International Journal of Computer Engineering and Data Science (ISSN: 2737-8543)
Volume 5–Issue 2, April - June 2026

The Development of Digital Transcription Device Through ESP32 Integration: An Assistive Communication Tool for the Hearing Impaired

RUBIX Bayla[1], ELISHA Gimpayan[1], PRECIOUS Bravante[1], GABRIEL Ballesteros[1], KARLYNNE Yapana[1], GJUAN Salvador[1], CYRUS Urbano[1], Dr. JULIE Real[1],

[1]Senior High School Department, Philippine School Doha, Qatar.
E-mail: bayla.rubix@psdqatar.com, gimpayan.elisha@psdqatar.com, bravante.precious@psdqatar.com, ballesteros.gabriel@psdqatar.com, yapana.karlynne@psdqatar.com, salvador.gjuan@psdqatar.com, urbano.cyrus@psdqatar.com, julieann.real@psdqatar.com

Article history

Received Feb 25, 2026
Revised Apr 9, 2026
Accepted Apr 20, 2026
Published Jun 30, 2026

ABSTRACT

Hearing impairment is a significant public health concern that continues to affect populations worldwide. The objective of this study is to provide hearing-impaired individuals with a real-time communication system that provides an additional communication support option, integrates artificial intelligence to ensure accurate communication, and addresses Sustainable Development Goal 10: reducing inequalities. The study aims to align with Qatar National Vision 2030 and current hearing impairment statistics in Qatar. This study employed a quantitative experimental research design to develop a Digital Transcription Device for the hearing-impaired as an assistive communication tool, integrating the ESP32. The study's results support the device's effectiveness and accuracy, with a rapid response time of 3.56-9.17 seconds for segmented conversations. The device achieved 100% accuracy with word counts ranging from 5 to 15. The effective distance of the Digital Transcription Device was found to achieve 100% accuracy at distances from 1 meter to 5 meters. Based on the findings, the Digital Transcription Device successfully enables accurate communication between hearing-impaired individuals by accurately recording the average time for words to display, the accuracy of the words displayed, and the maximum effective distance of the device, with minimal discrepancies.

Keywords: *Communication, ESP32, Hearing Impaired, Microcontroller, Transcription.*

I. INTRODUCTION

Hearing impairment on a global perspective remains a significant public health concern that affects how people of all ages engage, communicate, and perform in education. More than 1.5 billion people worldwide experience hearing impairment to some degree, underscoring its significant impact on global disability patterns [1]. Addressing hearing impairment is not only a health concern, but also a development priority since it relates to several Sustainable Development Goals (SDGs), including good health and well-being (SDG 3), and reduced inequalities (SDG 10), as people in low and middle-income countries are disproportionately affected by limited access to diagnosis and assistive communication technologies [2].

Hereditary hearing loss is a significant concern in the Middle East. Addressing hearing loss aligns with Qatar National Vision 2030 (QNV2030), which strongly promotes the inclusion and empowerment of people with disabilities and emphasizes the importance of accessible interventions and assistive technologies to improve social participation among individuals with hearing impairments. In Qatar, the prevalence of hereditary hearing impairment was estimated at 5.2% [3]. Beyond the medical diagnosis, untreated hearing loss has a profound impact on people's lives. In 2019, it affected an estimated 78.8 million individuals with hearing impairment in the Middle East, affecting not only health but also daily interactions, social connections, learning, and the ability to participate fully in society [4].

Despite advances in assistive technology, there remains a strong demand for portable transcription equipment, particularly among people with hearing impairments. In contrast to people with normal hearing, individuals with hearing impairments may rely more frequently on automatic speech recognition in everyday communication [5]. This is due to hearing impairment, which prevents the ability to distinguish speech. Consequently, this underscores the crucial need for an accessible, user-friendly tool for everyday purposes.

Current assistive technologies include hearing aids, implants, and smartphone apps. Although these technologies are advanced, they often rely on sophisticated digital processing and highly specialized biomedical equipment to function [6]. Current technologies are often costly and impractical for the average consumer because of the extensive research and development required to personalize each piece [7]. Additionally, manufacturing occurs at relatively low volumes compared with other electronics, resulting in fewer units produced overall.

This study aimed to develop a portable Digital Transcription Device using an ESP32 for people with hearing impairments. Unlike current software-based transcription applications, such as Voice Access and Live Transcribe, which require smartphones or laptops to function, these applications can limit accessibility for some users [8]. By combining dedicated hardware with the Deepgram API, a comprehensive voice AI platform, the device provides quick and accurate audio transcriptions to assist affected individuals. The device records speech via its microphone and converts it into text using the ESP32 microcontroller. The words are then transcribed onto a digital screen so the hearing-impaired person can read along with what is being said, supported by an artificial intelligence-based transcription system similar to those used in emerging speech-to-text technologies [9].

The prototype aimed to facilitate everyday conversation, particularly in situations where sign language is complex or where written communication would slow down communication. Through this device, means of communication for those in need are provided, enabling them to communicate more easily without sign language, via a portable solution. This study addresses a significant communication gap within the hearing-impaired community by providing a non-sign-language-dependent mode of communication. The objective of this research is to develop a viable Digital Transcription Device by leveraging the ESP32 microcontroller. The practical implementation serves as an essential reference for future technological advances and assistive communication researchers, and provides educators with insights into improving approaches [10].

Ultimately, this study is grounded in the principle of inclusion, as automatic transcription technology is instrumental in creating a more inclusive environment for individuals with hearing challenges [11]. By providing a non-intrusive and accessible tool, this device aligns with assistive technology goals, promoting natural communication by removing the burden of conversational adjustment. The study also addresses Sustainable Development Goal 10, which aims to reduce inequalities, ensuring that people with hearing impairments can access and participate in daily conversations.

A. *Research Questions*

The objective of this study is to develop a Digital Transcription Device through ESP32 integration as an assistive communication tool for the hearing impaired. Specifically, this research aims to answer the following questions:

- What is the average time for text to be displayed by the Digital Transcription Device?
- What is the accuracy of the Digital Transcription Device in displaying words?
- What is the maximum effective distance of the Digital Transcription Device in displaying words?

B. *Hypothesis*

H1: It is feasible to develop a Digital Transcription Device through the use of ESP32 as an assistive communication tool for the hearing impaired in terms of transcription speed, accuracy, and effective operating distance.

II. METHODOLOGY

This study employed a quantitative-experimental research design. The main goal of quantitative research is to use statistical, mathematical, and computer methods to quantify relationships, behaviors, or occurrences [12]. In the study, experimental methods were used to quantify results and calculate metrics to evaluate the effectiveness and accuracy of the Digital Transcription Device. Experimental research manipulates one or more factors to see how they affect a variable. In the study, the manipulated factors were the noise level, length of speaking and distance. The response variables tackled in the study assess latency, sentence accuracy, and the overall success rate of the device.

A. Schematic Diagram

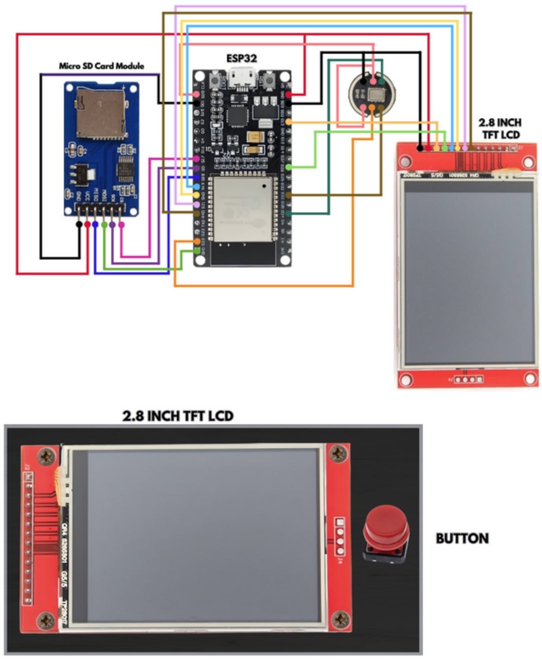

Figure 1: Schematic diagram of the Digital Transcription Device.

Figure 2: Prototype of the Digital Transcription Device.

B. Research Locale

The study was conducted in the Mesaimeer area of the State of Qatar. The researchers are students on campus and require resources, equipment, and facilities to develop the Digital Transcription Device. The researchers utilized the campus's technical support, specialized equipment, and the school's various laboratories to test the Digital Transcription Device. The school environment enabled the researchers to ensure the consistency and accuracy of the results by minimizing external interference.

C. Data Gathering Procedure

The procedure outlines the development of a Digital Transcription Device using ESP32 integration as an assistive communication tool for the hearing-impaired, and describes how its effectiveness was tested.

1. Creating the Digital Transcription Device

1. Design the device's casing and housing based on the schematics.
2. Cut a hole through the casing for inserting the button and TFT LCD.
3. Place the ESP32 on the casing and connect the MicroSD Module, with the MicroSD card inserted inside.
4. Attach the device's microphone to the ESP32 and firmly secure it to the casing, ensuring it cannot be displaced.
5. Connect the sensors to the appropriate GPIO pins on the ESP32 microcontroller.
6. Connect the TFT LCD to the ESP32, ensuring accurate connections to the MOSI and MISO pins per the schematics.

2. Programming the Device

1. Download the Arduino IDE software.
2. Boot the ESP32 and download the required libraries.
3. Program the ESP32 by determining the connection to the Deepgram API.
4. Calibrate the microphone and upload the code to the ESP32.

D. Testing Procedure

1. In testing the average time display, the time will be measured in seconds with a stopwatch by recording the time required for the Digital Transcription Device using ESP32 to display the words being spoken for different word amounts.
2. In testing accuracy, it will be measured by their ability to translate the text fully. The results will be calculated as percentages by dividing the number of words displayed by 3 different amounts (5, 10, 15 words) and multiplying by 100 to obtain the percentage accuracy across the trials.

3. In testing the maximum distance, it will be measured by its ability to transcribe the text fully at different predetermined distances (1 meter, 2 meters, 5 meters, and 12 meters), and checking whether the results of such trials still produce accurate percentages with the same number of words.

III. RESULTS

This section of the study presents the interpretations and results of the data collected during the procedures. This study aimed to develop a Digital Transcription Device integrated with an ESP32 as an assistive communication tool for the hearing impaired, exploring its effectiveness in terms of average time to display text, text display accuracy, and maximum effective distance.

Table 1: Average time for text to be displayed.

Words	Trial 1 time	Trial 2 time	Trial 3 time	Average
5 words	3.27s	3.59s	3.82s	3.56s
10 words	8.02s	7.23s	7.90s	7.72s
15 words	9.22s	10.35s	7.94s	9.17s

Table 1 displays the average display time of the Digital Transcription Device. To test the average time required for the Digital Transcription Device to display text, a stopwatch was used to measure the time for five-word, ten-word, and fifteen-word sentences. In each trial conducted, the sentence is spoken aloud, and the transcribe button is pressed by the device user; a stopwatch then recorded the time required for the text to appear on the Digital Transcription Device. The stopwatch is paused, and time is recorded when the Digital Transcription Device has displayed all spoken words. On average, the recorded times were 3.56 seconds for 5 words, 7.72 seconds for 10 words, and 9.17 seconds for 15 words, respectively. These results were determined by finding the summation of all trials for each sentence divided by three.

The results demonstrated the Digital Transcription Device's effectiveness in displaying words at varying rates, comparable to an Automatic Speech Recognition Device, which reported display times of approximately 3 seconds [13]. This illustrated response times consistent with existing mechanisms.

The average time to display text is a key factor in the effectiveness of the Digital Transcription Device, as it reflects the ESP32's response time and enables communication. These results are supported by a similar study that uses an ESP32 to develop a wearable-glasses transcription device [14], which achieved a display time of 3 seconds.

Table 2: Accuracy (%) in displaying words.

Words	Trial 1	Trial 2	Trial 3	Average
5 words	100%	100%	100%	100%
10 words	100%	100%	100%	100%
15 words	93.33%	100%	93.33%	95.56%

Table 2 presents the accuracy of the Digital Transcription Device in displaying different word amounts. To test the device's accuracy, a sentence is spoken aloud, and the button is pressed to display the text on the Digital Transcription Device. Then, the number of correct words displayed is compared with the original sentence to test the transcription's accuracy. Three trials were conducted to ensure reliability and consistency, and the average was calculated by dividing the sum of all values by the number of trials (3). To calculate the accuracy of each trial, the number of correct words was divided by the total number of words, then multiplied by 100%. The average accuracy percentage for the Digital Transcription Device when displaying 5 and 10 words was 100%. The last parameter of 15 words showcased an average of 95.56%.

The result of this trial aligns with the findings of a similar study on word error rate systems, which reported almost zero error rates for short, clean, and fluent speech tasks [7]. This study showed that the Digital Transcription Device aligns with current transcription mechanisms. These findings are consistent with the average number of words in an English conversation, which has a "turn" range of 8 to 15 words per segment, depending on the proficiency level of the user [15]. The Digital Transcription Device is consistent with prior studies, as shown in the table, which reports 100% accuracy in transcribing 5 and 10 words, and 95.56% accuracy at 15 words.

The accuracy of the Digital Transcription Device is a key indicator of its performance reliability, as it demonstrates the device's ability to display the correct words. The effectiveness of the Digital Transcription Device relies heavily on its ability to produce accurate results; thus, testing the artificial intelligence's ability to display words correctly is crucial to the study as a whole. Therefore, by maintaining 100% accuracy at 5 and 10 words, the Digital Transcription Device can operate accurately and reliably, while still having minimal inaccuracy at 15 words, with an accuracy percentage of 95.56%.

Table 3: Maximum effective distance.

Distance	Trial 1	Trial 2	Trial 3	Average
1 meter	5/5	10/10	15/15	100%
2 meters	5/5	10/10	15/15	100%
5 meters	5/5	10/10	15/15	100%
12 meters	1/5	3/10	2/15	21.11%

Table 3 exhibits the maximum effective distance of the Digital Transcription Device in displaying different word amounts. Two researchers stood apart at distances of 1, 2, 5, and 12 meters, with one speaking the sentences aloud and the other carrying the Digital Transcription Device. The device was tested at increasing distances to determine its farthest effective distance. After the person has spoken the sentences aloud, the Digital Transcription Device is activated, and the transcribed words are tested for accuracy in reproducing the spoken words. This was done by comparing the number of words the Digital Transcription Device correctly transcribed with the actual text, then calculating the percentage. Three trials were conducted to assess the accuracy and reliability of the results, and the average rate was calculated as the sum of all values divided by 3. In the first trial at 1 meter, the Digital Transcription Device accurately displayed all words. Similarly, in the second trial, it yielded the same result with 100% at 2 meters. In the third trial, the Digital Transcription Device achieved 100% accuracy at 5 meters. The fourth trial showed that the Digital Transcription Device failed at the maximum effective distance of 12 meters, yielding an accuracy of 21.11%.

However, these findings challenge a study that documented expected performance limits, reporting substantial word error rates for systems tested at a closer proximity of 0.3-0.5 meters [16]. This result demonstrates the effectiveness of the Digital Transcription Device in mitigating signal degradation, which commonly limits the performance of current solutions at 1 meter, indicating that the Digital Transcription Device can operate at a higher capacity than some current mechanisms.

The device's maximum effective distance is crucial for determining the accuracy and effectiveness of transcription. The proxemics of a conversation usually falls between a maximum range of 0.4 meters to 3.6 meters. This is based on the converser's intimacy, and effective communication occurs most often at these distances [17]. The results of the table showcase the Digital Transcription Device transcribing with 100% accuracy at a maximum effective distance of 5 meters. The Digital Transcription Device failed at a distance of 12 meters, indicating that it should be used in closer proximity to the speaker.

Table 4: Word error rate.

Words	Trial 1	Trial 2	Trial 3	Average
5 words	0 errors	0 errors	0 errors	0 errors
10 words	0 errors	0 errors	0 errors	0 errors
15 words	0 errors	0 errors	1 error	0.33 errors

Table 4 exhibits the word error rate of the Digital Transcription Device based on different word amounts. To test the word error rate, the amount of errors committed by the device was tallied based on the sentence length. In each trial conducted, a sentence is spoken aloud, and the transcription shown by the device is compared with the original text. Three trials were conducted to ensure reliability and consistency, and the average of the errors were calculated by dividing the sum of all values by the amount of trials (3). The average errors of the Digital Transcription Device at both 5 and 10 words were 0 errors. However the last parameter of 15 words yielded an average of 0.33 errors across the three trials as the third trial presented 1 error.

The results align with a study conducted that tracked word error rate and showcased that English segments that are short and clear are prone to showcasing minimal to zero error rate [7]. These findings are consistent with the average number of words in an English conversation, segmented amount of 8 to 15 words, depending on proficiency between users of the language [15]. The Digital Transcription Device showcases minimal word errors across the three word parameters with 0 errors at 5 and 10 words, while having 0.33 errors at 15 words.

The word error rate of the Digital Transcription Device is crucial to the testing procedure as it determines the accuracy of the device and the potential errors shown in transcription. The results of the table showcase the Digital Transcription Device transcribing with no errors at 5 and 10 words while minimal errors were shown at 15 words, showcasing 0.33 errors at the last parameter.

IV. DISCUSSION

This study focuses on the development of a Digital Transcription Device using ESP32 integration as an assistive communication tool, intended for the hearing-impaired, that turns spoken words into readable text. The Digital Transcription Device was studied as an assistive communication tool for the hearing impaired, with testing on display time, accuracy, and effective distance to ensure its reliability in daily communication.

Based on the Digital Transcription Device's effectiveness, the device displayed words in a range of 3.56 to 9.17 seconds, which is comparable to or better than current mechanisms, and it provided accurate results throughout segmented conversations. This study's findings demonstrated the effectiveness of the Digital Transcription Device in consistently displaying segmented conversations. This outcome aligns with a similar research highlighting the Deepgram API's ability to accurately transcribe words at a rapid rate on smart glasses [18]. This study indicated that the Digital Transcription Device effectively displayed transcribed, segmented conversations.

Additionally, the Digital Transcription Device has demonstrated precision in transcribing segmented conversations, maintaining 100% accuracy at different sentence lengths except for 15 words which showcases minimal accuracy at 95.56%, while failing at a distance of 12 meters, showcasing an accuracy percentage of 21.11%. The device communicated accurately across three word lengths that corresponded to the average length of a segmented conversation, which yielded an average accuracy of 100% for 5 and 10 words respectively while having an average accuracy percentage of 95.56%.

The effective distance of the Digital Transcription Device was tested over a range of 1 to 5 meters. The data showed that the device's maximum effective distance was 5 meters, and the text was displayed with 100% accuracy. The device failed at a distance of 12 meters, indicating that the microphone cannot properly receive the signal at that distance.

Lastly, the word error rate of the Digital Transcription device was tested. This parameter showcased 0 errors found at 5 and 10 words while 0.33 errors were validated at 15 words. This showcases the Digital Transcription Device's ability in transcribing segmented conversation.

Despite the study's results, it still has some limitations. The tests were conducted in a controlled environment with minimal noise and a single speaker under ideal conditions and may therefore not reflect the challenges of a real-world setting or noisy environments. Three trials were conducted at all parameters ensuring reliability of the results while speaking English with no accent. This poses limitations in a way where it may not reflect real world settings. Moreover, the system must remain connected to an active internet connection because it relies on the Deepgram API, thereby limiting offline usability.

A. Hypothesis

H1: The study's hypothesis, stating: *It is feasible to develop a Digital Transcription Device through the use of ESP32 as an assistive communication tool for the hearing impaired in terms of transcription speed, accuracy, and effective operating distance of 5 meters*, is accepted. Based on the results, the Digital Transcription Device can effectively transcribe and deliver quality communication.

V. CONCLUSION

In conclusion, the development of the Digital Transcription Device, which utilizes the ESP32 microcontroller, represents a step towards improvement in assistive technology, aimed at aiding the hearing-impaired by facilitating understanding of spoken information. The results confirm that the device is functioning, with response times ranging from 3.56 to 9.17 seconds. Based on the device's trials, it can be described as accurate in transcription, with the average accuracy across 5 and 10 words to be 100% while at 15 words an accuracy percentage of 95.56% was found. The Digital Transcription Device proved effective at a range of 1 to 5 meters. Furthermore, this study demonstrates the crucial role of the Digital Transcription Device in bridging the gap between individuals and the improvement of assistive technology. The Digital Transcription Device offers another way for hearing-impaired individuals to communicate effectively while implementing artificial intelligence.

It is recommended that future research investigate the device's technological limitations, such as the microphone's limited ability to detect audio at extended distances, the response time and how to decrease it, and the device's reliance on an online connection. The testing of the Digital Transcription Device in diverse environments and the effects of background noise is a limitation to be investigated. By addressing technological limitations, it is ensured that the device is reliable for everyday communication among the hearing-impaired.

ACKNOWLEDGMENT

The researchers wish to express their gratitude and appreciation to all those who supported and guided them throughout the study, especially to the following:

- Dr. Alexander S. Acosta, the Principal of the Philippine School Doha, for allowing the researchers to have a core memory experiencing the crafting of a research paper.
- Dr. Lorina S. Villanueva, the QAAD Vice Principal, and Dr. Noemi F. Formaran, the Senior High School Vice Principal, for giving the researchers the opportunity to conduct their study in Grades 11 and 12.
- Dr. Julie Ann B. Real, the researchers' Research Supervisor, for guiding, teaching, and advising the researchers throughout the study.
- Dr. Ismael A. Haguisan III, the researcher's Research Teacher, for teaching the researchers the value and process of writing a study.

- Engr. Arniel C. Gimpayan, for validating the internal electrical diagram of the device, ensuring through his electrical engineering expertise that the device is functional.
- Mr. Junrey R. Barde, for validating the code and programming of the device through his computer programming expertise by evaluating the efficiency of the device's programming through theoretical analysis.
- Mr. Aries L. Paco, for validating the prototype and schematic diagram of the device.
- Mr. and Mrs. Bayla, Mr. and Mrs. Gimpayan, Mr. and Mrs. Ballesteros, Mr. and Mrs. Espiritu, Mr. and Mrs. Yapana, Mr. and Mrs. Salvador, and Mr. and Mrs. Urbano for their support and encouragement of the researchers in completing the paper.

REFERENCES

[1] L. M. Haile, K. Kamenov, P. S. Briant, A. U. Orji, J. D. Steinmetz, A. Abdoli, and Rana, "Hearing loss prevalence and years lived with disability, 1990-2019: findings from the global burden of disease study," *The Lancet*, vol. 397, pp. 996–1009, 2019.

[2] S. A. Zarghami, "A decade of sustainable development goals: A cluster-based evaluation through four theoretical lenses," *Journal of Cleaner Production*, vol. 526, 2025.

[3] G. Girotto, M. Mezzavilla, K. Abdulhadi, D. Vuckovic, D. Vozzi, M. Khalifa Alkowari, and Badii, "Consanguinity and hereditary hearing loss in qatar," *Human Heredity*, vol. 77, no. 1-4, pp. 175–182, 2014.

[4] A. Alduais, H. Alfadda, and H. S. Alarifi, "Prevalence of hearing impairment in saudi arabia: Pathways to early diagnosis, intervention, and national policy," *In Healthcare*, vol. 13, no. 16, 2025.

[5] L. Pragt, P. Van Hengel, D. Grob, and J. W. A. Wasmann, "Preliminary evaluation of automated speech recognition apps for the hearing impaired and deaf," *Frontiers in Digital Health*, vol. 4, 2022.

[6] N. A. Alit, M. S. Ellias, and A. D. Ahmad, "Technology application in teaching and learning for hearing impaired students: A recent systematic review," *International Journal of Education*, vol. 10, no. 58, pp. 541–561, 2025.

[7] A. Ali and S. Renals, "Word error rate estimation for speech recognition: e-WER," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 20–24, Short Papers, 2018.

[8] J. A. Real, K.-E. Ceradoy, L. Fortuna, J. R. Gallarte, K. N. Soriano, A. F. Emperio, N. M. Carlos, and D. C. Camia, "The making of object recognition eyeglasses for the visually impaired using image AI," *International Journal of Innovative Science and Research Technology*, vol. 9, pp. 1012–1017, Apr. 2024.

[9] S. Di Leo, L. De Cicco, and S. Mascolo, "Real-time speech-to-text on edge: A prototype system for ultra-low latency communication with AI-powered NLP," *Information (Basel)*, vol. 16, p. 685, Aug. 2025.

[10] L. Z. Adornado, D. K. Latorre, A. I. Serrano, and a. L. K. L. Mohammad Elyjah Masukat, "The use of TensorFlow action recognition as the main component in making a sign language translator speaker for speech-impaired people," *International Journal of Innovative Science and Research Technology*, vol. 9, pp. 1203–1210, Apr. 2024.

[11] A. Apriyanto, A. Intes, S. R. Yahya, S. Hanim, and A. K. Alhamdani, "Supporting inclusivity through an automatic transcription application to improve hearing skills for the deaf," *Journal International of Lingua and Technology*, vol. 3, no. 2, pp. 425–440, 2024.

[12] M. E. McCue and A. M. McCoy, "Harnessing big data for equine health," *Equine Vet. J.*, vol. 51, pp. 429–432, July 2019.

[13] A. Valladares-Poncela, P. Fraga-Lamas, and T. M. Fernández-Caramés, "On-Device automatic speech recognition for IIoT and extended reality industrial metaverse applications," *Engineering Proceedings*, vol. 82, no. 1, 2024.

[14] G. Nooka Raju, M. Sreedhar, and P. M. K. Prasad, "Speech-to-text conversion wearable smart glasses for the hearing-impaired using esp 32," *DIBH*, vol. 53, pp. 187–194, Apr. 2025.

[15] N. Zoghlami, "Insights into L2 connected speech segmentation: A gating experiment with listeners of different english proficiency levels," *Australian Journal of Applied Linguistics*, vol. 6, no. 2, pp. 94–113, 2023.

[16] M. I. Mandel and J. Barker, "Multichannel spatial clustering for robust far-field automatic speech recognition in mismatched conditions," in *Interspeech 2016*, (ISCA), ISCA, Sept. 2016.

[17] D. Nolte, R. Hjoj, T. Sánchez Pacheco, A. Huang, and P. König, "Investigating proxemic behaviors towards individuals, pairs, and groups in virtual reality," *Virtual Reality*, vol. 29, no. 2, 2025.

[18] D. Joshy, J. Unnikrishnan, N. Yadhu, and V. Krishnaveni, "ECHOLENS: Smart glasses for real-time speech display for deaf individuals," *International Journal of Advanced Research and Computer Communication Engineering*, vol. 13, June 2024.