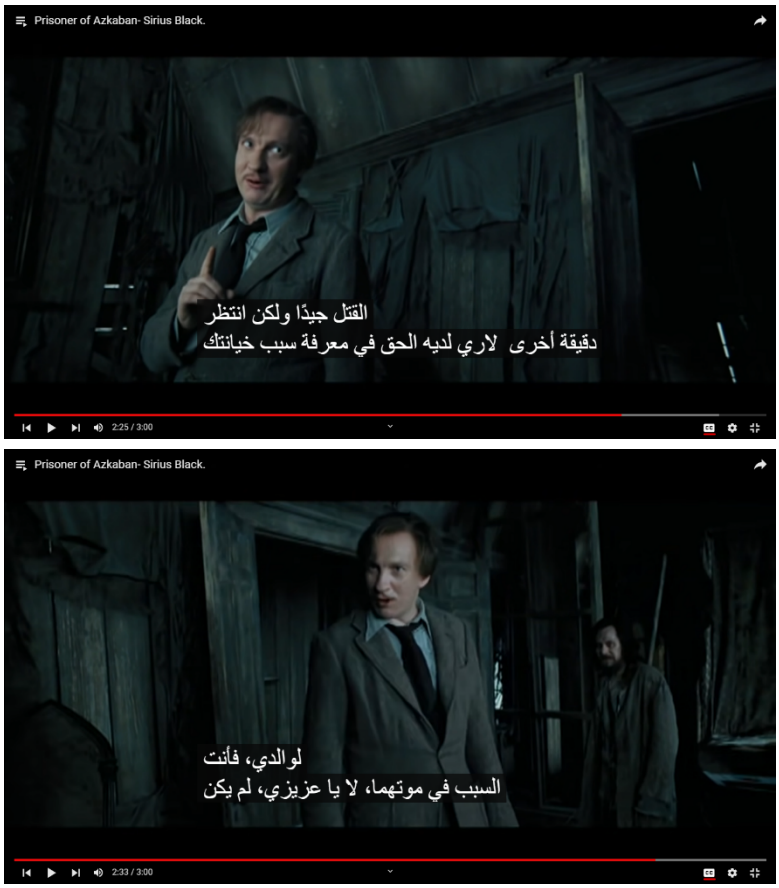

Figure 11. Examples of segmentation and spotting.

(8) Punctuation and Graphics

"Punctuation in subtitling is more important than in other texts"[12]. The use of italics to indicate 'irrealis' is a noteworthy illustration. Italics serve as a typographic convention in several countries to indicate the presence of a voice or text that is deemed to be absent. This encompasses various contexts, such as telephone conversations, television broadcasts, public address systems, dreams, internal thoughts, and recollections, among others. The widespread use of this practice in many contexts has established it as a customary norm, integrating it into the structure of the contract of illusion. Consequently, any incorrect utilization of this practice should be seen as a standard mistake. Similarly, the use of dashes follows the same principle. There is a considerable degree of variety in the use of dashes. Speaker indication is one of the primary functions of subtitles, facilitating the identification of the speaker.

Additionally, they are used to maintain continuity between utterances. In more infrequent cases, subtitles may also be utilized to signify the speaker's engagement with a different person[12]. It is important to acknowledge that none of the aforementioned elements have been used in YMT. We have focused on the punctuation in-text.

Example (12)

ST: Dementors force us to relive our very worst memories. Our pain becomes their power.

TT: يجبر الديمنتورز على استعادة أسوأ ذكرياتنا، ويصبح ألمنا همهم. القوة،

(**BT:** Dementors force us to relive our very worst memories. Our pain becomes their concern, the power).

In example (12), Professor Lupin explains to Harry about the Dementors and how they feed on humans' good memories and leave them the worst ones to live with. YMT translates it in a good way, almost conveying the meaning correctly, but the machine puts the point between the words "their" and "power," which affects the translation and disturbs the viewers (See **Figure 12**). The error is classified as standard. Suggested translation: على إحياء أسوأ ذكرياتنا، وتصبح آلامنا قوتهم. يجبرنا الديمنتورز (BT: Dementors force us to relive our worst memories, and our pains become their power).

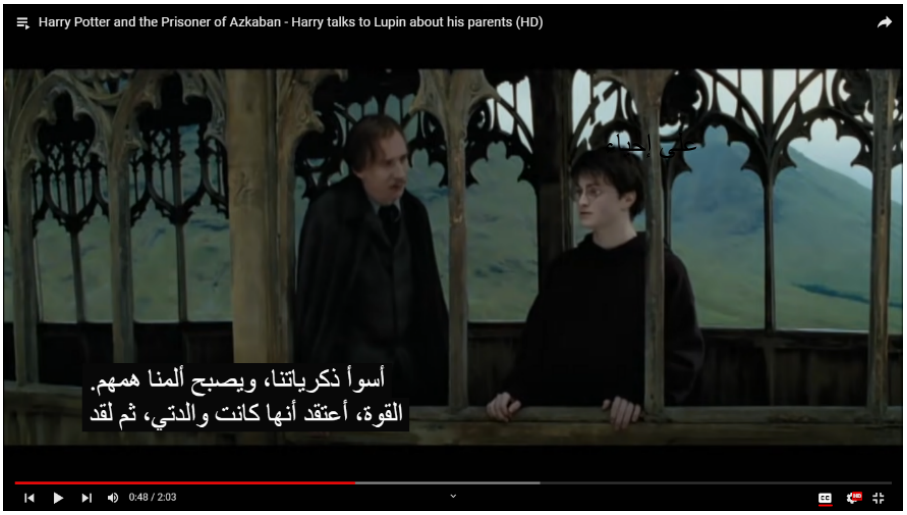

Figure 12. An example of punctuation and graphics.

5. Conclusions

The YMT system exhibited limitations in accurately conveying meaning due to its inability to select appropriate words that align with the given context. At times, the translation process resulted in a literal word-for-word rendition, causing the translated text to deviate from the intended meaning of the source text. Semantic analysis is a fundamental aspect and prerequisite for accurately understanding the grammatical structure of a phrase. Nevertheless, there is an absence of substantial research advancements in this particular domain, hence demanding a heightened focus on its practical implementation. According to grammar, YMT did not keep the tenses of the ST sometimes. YMT resorted to omitting and untranslated some parts of the ST, whether content word omissions or deleting the utterance completely; that is, 34.3% of the standard errors are untranslated examples. Regarding punctuation, YMT most often uses the comma in a lot of instances, even instead of the question marks, and almost 53% of punctuation errors put the comma at the end of the sentence instead of the point.

In the end, there is a need for enhancing the quality of speech recognition systems and optimizing the functionality of segmentation, whether at the micro or macro level. As for segmentation errors, this is attributed to the fact that YouTube is as close as possible to being a speech-to-text engine. Also, despite significant advancements in the technological maturity of speech recognition, achieving 100% accuracy remains unachievable. Significant variations in recognition rates across various scenarios are the result of a number of

factors, including background noise and speech pace. The main drawback of the lack of speech recognition is its inability to modify text based on contextual cues. Additionally, inadequate semantic understanding is a major obstacle in this field. In order to address this issue, it is imperative to enhance the algorithm and acquire a substantial volume of dependable data for algorithmic training. Nevertheless, it is important to acknowledge that machine-translated subtitles, although not yet meeting the necessary standards for direct market use, exhibit accuracy in certain instances, devoid of grammatical errors or omissions. The errors that occurred can be attributed to the complex diversity of structures, polysemy, and vocabulary present in the Arabic language, as well as the limitations of Arabic machine translation and its database. However, the automated generation of timetables serves to save time for subtitlers and improve overall efficiency. Further studies should be undertaken to assess other forms of machine translation that employ distinct operational systems on different platforms.

Author Contributions

Conceptualization, A.M.A.-H. and R.H.R.; methodology, A.M.A.-H.; software, A.M.A.-H.; validation, A.M.A.-H., and R.H.R.; formal analysis, A.M.A.-H.; investigation, A.M.A.-H.; resources, R.H.R.; data curation, R.H.R.; writing—original draft preparation, A.M.A.-H.; writing—review and editing, R.H.R.; visualization, A.M.A.-H. and R.H.R.; supervision, A.M.A.-H.; project administration, A.M.A.-H.; funding acquisition, A.M.A.-H. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

No new data were created.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Maučec, M.S., Donaj, G., 2019. Machine translation and the evaluation of its quality. Recent Trends In Computational Intelligence. 143, 1–20.

[2] Matkivska, N., 2014. Audiovisual translation: Conception, types, characters' speech and translation strategies applied. Studies About Languages. 25, 38–44.

[3] Yao, G., 2022. Evaluation of machine translation in English-Chinese automatic subtitling of ted talks. Modern Languages, Literatures, and Linguistics. 1(01), 12–22.

[4] Hagström, H., Pedersen, J., 2022. Subtitles in the 2020s: The influence of machine translation. Journal of Audiovisual Translation. 5(1), 207–225.

[5] Karakanta, A., 2022. Experimental research in automatic subtitling: At the crossroads between machine translation and audiovisual translation. Translation Spaces. 11(1), 89–112.

[6] VARGA, C., 2021. Online Automatic Subtitling Platforms and Machine Translation. An Analysis of Quality in AVT. Scientific Bulletin of the Politehnica University of Timisoara Transactions on Modern Languages, 20(1), 37–49

[7] Matusov, E., Wilken, P., Georgakopoulou, Y., et al., 2019. Customizing neural machine translation for subtitling. Proceedings of the Fourth Conference on Machine Translation; Florence, Italy, 19–23 August 2019. pp. 82–93.

[8] Song, H.J., Kim, H.K., Kim, J.D., 2019. Inter-sentence segmentation of YouTube subtitles using long-short term memory. Applied Sciences. 9(7), 1504.

[9] Gupta, P., Sharma, M., Pitale, K., et al., 2019. Problems with automating translation of movie/tv show. Available from: https://arxiv.org/abs/1909.05362 (cited 4 September 2019).

[10] Hiraoka, Y., Yamada, M., et al., 2019. Pre-editing plus neural machine translation for subtitling: effective pre-editing rules for subtitling of TED Talks. Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks; Dublin, Ireland, 19–23 August 2019. pp. 64–72.

[11] Athanasiadi, R., 2019. Exploring the potential of machine translation and other language assistive tools in

subtitling: A new era? In: Deckert, M. (ed.). Audiovisual Translation: Research and Use. Peter Lang: Berlin, Germany. pp. 29–49.

[12] Pedersen, J., 2017. The FAR model: Assessing quality in interlingual subtitling. The Journal of Specialized Translation. 28, 210–229

[13] d'Ydewalle, G., Warlop, L., Van Rensbergen, J., 1989. Television and attention: Differences between young and older adults in the division of attention over different sources of TV information. Medienpsychologie: Zeitschrift für Individual- und Massenkommunikation. 1, 42–57.