

PAPER

An AI-Powered Interactive Mobile Application for Multilingual Conference Accessibility: Design, Deployment, and Field Evaluation

Ameen Shaheen  (✉),
Wael Alzyadat,
Aysh Alhroob 

Al-Zaytoonah University of
Jordan, Amman, Jordan

a.shaheen@zuj.edu.jo

ABSTRACT

Global academic conferences often face accessibility challenges for multilingual participants, especially during live sessions where real-time comprehension is important. This study presents the design, deployment, and evaluation of an AI-powered multilingual mobile application developed to support conference accessibility. The application integrates real-time speech translation, agenda management, venue navigation, and cultural information in a single mobile platform. It was deployed during the International Conference on Information Technology (ICIT 2025) and evaluated with 200 participants across four supported languages: Arabic, English, French, and Spanish. Using a design-based research (DBR) approach, the evaluation combined objective performance metrics with user-reported usability measures. The system achieved 92.5% translation accuracy, a word error rate (WER) of 7.2%, and system uptime of 99.1%. The usability results showed high participant satisfaction, with a mean score of 4.6 out of 5 and strong internal reliability (Cronbach's $\alpha = 0.91$). Open-ended feedback indicated that participants valued translation usefulness, ease of use, and the integration of conference services, while also suggesting additional language support and offline functionality. The findings suggest that integrating AI-based multilingual translation with mobile conference services can help reduce language barriers and improve participant experience in international academic events.

KEYWORDS

artificial intelligence (AI), interactive mobile application, conference accessibility, real-time speech translation, multilingual communication, usability evaluation

1 INTRODUCTION

International academic conferences are a vital platform for exchanging knowledge, collaborating and socializing with other researchers, practitioners and students around the world [1] [2]. The international nature of the event does create

Shaheen, A., Alzyadat, W., Alhroob, A. (2026). An AI-Powered Interactive Mobile Application for Multilingual Conference Accessibility: Design, Deployment, and Field Evaluation. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(18), pp. 138–157. <https://doi.org/10.3991/ijim.v20i18.62706>

Article submitted 2026-06-21. Revision uploaded 2026-07-21. Final acceptance 2026-07-22.

© 2026 by the authors of this article. Published under CC-BY.

some challenges of accessibility, particularly where there is diversity of languages and cultures [3]. Language differences can make it difficult for those attending the event to interpret or understand what is being said, participate in discussions, ask questions and participate in social networking events [4]. Therefore, providing support for multilingual communication has become one of the most important requirements for having inclusive academic conferences.

This challenge is even greater at technical and science conferences due to the fact that many of the presentations will contain specific terms related to the subject matter and a lot of interaction between presenters and attendees [5]–[7]. The participants who are not fluent in the primary language used by the majority of the conference attendees could be limited in their ability to fully comprehend the information provided during the presentation and as such would have reduced real time participation in the overall process of learning from their peers.

Recent advancements in artificial intelligence (AI), natural language processing (NLP) [8], [9] and speech-to-text capabilities are enabling real time translation services to be used to aid in overcoming language barriers. Most current AI based translation platforms utilize deep learning techniques to train on large volumes of multilingual data sets; therefore, they are able to generate translations that represent both the literal context of speech and the intended meaning behind it [10], [11]. Platforms such as Google Translate and Microsoft Translator demonstrate how AI technology can be utilized for multilingual interaction. Although there is great promise for AI in aiding in multilingual communication, its adoption in real time applications such as live academic conferences has been limited due to issues such as delay, lower than expected levels of accuracy for conference-specific terms, and no integration into existing event related applications [12], [13].

Most current AI-powered translation solutions are focused on translation and therefore lack full conference support. Attendees need to access not only the schedule for presentations but also speaker bio's, maps of the event space and local assistance especially when attending an event in another city [14]. Since there is no single mobile application that can integrate all of these features, attendees usually end up using a number of separate applications or printed material. This adds further cognitive load and negatively impacts overall user experience [15].

This study aims to develop and test an AI-based mobile app that is available in multiple languages to facilitate communication among participants of ICIT 2025, a major international conference hosted by the Faculty of Science and Information Technology at Al-Zaytoonah University of Jordan. The app is designed to provide a real-time translation service in four languages: Spanish, French, English, and Arabic. It will provide conference attendees with an easy-to-use interface to view the current events that are taking place, get speaker bios, use interactive maps to navigate the venues, and find out about local popular tourist attractions and cultural destinations.

The significance of this study lies in its development of an application based upon the needs of conference attendees; it was intended to be used as a venue-specific translation tool within an academic conference environment. Following implementation into the conference environment, usability, clarity of translations, perceived responsiveness, and overall user satisfaction were assessed using a bilingual English/Arabic survey.

The contribution of this study is mainly in the design and field evaluation of an integrated conference application. The application combines real-time translation with services needed during the event, such as schedules and conference information. The study also uses feedback from actual attendees to understand how the

application supported access and participation, and it identifies design points that can inform future applications in similar settings.

The main objectives of this study are thus to:

1. Evaluate the effectiveness and usefulness of the AI-driven live translation function in reducing language barriers at ICIT 2025.
2. Assess the role of integrated conference support features in improving the overall participant experience; and
3. Provide evidence-based design recommendations for future AI-driven conference mobile applications.

2 RELATED WORK

2.1 Artificial intelligence-powered mobile translation applications for live events

A rise in attendance at conferences has resulted in a greater demand for multilingual communication tools capable of translating in real time; therefore, using AI-technology to create mobile translation applications has provided an opportunity to assist with face-to-face communication [16]. AI regulation has also received research attention [17]. Many early versions of mobile translation tools were based upon typing, or optical character recognition (OCR), whereas current versions utilize automatic speech recognition (ASR) and neural machine translation (NMT) which enable voice-to-text as well as voice-to-voice translation [18], [19].

Prior research indicates that translation accuracy alone does not guarantee a favorable user-experience. For example, [20] investigated auto-generated captions used for deaf or hard-of-hearing users and determined that both the readability of the captions, the amount of time it took to generate them, and the ability of the user to adjust their captioning options greatly affected how they would be able to use the application. Similarly, [21] assessed the usability of a mobile translation application being utilized in a medical communications course and noted that in order to provide an adequate user-experience, the application should have easy to use navigation options, flexibility with regard to controlling how to access features offered by the app, and sufficient languages available. In this respect, research assessing usability issues associated with the aforementioned apps will serve as a starting point when developing the ICIT 2025 mobile conference-support application which utilizes AI-based technology to translate in real time from Arabic, English, French, and Spanish.

The ICIT 2025 application builds on these findings by integrating real-time AI-powered translation in Arabic, English, French, and Spanish within a broader mobile conference-support system. Unlike standalone translation tools, the application combines linguistic support with event-related services, including schedules, venue guidance, and cultural information.

2.2 Real-time speech translation

The quality of the translation for live events is as fast as a response time [22]. Simultaneous speech translation studies examined the ability to achieve both high levels of accuracy and low delays through the use of either a wait-k policy

or an adaptive approach to translate [23]–[25]. The most current speech-to-speech translation systems also utilize combinations of ASR, NMT, and TTS functions to provide nearly real-time multilingual capabilities [26].

Prior studies focused mainly on computer based measures such as avg. latency, or translated words. Users perception of delay can also depend upon internet connectivity, devices capabilities, and application reaction speed [27]. Therefore, delay was examined in this study with the users' point of view in addition to the measured system latency via the "perceived responsiveness" items in the post use survey.

2.3 Multilingual ASR as an enabling technology

Robust ASR is an important element in mobile translation systems since it has to understand spoken language across all languages, accents, and backgrounds with a lot of noise and specific terms used for each domain [28]. More recent ASR models are improving performance when recognizing spoken words in noisy environments and with multiple languages being spoken at once [29]–[31]. Research studies about automated captions are showing that transcription quality needs to be aided by good display design and user control features to create a successful accessible experience [32]–[34].

ICIT 2025's deployment will extend the scope of this area of research, since we evaluate ASR and translation as embedded into a mobile conference app. The conditions of this study include many technical terms, various speakers, and real-time interactions between presenters and participants. Thus, it can test how well the integrated system functions.

2.4 Mobile application usability in event-like contexts

Mobile usability studies have emphasized simplicity of navigation, simplicity of information architecture and a minimal amount of cognitive load on users [35]. For example, tourism related applications, museum applications and location-based applications, demonstrate how mobile systems can be used to enhance user experiences with mobile systems by using contextualized and layered ways to present information [36]–[38]. It seems reasonable to conclude that many of the same concepts will apply to conference based applications, which require attendees to switch back and forth between various functions such as translation, schedules, maps, and speaker information.

While most available applications and studies focus upon a single major function or service area such as translation, captioning, navigation, or cultural information. In contrast, the ICIT 2025 application integrates these functions in a single mobile interface. The evaluation, therefore, examines both specific functions, such as translation clarity, and overall system usability, such as ease of navigation and switching between services.

Previous studies have made significant progress in several areas including AI translation, real-time speech recognition and transcription, multilingual ASR and mobile usability. Unfortunately, while each of the aforementioned advances has been separately explored, they typically have not been examined together. Many commercially available solutions such as Google Translate and iTranslate offer generalized translation capabilities but they are not designed around the workflow

of academic conferences and do not integrate conference services such as agenda management, speaker information, and venue navigation [39], [40].

Similarly, prior research on providing captions or translating spoken content has generally focused upon educationally controlled environments or environments with accessibility-orientated goals [20], [21]. This study addresses this gap by presenting and evaluating an integrated AI-powered mobile application designed for a live academic conference. Table 1 summarizes the position of this work in relation to commercial translation tools and representative prior studies.

Table 1. Comparison with existing translation and accessibility solutions

Factor	Commercial Apps [39], [40]	Prior Studies [20], [21]	This Work
Purpose	General translation	Captioning/translation usability	Conference accessibility
Function	Standalone translation	Component-level evaluation	Real-time translation with conference services
Integration	Not conference-oriented	Limited	Agenda, speakers, maps, voting, cultural info
Context	Product-level use	Controlled settings	Field deployment at ICIT 2025
Evidence	Usually not research-based	Usability/feasibility evidence	200 users, logs, questionnaires, accuracy, WER, uptime

Table 1 shows that the main distinction of this study is not the translation function alone, but the integration of AI-powered multilingual translation with conference-support services and its evaluation in a live academic conference setting.

3 METHODOLOGY

This study followed a design-based research (DBR) approach and included a usability evaluation of the multilingual mobile application. DBR was appropriate because the application was developed for an actual conference setting, where participants used it during the event while attending sessions and accessing conference services [41]. It followed a structured process that involved participant recruitment, application deployment, usage logs, data collection from post-use questionnaires, and both qualitative and quantitative analyses.

3.1 System architecture and AI translation pipeline

The proposed was developed for Android-based smartphones that run on version 10 or later versions. Access to the app was provided through both the official ICIT 2025 website and by scanning a QR code available at the Registration Desk and Session Venues. The app utilized a pipeline of ASR, NMT and text rendering as part of its real-time translation process. User speech was captured via their smartphone's built-in microphone and then processed through the Google Cloud Speech-to-Text Streaming API. This resulted in a transcription of user speech into text. Google Cloud Translation API was then used to translate this text from one source language into the user specified target language and display it back to the user in almost real time.

Arabic, English, French and Spanish were supported languages during the duration of the conference. During the conference, the system collected logs containing the output of translations performed, errors related to those translations, and how long the system had been operating.

3.2 Participants and demographics

A total of 200 ICIT 2025 attendees participated in the evaluation. Attendees were selected to participate as they registered either at the desk during registration or logged onto the conference portal. The demographics of the attendees were diverse including students, university faculty and international attendees from various universities with different language skills. Table 2 summarizes the demographic profile of the participants.

To be eligible for participation, attendees had to have an Android smartphone compatible with the application. In addition, attendees had to have a basic knowledge of one of the languages that the application supported and provide their informed consent to participate. Eligible individuals with significant visual/hearing impairments that would limit the attendee's ability to use the application were excluded. post-hoc power analysis for a paired t-test using $\alpha = 0.05$ and a medium effect size ($d = 0.5$) indicated statistical power greater than 0.80 for $N = 200$ [42].

Table 2. Demographic characteristics of the participants (N = 200)

Demographic Variable	Category	Number (n)	Percentage (%)
Gender	Male	98	49%
	Female	102	51%
Age Range	18–30	84	42%
	31–40	78	39%
	40 and above	38	19%
Academic Major	Software Engineering	66	33%
	Artificial Intelligence	45	22.5%
	Cyber Security	36	18%
	Mathematics	28	14%
	Multimedia	25	12.5%

3.3 Research instruments

Four assessment techniques were used. First, a pre-test questionnaire was filled out by the participants to gather information about their previous language abilities, their experiences using translation devices, and their expectations from the use of an AI-based translation device. Next, application use logs recorded objective measures including: Translation accuracy, word error rate (WER), and system uptime. Afterward, bilingual Arabic/English post-use questionnaires were submitted by the participants to assess their degree of ease of use, clarity of translation, perceived

responsiveness, and satisfaction, as well as likelihood of recommending the tool to other users via a five-point Likert scale. Lastly, open-ended feedback forms permitted participants to describe the area in which the tool was effective, ineffective or provide suggestions on improving the device.

The questionnaire was examined by three subject matter experts; two in human computer interaction and one in linguistics to determine clarity, content validity and appropriateness for use among native speakers of both Arabic and English. Cronbach's alpha was calculated to assess internal reliability of the measures of each construct. All constructs met the acceptable minimum value of 0.70 [44].

3.4 Data collection procedure

The field evaluation was performed during the 12th International Conference on Information Technology (ICIT 2025) at Al-Zaytoonah University of Jordan. Participants downloaded the Application from the conference portal or via QR Codes that were placed throughout the Venue. Once installed, participants selected their desired interface language and used the Application in real time throughout each session to obtain live interpretation, conference schedule, speaker information, navigation to venue Areas, and cultural information.

For translation evaluation, 50 randomly selected segments for translation were obtained from different conference sessions and tracks after completion of the Event. Each segment was approximately 20–40 seconds long with varying levels of technical and general academic speech. Two bilingual annotators worked separately to create reference transcripts of the segments by transcribing them and resolving all discrepancies discussed among themselves. Translation accuracy was calculated as the proportion of correctly translated words, while WER was computed using the standard SCLITE scoring procedure [43]:

$$WER = \frac{\text{Substitutions} + \text{Deletions} + \text{Insertions}}{\text{Total Words in Reference}} * 100 \quad (1)$$

Translation accuracy was assessed at the semantic word/phrase level rather than by exact lexical matching. A translated unit was marked as correct when it preserved the intended meaning of the source utterance in the target language. Disagreements between the two bilingual annotators were discussed until consensus was reached. WER was used to evaluate recognition and transcription- level errors against the manually prepared reference transcript, while translation accuracy was used to evaluate the end-to-end translated output. The sampled segments were distributed across the main supported language pairs according to the actual usage patterns observed during the conference.

3.5 Performance and usability metrics

The evaluation combined objective system indicators and subjective user-reported measures. Translation accuracy, WER, and system uptime were obtained from system logs and manual scoring of sampled segments. Since direct system-level latency was not captured in the logs, delay was evaluated as perceived responsiveness through questionnaire items. Table 3 summarizes the main evaluation metrics.

Table 3. Evaluation metrics

Metric	Measurement Source
Translation Accuracy	Scored sampled translation segments
Word Error Rate (WER)	SCLITE-based comparison with reference transcripts
System Uptime	Application usage logs
Perceived Responsiveness	Post-use questionnaire
Ease of Use	Post-use questionnaire
Translation Clarity	Post-use questionnaire
Satisfaction	Post-use questionnaire
Recommendation Likelihood	Post-use questionnaire
User Feedback Themes	Open-ended responses

3.6 Data Analysis

Quantitative evaluation included descriptive statistics such as means, standard deviations, and percentage distributions for demographic variables and Likert-scale responses [45]. A paired-sample t-test was used to evaluate whether there was a statistically significant difference between the participants' confidence in using the application prior to versus subsequent to using it, while a one-way ANOVA was conducted to determine if there were any statistically significant differences in terms of participant satisfaction across academic majors [46]. Additionally, correlation coefficients were calculated to assess relationships between language proficiency and perceived usability.

Qualitative feedback from open-ended responses were coded thematically using NVivo software to provide an overview of thought patterns that respondents used when commenting about strengths, weaknesses and suggested improvement areas relating to the use of the application [47] and in conjunction with system log data and qualitative comments provided a balanced evaluation of both user experience and technical performance.

The study followed the ethical guidelines of Al-Zaytoonah University of Jordan. Participation was voluntary, and informed consent was obtained before data collection. No identifying personal information was collected. Questionnaire responses and system logs were anonymized before analysis and stored securely, with access limited to the research team.

4 RESULTS

The data provided in this section represent the empirical findings from the field evaluation study that occurred at ICIT 2025. Technical performance is examined as well as usability assessment, perceived responsiveness, translation quality, qualitative open-ended comments from users, and inferential statistics.

4.1 Technical performance metrics

System logs and sampled segments of translations were analyzed to assess the quality of translations, WER, and system uptime. Direct latency could not be measured using logs; therefore, a measure of time was used to evaluate delay based

on user-perceived responsiveness questions included in the post-use questionnaire. Table 4 shows an overview of technical performance results.

Table 4. Technical performance metrics of the translation system

Metric	M	SD	95% CI
Translation Accuracy (%)	92.5	3.1	[92.1, 92.9]
Word Error Rate (WER, %)	7.2	1.4	[7.0, 7.4]
System Uptime (%)	99.1	0.3	[99.06, 99.14]

Results clearly demonstrated how effectively the application could produce extremely accurate translations and low WER for live conferencing. The results showed on average that the translations had an accuracy rate of 92.5%. This indicates that in almost every single one of the sample sections, the translations accurately represented the source speech as well as the original content. Moreover, the system's WER score of 7.2% indicated that the number of errors in both transcribing and translating was minimal. Thus, these errors did not greatly impact understanding or knowledge retention of information presented during the conference. In addition, the system uptime was at a level of 99.1% which demonstrates that the system functioned continuously throughout the entire duration of the conference. Figure 1 shows the live translation screen used during the conference. Participants could choose the source and target languages and view the translated text in real time.

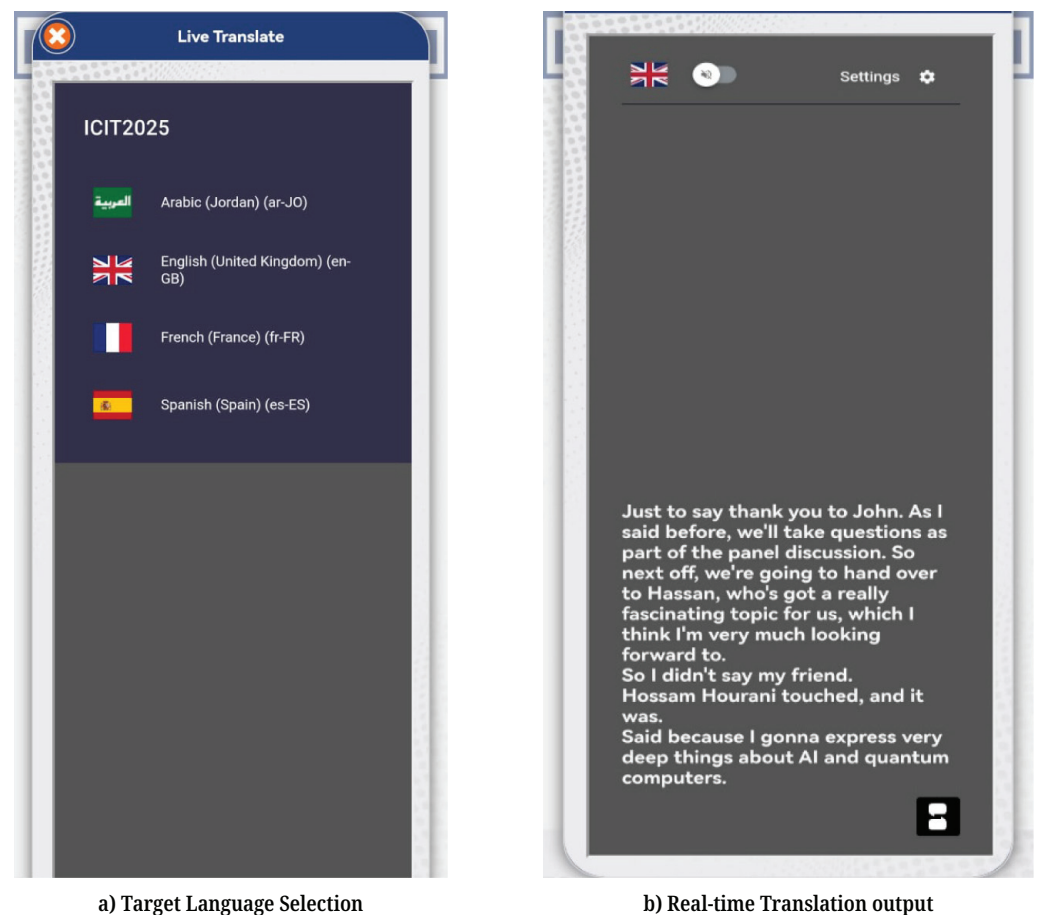


Fig. 1. Live translation interface: (a) language selection screen; (b) real-time translated output during live sessions

As shown in Figure 2, translation accuracy varied slightly across the evaluated language pairs. Arabic-English had the highest accuracy, while French-English had the lowest.

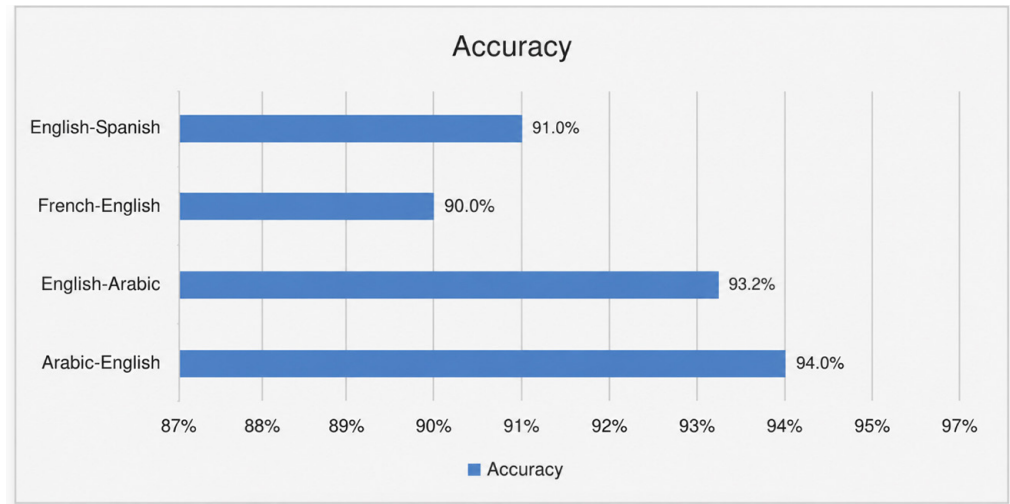


Fig. 2. Translation accuracy by language pair

The sampled segments were also reviewed for common translation errors. Table 5 shows the main error types and how they may affect comprehension.

Table 5. Classification of translation error types and impact on comprehension

Error Type	Example	Impact
Lexical Error	“Natural language processing” → “Natural language handling”	Low
Semantic Error	“Blockchain consensus mechanism” → “Blockchain agreement mechanism”	Medium
Technical Jargon Error	“Prompt engineering” → “Question engineering”	High
Sentence Truncation	Incomplete rendering of a long technical sentence	Medium
Idiomatic/Cultural Error	“break new ground” → “break new soil”	Low–Medium
Grammar/Syntax Error	“Researchers encourage to submit ...”	Low
Ambiguity Error	“Virtual and Augmented Reality” → “Virtual or Augmented Reality”	Medium
Domain-Outside Vocabulary	Incorrect translation of specialized biometric terms	High

4.2 Usability evaluation

The application provided real-time translation and the meeting agenda of speakers, navigation to the location, support contact, and culture. All these options were integrated into a single user interface for participant movement through different services at the same time as the conference. Figure 3 shows the main app displays that participants used during ICIT 2025.



Fig. 3. Application overview and navigation: (a) feature dashboard; (b) tracks screen; (c) support menu; (d) venue map

Figure 4 presents an overview of the agenda and participation tools that were contained in the app. The Agenda Screen was used by participants to check session times, venues, and related details during the conference. The Voting Screen was a different way for participants to participate in event activity through the app at the time they attended ICIT 2025.

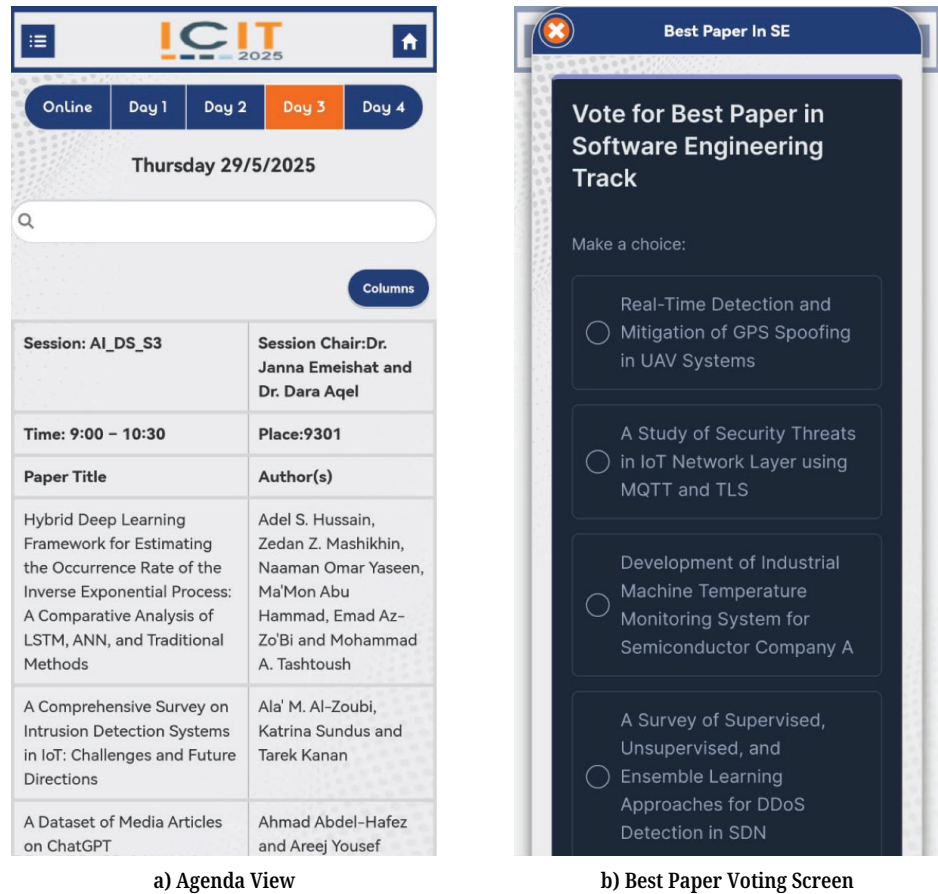


Fig. 4. Agenda and participation tools: (a) interactive agenda view; (b) best paper voting screen

The usability of the app was measured through the post-use questionnaire, on a 5-point Likert scale. Table 6 presents the descriptive statistics for each usability measure. Overall, the post-use questionnaire reported positive results to indicate that the system is useful. The system achieved its highest scores in terms of the users' perceptions about the system's usefulness; this was based upon recommendation likelihood ($M = 4.67$) then the user's perception of how easy it was to use the system ($M = 4.60$), and finally the perceived clarity of translations ($M = 4.55$) and overall user satisfaction ($M = 4.50$). All Cronbach's α values were over 0.86, indicating excellent reliability. Figure 5 shows the usability results.

Table 6. Descriptive statistics for post-use questionnaire constructs ($N = 200$)

Construct	α	M	SD	95% CI
Ease of Use	0.89	4.60	0.52	[4.53, 4.67]
Translation Clarity	0.86	4.55	0.58	[4.47, 4.63]
Satisfaction	0.91	4.50	0.50	[4.43, 4.57]
Recommendation	0.88	4.67	0.46	[4.60, 4.74]

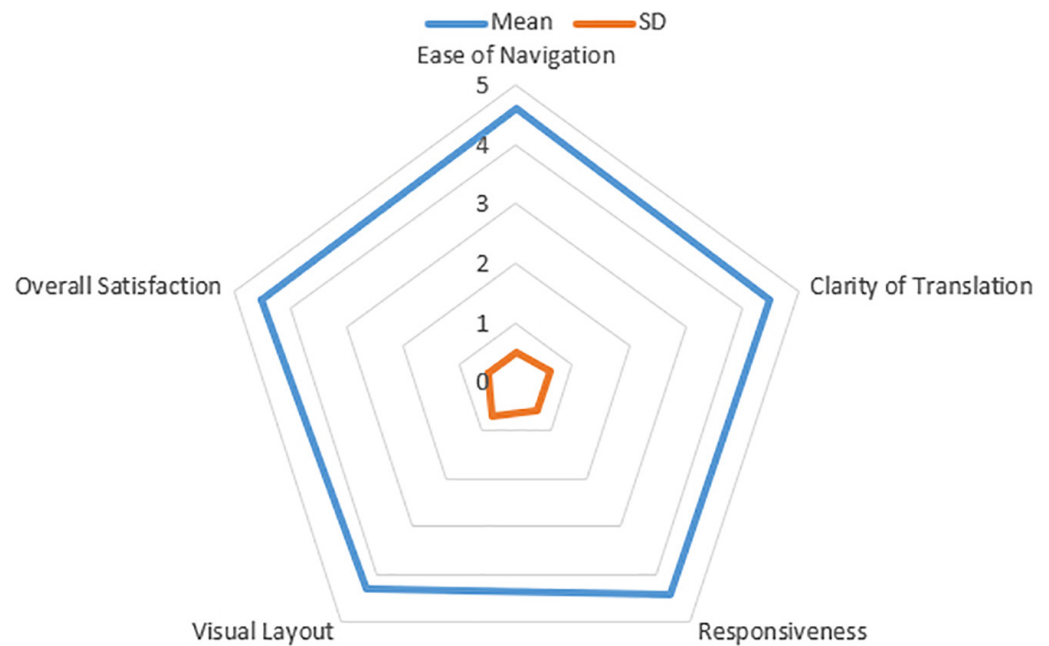


Fig. 5. Usability dimensions based on mean scores

4.3 Perceived responsiveness and translation quality

Participants generally were able to keep up with real-time session activity of the app. The application's responsiveness is reflected through participant feedback as opposed to actual measured technical latency that was logged within the system. Table 7 compares objective translation accuracy with subjective clarity ratings across representative supported language pairs.

Table 7. Descriptive statistics for post-use questionnaire constructs (N = 200)

Language Pair	Objective Accuracy (%)	Subjective Clarity (Mean Likert)
Arabic-English	94.0	4.6
English-Arabic	93.2	4.5
French-English	90.0	4.3
English-Spanish	91.0	4.4

The results in Table 7 show a consistent relationship between objective translation quality and participants' perceived clarity. Language pairs with higher objective accuracy also received higher subjective clarity ratings.

4.4 Thematic analysis of open-ended feedback

Open-ended feedback was organized into four categories: Positive User Experience, Enhancements, Interface Design Ideas, and Contextual Integration. As shown in Figure 6, Positive User Experience represented the largest number of responses because several users indicated that they valued the ability to translate information through this service, found it very easy to use, and appreciated having

access to multiple services for a single event from one application. Responses for enhancements included requests for additional language support and to enable offline access. Additional responses included increased text sizes, the option to customize displays, providing the ability to take notes while using the application, and integrating real-time question/answer capabilities.

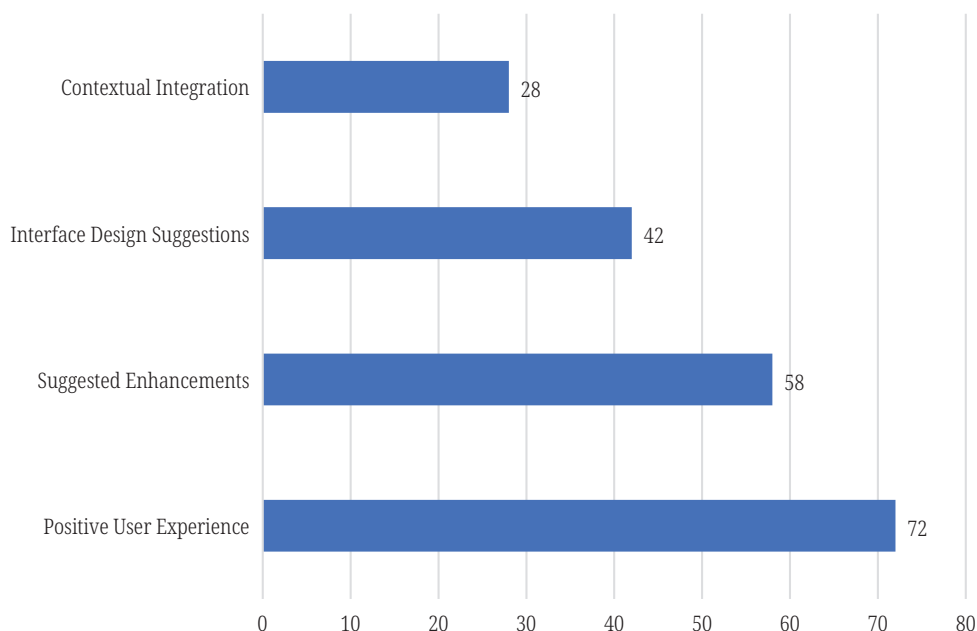


Fig. 6. Frequency of themes in open-ended feedback

4.5 Inferential statistical analysis

Tests of inference assessed whether confidence levels changed for all participants and whether there was a difference among the participants' groups. The paired samples t-test demonstrated an increased level of confidence from participants regarding their ability to use translation tools with the application compared to prior to its use ($M_{Pre} = 3.1$, $SD = 0.8$; $M_{Post} = 4.4$, $SD = 0.6$); $t(199) = 21.7$; $p < .0001$; Cohen's $d = 1.53$. The results indicate a substantial effect size and suggest that the application has increased the level of comfort of participants when addressing multilingual conference material.

The results from a one-way ANOVA indicated there were no statistical differences among academic major groups concerning their levels of satisfaction with the application, ($F[4,195] = 1.42$, $p = .23$; $\eta^2 = .02$). In addition to these findings, independent samples t-tests revealed no statistical difference existed between males and females on overall satisfaction with the application, ($t[198] = .87$; $p = .38$). These results suggest that the application was perceived positively across different participant groups.

A strong positive relationship existed between a user's self-assessment of their level of proficiency in English and how easy or difficult they perceived using this Application to be; $r(46) = .41$, $p < .01$. Users who self-identified as having greater levels of English ability indicated an even slighter margin in terms of usability. The fact that many users with low levels of English proficiency were satisfied with their experience indicates that the Application is both highly accessible and usable by individuals with differing levels of linguistic proficiency.

Overall, this study found the mobile app has high technical performance, is usable, and received very positive user responses to its testing in the field at the ICIT 2025 conference. Results from objective measures, questionnaire data, and open-ended comments indicate that providing multilingual translation along with support services provided by conferences through one mobile platform provides value.

5 DISCUSSION

The ICIT 2025 field test participants expressed a belief that the App would provide an efficient way to access conference event services. It provided a convenient location for users to obtain translations of conference sessions, view session details, and receive other required information. The convenience of having this information accessible made it easier for users to navigate through the language barriers that may exist at the conferences.

5.1 Effectiveness of AI-powered translation

The translation accuracy for the sample conference speech was found to be 92.5% with a WER of 7.2%. The data collected indicates that the system performed well on most of the sample speech in terms of translation quality. A review of the sample segments showed that the translated text was generally accurate and conveyed the major content of the original spoken academic information.

These findings extend previous work on captioning and mobile translation applications [20], [21] by evaluating AI-based translation in a live academic conference rather than in a controlled or classroom-based setting. However, the results also show that translation quality alone is not sufficient. Participants valued clarity, ease of navigation, and the ability to access conference information within the same mobile interface. This supports previous usability findings that accessibility tools must combine technical performance with simple and user-centered interaction design [33].

5.2 Usability and user experience

The study's results showed that all aspects of usability were extremely similar among each of the usability constructs, including ease of use, translation clarity, overall satisfaction with the program, and likelihood of recommending the app. Therefore, these findings show that by providing users with an integrated interface with the ability to perform tasks related to translation, schedule events, navigate the site, and access cultural information through one single app, it does not hinder user usability. Rather than hindering usability, the app's combined design appears to provide a better experience for the end user.

The inferential analysis results also indicate the usability of the application for the different groups. Satisfaction among all participants did not differ significantly based on gender or academic area. There was a positive correlation between how proficient a participant's English is and how easy they perceive the app to be to use. Nevertheless, participants with lower language proficiency also reported high satisfaction, suggesting that the translation feature helped reduce language-related barriers.

5.3 Confidence and engagement outcomes

The paired-samples t-test showed a significant increase in participants' confidence after using the application, with a large effect size. This indicates that the application not only supported immediate translation needs but also increased users' confidence in dealing with multilingual conference content.

Previous studies have also shown that captions and translation tools can help non-native users understand spoken content [32]. In this study, participants reported that real-time translation helped them follow the sessions and participate in conference activities. This shows that the application helped participants understand the content and take part in the event.

5.4 Design implications

Participants' comments also highlighted a number of aspects that could be addressed in future versions of the application. A few participants were pleased with having all of the translation features, schedule, venue information, and other conference services in the same application as part of this app. It was easier to use the app at conferences. Some participants also requested additional language support, to allow larger fonts, greater user control on how they wanted things displayed, and simply an easy way to take notes or submit questions during the session.

The results indicate that translation accuracy is a small aspect of an overall conference experience. Users utilized this app as much for translation as they did to look at the schedule, gather info, and travel between activities. Therefore, future versions should be designed with improved links between translation and the other features of the app, as well as provide better support for offline access and screen sizes.

5.5 Contributions, limitations, and future work

The study shows that a mobile app can be used for an academic conference to include participants from many languages. It includes two ways of using technology at conferences: Live Translation and Basic Conference Services. The evaluation was conducted during ICIT 2025 with 200 participants and included both system measures, such as translation accuracy, WER, and uptime, and user measures related to usability and satisfaction.

This may be useful in events where participants come from different language backgrounds and interpretation services are not available for all sessions. The same need can also appear in exhibitions, official forums, cultural events, and professional meetings.

Despite the positive results, several limitations should be noted. The evaluation was conducted in a single academic conference, which may limit generalizability to other event types or audiences. The application was also tested only on Android devices, and direct system-level latency was not recorded; therefore, delay was evaluated through perceived responsiveness. In addition, the system supported only four languages, although many international events may require broader language coverage. The study also did not include a direct baseline comparison with standalone translation tools or conventional conference applications, which should be considered in future evaluations. Long-term effects, such as continued

use of translation tools or improved cross-lingual collaboration after the event, were not examined.

Future work should assess how effectively it can be applied across multiple conferences with different audiences, languages, and technical conditions. Further development could also include cross-platform support, offline and edge-AI capabilities, larger language coverage, improved handling of technical terminology, in addition to other features that enhance accessibility for participants who are deaf or hard-of-hearing through automatically generated captions. Future studies may also compare this proposed integrated system to standalone translation applications or traditional conference-support applications to better isolate the value added by conference-specific integration.

6 CONCLUSION

This study presented an AI-powered multilingual mobile application designed to support conference accessibility and participation. The application integrated real-time speech translation, agenda management, venue navigation, and cultural information into a single mobile platform and was evaluated during ICIT 2025. The results showed strong technical and usability performance, with translation accuracy of 92.5%, system uptime of 99.1%, and high participant satisfaction. The findings indicate that integrating AI-driven translation with conference support services can help reduce language barriers and improve the participant experience in multilingual academic events. The study also provides practical design insights for developing future mobile accessibility systems for conferences and similar professional settings. Future work should extend the application to more languages, cross-platform deployment, offline support, and additional multimodal accessibility features.

7 REFERENCES

- [1] T. Rogers and P. Wynn-Moylan, "Conferences and Conventions," *Routledge*, p. 378, 2022. <https://doi.org/10.4324/9781003298953>
- [2] A. U. Patel, Q. Gu, R. Esper, D. Maeser, and N. Maeser, "The crucial role of interdisciplinary conferences in advancing explainable AI in healthcare," *BioMedInformatics*, vol. 4, no. 2, p. 13631383, 2024. <https://doi.org/10.3390/biomedinformatics4020075>
- [3] Y. Xia, S.-Y. Shin, and J.-C. Kim, "Cross-cultural intelligent language learning system (CILS): Leveraging AI to facilitate language learning strategies in cross-cultural communication," *Applied Sciences*, vol. 14, no. 13, p. 5651, 2024. <https://doi.org/10.3390/app14135651>
- [4] L. Li, G. Wang, and G. Freeman, "'It's about Research. It's Not about Language': Understanding and designing for mitigating non-native english-speaking presenters' challenges in live Q&A sessions at academic conferences," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, p. 115. <https://doi.org/10.1145/3706598.3713124>
- [5] R. E. Prasetya, "Exploring English language instruction through synchronous virtual meetings in Indonesian EFL context: Pedagogical approaches, opportunities and challenges, and multimedia integration," *Journal on English as a Foreign Language*, vol. 14, no. 1, pp. 179–208, 2024. <https://doi.org/10.23971/jefl.v14i1.7532>
- [6] J. Bernatavičienė, "15th conference on data analysis methods for software systems. November 28–30, 2024," in *Vilnius University Proceedings*, vol. 52, 2024, pp. 1–136. <https://doi.org/10.15388/DAMSS.15.2024>

- [7] M. R. Haupt, M. Chiu, J. Chang, Z. Li, R. Cuomo, and T. K. Mackey, "Detecting nuance in conspiracy discourse: Advancing methods in infodemiology and communication science with machine learning and qualitative content coding," *PLoS ONE*, vol. 18, no. 12, pp. e0295414–e0295414, 2023. <https://doi.org/10.1371/journal.pone.0295414>
- [8] N. Arman, F. Khamayseh, and E. Awad, "A semi-automated approach for classifying non functional arabic user requirements using NLP tools," *International Journal of Advances in Soft Computing and its Applications*, vol. 17, no. 1, 2025. <https://doi.org/10.15849/IJASCA.250330.15>
- [9] M. Lafi, B. Hawashin, and S. A. Bi, "Eliciting requirements from stakeholders' responses using natural language processing," *Computer Modeling in Engineering & Sciences*, vol. 127, no. 1, pp. 99–116, 2021. <https://doi.org/10.32604/cmescs.2021.013026>
- [10] Y. A. Mohamed, A. Khanan, M. Bashir, A. H. H. M. Mohamed, M. A. E. Adiel, and M. A. Elsadig, "The impact of artificial intelligence on language translation: A review," *IEEE Access*, vol. 12, pp. 25553–25579, 2024. <https://doi.org/10.1109/ACCESS.2024.3366802>
- [11] A. Genovese *et al.*, "Artificial intelligence in clinical settings: A systematic review of its role in language translation and interpretation," *Annals of Translational Medicine*, vol. 12, no. 6, pp. 117–117, 2024. <https://doi.org/10.21037/atm-24-162>
- [12] S. Jaballi, M. J. Hazar, S. Zrigui, A. Mahjoubi, H. Nicolas, and M. Zrigui, "Decoding multilingual topic dynamics and trend identification through ARIMA time series analysis on social networks: A novel data translation framework enhanced by LDA/HDP Models," *Journal of Electrical and Computer Engineering*, vol. 2024, no. 1, 2024. <https://doi.org/10.1155/2024/6669491>
- [13] M. Amini, L. Ravindran, and K-F. Lee, "Implications of using AI in translation studies: Trends, challenges, and future direction," *Asian Journal of Research in Education and Social Sciences*, vol. 6, no. 1, 2024. <https://doi.org/10.55057/ajress.2024.6.1.67>
- [14] D. R. Seals, "The complete guide to hosting a guest speaker," *Advances in Physiology Education*, vol. 48, no. 2, 2024. <https://doi.org/10.1152/advan.00187.2023>
- [15] F. Federici, "Translating hazards: Multilingual concerns in risk and emergency communication," *Translator*, vol. 28, no. 4, pp. 375–398, 2022. <https://doi.org/10.1080/13556509.2023.2203998>
- [16] O. Oliinyk, N. Bezv, I. Shuliakov, and I. Derevianko, "The impact of multilingual education on developing communication skills for university students within foreign language learning," *Taduction Et Langues*, vol. 23, no. 1, pp. 125–142, 2024. <https://doi.org/10.52919/translang.v23i1.973>
- [17] S. G. Yaseen, Q. H. A. Albadrany, D. Dajani, and M. M. Al-Afaishat, "Artificial intelligence regulation: A bibliometric analysis," *International Journal of Advances in Soft Computing and its Applications*, vol. 17, no. 1, 2025. <https://doi.org/10.15849/IJASCA.250330.12>
- [18] R. O. Ogundokun, J. B. Awotunde, S. Misra, T. Segun-Owolabi, E. A. Adeniyi, and V. Jaglan, "An android based language translator application," *Journal of Physics: Conference Series*, vol. 1767, no. 1, p. 012032, 2021. <https://doi.org/10.1088/1742-6596/1767/1/012032>
- [19] D. Hu and J. Li, "Contrastive learning for robust neural machine translation with ASR errors," in *Lecture Notes in Computer Science*, vol. 13551, 2022, pp. 81–91. https://doi.org/10.1007/978-3-031-17120-8_7
- [20] S. Kafle and M. Huenerfauth, "Evaluating the usability of automatically generated captions for people who are deaf or hard of hearing," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, 2017, pp. 165–174. <https://doi.org/10.1145/3132525.3132542>
- [21] A. Herrmann-Werner, T. Loda, S. Zipfel, M. Holderried, F. Holderried, and R. Erschens, "Evaluation of a language translation app in an undergraduate medical communication course: Proof-of-concept and usability study," *JMIR mHealth and uHealth*, vol. 9, no. 12, p. e31559, 2021. <https://doi.org/10.2196/31559>

- [22] K. Shehadeh, N. Arman, and F. Khamayseh, "Classification of arabic user requirements: A semi-automated approach using NLP tools," *International Journal of Advances in Soft Computing and its Applications*, vol. 16, no. 3, 2024. <https://doi.org/10.15849/IJASCA.241130.01>
- [23] L. Zhao et al., "Adaptive policy with wait-k model for simultaneous translation," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, vol. 1, 2023, pp. 4816–4832. <https://doi.org/10.18653/v1/2023.emnlp-main.293>
- [24] L. Zhang, K. Fan, J. Bu, and Z. Huang, "Training simultaneous speech translation with robust and random Wait-k-Tokens Strategy," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2024, pp. 7814–7831. <https://doi.org/10.18653/v1/2023.emnlp-main.484>
- [25] M. Li, W. Liu, and C. Chen, "Adaptive financial literacy enhancement through cloud-based AI content delivery: Effectiveness and engagement metrics," *Annal. App. Sci.*, vol. 5, no. 1, 2024.
- [26] M. Gupta, M. Dutta, and M. C. Kumar, "Direct speech-to-speech neural machine translation: A survey," *arXiv preprint arXiv:2411.14453*, 2024. <https://doi.org/10.48550/arxiv.2411.14453>
- [27] L. Wang et al., "Benchmarking and improving long-text translation with large language models," *ACL Anthology*, pp. 7175–7187, 2024. <https://doi.org/10.18653/v1/2024.findings-acl.428>
- [28] S. Katkov, A. Liotta, and A. Vietti, "ASR systems under acoustic challenges: A multilingual study," in *Lecture Notes in Computer Science*, vol. 15450, 2024, pp. 200–213. https://doi.org/10.1007/978-3-031-80607-0_16
- [29] K. Mølnevik, "Automatic speech recognition (ASR) performance in noisy conditions: A study on norwegian police interviews," *Handle.net*, 2024.
- [30] J. R. Batista, *Learn OpenAI Whisper: Transform Your Understanding of GenAI Through Robust and Accurate Speech Processing Solutions*. Packt Publishing Ltd., 2024.
- [31] Z. Ma et al., "Speech recognition meets large language model: Benchmarking, models, and exploration," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 23, pp. 24840–24848, 2025. <https://doi.org/10.1609/aaai.v39i23.34666>
- [32] S. Malakul and I. Park, "The effects of using an auto-subtitle system in educational videos to facilitate learning for secondary school students: Learning comprehension, cognitive load, and satisfaction," *Smart Learning Environments*, vol. 10, no. 1, 2023. <https://doi.org/10.1186/s40561-023-00224-2>
- [33] N. Giovanni and W. Nurhayat, "The effectiveness of automatic speech recognition in the capcut application for developing inclusive learning media," *JKTP: Jurnal Kajian Teknologi Pendidikan*, vol. 8, no. 2, pp. 157–169, 2025. <https://doi.org/10.17977/um038v8i22025p157>
- [34] E. J. McDonnell, P. Liu, S. M. Goodman, R. Kushalnagar, J. E. Froehlich, and L. Findlater, "Social, environmental, and technical: Factors at play in the current use and future design of small-group captioning," in *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, 2021, pp. 1–25. <https://doi.org/10.1145/3479578>
- [35] J. Cao, K.-Y. Lam, L.-H. Lee, X. Liu, P. Hui, and X. Su, "Mobile augmented reality: User interfaces, frameworks, and intelligence," *ACM Computing Surveys*, vol. 55, no. 9, 2022. <https://doi.org/10.1145/3557999>
- [36] L.-H. Lee and Z. Lin, "Danger, nuisance, disregard: Analyzing user-generated videos for augmented reality gameplay on hand-held devices," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, pp. 1–33, 2024. <https://doi.org/10.1145/3677063>
- [37] H. Stefanidi, M. Tatzgern, and A. Meschtscherjakov, "Augmented reality on the move: A systematic literature review for vulnerable road users," in *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, 2024, pp. 1–30. <https://doi.org/10.1145/3676490>

- [38] R. Ivanov and V. Velkova, "Analyzing visitor behavior to enhance personalized experiences in smart museums: A systematic literature review," *Computers*, vol. 14, no. 5, pp. 191–191, 2025. <https://doi.org/10.3390/computers14050191>
- [39] Google LLC, "Google Translate," 2026. <https://translate.google.com>
- [40] iTranslate GmbH, "iTranslate Voice & Text Translation App," 2026. <https://itranslate.com>
- [41] L. Tinoca, J. Piedade, S. Santos, A. Pedro, and S. Gomes, "Design-based research in the educational field: A systematic literature review," *Education Sciences*, vol. 12, no. 6, p. 410, 2022. <https://doi.org/10.3390/educsci12060410>
- [42] B. Langenberg, M. Janczyk, V. Koob, R. Kliegl, and A. Mayer, "A tutorial on using the paired *t* test for power calculations in repeated measures ANOVA with interactions," *Behavior Research Methods*, vol. 55, no. 5, pp. 2467–2484, 2022. <https://doi.org/10.3758/s13428-022-01902-8>
- [43] NIST, "The NIST scoring toolkit (SCTK)," 2026. <https://www.nist.gov/itl/iad/mig/tools>
- [44] C. P. Malkewitz, P. Schwall, C. Meesters, and J. Hardt, "Estimating reliability: A comparison of Cronbach's α , McDonald's ω_i and the greatest lower bound," *Social Sciences & Humanities Open*, vol. 7, no. 1, p. 100368, 2023. <https://doi.org/10.1016/j.ssaho.2022.100368>
- [45] I. Kusmaryono, D. Wijayanti, and H. R. Maharani, "Number of response options, reliability, validity, and potential bias in the use of the likert scale education and social science research: A literature review," *International Journal of Educational Methodology*, vol. 8, no. 4, pp. 625–637, 2022. <https://doi.org/10.12973/ijem.8.4.625>
- [46] D. Lakens, "Concerns about replicability across two crises in social psychology," *International Review of Social Psychology*, vol. 38, no. 1, 2025. <https://doi.org/10.5334/irsp.1036>
- [47] J. Elliott, "The craft of using NVivo12 to analyze open-ended questions: An approach to mixed methods analysis," *The Qualitative Report*, vol. 27, no. 6, pp. 1673–1687, 2022. <https://doi.org/10.46743/2160-3715/2022.5460>

8 AUTHORS

Ameen Shaheen is an Assistant Professor of Computer Science at the Department of Software Engineering, the Faculty of Science and Information Technology, Al-Zaytoonah University of Jordan, Jordan. He received his Ph.D. in Computer Science from the University of Jordan in 2019. His research interests are in the fields of Software Engineering, artificial intelligence, cloud computing, algorithms, parallel computing, and distributed systems (E-mail: a.shaheen@zuj.edu.jo).

Wael Alzyadat is currently an Associate Professor of software engineering. He also works with Al-Zaytoonah University, Jordan. His research area encompasses the area of software analysis, intelligence systems, streaming data, and big data. Moreover, he has published more than 60 articles and achieved two copyrights (E-mail: wael.alzyadat@zuj.edu.jo).

Aysh Alhroob currently works as a Dean, the Faculty of Science and Information Technology and a full Professor at the Department of Software Engineering, Al-Zaytoonah University, Jordan. He has done his PhD at the University of Bradford, UK. He has a number of published articles in various computer science and software engineering topics in international journals and conferences. In addition, he published his first book on software testing in 2010 (E-mail: aysh@zuj.edu.jo).