Open Access

Check for updates

RESEARCH ARTICLE   **Section:** *Literature, Linguistics & Criticism*

Audience responses to cultural and linguistic gaps in English–Arabic auto-subtitles on YouTube

Abdullah Al-Momani[1], Ahmad S Haider[1] iD **, Abdulazeez Jaradat[2], Wala' Mohammad Akasheh[1], Mohammed Dagamseh[3] & Ronza N. Abu Rumman[4]**

[1]Department of English Language and Translation, Applied Science Private University, Amman, Jordan

[2]Department of English and Literature, Yarmouk University, Irbid, Jordan

[3]Department of Linguistics, University of Canterbury, Christchurch 8140, New Zealand

[4]Department of English Language and Literature, University of Jordan, Amman, Jordan

*Correspondence: a_haidar@asu.edu.jo

ABSTRACT

YouTube is considerably used by millions of people around the world. Auto-generated subtitling is one of YouTube features that is intensively utilized to increase users' understanding of different videos in various fields, including media, education, and entertainment. This study examines the responses of 4500 participants to the quality of this service with a special focus on linguistic accuracy, cultural adaptation, viewing experience, and attitudes. The findings showed that YouTube is frequently used by global users on a daily basis. Subtitles were enabled by most respondents for several reasons, which include language learning, accent navigation, and clarifying unclear audio. It is observed that some participants reported several incidents of mistranslations and cultural misrepresentations which affected comprehension and reduced enjoyment. A large proportion of participants preferred the subtitles created by human subtitlers to the auto-generated ones. Content type affected the attitudes of users with low trust in the auto-generated subtitles for formal and educational content. The study recommends conducting further research on the use of autogenerated subtitles across different language pairs beyond Arabic and English. The results of this study are likely to be useful for AI developers, translators, and language learners.

KEYWORDS: audience reception, auto-generated subtitles, cultural gaps, English–Arabic translation, media accessibility, YouTube

Introduction

Recent scholarship has increasingly highlighted the intersection of language, culture, and technology in diverse contexts (Al-Dhuhli, Alkindi, & Al-Taani, 2022; Ammari & Al-Ahmad, 2023; Fraihat, 2024; Ja'afreh, 2023; Khatatbah & Ashour, 2024; Masoud, 2025; Meqdad, Al-Bayyari, & Al-Taher, 2023; Naib, 2025). Translation is defined as rendering a text written in one source language into another target language. Nida and Taber (1982) stated that "translation consists in reproducing in the receptor language the closest natural equivalent of the source language message, first in terms of meaning and secondly in terms of style" (p.12).

In the last two decades, the creation of audiovisual (AV) content has undergone a revolutionary transformation, driven by the widespread availability of various streaming platforms (Al-Darabee, 2024; Du & Lu, 2024; Haider, Alzghoul, & Hamadan, 2023; Hashish & Hussein, 2022). This increased the popularity of creating different types of AV content in different places around the world, making a shift in the way we communicate, share stories, and connect (Alkhatib & Haider, 2024; Haider & AlKhatib, 2025; Haider & Shohaibar, 2024). One of the benefits of YouTube is the fact that filming people in unreachable areas is way easier, and their cultural and linguistic characteristics can be much more accessible to other groups if they are filmed and published online. Short documentary videos have emerged as powerful resources for capturing the essence of real-life narratives in an impactful manner.

Technological development has also played a significant role in making videos, sometimes with their translation, accessible to groups from different parts of the world (Ahmad, Haider, & Saed, 2025; Haider et al., 2025a; Haider, Tair, & AlBkour, 2025b; Hassan & Haider, 2024; Saideen et al., 2024). It provided user-friendly platforms where individuals can share AV content and provide within-platform features so people can share their ideas and points of view; providing automated translations that can help people communicate, and generating intralingual auto-generated subtitles, which is a very important technology and helps in providing written subtitles that are written in the language of the speaker in these videos (Allam, 2023; Alzaabi & Rabab'ah, 2023).

Dharmale and Patil (2019) mention that "Automatic Speech Recognition is a technology which permits the machine to take out oral content from a speech signal and produce a text message by using feature extraction and classification techniques." As part of Spoken Language Translation, Machine Translation (MT) can be defined as the subfield of computational linguistics concerned with using software in translation across human languages (Almahasees, 2017). Besides improving access to spoken words for all individuals, ASR technology offers many benefits for its users, including facilitating tasks for transcribers and improving pronunciation and fluency for language learners (Akasheh et al., 2024; Romero-Fresco & Fresno, 2023). Still, as Guskaroska (2019) mentions, "the accuracy of ASR for language learning applications remains uncertain" (P.3). On the other hand, ASR has a broader concept that involves the identification and interpretation of various sounds, not limited to speech, and encompasses the recognition of any sound or noise in an environment. These technologies are used across many online AV platforms and help generate intralingual subtitles.

In this era of digital learning, entertainment and media, accessibility features matter. This is because viewers are expected to have difficulty understanding diverse video content on digital platforms, including YouTube (Altunisik, Firat, & Keceli, 2022; Pfender, Wanzer, & Bleakley, 2024; Rahmatika, Yusuf, & Agung, 2021; Saed et al., 2021). The role of accessibility features is to enhance viewers' understanding of such content. Auto-generated subtitling is one of the features on YouTube that is intensively used among viewers to understand videos that are related to different fields, including media, education, and entertainment (Allam, 2023; Altunisik et al., 2022; Rahmatika et al., 2021; Saed et al., 2021; Xie, 2022a) Recently, there has been increasing interest in evaluating the effectiveness of auto-generated subtitling in helping viewers gain a deeper understanding of different kinds of video content and viewing preferences, whether educational, entertaining, or media-related.

Upon the revision of the most recent literature on this issue within the English-Arabic context, it is clear that it is gaining more attention, as linguistic inaccuracies and insensitivity to cultural aspects and differences are potential (Alrousan & Haider, 2022; Banikalef & Abu Naser, 2019; Farghal, 2019; Haider, Saideen, & Hussein, 2023; Khalil, 2022; Obeidat, 2023). Audience reception studies underscore that viewers are not passive consumers of subtitles but active negotiators of meaning (Al-Abbas, Haider, & Saideen, 2022; Aldualimi & Almahasees, 2022).

The present study uses a 6-construct questionnaire and aims to investigate the responses of 4500 users

on the quality of YouTube auto-generated subtitles, with special focus on the linguistic and cultural aspects of these subtitles. The present study aims to answer the following question:

- In this study, we explore the reactions of 4500 participants to the use of YouTube autogenerated subtitles with a focus on linguistic and cultural aspects. It attempts to answer the following research questions:

  - How are YouTube auto-generated subtitles between Arabic and English perceived by Arabic-speaking and bilingual audiences?

Literature Review

Artificial Intelligence (AI) powered subtitles use machine learning algorithms and natural language processing (NLP) techniques to generate audio or video content subtitles. This approach involves using AI to analyse the audio or video content, transcribe the spoken words, sign language, or paralinguistic elements and generate accurate and synchronized subtitles.

When looking at the types of subtitles, different researchers have proposed many classifications. In one of the attempts to classify the subtitles, Gottlieb (1997) has characterized subtitles from a linguistic perspective that can be either *'intralingual,'* which is within the same language, or *'interlingual,'* which is between two languages. Later on, Diaz-Cintas and Remael (2014) added a third type that displays verbal and nonverbal signs in two or more languages and called it bilingual. The three types, as proposed by Diaz-Cintas and Remael (2014), are shown in Figure 1 below.
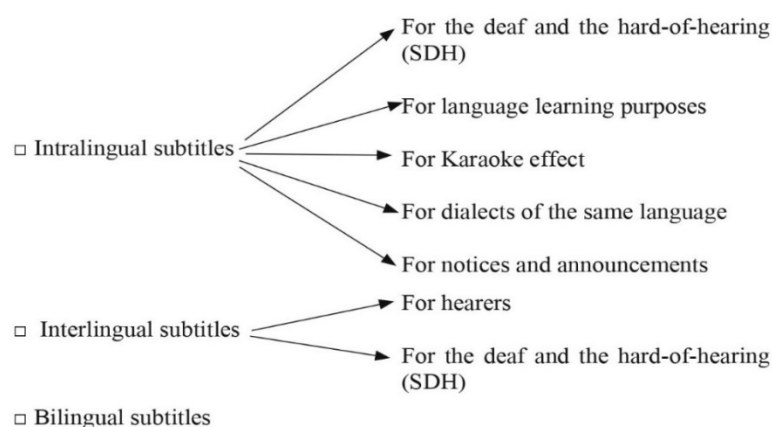

Figure 1: Classifications of subtitles based on linguistic dimension (Diaz-Cintas & Remael, 2014, p. 14)

Moreover, Diaz-Cintas and Remael (2014) categorized subtitles from another perspective, which is the time available time for preparation, as shown in Figure 2.
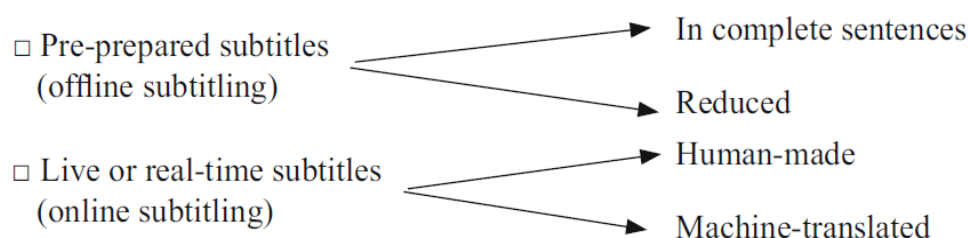
Figure 2: Classifications of subtitles based on time available for preparation (Diaz-Cintas & Remael, 2014, p. 19)

Normally, the process of subtitling is done by humans, yet technology's role is increasing(Díaz-Cintas, 2005). Such a development has led to auto-generated subtitles, which do not depend on humans; they depend on software. One of the technologies that can generate subtitles is Automatic Speech Recognition (ASR), which enables the recognition of spoken utterances and turns them into text by computers (Almahasees, 2017, 2018, 2021; Almahasees & Jaccomard, 2020). These subtitles that depend on speech recognition are automatically

generated and are not pre-prepared.

Generating captions or subtitles for audio and video content has been a topic of interest in recent years, as it provides a fast way to improve accessibility and saves time and money compared to human transcribers. Anusuya and Katti (2009, p. 181) argue that the main focus of speech recognition technology is to convert human speech into written form. This revolutionary technology has come a long way since its inception. The roots of ASR can be traced back to the 1950s, when researchers began experimenting with machine-based methods for transcribing spoken words, and it can be witnessed in many papers and studies that tracked the history of Speech recognition (Baker et al., 2007; Hannun, 2021; Juang, Rabiner, & Barbara, 2005).

Peacocke and Graf (1995, p. 27) state that "The 1960s saw advances in the automatic segmentation of speech into units of linguistic relevance (such as phonemes, syllables, and words) and on new pattern-matching and classification algorithms." One can ask about the most recent topics that have been addressed in ASR research and what can be added to the literature. This unique technology has been the basis of auto-generated subtitles for many platforms when generating intralingual subtitles that are based on machines and without any human assistance. Chan, Kruger, and Doherty (2019, p. 239) state that "Automatic subtitles are advantageous because their creation is faster, more cost-effective, and less time-consuming."

Today, auto-generated subtitles are provided by many online content-sharing platforms such as YouTube. These platforms may also use Machine translation technology, which is concerned with using computer programs to translate human languages (Almahasees, 2017). This can be related to the fact that the increasing need for multilingual content and the widespread availability of MT have brought more attention to it, particularly with the advent of automated subtitle translation due to the integration of Google Translate into YouTube (Xie, 2022b). Also, the huge amount of information shows the importance of MT (Almahasees & Jaccomard, 2020). Some online video editing platforms, such as Veed, that use artificial intelligence (AI) also provide auto-generated subtitles for their users. Still, the capability of these systems to recognise all spoken utterances is facing many challenges. One of the challenges that these systems may face is the code-switching phenomenon in bilingual and multilingual societies. Mustafa et al. (2022) mention that "While there are established databases for monolingual speech recognition that can be used in bilingual and multilingual speech recognition, the availability of the CS corpus is limited due to the various combinations of languages in CS" (P.5).

Arabic is the primary language that is spoken by millions of people across the Arab world, and it has many dialects (Abulawi, Al Salman, & Haider, 2022; Haider & Hussein, 2022). Its formal script is written from right to left, in which the writing system used is Abjad, where consonants are represented, and the representation of vowels is optional since they can be inferred by the reader. Versteegh (2014) mentions that "the richness of the Arabic lexicon is far beyond that of other languages. The latter has, for instance, only one word for 'lion' or 'sword,' whereas Arabic has dozens of synonyms for them" (P.85). Versteegh (2014) added that "In addition to this lexical abundance.

Many spoken dialects are being used as a medium of communication, and there are considerable differences in lexical and phonological levels. These differences are often related to geographical and social considerations. However, Almahasees (2021) pointed out that "written Arabic is similar in all Arabic-speaking countries because each uses MSA as the language of official communication" (P.5).

Watson (2012) concludes that "A comparison of dialect material across widely geo-graphically separated areas shows both long-distance effects due to population movement and local effects due to interaction with the original local languages" (P.868). Arabic speakers use two distinct varieties of Arabic that coexist in the linguistic situation of almost all countries that speak Arabic. This diglossic situation reflects a complex relationship between the standard language and the dialectal variety. Ferguson (1959, p. 330) says that "Arabic diglossia seems to reach as far back as our knowledge of Arabic goes, and the superposed 'Classical' language has remained relatively stable."

The main aim of the ASR systems is to accurately transcribe spoken utterances. This may be challenging and needs special focus and attention. Liao et al. (2023) presented a new process to improve the readability of the ASR output. Their goal was to correct grammatical mistakes and disfluencies to make it more readable to humans. There are many challenges that ASR systems may face, and code-switching (CS) is one of the challenges that ASR models may face when transcribing utterances. Yu et al. (2023) highlighted the benefits of using external data by generating code-switching text and using cross-modality learning methods to inject it.

Their research and experiment are very important since they would help to enhance code-switching Mandarin-English ASR performance. Another research guided by Hussein et al. (2023) proposed a new pipeline for Arabic-English code-switching text generation to improve speech modules. Their study showed that morphological segmentation is important for correct alignments to synthesize more realistic CS and confirmed that the best approach in their zero-shot learning scenario is achieved with morphological segmentation and random lexical replacements. Moreover, Dhawan, Rekesh, and Ginsburg (2023) investigated the training of bilingual and code-switching models using monolingual datasets, where they proposed two new techniques to train CS ASR models. The investigation focused on two language pairs, English-Hindi and English-Spanish, yet the results recommended that these methods could be extended to additional languages. Liu, Niehues, and Spanakis (2020) investigated the task of compressing ASR outputs to improve subtitle readability. Their experiment showed that with limited data, adapting a Transformer-based ASR model to incorporate both transcription and compression capabilities is possible.

In general, ASR systems can recognize spoken utterances in high-resource languages due to the availability of data and the focus of research and development on these languages. Yet, some recent attempts focused on low-resource languages. Tsoukala et al. (2023) collected recordings of the Pomak language, which is a low-resource language that is spoken in Greece and is considered to be a Southeast Slavic language variety. They proposed a pipeline for data processing and ASR model training for low-resource languages based on the language family, which, as per their study, is considered to be the first Pomak ASR model. Also, Liu, Spence, and Prud'Hommeaux (2023) investigated the impact of language model size on low-resource ASR for five widely spoken low-resource languages and one endangered Native American language. Their findings proved that having larger language models for endangered languages may not always lead to better performance.

In most cases, auto-generated subtitles rely on ASR technology to convert spoken language into written text, forming subtitles that are displayed for viewers. This topic captured the interest of many researchers in the field.

This current study examines the evaluation of 4500 participants on the quality of YouTube autogenerated subtitles, which are generated in one way or another using ASR technology, with special focus on linguistic accuracy, cultural adaptation, viewing experience, and attitudes.

Methodology

The study surveyed 4,500 YouTube viewers from Arabic-speaking and bilingual communities across multiple regions to explore perceptions of English–Arabic auto-subtitles. A two-part questionnaire was used: the first gathered demographic data, and the second, comprising 29 items across six themes (see Table 1), measured responses on a 3-point Likert scale. Data were collected via an online survey over three months, following a pilot test with 50 participants to ensure clarity.

To assess the internal consistency of the six thematic constructs, Cronbach's alpha reliability test was calculated. All constructs achieved acceptable reliability, with alpha values ranging between 0.72 and 0.81 (see Table 1).

Table 1: Cronbach's Alpha Reliability Coefficients for Thematic Constructs (Cronbach, 1951)

Thematic Construct	Cronbach's Alpha (α)
Media Consumption Habits	0.74
Use of Subtitles	0.77
Motivations for Subtitle Use	0.75
Reception of Linguistic Gaps	0.79
Reception of Cultural Gaps	0.72
Trust in Auto-Generated Subtitles	0.81

The reliability scores indicated good internal consistency across constructs.

It is worth noting that the study was conducted in full compliance with the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board (IRB) of The Applied Science Private University, Jordan (Approval No. FOAH 3/2024, dated 17/01/2024).

Results and Discussion

Sample Characteristics

Demographic data of the 4500 participants are presented in Table 2.

Table 2: Demographic Summary of Participants (N = 4,500)

Variable	Category	Count	Percentage (%)
Age Group	18–24	1,530	34.00%
	25–34	1,395	31.00%
	35–44	945	21.00%
	45–54	450	10.00%
	55+	180	4.00%
Gender	Male	2,214	49.20%
	Female	2,187	48.60%
	Prefer not to say	99	2.20%
Primary Language(s)	Arabic	2,070	46.00%
	English	945	21.00%
	Both Arabic and English	1,485	33.00%
YouTube Viewing Frequency	Daily	2,520	56.00%
	A few times per week	1,350	30.00%
	A few times per month	405	9.00%
	Rarely	225	5.00%
Proficiency in English	Native/Fluent	1,170	26.00%
	Advanced	1,755	39.00%
	Intermediate	1,215	27.00%
	Beginner	360	8.00%
Proficiency in Arabic	Native/Fluent	2,835	63.00%
	Advanced	990	22.00%
	Intermediate	495	11.00%
	Beginner	180	4.00%
Device Most Frequently Used	Smartphone	2,925	65.00%
	Laptop/Computer	990	22.00%
	Tablet	315	7.00%
	Smart TV	225	5.00%
	Other	45	1.00%

As Table 2 shows, the 4,500 participants represented a young, digitally active, and gender-balanced audience, with most aged 18–34 (65%) and nearly equal numbers of males and females. Linguistically, 46% were Arabic speakers and 33% bilingual, with high proficiency in Arabic (85%) and strong English skills (65%). Media habits showed that 56% used YouTube daily, and smartphones (65%) were the primary viewing device, reflecting modern, mobile-centered media engagement.

Analysis of the Questionnaire's Items

In Table 3, frequencies and percentages of the questionnaire items are calculated.

Table 3: Descriptive Analysis of the questionnaire's items

#	Questionnaire Item	Agree (%)	Neutral (%)	Disagree (%)
	Media Consumption Habits			
1	I watch YouTube regularly as part of my daily routine.	72.4	17.3	10.3

2	I mostly watch entertainment content such as movies, music, or comedy on YouTube.	64.1	21.6	14.3
3	I mainly use YouTube for educational purposes such as tutorials, lectures, or language learning.	41.7	29.5	28.8
4	I often enable subtitles while watching YouTube videos.	59.8	23.1	17.1
Use of Subtitles				
5	I prefer Arabic subtitles when watching YouTube videos.	44.5	28.1	27.4
6	I prefer English subtitles when watching YouTube videos.	39.2	31.4	29.4
7	I usually keep subtitles turned on when watching YouTube videos.	55.7	25.6	18.7
8	I often switch between Arabic and English subtitles depending on the content.	48.9	27.7	23.4
9	I rarely watch YouTube videos without subtitles.	37.4	26.8	35.8
Motivations for Subtitle Use				
10	I use subtitles to better understand content in a language I am not fluent in.	62.5	24.7	12.8
11	I use subtitles to clarify unclear or poor-quality audio.	68.2	20.3	11.5
12	I use subtitles to follow speakers with different accents or dialects.	61.9	26.4	11.7
13	I use subtitles to learn or practice a foreign language.	54.3	28.1	17.6
14	I use subtitles when watching videos in noisy environments.	65.4	21.2	13.4
Reception of Linguistic Gaps				
15	I often notice mistranslations in English–Arabic auto-generated subtitles on YouTube.	71.1	18.6	10.3
16	I usually ignore subtitle errors if I can understand the context.	44.8	28.7	26.5
17	I find it frustrating when subtitles do not match the spoken dialogue.	68.9	19.1	12
18	Linguistic mistakes in subtitles reduce my overall enjoyment of the video.	63.7	22.4	13.9
19	I sometimes stop watching a video because of inaccurate subtitles.	36.2	27.5	36.3
Reception of Cultural Gaps				
20	I often notice when cultural expressions (e.g., idioms, sayings) are mistranslated in subtitles.	66.5	22.7	10.8
21	The loss of cultural meaning in subtitles makes it harder for me to fully understand the video.	62.1	25.6	12.3
22	When humor or jokes are mistranslated, it reduces my enjoyment of the video.	70.4	18.9	10.7
23	I sometimes miss important cultural references because they are not conveyed in the subtitles.	59.8	28.5	11.7
24	I feel that auto-generated subtitles often fail to capture the cultural context of the video.	64.2	24.6	11.2
Trust in Auto-Generated Subtitles				
25	I generally trust the accuracy of YouTube's auto-generated subtitles.	41.6	29.2	29.2
26	I believe auto-generated subtitles provide a reliable understanding of the video content.	45.3	31.4	23.3
27	I only use auto-generated subtitles when no better option is available.	54.7	25.6	19.7
28	I trust human-translated subtitles more than auto-generated subtitles.	72.8	18.4	8.8
29	My trust in subtitles depends on the type of content (e.g., educational vs. entertainment).	68.5	21.2	10.3

The analysis of responses from 4500 participants revealed that the majority of them used YouTube regularly on

a daily basis. 64.1% reported that they used YouTube for entertainment purposes i.e. to watch movies, comedy shows, and series in addition to listening to music. 41.7% use this platform for educational and learning purposes, i.e., to attend lectures and learn languages. Of the 4500 respondents, 59.8% stated that they enable autogenerated subtitles while using YouTube.

In the second construct concerning the use of subtitles, 44.5% of the study population preferred choosing Arabic subtitles, while 39.2% tended to select English subtitles when watching videos on the platform. 48.9% reported that they switched between the languages based on the content of the videos. 55.7% stated that they keep the subtitles on while watching YouTube videos, while 37.4% mentioned that they rarely watch audiovisual content on YouTube without subtitles.

The survey results also indicate several reasons why participants choose to use subtitles when they watch videos on YouTube. Around 62.5% have reported that they use subtitles to improve their understanding of the content presented in a language they are not very fluent in. Additionally, 68.3% of the respondents stated that the subtitles help them clarify incomprehensible or poor-quality audio, while 61.9% stated that they use subtitles to follow speakers with different accents or dialects. Moreover, 54.3% have mentioned that they use subtitles to better learn or practice a foreign language, and 65.4% have reported that doing so in noisy environments was manageable.

In terms of receiving the linguistic gaps, around 71.1% of the participants have reported that they notice mistranslations in these subtitles. Despite this, 44.8% stated that they usually ignore such subtitle errors as long as they can still understand the overall context. However, 68.9% of the participants expressed deep frustration when the displayed subtitles did not match the spoken dialogue that they listened to, and 63.7% justified that such errors in subtitles reduce their overall enjoyment of the video. In addition, 36.2% have stated that they sometimes stop watching the video when the subtitles are both inaccurate and imprecise.

As for the respondents' reception of the cultural gaps, the results suggest that viewers are highly aware of whether the cultural sensitivity elements were discarded in subtitles. Almost 66.5% of the respondents have reported that they usually can notice when culturally loaded expressions, such as idioms, proverbs, and sayings, are being mistranslated. Therefore, 62.1% have maintained that such mistranslations complicate things and make it too difficult for them to fully understand the video. Likewise, 70.4% of the participants have stated that their enjoyment gradually fades away when humor or jokes are mistranslated. Moreover, 59.8% of the respondents mentioned that they sometimes miss important cultural references because these are discarded in the subtitles, and around 64.2% of the respondents agreed that auto-generated subtitles often fail to capture the fullness of the cultural context of the video.

Finally, in terms of their trust in auto-generated subtitles, the findings show a wide array of mixed levels of trust in YouTube's auto-generated subtitles. Almost 41.6% of the respondents have shown that they generally trust the accuracy and precision of these subtitles, and almost 45.3% have noted that auto-generated subtitles provide a reliable understanding of the content of the video. Around 54.7% of the respondents have reported that they use auto-generated subtitles only when there are no other better options. A significant majority that constitutes a total of 72.8% have explicitly mentioned that they trust human-translated subtitles more than any auto-generated ones. Furthermore, around 68.5% of the respondents have stated that their level of trust varies depending on the type of content (i.e., whether the video is educational or for entertainment purposes).

Conclusion

The study examined the responses of 4500 participants from different age groups, as well as linguistic and cultural backgrounds to the use and effectiveness of YouTube auto-generated subtitles. It explores whether these subtitles are linguistically and culturally appropriate from their point of views. The respondents reported that although they have some concerns about the reliability of translation, they still use them because they make the content of videos accessible and comprehensible. The participants stated that when it comes to religious and culturally sensitive videos as well as the ones with educational content, they prefer resorting to subtitles created by humans over the ones created by the platform. Overall, the respondents believe that the use of YouTube generated subtitles is useful and helpful but they still need some improvement to better meet the expectations of users.

The study recommends that further research be conducted on different Arabic vernaculars to check

the accuracy of YouTube auto-generated subtitling services, when processing videos produced with different dialects. Researchers can also investigate different languages or language pairs to check the quality of the service when examining Latin languages when compared to Semitic languages. Further studies can also compare and contrast the translations created by humans with those generated by the machine, with special focus on how humour and figurative language, for example, are rendered. The study recommends that linguists and translators work hand in hand with AI developers to build better auto-subtitling systems or improve the existing ones.

References

Abulawi, F., Al Salman, S., & Haider, A. S. (2022). Modern Standard Arabic vs. Egyptian Vernacular in dubbing animated movies: A case study of A Bug's Life. *The International Journal of Communication and Linguistic Studies, 21*(1), 125-141. doi:https://doi.org/10.18848/2327-7882/cgp/v21i01/125-141

Ahmad, M., Haider, A. S., & Saed, H. (2025). Assessing AI-driven dubbing websites: Reactions of Arabic native speakers to AI-dubbed English videos in Arabic. *Research Journal in Advanced Humanities, 6*(1), 1-20. doi:https://doi.org/10.58256/64dz1c72

Akasheh, W. M., Haider, A. S., Al-Saideen, B., & Sahari, Y. (2024). Artificial intelligence-generated Arabic subtitles: insights from Veed. io's automatic speech recognition system of Jordanian Arabic. *Texto Livre, 17*, e46952-e46952. doi:10.1590/1983-3652.2024.46952.

Al-Abbas, L. S., Haider, A. S., & Saideen, B. (2022). A quantitative analysis of the reactions of viewers with hearing impairment to the intralingual subtitling of Egyptian movies. *Heliyon*, e08728. doi:https://doi.org/10.1016/j.heliyon.2022.e08728

Al-Darabee, M. (2024). *Human and Machine-generated Arabic Subtitles: A case study of Taboo Expressions in the English Movie "The Wolf of Wall Street"*. (Unpublished MA Thesis). Applied Science Private University, Amman-Jordan.

Al-Dhuhli, R., Alkindi, L., & Al-Taani, W. (2022). Digital Technology and its Role in Achieving Sustainable Development Requirements in Basic Education Schools in the Sultanate of Oman. *Jordan Journal of Applied Science Humanities Science Series (JJOAS), 31*(1), 43-62. doi:https://doi.org/10.35192/jjoas-h.v31i1.282

Aldualimi, M., & Almahasees, Z. (2022). A Quantitative Analysis of the Reactions of Viewers to the Subtitling the Arabic Version of Ar-Risalah's Movie" The Message" in English. *Journal of Southwest Jiaotong University, 57*(6), 1053-1063. doi:https://doi.org/10.35741/issn.0258-2724.57.6.91

Alkhatib, R., & Haider, A. S. (2024). Subtitling Legal Expressions in English Series into Arabic by Netflix, Machine, and Artificial Intelligence. *Pakistan Journal of Criminology, 14*(4), 513-528. doi:https://doi.org/10.62271/pjc.16.4.513.528

Allam, A. (2023). Quality Assessment of Interlingual YouTube Auto-generated Closed Captions in Some Crime Narratives Applying the NTR Model. *Textual Turnings: An International Peer-Reviewed Journal in English Studies, 5*(2), 59-83.

Almahasees, Z. (2017). Machine Translation Quality of Khalil Gibran's The Prophet. *AWEJ for translation & Literary Studies, 1*(4), 151-159. doi:10.2139/ssrn.3068518

Almahasees, Z. (2018). Assessment of Google and Microsoft Bing translation of journalistic texts. *International Journal of Languages, Literature and Linguistics, 4*(3), 231-235.

Almahasees, Z. (2021). *Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr*. London: Routledge.

Almahasees, Z., & Jaccomard, H. (2020). Facebook Translation Service (FTS) Usage among Jordanians during COVID-19 Lockdown. *Advances in Science, Technology and Engineering Systems Journal, 5*(6), 514-519. doi:10.25046/aj050661

Alrousan, F., & Haider, A. S. (2022). Dubbing television advertisements across cultures and languages: A case study of English and Arabic. *Language Value, 15*(2), 54-80. doi:https://doi.org/10.6035/languagev.6922

Altunisik, E., Firat, Y. E., & Keceli, Y. K. (2022). Content and quality analysis of videos about multiple sclerosis on social media: the case of YouTube. *Multiple Sclerosis and Related Disorders, 65*, 104024.

Alzaabi, S., & Rabab'ah, G. (2023). Conceptual metaphors in YouTube auto-generated subtitles: BBC travel show as a case. *The International Journal of Humanities Education, 22*(1), 57.

Ammari, R., & Al-Ahmad, W. (2023). A corpus-based study of English synonyms: Small, little, tiny, and petite. *Jordan Journal of Applied Science-Humanities Series, 36*(2), 1-20. doi:https://doi.org/10.35192/jjoas-h.v36i2.542

Anusuya, M. A., & Katti, S. K. (2009). Speech recognition by Machine: A Review. *(IJCSIS) International Journal of Computer Science and Information Security, 9*. Retrieved from https://arxiv.org/ftp/arxiv/papers/1001/1001.2267.pdf

Baker, J., Deng, L., Khudanpur, S., Lee, C., Glass, J., & Morgan, N. J. M. R. o. t. S. U. W. G., NIST. (2007). Historical development and future directions in speech recognition and understanding. *2006*.

Banikalef, A. E., & Abu Naser, J. (2019). The Difficulties in Translating Culture-Specific Expressions from Arabic into English. *Journal of Education and Practice, 10*(9), 78-82.

Chan, W. S., Kruger, J.-L., & Doherty, S. J. L. A., New Series–Themes in Translation Studies. (2019). Comparing the impact of automatically generated and corrected subtitles on cognitive load and learning in a first-and second-language educational context. *18*.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334. doi:https://doi.org/10.1007/bf02310555

Dharmale, G. J., & Patil, D. D. (2019). Evaluation of Phonetic System for Speech Recognition on Smartphone. *International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8*(10), 3354-3359. doi:10.35940/ijitee.J1215.0881019

Dhawan, K., Rekesh, K., & Ginsburg, B. (2023). *Unified Model for Code-Switching Speech Recognition and Language Identification Based on Concatenated Tokenizer.* Paper presented at the Proceedings of the 6th Workshop on Computational Approaches to Linguistic Code-Switching.

Diaz-Cintas, J., & Remael, A. (2014). *Audiovisual translation: subtitling.* London: Routledge.

Díaz-Cintas, J. J. T. T. (2005). Audiovisual Translation Today–A question of accessibility for all. *4*, 3-5.

Du, J., & Lu, J. (2024). *A Comparative Study on the Translation Quality between Human and Machine-Generated Subtitles.* Paper presented at the 2024 6th International Conference on Natural Language Processing (ICNLP).

Farghal, M. (2019). *Jordan's Proverbs as a Window into Arab Popular Culture: The Fox in the Blackberries.* Cambridge: Cambridge Scholars Publishing.

Ferguson, C. A. (1959). Diglossia. *WORD, 15*(2), 325-340. doi:10.1080/00437956.1959.11659702

Fraihat, B. (2024). The Effect of Implementing a Remedial Plan Based on Reading Strategies on Improving the Reading Comprehension Achievement of Jordanian EFL Eleventh-Grade Students in the Vocational Stream. *Jordan Journal of Applied Science-Humanities Series, 39*(1), 18-26. doi:https://doi.org/10.35192/jjoas-h.v39i1.729

Gottlieb, H. (1997). *Subtitles, Translation & Idioms.* University of Copenhagen., Copenhagen.

Guskaroska, A. (2019). *ASR as a tool for providing feedback for vowel pronunciation practice.* Iowa State University,

Haider, A. S., & AlKhatib, R. (2025). Evaluating Netflix, Machine, and Artificial Intelligence Subtitling of Arabic Legal Expressions into English. *Comparative Legilinguistics*.

Haider, A. S., Alzghoul, O., & Hamadan, Y. (2023). Creating and Experimenting a New Parallel Corpus of English-Arabic Subtitles of Culinary Shows: A Useful Guide for Translating Food Across Cultures. *World Journal of English Language, 13*(8), 358-372. doi:https://doi.org/10.5430/wjel.v13n8p358

Haider, A. S., & Hussein, R. F. (2022). Modern Standard Arabic as a Means of Euphemism: A Case Study of the MSA Intralingual Subtitling of Jinn Series. *Journal of Intercultural Communication Research, 51*(6), 628-643. doi:https://doi.org/10.1080/17475759.2022.2106289

Haider, A. S., Sahari, Y., Saed, H., & Tair, S. A. (2025a). Enhancing Accessibility for Saudi Deaf and Hard-of-Hearing Audiences: Sound Representation in SDH Subtitling of Saudi Series. *Journal of Disability Research*.

Haider, A. S., Saideen, B., & Hussein, R. F. (2023). Subtitling taboo expressions from a conservative to a more liberal culture: The case of the Arab TV series Jinn. *Middle East Journal of Culture Communication Research, 16*(4), 363–385. doi:https://doi.org/10.1163/18739865-tat00006

Haider, A. S., & Shohaibar, R. S. (2024). Netflix English subtitling of idioms in Egyptian movies: challenges and strategies. *Humanities & Social Sciences Reviews, 11*(1), 1-13. doi:https://doi.org/10.1057/s41599-024-03327-4

Haider, A. S., Tair, S. A., & AlBkour, N. (2025b). From sight to sound: Characterising physical features and body shapes in Netflix audio description *Training, Language and Culture*.

Hannun, A. J. a. p. a. (2021). The history of speech recognition to the year 2030.

Hashish, R., & Hussein, R. F. (2022). Strategies Subtitlers Use in Rendering English Slang Expressions Into

Arabic. *Theory and Practice in Language Studies, 12*(4), 752-762. doi:https://doi.org/10.17507/tpls.1204.16

Hassan, S., & Haider, A. S. (2024). Options for Subtitling English Movie Lyrics into Arabic. *Studies in Linguistics, Culture and FLT, 12*(1), 82-104. doi:https://doi.org/10.46687/LHTD6033

Hussein, A., Chowdhury, S. A., Abdelali, A., Dehak, N., Ali, A., & Khudanpur, S. (2023). *Textual data augmentation for Arabic-English code-switching speech recognition*. Paper presented at the 2022 IEEE Spoken Language Technology Workshop (SLT).

Ja'afreh, N. (2023). A pragmatic contrastive analysis of apology strategies in Jordanian arabic and english language. *Jordan Journal of Applied Science-Humanities Series, 37*(2), 69-86. doi:https://doi.org/10.35192/jjoas-h.v37i2.601

Juang, B.-H., Rabiner, L. R. J. G. I. o. T. A. R. U., & Barbara, t. U. o. C. S. (2005). Automatic speech recognition–a brief history of the technology development. *1*, 67.

Khalil, H. E. (2022). Translation of Culture-Specific Fixed Expressions and Swear Words in Some Arabic-into-English Subtitled Films. *Journal of the Faculty of Arts Port Said University, 19*(19), 1-46.

Khatatbah, D., & Ashour, M. (2024). The Role of School Administration in Activating the Use of Information Technology at Schools in Irbid Governorate. *Jordan Journal of Applied Science-Humanities Series 38*(2), 121-137. doi:https://doi.org/10.35192/jjoas-h.v38i2.718

Liao, J., Eskimez, S., Lu, L., Shi, Y., Gong, M., Shou, L., . . . Processing, L.-R. L. I. (2023). Improving readability for automatic speech recognition transcription. *22*(5), 1-23.

Liu, D., Niehues, J., & Spanakis, G. (2020). Adapting end-to-end speech recognition for readable subtitles. *arXiv preprint arXiv:.12143*.

Liu, Z., Spence, J., & Prud'Hommeaux, E. (2023). *Studying the impact of language model size for low-resource ASR*. Paper presented at the Proceedings of the Sixth Workshop on the Use of Computational Methods in the Study of Endangered Languages.

Masoud, S. (2025). Investigating Teachers' Attitudes toward Using Assistive Technology in Teaching Students with Hearing Impairments. *Jordan Journal of Applied Science-Humanities, 43*(2), 157-171. doi:https://doi.org/10.35192/jjoas-h.v43i2.1693

Meqdad, S., Al-Bayyari, M., & Al-Taher, M. (2023). Semantic Disparity between Amr (Command) and Nahy (Negative Command) in the Noble Quran: An Interpretive/Contextual Study. *Jordan Journal of Applied Science-Humanities Series, 36*(1), 1-10. doi:https://doi.org/10.35192/jjoas-h.v36i1.516

Mustafa, M. B., Yusoof, M. A., Khalaf, H. K., Rahman Mahmoud Abushariah, A. A., Kiah, M. L. M., Ting, H. N., & Muthaiyah, S. J. A. S. (2022). Code-Switching in Automatic Speech Recognition: The Issues and Future Directions. *12*(19), 9541.

Naib, A. (2025). Algerian Viewers' Attitudes toward Taboo Language: A Case Study of Mourad Oudia's Humorous YouTube Content. *Jordan Journal of Applied Science-Humanities Series, 44*(1), 167-191. doi:https://doi.org/10.35192/jjoas-h.v44i1.1830

Nida, E. A., & Taber, C. R. (1982). *The Theory and Practice of Translation* (Vol. 8). Leiden, Boston: Brill.

Obeidat, M. M. (2023). Translating Culture in the Jordanian TV Comedy Series "al jar gabl al dar"(My American Neighbor) Into English. *SAGE Open, 13*(3), 21582440231197011. doi:https://doi.org/10.1177/21582440231197011

Peacocke, R. D., & Graf, D. H. (1995). An introduction to speech and speaker recognition. In *Readings in Human–Computer Interaction* (pp. 546-553): Elsevier.

Pfender, E. J., Wanzer, C., & Bleakley, A. (2024). A content analysis of social media influencers' "What I Eat in a day" Vlogs on YouTube. *Health Communication, 39*(11), 2244-2255.

Rahmatika, R., Yusuf, M., & Agung, L. (2021). The effectiveness of YouTube as an online learning media. *Journal of Education Technology, 5*(1), 152-158.

Romero-Fresco, P., & Fresno, N. (2023). The accuracy of automatic and human live captions in English. *Linguistica Antverpiensia, New Series–Themes in Translation Studies, 22*, 114–133. doi:https://doi.org/10.52034/lans-tts.v22i.774

Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking skills of Jordanian EFL university students. *Heliyon, 7*(7).

Saideen, B., Thalji, A., Lababaneh, A., Tartory, R., Haider, A. S., & Obeidat, M. (2024). Netflix English subtitling of the Jordanian movie 'The Alleys': Challenges and strategies. *Research Journal in Advanced Humanities, 5*(1), 157-177. doi:https://doi.org/10.58256/sptxps94

Tsoukala, C., Kritsis, K., Douros, I., Katsamanis, A., Kokkas, N., Arampatzakis, V., . . . Pavlidis, G. (2023). *ASR pipeline for low-resourced languages: A case study on Pomak*. Paper presented at the Proceedings of the Second Workshop on NLP Applications to Field Linguistics.

Versteegh, K. (2014). *Arabic language*: Edinburgh University Press.

Watson, J. C. E. (2012). 50. Arabic Dialects (general article). In W. Stefan (Ed.), *The Semitic Languages* (pp. 851-896). Berlin, Boston: De Gruyter Mouton.

Xie, B. (2022a). *A comparative study of machine translated subtitles based on the user-centered approach: a case study between Bilibili and YouTube*. Research Square. Durham, North Carolina. Retrieved from https://www.researchsquare.com/article/rs-2179598/v1

Xie, B. (2022b). A comparative study of machine translated subtitles based on the user-centered approach: a case study between Bilibili and YouTube.

Yu, H., Hu, Y., Qian, Y., Jin, M., Liu, L., Liu, S., . . . Zeng, M. (2023). *Code-Switching Text Generation and Injection in Mandarin-English ASR*. Paper presented at the ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).