Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 3 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i03.46811

PAPER

Real-Time Transcriptionist Based on Artificial Intelligence to Facilitate Learning for People with Hearing Disabilities in Virtual Classes

Christian Ovalle[1](✉), Isaac Leonardo Vallejos García[2], Franco Rafael Zapata Berrios[2]

[1]Universidad Tecnológica del Perú, Lima, Peru

[2]Universidad Privada del Norte, Trujillo, Peru

c19349@utp.edu.pe

ABSTRACT

Schools have historically been ill-prepared to cater to the needs of deaf students at the elementary and secondary levels. This leads to communication difficulties that impact the learning process for each individual. During the recent COVID-19 pandemic, educational institutions for deaf students faced difficulties in providing effective teaching to children and youth. It is important to emphasize that education is fundamental for all individuals, without exception, as acquiring literacy skills enables them to lead a more fulfilling life. In this context, our research aims to investigate how the use of a computer tool can enhance communication for deaf students in a virtual environment. The methodology used involved the use of a checklist to gather data from each participant's evaluation. The post-test yielded favorable results, thanks to the statistical analysis employed in the research. In conclusion, it has been determined that a real-time transcriber facilitates learning, leading to improved educational outcomes for deaf students.

KEYWORDS

learning, education, deaf students, transcriptionist, artificial intelligence, hearing impairment

1 INTRODUCTION

The World Health Organization (WHO) World Hearing Report 2021 projects that by the year 2050, approximately 2.5 billion people will be living with some degree of hearing loss, with more than 700 million experiencing disabling hearing loss [1]. At the time of publication, over 1.5 billion people were affected by hearing loss, and among them, 430 million needed rehabilitation services, with 62.8 million in the American region [2], [3]. Furthermore, it has been revealed that there is a service gap of 83%, indicating that only 17% of individuals who could benefit from hearing aids actually use them. Globally, this gap ranges from 77% to 83% [2], [3].

Ovalle, C., Vallejos García, I.L., Zapata Berrios, F.R. (2024). Real-Time Transcriptionist Based on Artificial Intelligence to Facilitate Learning for People with Hearing Disabilities in Virtual Classes. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(3), pp. 75–88. https://doi.org/10.3991/ijoe.v20i03.46811

Article submitted 2023-10-16. Revision uploaded 2023-11-21. Final acceptance 2023-11-25.

The *Universal Declaration of Human Rights* [4] establishes that all people have equal rights before the law, implying equal protection against discrimination. In line with this, the International Convention on the Rights of Persons with Disabilities was established to enhance the human rights conditions of individuals with disabilities and guarantee their full and equal enjoyment. This convention includes several articles (12, 23, 24, 29, etc.) [5], which require governments to promote the learning of sign language to enhance the identity of the deaf community and facilitate their participation in education and society.

The World Federation of the Deaf has launched an agenda for 2030 in order to contribute to this initiative and issue a global call to action. In Peru, according to the National Institute of Statistics and Informatics, approximately 5.2% of the population (1,575,000 people) has some type of disability, and 262,000 people have permanent limitations in communication, despite the use of sign language and other methods. The primary causes of this disability include genetic/congenital factors (50.8%), advanced age (13.0%), chronic disease (10.4%), common disease (4.6%), and medical negligence (2.4%) [6].

In Peru, there are only three schools for children with hearing impairments, all located in the department of Lima. Only one of these schools is considered a special school because it teaches in Peruvian sign language (LSP), a language of Peruvian origin endorsed by Supreme Decree (DS) No. 006-2017-MIMP [7]. This decree approves the Regulation of Law No. 29535 [8], which grants official recognition to LSP. This law aims to ensure that individuals with hearing impairments have access to public services provided by both public and private entities. It also aims to guarantee that children and young people can access high-quality educational content using LSP at all levels of the educational system. Additionally, the law seeks to promote the training of interpreters in institutes, higher education schools, and universities. Law No. 28044 [9], Peru's General Education Law, in Article 20-A, upholds the right of individuals with disabilities to access high-quality inclusive education, irrespective of their circumstances. The state promotes and ensures the inclusion of individuals with disabilities in educational institutions at all levels and in all forms, guaranteeing appropriate infrastructure and the presence of qualified teachers.

In [10], research was conducted to assist individuals with autism, facial disabilities (facial paralysis), and deaf-mutes in the communication process. The project, referred to by the authors as a virtual human social system called "Avatar to Person," aimed to address these communication challenges. The aim of this project was to enhance communication effectiveness for people with disabilities within a virtual environment by proposing a system to bridge the gap between virtual and in-person interactions. Of the three target groups, only deaf-mute and autistic individuals were able to participate. His multidisciplinary proposal, which encompasses the fields of communication, artificial intelligence, and animation, was executed in a 3D simulation environment enriched with artificial intelligence. It was tested on multiple participants in real time, yielding positive results and significantly enhancing communication effectiveness within certain participant groups.

Hearing impairment causes various challenges for individuals affected by the condition, and previous research has been conducted in highly specific, limited, and controlled environments. When used in real conversations with different participants, its effectiveness is considerably reduced. In [11], an application called "Deaf Chat" was developed. This mobile application aims to facilitate communication between people with hearing impairments and those who can hear uncontrolled environments, eliminating the need for one person to listen to the other. Through a chat interface that enables a hearing-impaired individual to write and read messages to a hearing person and allows a hearing person to speak and read to a hearing-impaired individual, the goal is to facilitate the inclusion of hearing-impaired individuals in conversations with various participants. The research results indicate a high level of user acceptance, with 83.7% of users recommending the use of the application. The accuracy of transcription, as reported by users, is 86%.

During the COVID-19 pandemic, issues emerged regarding real-time communication access for deaf students in virtual environments. This led to poor learning outcomes, communication challenges, and the limited inclusion of hearing-impaired students in videoconferencing. The deployment of a real-time transcriber [12] to facilitate communication for the deaf community would enable these individuals to communicate more effectively with those who do not know sign language. By automatically transcribing audio to text, deaf individuals can read real-time speech, facilitating communication at public events such as meetings, conferences, and presentations. This tool is applicable in digital environments, such as video-conferencing platforms, which would enable deaf individuals to more easily participate in virtual meetings and have more equitable access to distance education. In addition, employing a real-time transcriber would remove barriers to inclusion and greatly enhance access to high-quality educational information for individuals with hearing impairments, thereby promoting equal opportunities.

An artificial intelligence-based Arabic sign language translator [13] has been developed to assist individuals with hearing impairments in becoming productive members of society without the need to purchase traditional tools like assistive listening devices or hearing aids. This proposed system can recognize the gestures made by users of Arabic sign language in real time. The goal is to use a snapshot photograph captured by the same system to identify the most effective artificial intelligence algorithm for distinguishing sign language gestures from images. Models were created using various algorithms [14], [15], including decision trees, classification, regression, and random forests. As a result, it was found that the random forest classification is best suited for the research environment, achieving an accuracy of over 90%.

On the other hand, there is interdisciplinary research related to information and communication technologies (ICT) aimed at facilitating communication between deaf-mute individuals and non-deaf-mute individuals through sign language without requiring the latter to learn sign language [16]. In this system, a leap motion device is utilized to detect hand and finger movements as input for tracking and prediction. In the research, gestures are recognized and processed using deep learning techniques, such as a convolutional neural network, as demonstrated in [3]. Additionally, another supervised machine learning algorithm is employed to process the gestures and convert them into speech. The dataset includes different sign language datasets: American, Pakistani, and Spanish. This research does not mention whether the system works in real time. It only states that the gestures are processed and transformed into audio messages to be sent at a later time. The results were positive, as the model achieved an accuracy of over 95%.

In reference [17], human-computer interaction (HCI) and Internet of Things (IoT) applications were utilized and integrated to develop a gesture recognition system on a cloud server. The purpose of this system was not only to represent letters of the alphabet or numbers but also to recognize the complete sign language gesture, including the movement of the arms and hands, the alignment of the fingers, and the position of the hands in relation to the body. Machine learning was utilized for sequential gesture recognition, specifically employing a classification algorithm to train the model, which was subsequently deployed on the cloud server. Despite encountering various challenges in selecting the ICH application, such as recognizing different gesture patterns or complete sign language gestures, we were able to achieve an accuracy of 97% and an efficiency of 95%.

In summary, the present research is to develop a real-time transcriber using artificial intelligence to facilitate communication for students in the deaf community within a virtual environment. This real-time transcription tool is characterized by its high accuracy, low latency, and technological accessibility. These features will allow it to have a significant impact on society and pave the way for future studies aimed at improving the educational conditions of the hearing impaired and promoting their full participation in society.

2 METHODOLOGY

This research emphasizes the significance of facilitating the education of students with hearing impairments and outlines the proposed solution for implementation. The main objective of this research is to implement a real-time transcriber in artificial intelligence to enhance the learning experience of individuals with hearing impairments in virtual classes. The specific objectives are to investigate the impact of a real-time transcriber on the learning of students from the deaf community in society and to assess how a real-time transcriber affects the learning of students from the deaf community. Finally, we propose implementing a real-time transcriber in artificial intelligence to enhance the learning experience of hearing-impaired individuals in virtual classrooms.

In this manner, in Figure 1 a speech recognition model utilizes trained acoustic and linguistic models to make predictions. The neural network plays an immediate role in producing a result based on the context, resulting in a precise sentence with words contextualized according to the linguistic structure of the sentence.

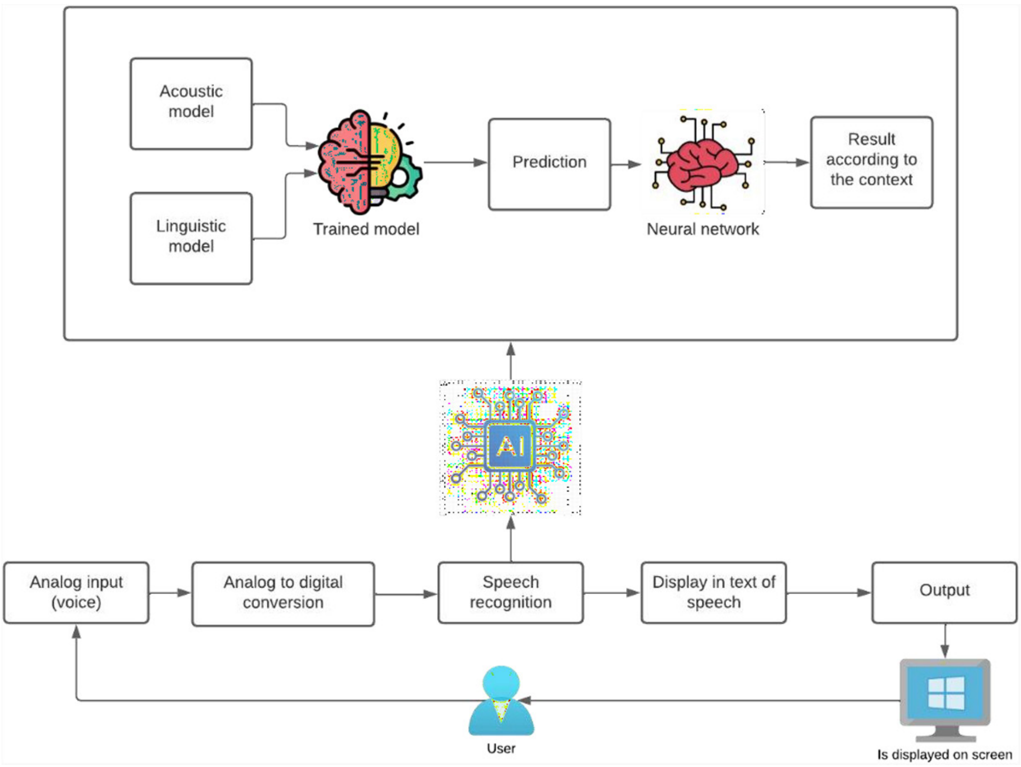

Fig. 1. Speech recognition model with "speech recognition"

The current research can be categorized as applicative and falls within the experimental field with a pre-experimental design. It focuses on emerging technologies based on artificial intelligence and adopts a quantitative approach to the target population, specifically students with hearing impairments. Continuing, the current research is an applied study [2] that is characterized by its ability to transform scientific or empirical knowledge into ideas, products, and/or prototypes that can be tailored to a specific sector or used to solve a particular problem. From the perspective of this research, it is proposed to develop a program that enables the interaction between the teacher and an intelligent system capable of recognizing and transcribing speech into text.

The process in Figure 2 that follows the interaction with the research transcriber involves an infinite loop in which the microphone continuously captures audio while adjusting the volume in response to background or ambient noise. This enables the system to detect whether the user is speaking or not. If no noise is detected, there will be no text output, and the loop will start again. Otherwise, the method or portion of code responsible for translating speech to text will be called.

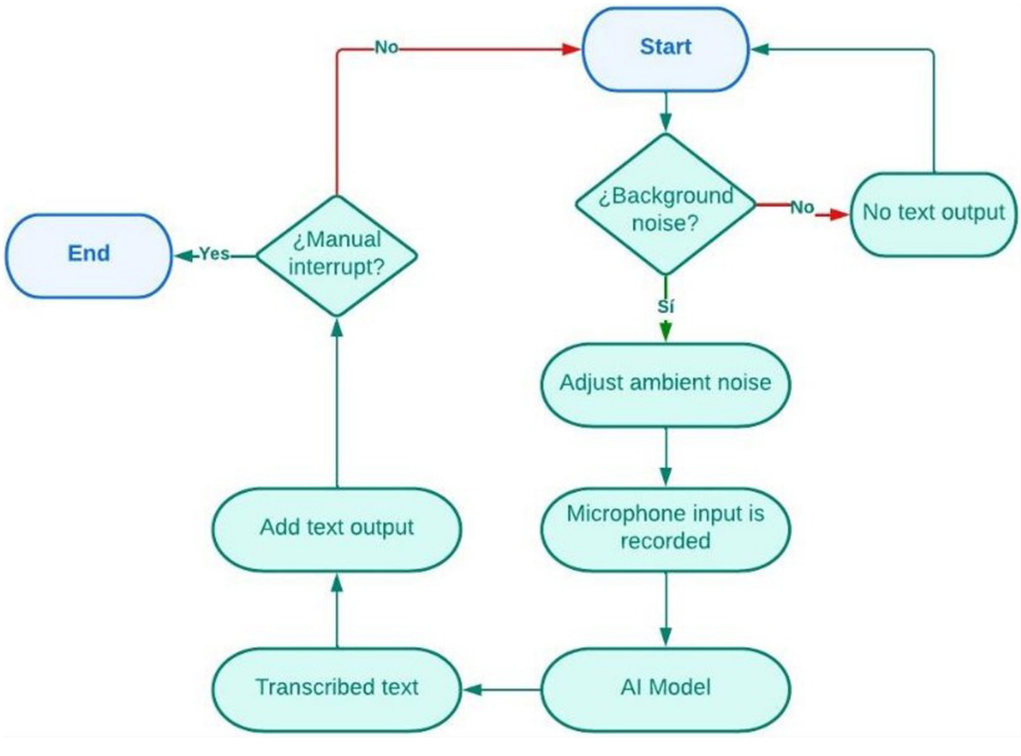

Fig. 2. Flow chart of the real-time transcriptome

The function or portion of code described above will be responsible for retrieving the text from the audio captured by the microphone. The text was originally a series of words identified by their sound based on the sound spectrum of each word. When transcribing speech to text, the trained model may encounter certain difficulties, such as homophonic words that sound similar but are spelled differently. At that point, the neural networks intervene to replace these types of words based on the surrounding words. In this way, each homophone word is assigned a probability, allowing the selection of the word with the higher probability and resulting in a correctly contextualized sentence.

Table 1 shows that the source code editor used was Visual Studio, a tool that allows us to debug, compile, and edit code at runtime in different languages, thanks to the wide variety of extensions available. Finally, one of the most important components is the speech recognition library, which offers various useful methods for speech recognition in Python. This library can be obtained and installed through the PyPI package manager, which also serves as a repository for third-party Python applications. The speech recognition library can be installed using the following command: "pip install Speech Recognition" from PyPI.

A) Software tools

Table 1. Software tools

Programming language	Python 3
Code editor	Visual Studio Code
AI Libraries	Speech Recognition, torch, transformers
Package manager	Python Package Index or PyPI

In order to be able to land and materialize the idea of the speechto text transcriber, Figure 3 shows all the inputs that were used for the development of the project. Python was used as the programming language, a high level language that is widely known for its variety of libraries that facilitate the development and/or implementation of artificial intelligence algorithms (machine learning, deep learning, etc.) such as: regression, trees, classification, neural networks, etc.

B) Program Design

Different alternatives and available open source resources were studied in order to implement them in a Python solution or API that could be easily used by teachers as a teaching tool. One of the factors taken into account when implementing the transcriber is that it should work without the need to be connected to the Internet, so we decided to use an artificial intelligence model based on the "wav2vec 2.0" frame work. Figure 4 is an illustration of how this model can be obtained in Python.

```
def texto(self):
    self.textEdit.insertPlainText("Say something...\n")
    with sr.Microphone(sample_rate=16000) as source:
        while True:
            r.adjust_for_ambient_noise(source, duration=1)
            audio = r.listen(source)
            data = io.BytesIO(audio.get_wav_data())
            clip = AudioSegment.from_wav(data)
            x = torch.FloatTensor(clip.get_array_of_samples())

            inputs = tokenizer(x, sampling_rate=16000, return_tensors='pt', padding='longest').input_values
            logits = model(inputs).logits
            tokens = torch.argmax(logits, axis=-1)
            text = tokenizer.batch_decode(tokens)
            #print('You said: ', str(text).lower())
            self.textEdit.insertPlainText("You said: ")
            self.textEdit.insertPlainText(str(text).lower())
            self.textEdit.insertPlainText("\n")
```

Fig. 3. Real-time transcriber source code

```
tokenizer = Wav2Vec2Processor.from_pretrained('facebook/wav2vec2-base-960h')
model = Wav2Vec2ForCTC.from_pretrained('facebook/wav2vec2-base-960h')
```

Fig. 4. Declaring wav2vec2 model and tokenizer

In order for the code to compile without errors, the speech recognition, torch, transformers, pyAudio and pydub libraries must first be installed with the help of the PyPI package manager. From the Visual Studio Code terminal or from the Python terminal you can execute the "pip install" instruction and the respective library names in order to install the necessary libraries that will be input to the source code. In which, the microphone volume is adjusted according to the background noise picked up by the microphone, then the sound retrieved by the microphone is recorded and a wav file is temporarily generated and then transformed into a pyAudio object.

To compile the code without errors, the speech recognition, torch, transformers, pyAudio, and pydub libraries must be installed using the PyPI package manager. From the Visual Studio Code terminal or the Python terminal, you can execute the "pip install" command followed by the respective library names to install the necessary libraries required for the source code.

The microphone volume is adjusted based on the background noise picked up by the microphone. Subsequently, the sound captured by the microphone is recorded, and a temporary WAV file is generated and then transformed into a pyAudio object. The voice data, which is a sentence, is divided into a list or array of words, labeled. For each element of this list, a transcribed text is obtained from the artificial intelligence model, and then displayed in the following interface.

Then, this object, which is voice data, i.e. a sentence, is divided into a list or array of words, labeled and for each element of this list a transcribed text given by the artificial intelligence model is obtained, which is then displayed in the user interface.

C) Model of artificial intelligence

The artificial intelligence model is based on the "wav2vec 2.0" model, a self-supervised learning framework for speech representations shown in Figure 5. This model consists of a multilayer convolutional neural network that inputs a raw audio (X) and outputs a latent speech representation (Z) for each fragment of this audio. Then, these representations (Z) pass through a transformer to obtain a contextualized representation (C) by compiling the complete raw audio information. That is, individual representations are first obtained separately and then the surrounding words are analyzed to obtain a contextualized output.

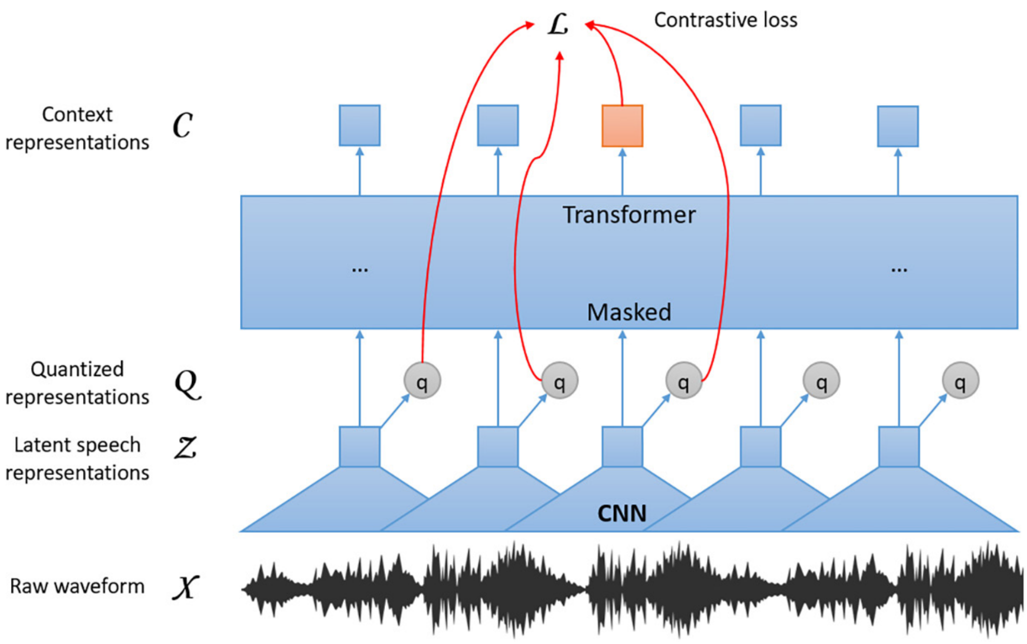

Fig. 5. Wav2vec2 framework architecture

The outputs from the convolutional neural network (Z) are passed through a module (Q) that quantifies the individual target representations (q) to achieve self-supervised learning.

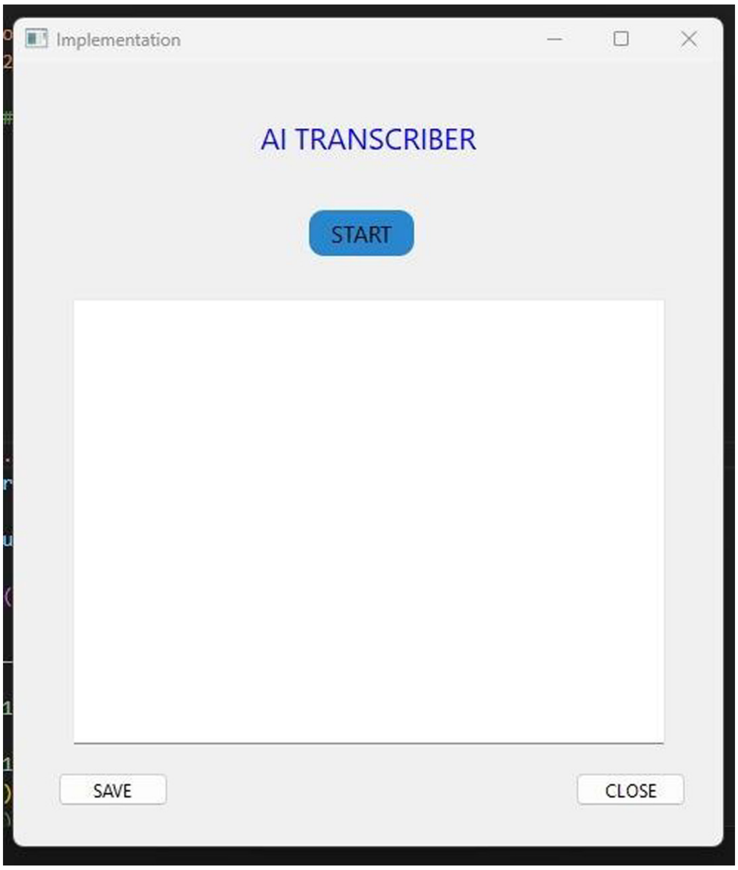

Fig. 6. Real-time transcriptor interface design

Mainly, the user interface design should be simple and intuitive. That is why the following elements shown in Figure 6 were considered: a button to initiate text transcription, an area to display the transcribed text, and a button to save the text in a file (notepad). All the elements included in the interface are crucial for the proper functioning of the real-time transcriber. The addition of the save button allows users to maintain a history of transcriptions, which can be beneficial for children in the early stages of learning and for teachers to track their class sessions effectively.

Subsequently, the software was tested on students with hearing impairments, and the supervising teacher used the transcriber in real time to facilitate the students' learning of the Spanish language and enhance their overall learning experience, as depicted in Figure 7.

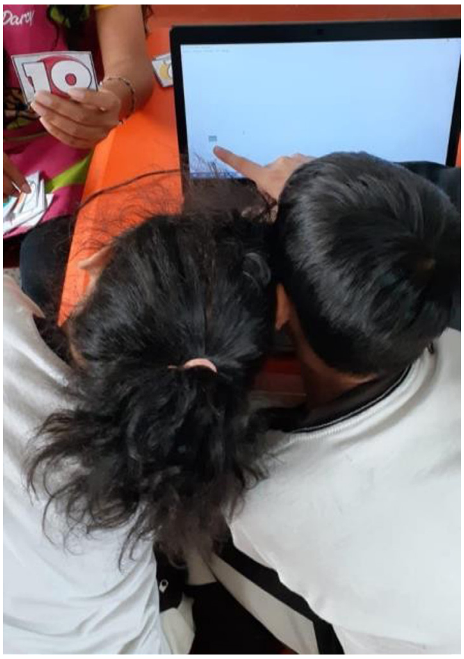

Fig. 7. Students interacting with the real-time transcriptor

Finally, after coordinating with the teachers and authorities, visits were made to the institution to observe the reality of CEBE. Subsequently, an instrument was developed to measure student learning in accordance with the evaluation criteria (see Table 2).

Table 2. Evaluation criteria stipulated in the data collection instrument

Criteria	
Oral Expression/Sign Language	**Oral Expression/Sign Language**
Uses frequently used words, sounds and phrases to express needs and communicate.	Understand when you are told "no" and explain why you cannot act. E.g., "Do not touch the plug because it is dangerous."
Combines two or three words into a meaningful sentence (even if not pronounced correctly).	Understands the meaning of familiar words (dada, mama, milk, water, etc.).
Use pronouns such as "I" and "you", or the possessive "my" or "mine".	Increases understanding of the meaning of words (has a growing vocabulary).
Responds with everyday words or sign language to questions asked of him/her.	Reacts when called by first and last name.
Seek information by asking questions.	Recognizes the names of family members and their roles. E.g., Recognizing that Maria is her mom, etc.

3 RESULTS

A real-time transcription system based on new technologies, specifically artificial intelligence, was developed with the objective of assisting deaf students in their learning during virtual classes.

3.1 Results obtained in the evaluation of the learning session

Evaluations of the learning session were conducted with deaf children at the "CEBE (Special Basic Education Centre) VIRGEN DEL CARMEN" using an instrument designed to collect pre- and post-test data. This data was then used to perform the necessary statistical inferential analyses. The presentation of the tests conducted will be divided into two phases based on the variables presented below.

Phase 1: Pretest learning assessment. Based on the research objectives, a pretest was conducted to evaluate the learning session of students with hearing impairment. The results are presented in Table 3.

Table 3. Learning assessment pretest

Participants	Criteria		
	Oral Expression/ Sign Language	Oral Comprehension	Total
Participant 1	3	4	7
Participant 2	3	5	8
Participant 3	2	2	4
Participant 4	3	5	8
Participant 5	2	5	7
Participant 6	5	3	8
Participant 7	3	5	8
Participant 8	2	4	6
Participant 9	2	5	7
Participant 10	3	5	8

It was found that the average score during the learning session under regular conditions was 7.1 out of 10.

Phase 2: Post-test learning assessment. Subsequently, in line with the specific objectives outlined in the research, the post-test was conducted to assess the learning session of students with hearing impairments, and the results are presented in Table 4.

Table 4. Learning assessment post-test

Participants	Criteria		
	Oral Expression/Sign Language	Oral Comprehension	Total
Participant 1	5	5	10
Participant 2	5	5	10
Participant 3	3	5	8

(Continued)

Table 4. Learning assessment post-test *(Continued)*

Participants	Criteria		
	Oral Expression/Sign Language	Oral Comprehension	Total
Participant 4	5	5	10
Participant 5	5	5	10
Participant 6	5	5	10
Participant 7	4	5	9
Participant 8	5	4	9
Participant 9	4	5	9
Participant 10	5	5	10

3.2 Statistical inferential analysis

Descriptive and inferential statistical analyses were conducted with the help of a statistical expert, yielding the following results:

To assess the reliability of the instrument used to measure learning, a Cronbach's alpha reliability test was conducted with the participation of 10 students. The results presented in Table 5 indicate a moderate reliability of the instrument ($\alpha = 0.889$).

Table 5. Cronbach's alpha reliability test results

Participants	Variable	Cronbach's Alpha
1	V1: Pretest Learning	77.3%
10	V2: Post-test learning	88.9%

Subsequently, a Shapiro-Wilk normality test was performed on the two variables using the samples collected from each deaf student. The Shapiro-Wilk test was chosen due to the sample size being less than 50. The results obtained in Table 6 for both measurements showed that the significance level in the "Sig." and "Shapiro-Wilk" columns was less than 5%. Therefore, it is concluded that the variable does not follow a normal distribution, and it was decided to apply a nonparametric statistical test. For hypothesis testing, the Wilcoxon test was utilized.

Table 6. Results of the Shapiro-Wilk normality test

	Shapiro-Wilk		
	Statistician	gl	Sig.
Pre-test learning	.781	10	.008
Post-test learning	.731	10	.002

Table 7 shows the descriptive values of the statistical parameters of the Wilcoxon test, in which a significant difference is observed between the post-test and the pretest.

Table 7. Result of the Wilcoxon signed-rank test

		N	Average Range	Sum of Ranks
Post- and Pre-test learning	Negative ranges	0[a]	.00	.00
	Positive ranges	10[b]	5.50	55.00
	Ties	0[c]		
	Total	10		

Notes: [a]Post-test learning < Pretest learning; [b]Post-test learning > Pretest learning; [c]Post-test learning = Pretest learning.

In Table 8, the results of the Wilcoxon test show a significance value of 0.004. Therefore, it is concluded that the alternative hypothesis should be accepted, since the value does not exceed 5% (0.005). This indicates that there is a difference between the pretest and posttest, and they cannot be considered equal.

Table 8. Wilcoxon test results

	Post-Test Learning – Pre-Test Learning
Z	–2.877[a]
Sig. asin. (bilateral)	.004

Note: [a]It is based on negative ranges.

4 DISCUSSION

Considering the research of a chat application for hearing impaired people, both researches innovatively address the problem of communication for hearing impaired people, although they present different approaches. The research in this paper focus-es on a real-time transcription system based on artificial intelligence, while on the other hand it focuses on a chat application called "Deaf Chat". The strengths and limitations of each approach, as well as the implications for the improvement of the learning process and social inclusion of the hearing impaired, will then be discussed. The research that implemented a real-time transcription system based on artificial intelligence is distinguished by its focus on the educational domain, especially on the learning process in virtual classrooms. The application of pre- and post-test criteria provides a robust quantitative assessment of the impact of the tool on academic prog-ress. The results reveal a significant increase of 24% in the academic performance of the participants, indicating that the implementation of this technology can have a substantial positive impact on the learning of people with hearing impairment.

On the other hand, the research presenting the "Deaf Chat" application stands out for its focus on facilitating communication in uncontrolled environments. The application addresses the oral communication barrier by allowing hearing impaired people to actively participate in conversations with hearing people through a chat. High user acceptance, with 83.7% recommending the application, supports the perceived use-fulness of the tool. In addition, transcription accuracy of 86% indi-cates solid performance in transforming auditory information to text. Despite the strengths of both approaches, it is essential to consider the inherent limitations. The real-time transcription system can be highly dependent on the quality of speech recognition and the availability of advanced technological resources, limiting its applicability in resource-limited contexts. In contrast, the "Deaf Chat" application may face challenges in noisy environments or in situations where written communication is not the most efficient or natural form of interaction. In conclusion, both investigations represent significant advances in the field of communication for the hearing impaired. The real-time transcription system excels in the educational setting, enhancing the learning process, while the "Deaf Chat" application focuses on social inclusion and communication in everyday environments. Integrating these approaches could provide a more complete solution, addressing both the educational and social needs of the hearing impaired.

5 CONCLUSIONS

A real-time transcription system based on Artificial Intelligence was success fully implemented with the purpose of improving the learning process of hearing impaired people in virtual classrooms. To evaluate the effectiveness of the system, a study was conducted using pre and post-test criteria, using an instrument developed specifically to measure learning progress.

The results obtained revealed that the variable analyzed does not follow a normal distribution, which leads to the rejection of the null hypothesis. Consequently, it was decided to use a nonparametric statistical test. It is important to note that, during the development of the research, information limitations related to the population of deaf students in Peru were encountered, due to the lack of specific educational records.

Learning was measured before and after applying the tool as a support in the classroom sessions, in the pretest an average of 7.1 points was obtained and in the post-test an average of 9.5 points out of 10. Thus, the real-time transcriber based on artificial intelligence improved the learning process by 24% with respect to the initial results of the research.

Finally, in the present research the implementation and testing of a real-time transcriber based on Artificial Intelligence was developed to facilitate the learning of students from the deaf community in virtual classes. Applying the data collection with the students of SBEC Virgen del Carmen in the elaborated instrument, it was statistically determined that there is a substantial improvement in the learning process of the students of the deaf community, that is to say, that the use of a real-time tran-scriber based on artificial intelligence positively influences the learning process of the students of the deaf community. It is recommended to apply a teaching methodology in which the real-time transcriber is used as a support tool together with the intervention of the respective teacher, but not as a complement or an additional tool, since the students themselves have no knowledge or notions of the functioning of the tool in question.

6 REFERENCES

[1] "Informe Mundial Sobre La Audición Resumen Ejecutivo," 2021. https://www.who.int/es/publications/i/item/9789240021570

[2] A. Sultan, W. Makram, M. Kayed, and A. A. Ali, "Sign language identification and recognition: A comparative study," *Open Computer Science*, vol. 12, no. 1, pp. 191–210, 2022. https://doi.org/10.1515/comp-2022-0240

[3] M. M. Balaha *et al.*, "A vision-based deep learning approach for independent-users Arabic sign language interpretation," *Multimed Tools Appl*, vol. 82, no. 5, pp. 6807–6826, 2023. https://doi.org/10.1007/s11042-022-13423-9

[4] M. X. É I C O and M. Iván del Toro Huerta, "DERECHOS HUMANOS La Declaración Universal de Derechos Humanos: un texto multidimensional."

[5] "Convention on the Rights of Persons with Disabilities and Optional Protocol UNITED NATIONS."
[6] "'En el Perú 1 millón 575 mil personas presentan algún tipo de discapacidad', INEI, n°178, pp. 2, diciembre 2013."
[7] "Ministerio de la Mujer y Poblaciones Vulnerables (14, ago. 2017). Decreto Supremo N° 006-2017-MIMP, Aprueba el Reglamento de la Ley N° 29535."
[8] E. S. Peruana de Servicios Editoriales - Editora Peru, "Congreso de la República del Perú. Ley N° 29535, Ley que otorga reconocimiento ofi-cial a la Lengua de Señas Peruana." [Online]. Available: www.minjus.gob.pe.
[9] "N° 038-2022-MINEDU, Resolución Ministerial, Lima-Perú, 2022, p. 2."
[10] Z. Guo, Z. Wang, and X. Jin, "'Avatar to Person' (ATP) virtual human social ability enhanced system for disabled people," *Wirel Commun Mob Comput*, vol. 2021, 2021. https://doi.org/10.1155/2021/5098992
[11] M. Shezi and A. Ade-Ibijola, "Deaf chat: A speech-to-text communication aid for hearing deficiency," *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 5, pp. 826–833, 2020. https://doi.org/10.25046/aj0505100
[12] D. Li *et al.*, "Transcribing natural languages for the deaf via neural editing programs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 11, 2022, pp. 11991–11999. https://doi.org/10.1609/aaai.v36i11.21457
[13] G. Latif, J. Alghazo, N. Mohammad, and R. Alghazo, "Communicating with the deaf and hard of hearing through automatic Arabic sign language translator," *Journal of Physics: Conference Series*, 2021. https://doi.org/10.1088/1742-6596/1962/1/012055
[14] I. Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras, "Artificial intelligence technologies for sign language," *Sensors*, vol. 21, no. 17, p. 5843, 2021. https://doi.org/10.3390/s21175843
[15] D. Bragg *et al.*, "The FATE landscape of sign language AI datasets: An interdisciplinary perspective," *ACM Trans Access Comput*, vol. 14, no. 2, pp. 1–45, 2021. https://doi.org/10.1145/3436996
[16] M. I. Saleem, A. Siddiqui, S. Noor, M. A. Luque-Nieto, and E. Nava-Baro, "A machine learning based full duplex system supporting multiple sign languages for the deaf and mute," *Applied Sciences (Switzerland)*, vol. 13, no. 5, p. 3114, 2023. https://doi.org/10.3390/app13053114
[17] S. Adhikari *et al.*, "A novel machine learning–based hand gesture recognition using HCI on IoT assisted cloud platform," *Computer Systems Science and Engineering*, vol. 46, no. 2, pp. 2123–2140, 2023. https://doi.org/10.32604/csse.2023.034431

7 AUTHORS

Christian Ovalle is a Systems Engineer graduated from the Universidad Peruana Union belonging to the top third, Master in Management and Management of Information Technology, with a Diploma in Process Management and Specialty in Operations and Logistics from ESAN University, with extensive experience in business projects, and with a passion for research (E-mail: c19349@utp.edu.pe).

Isaac Leonardo Vallejos García is a graduate from the Universidad Privada del Norte, Faculty of Computer Systems Engineering, belonging to the tenth grade, with Peruvian nationality, identified with ORCID: 0009-0000-4377-5816, in the department of Lima, residing in Jr. San Martin 3415 – 15106 (E-mail: n00238100@upn.pe).

Franco Rafael Zapata Berrios, Graduate of the Universidad Privada del Norte, Faculty of Computer Systems Engineering, belonging to the tenth grade, with Peruvian nationality, identified with ORCID: 0009-0007-3664-6363, in the department of Lima, residing in Unidad Vecinal 3 – Colonial (E-mail: n00225566@upn.pe).