SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I3P119

Volume 11 Issue 3, 231-239, March 2024


*Original Article*

# Design of a System of Response Options Using a Neural Network for People with Communication Difficulties

Giancarlo Apaza Zurita[1], Sebastian Miguel Cáceres Huamán[2], Jonathan Jesús Mendoza Núñez[3], Talavera-Suarez, Jesus[4], Joseph, G. M.[5]

[1,2,3,4,5]*Universidad Nacional de San Agustin de Arequipa, Arequipa, Perú.*

[5]*Corresponding Author : guevajoan2@gmail.com*

Received: 12 January 2024 Revised: 10 February 2024 Accepted: 11 March 2024 Published: 25 March 2024

***Abstract -*** *The paper proposes a communication system aimed at improving the interaction of people who have lost the ability to speak due to various circumstances, such as accidents, facial paralysis, or vegetative states. The system consists of speech recognition and a neural network with Bidirectional Encoder Representations from Transformers (BERT) structure, which was trained to provide response options in order to establish fluid and meaningful interactions and maintain a conversation between the patient and the environment. Its design seeks to improve the quality of life of those with substantial communication challenges; for this purpose, we propose an interface consisting of two modes, one for people who have some mobility in their hands and another for people in a completely vegetative state using Electrooculogram (EOG) systems to configure the signals with our proposed system. The effectiveness of the system was evaluated using the subjective Scale Utility System (SUS) method, yielding promising positives that underline its effectiveness and perceived usefulness for real environments.*

***Keywords -*** *Artificial Intelligence, Assisted communication, Assistive technology, EOG, Voice recognition.*

## 1. Introduction

Today, it is essential for people to talk to each other to socialize or express themselves. Suppose this ability is lost for a variety of reasons, such as an accident, facial paralysis or a vegetative state. In that case, a person may no longer be able to interact with others. To help those with communication problems, there has been increasing interest in developing alternative communication technologies. To help these people communicate more easily and lead a better life, the present study focuses in this context on the construction of a responsive possibility system using artificial intelligence.

Since artificial intelligence can learn complicated patterns and provide responses adaptively, using them in the design of response systems is a potential alternative. It is possible to create a system that understands and translates the needs and desires of people with communication problems into understandable and contextual responses by training artificial intelligence with relevant data.

Technology is advancing in giant steps; these advances in technology and artificial intelligence have opened new possibilities to help these people recover their ability to communicate and to support us in a myriad of tasks, both insecure as well as also applied to achieve support in medicine, achieving significant breakthroughs. This paper presents an artificial intelligence-based communication system that aims to support people who have lost their ability to speak. Our system uses an intuitive and easy-to-use user interface, which provides response options so that the user can select the one that best fits his or her responses to communicate. Each user response is analyzed and used to generate new response options that fit the conversation, allowing for smoother communication.

This article is divided as follows: In the Related Works Section, the works related to this research are presented; the methodology is developed in the Methodology Section; in the Proposed System Development Section, the description of the speech recognition algorithm, the proposed trained artificial intelligence and EOG design and processing with the two developed interfaces is performed; the description of the developed experimentation is explained in the results Section with the obtained results and discussions, and the conclusions of this research are presented in the Conclusions Section with the future work we will perform.

## 2. Related Works

In [1], they designed a system for patients with Amyotrophic Lateral Sclerosis (ALS) who have lost control of all muscles except eye movements. The system uses the NeuroSky MindWave Mobile device to obtain coded blink signals, which are then analyzed using a K-Nearest Neighbors algorithm to determine the letter corresponding to the blink sequence code and vocalize the resulting word. In [2], we are presented with an EEG-based spelling system and algorithms via a Brain-Computer Interface (BCI). The system uses a flexible visual stimulus mechanism to exploit the P300 wave generated in response to unpredictable stimuli. They analyze the EEG signal processing and classification algorithms used for automated online decision-making on the character pointed by the subject, using a classifier based on Bayesian Linear Discriminant Analysis (BLDA) and a greedy approach to increase the spelling rate.

In [3], they propose a new P300-based spelling Brain-Computer Interface (BCI) paradigm that uses initial character typing with word suggestions and a new P300 classifier to increase the speed and accuracy of word typing. The modified text on 9 keys (T9) interface allows users to type initial characters using a 3x3 matrix interface and an integrated custom dictionary that suggests candidate words. The authors adopt a random forest classifier that significantly improves the accuracy of P300 classification by combining multiple decision trees.

In [4], we describe the design and implementation of an Electrooculogram (EOG)-based Human-Computer Interface (HCI) system that allows people to communicate using only eye movements. The system uses a virtual keyboard and classifies horizontal and vertical EOG signals with a nearest-neighbor algorithm. The system has a high common mode rejection rate and 95% classification performance.

In [5], a novel and effective method of gaze-based text input for disabled people is presented. This method divides the human gaze into nine directions and uses flicker for text input. The authors built a Convolutional Neural Network (CNN) model for nine-way gaze estimation and used a nine-key T9 input method based on bar phones. They created a large-scale dataset to train the model, which accurately estimated the gaze of different people in various lighting conditions. The method does not require a complex calibration process and can be run on-screen or off-screen.

In [6], he shows us a novel method of linguistic communication called “eye-writing”, which involves detecting eye movement traces to determine the corresponding symbols by pattern recognition. The study used Electrooculography (EOG) to measure eye movements and evaluated the system with a set of symbols consisting of 10 Arabic numerals and 4 mathematical operators. The study found that the recognition rate ranged from 50% to 100%.

In [7], they have designed a low-cost device that reads and converts patient blinks into a universally accepted communication code: Morse code. They also tell us that some customized Augmentative and Alternative Communication (AAC) devices have been developed to use signals from the patient and convert them into some kind of data that can be communicated. However, these devices are very expensive and are practically out of reach for most people.

In [8], he discusses how tongue bar piercing can be used as an assistive device for people with severe disabilities such as tetraplegia, allowing them to control their environment through tongue movement. A narrative review of the existing scientific literature was conducted and found that these assistive devices have a high degree of acceptance and performance when integrated with mechatronics tools.

The study [9] examines the use and awareness of assistive technologies among people with physical disabilities in the United Arab Emirates. Through a survey of 50 participants, it was found that only 40% currently use these technologies, while 61.2% are aware of their existence. Wheelchairs are the most recognized technologies, while head movement-sensitive systems are less known. Although only 10% are familiar with emerging technologies, 70% would be willing to adopt them. In addition, 60% receive government assistance, 40% rely on philanthropic or family support, and only 2% are fully employed. These findings underscore the need for awareness initiatives to improve the use of assistive technologies among people with physical disabilities.

In [10], it aims to explore the current status of Self-Help Devices (SHDs) for people with quadriplegia. After a comprehensive review of 222 articles, 75 were selected for further review. A high dropout rate of SHDs in the current literature is highlighted, especially in India, where people with spinal cord injuries face additional challenges due to a lack of education and financial resources. It highlights the need to develop devices more focused on meeting the specific needs of people with quadriplegia. This could increase the effectiveness and reduce the dropout rate of these devices, thus providing greater independence and quality of life for this group of patients.

In [11], the performance of the Head-Tongue Controller (HTC), a multimodal interface designed for people with quadriplegia, which combines tongue and head tracking to provide discrete and proportional controls in a single device, is evaluated. Seventeen patients with quadriplegia participated in tests of simple driving or advanced manoeuvres using their Personal Alternative Controller (PAC) and different combinations of driving modalities with the HTC. It was found that, for some simple tasks, the HTC improved completion time compared to the PAC, while, for advanced tasks, the HTC showed completion times comparable to the PAC. Participants expressed interest in HTC and recognized its usefulness for this population, according to post-study questionnaire results.

In [12], investigated the feasibility of adapting cognitive assessments for children with Cerebral Palsy (CP) using switches, a method rarely formally employed. Accessibility to receptive vocabulary and nonverbal reasoning assessments was successfully tested, although challenges persisted in executive functions and visual perception. Although the adaptations significantly increased assessment time, the user experience was positive in terms of usability and cognitive load. More research is needed to improve the accessibility of assessments for children with motor and/or speech disabilities.

In [13] examines the feasibility of recognizing dynamic Chinese Sign Language (CSL) using millimeter wave (mmWave) radar sensors to improve Human-Computer Interaction (HCI) for people with Deafness and Hearing Loss (D&HL). Fundamental challenges are addressed, and preliminary solutions are proposed, including a case study employing a lightweight Convolutional Neural Network (CNN) to classify Chinese sign words based on radar spectrograms. The article highlights the importance of interdisciplinary collaboration and continued progress in this emerging area.

## 3. Methodology

In this article, Figure 1 details the operating cycle of the proposed system. First, the training stage consists of two parts:

- Speech recognition: it will be in charge of listening and transcribing to text the words that User A wants to communicate.
- Artificial Intelligence: We worked with neural networks to provide at least 4 possible response options.

In the training stage, it was possible to unite both parts mentioned above and train them to correctly and fluently provide the options of possible answers to be chosen by User B.

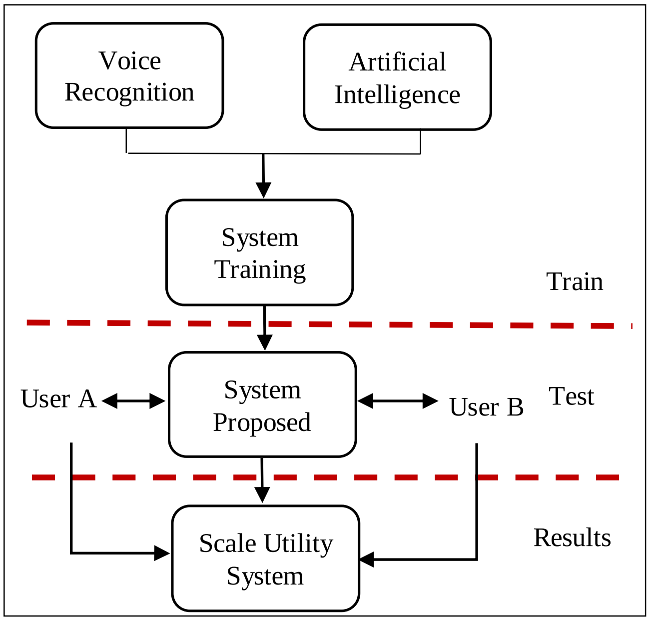


**Fig. 1 Methodology**

The system was tested in the laboratory, with two types of interfaces, one for people who have mobility to be able to choose and another one simulating people in a vegetative state. Tests were carried out with 5 pairs of people. In order to analyze the effectiveness of the proposed system, subjective tests will be applied to measure the usability of the system for users, such as Scale Usability System (SUS) and Microsoft Reactions Cards.

## 4. Proposed System

The system proposed in this article is the first system that we want to develop for people with communication difficulties; as briefly explained in the methodology, the system has 4 stages, which can be seen in Figure 2, and we will explain them below:

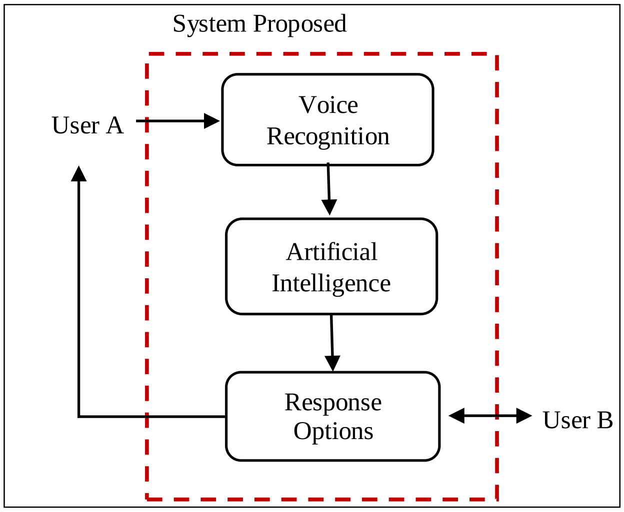


**Fig. 2 System proposed**

### *4.1. Voice Recognition*

The speech recognition system was designed with a circuit using a MEMS omnidirectional microphone model SPW2430. This type of microphone is ideal for picking up voice signals due to its high sensitivity, low distortion and linear frequency response in the voice frequency range.

The microphone was connected to an instrumentation amplifier, which provides adjustable gain and low noise levels to improve the quality of the speech signal. Subsequently, the amplified voice signal was sent to a 16-bit Analog-to-Digital Converter (ADC), which converts the analog signal into a digital signal so that it can be processed by the Raspberry Pi 4B+ 8GB (See Figure 3).

An algorithm developed in the Python language is used, using speech recognition libraries, such as "speech_recognition" supported by the recognize_google class, thus using the Google API to obtain in the text the audio obtained by the microphone of User A. This text obtained will go to the next stage of the system to be processed.

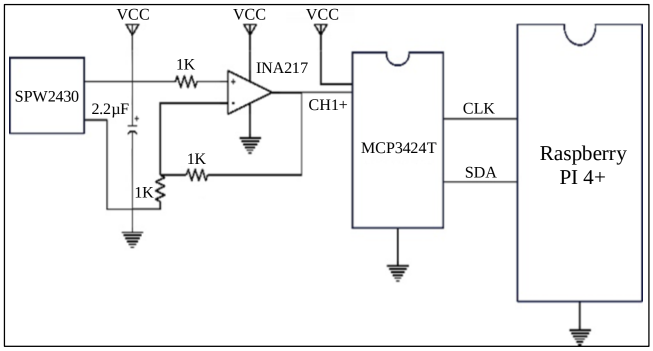


**Fig. 3 Speech recognition schematic diagram**

### *4.2. Artificial Intelligence*

The obtained text is the input for the artificial intelligence, in this case, working with a neural network with BERT (Bidirectional Encoder Representations from Transformers) architecture, a neural network based on Transformers, to understand the semantic context of the questions and generate relevant answers. In the training process, the BERT model is exposed to a dataset containing paired questions with multiple possible answers. The fundamental objective is to provide the model with the ability to rank the correct answer among the provided options, thus enabling contextually relevant answer generation.

During the data preprocessing phase, a BERT tokenizer is employed to split the text into tokens, and the input format is adjusted to incorporate attention masks and segments. This stage is crucial to ensure that the model can effectively capture the semantic complexity of the questions. The output layer of the neural network is designed with a softmax activation function. This enables the model to generate up to four possible answers for each question, thus providing the user with a variety of options to select from visually (see Figure 4).

During the training process, cross-categorical entropy is used as a loss function, and the Adam optimizer is used to adjust the model weights. Performance evaluation is carried out using metrics such as accuracy, and the early stopping technique is incorporated to mitigate the risk of overfitting and ensure a generalizable model (See Figure 4).

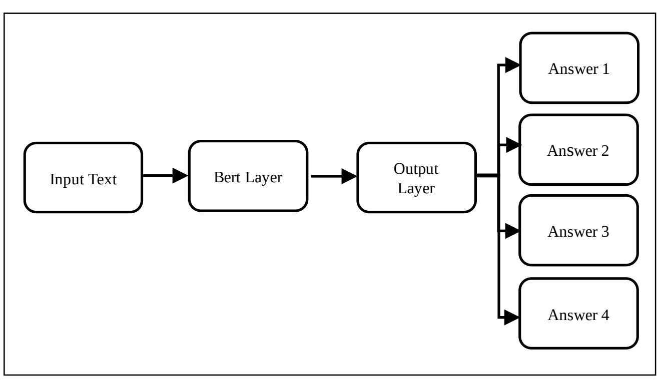


**Fig. 4 Neural network structure**

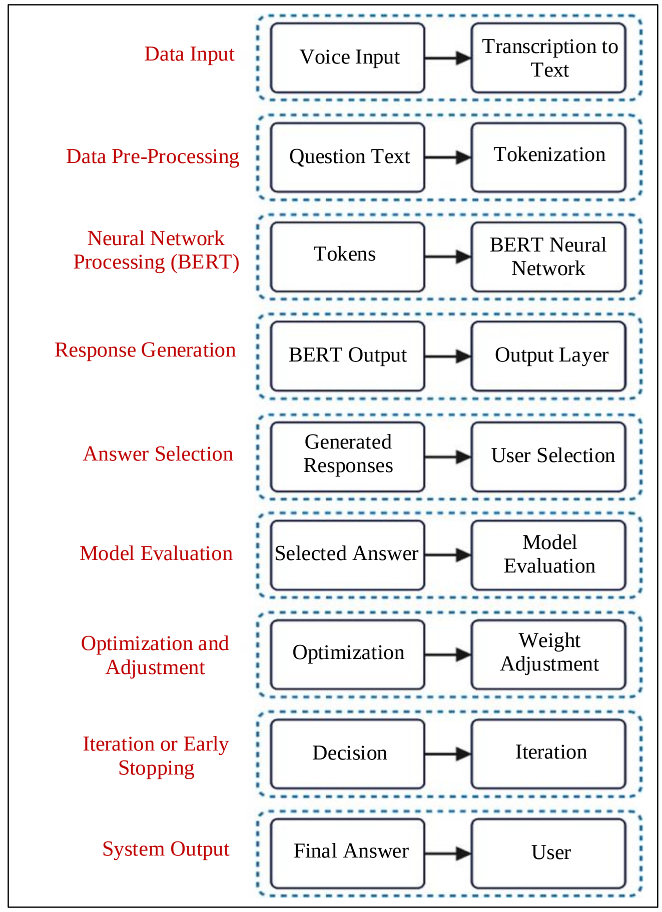


**Fig. 5 Process flow from data input to final response presentation**

### *4.3. EOG Design and Processing*

Eye movement is controlled by six external muscles surrounding the eyeball, allowing movements in different directions. The Electrooculogram (EOG) records eye movements by analyzing the potential difference in the eyeball using electrodes placed on the side of the eyes for horizontal shunting and another pair of electrodes above and below the eyes for vertical shunting (See Figure 6).

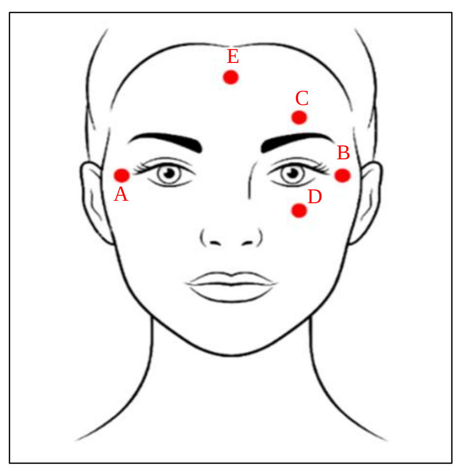


**Fig. 6 Electrode locations, point A (left), point B (right), point C (top), point D (bottom), and point E (reference)**

The retina, composed of cones and rods, exhibits different electrical sensitivities to light. The outer layer of the eye has a negative potential with respect to the synaptic termination in the retina, generating an electric field. The eye is considered an electric dipole, with the posterior part being negative compared to the cornea, which is positive.

Measurement of the potential difference between the cornea and the back of the eye provides a corneo-retinal potential of 6 mV. This potential, aligned with the optical axis and rotated with gaze, is measured by surface electrodes, recording actual potentials between 15 and 200 UV (see Figure 7). The proposed system consists of different stages that amplify and filter the biosignals captured by the electrodes (see Figure 8).

The Raspberry Pi 4 oversees receiving the signals and identifies which are up, down, left, right, and blinking. The signals picked up by the electrodes have magnitudes on the order of microvolts and are susceptible to noise, so it is essential to use an instrumentation amplifier. These amplifiers are used to measure extremely small voltages in the presence of interference. Such amplifiers possess the following key characteristics: a stable and linear selectable gain, expressed as “Ad”; a differential input with a high Common Mode Rejection Ratio (CMRR), and a high input impedance coupled with a very low output impedance.

Since the potential difference between the electrodes and the reference has a DC current component, a high-pass filter with a cutoff frequency of 0.15 Hz is implemented to remove this component from the signals. To reduce the high-frequency noise in the signal, a second-order low-pass filter of the Sallen-Key type with a cutoff frequency of 40 Hz and a gain of 2 was used.

Despite the application of these two filters, it was observed that the signal is still susceptible to 60 Hz interference from the mains. To address this problem, a Notch or band reject filter is used to eliminate this interference (see Figure 9).

After applying filters to the signals, a 4-channel Analog-to-Digital Converter (ADC) is used to feed the data into the Raspberry Pi. This data is then processed to derive outputs corresponding to up, down, right, and left movements. These outputs are sent to the interface to enable interaction with the user during conversation (See Figure 10).

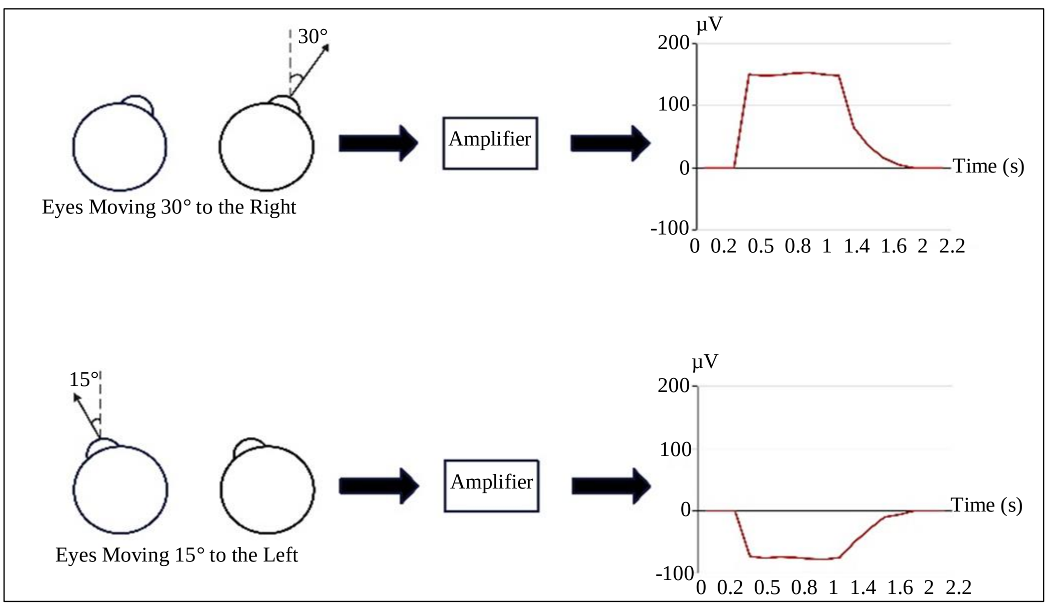


**Fig. 7 Electrooculogram (EOG) signal produced while the eyeball is rotated horizontally**

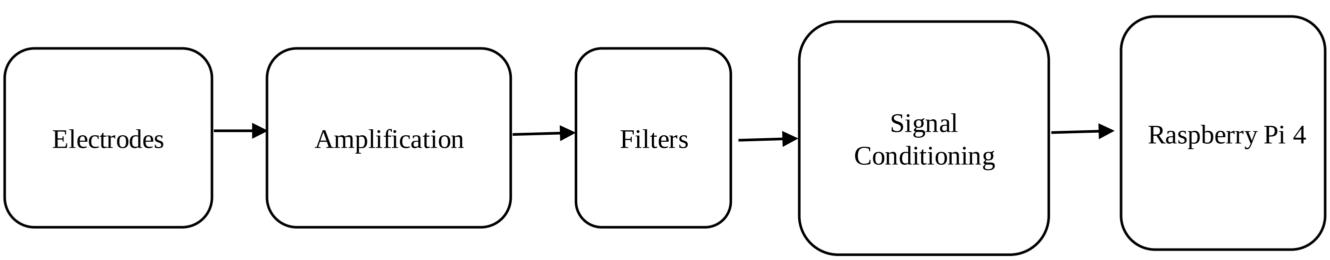


**Fig. 8 Block diagram of the proposed system**

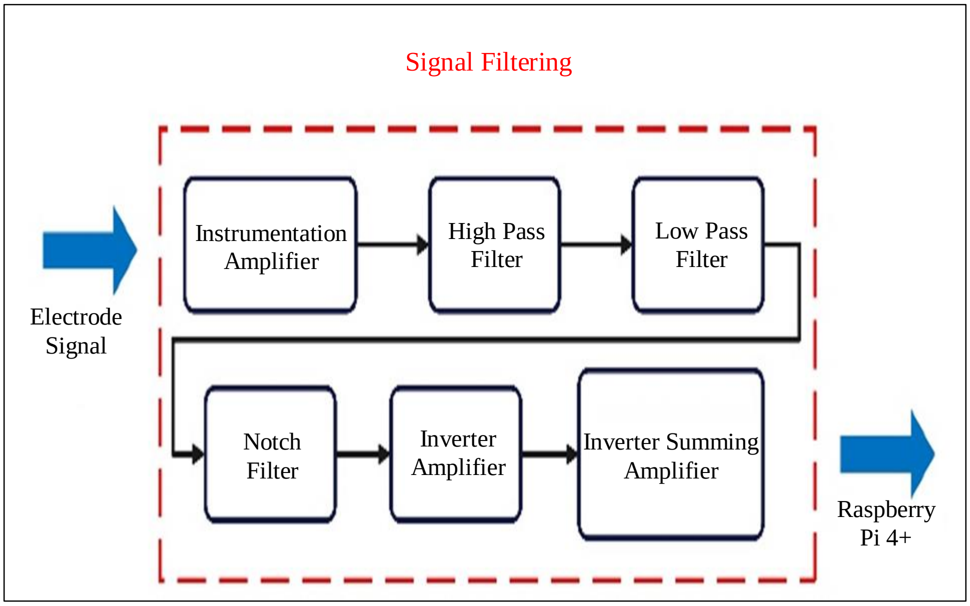


**Fig. 9 Signal filtering and conditioning steps**

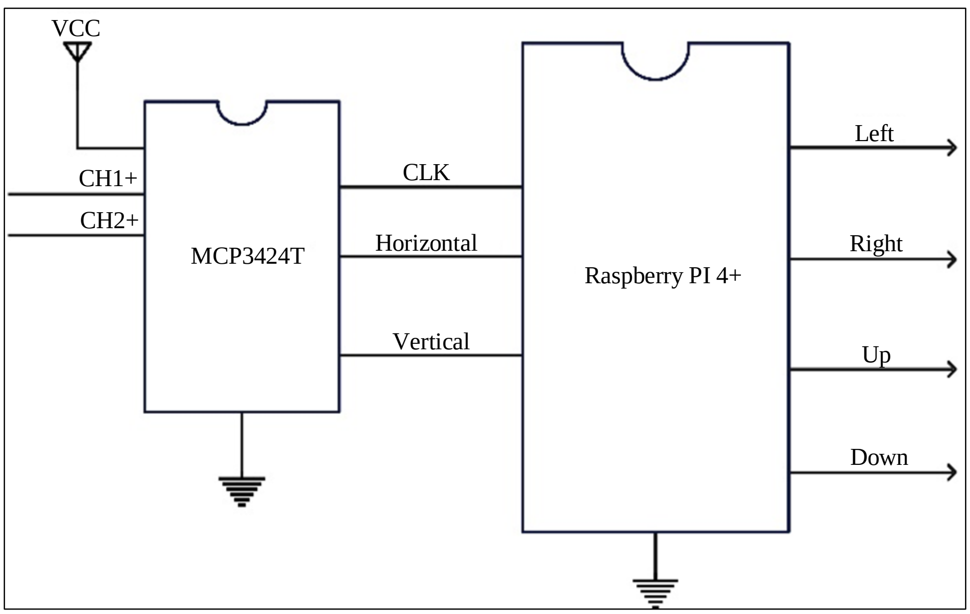


**Fig. 10 ADC communication with Raspberry Pi and output to GUI interface**

### *4.4. Options Response*

The answers proposed by the trained neural network proceeded to put in a Graphical User Interface (GUI) to display these options and be able to choose one and continue with the conversation. Two GUI interfaces were developed and applied to each user:

1. The first GUI interface was focused on people who have facial paralysis or who lost their speech due to an accident and have mobility of any limb. The GUI interface consists of an APK, which displays the response options provided by the neural network and 5 buttons to choose the alternative response that you want to communicate to User A (See Figure 11).

2. The second GUI interface developed was especially for people in a vegetative state, i.e. who only have eye movement. This interface consists of an APK that shows only 4 options with the distribution shown in Figure 12 to choose the alternatives to be communicated. This distribution was chosen so that the user can select with an EOG system. The eye movement in the upper direction selects alternative 1, in the left direction selects alternative 2, in the right direction selects alternative 3, and in the lower direction selects alternative 4.

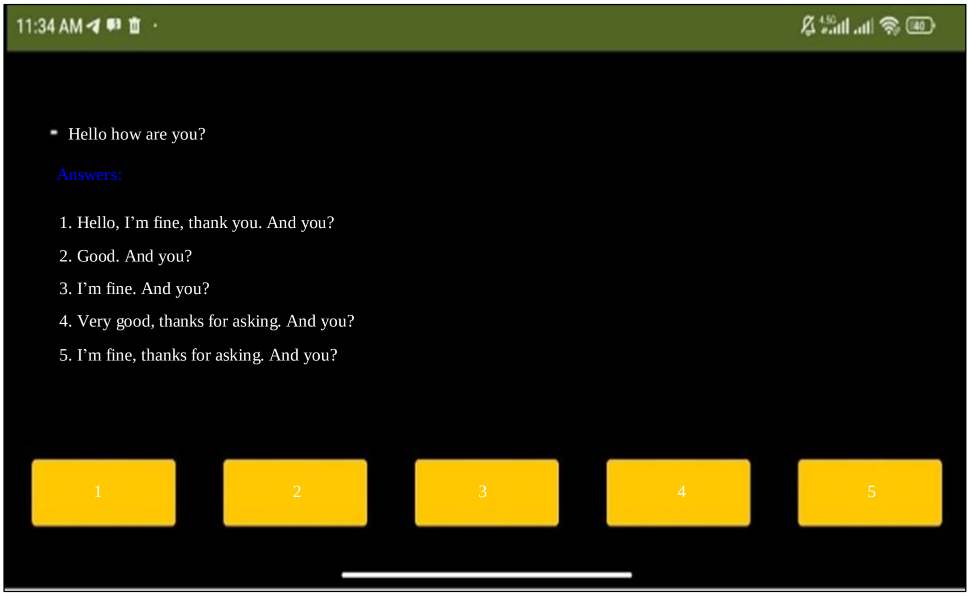


**Fig. 11 First GUI designed**

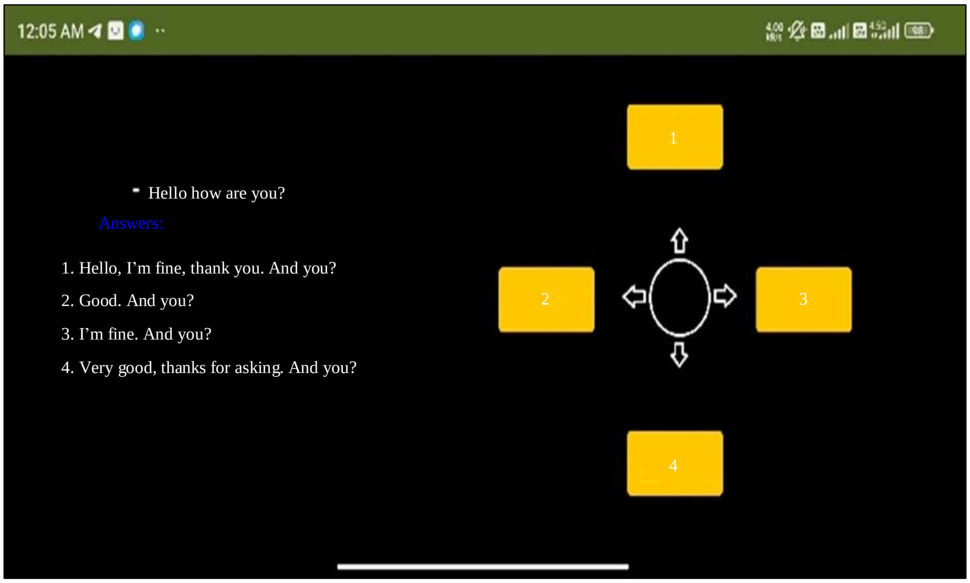


**Fig. 12 Second GUI designed**

## 5. Experimentation

For the first part of this communication system for people who cannot communicate, we tested 5 pairs of people. Each test consisted of 10 communication interactions between Users A and B and was performed in this way to evaluate two points:

- Conversation fluency: It was evaluated that the response options provided have relevance and agreement with what they want to communicate. In this way, redundancy was avoided, and the conversation was deeper and deeper.
- Intuitive, comfortable, and non-invasive: It was evaluated that the system is the most comfortable for the users.

To evaluate the two points mentioned above, we applied subjective tests to User A, the SUS method, which will indicate the usability of the proposed system. To evaluate User B, in the options of the developed application, we showed possible reactions to the system; this was done based on the method of Microsoft Reactions Cards; in this way, we are not invasive to User B obtaining their appreciation for the developed system.

## 6. Results and Discussion

Figure 13 shows the results of the SUS method applied to User A, where we have a great acceptance of the system, with an average of 99.2%, which means that our system is recommended in support of people's communication.

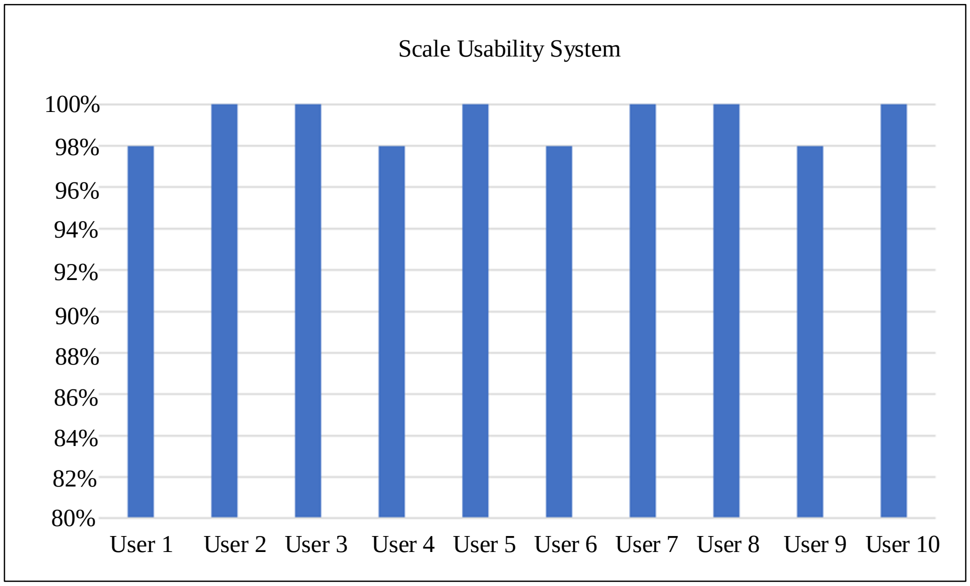


**Fig. 13 SUS applied to user A**

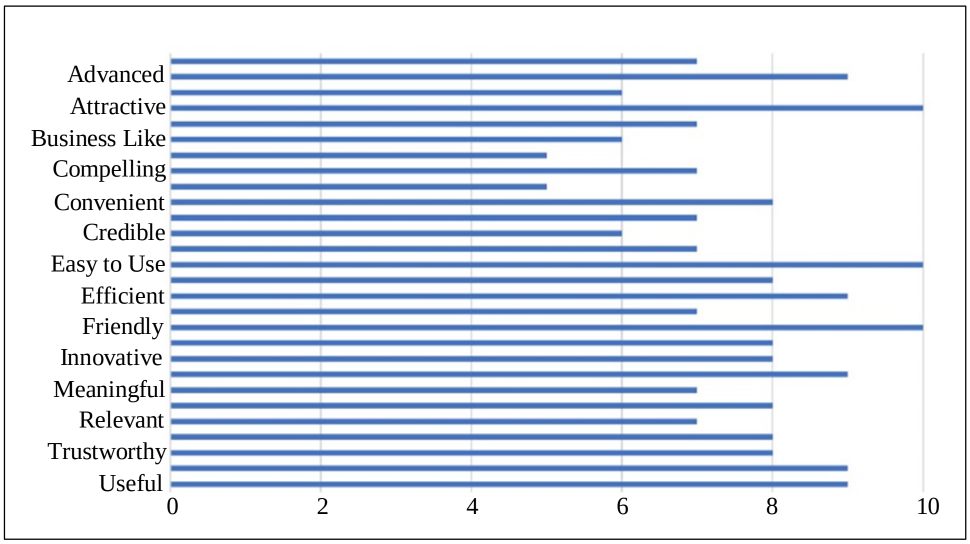


**Fig. 14 Microsoft reactions cards applied to user B**

Among the appreciations of User B, the most outstanding are the following (See Figure 14) among the most indicated are "easy to use", "friendly", "comfortable", and "collaborative". These results confirm that the system developed is effective and intuitive for creating a conversation.

## 7. Conclusion

We have presented a system using artificial intelligence, both voice recognition and trained neural networks, to achieve a conversation with people with communication difficulties. Our results suggest that the use of this system achieves a fluid conversation between users, with 99.2% accuracy and acceptance by users indicating that the system serves as a support to communicate with people.

The tests performed in the laboratory indicated that the system is ready to be tested with real patients.We tried to design the system as intuitive as possible for people in a vegetative state since, according to the literature reviewed, the methods developed can be invasive or exhausting, so we rely on artificial intelligence to release that burden. Future work will focus on real tests, i.e. no longer in a laboratory; the challenge will be to test it directly with people in a vegetative state, trying to be as minimally invasive as possible and to compare our proposed system with other systems developed today.

## Acknowledgments

The Universidad Nacional de San Agustin de Arequipa is gratefully acknowledged by the writers.

## References


[1] Gunes Ekim, Nuri Ikizler, and Ayten Atasoy, "A Study on Eye-Blink Detection-Based Communication System by Using K-Nearest Neighbors Classifier," *Advances in Electrical and Computer Engineering*, vol. 23, no. 1, pp. 71-78, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Armagan Amcalar, and Mujdat Cetin, "A Brain-Computer Interface System for Online Spelling," *2010 IEEE 18th Signal Processing and Communications Applications Conference*, Diyarbakir, Turkey, pp. 196-199, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[3] Faraz Akram, Seung Moo Han, and Tae-Seong Kim, "An Efficient Word Typing P300-BCI System Using a Modified T9 Interface and Random Forest Classifier," *Computers in Biology and Medicine*, vol. 56, pp. 30-36, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[4] Ali Bulent Usakli, and Serkan Gurkan, "Design of a Novel Efficient Human–Computer Interface: An Electrooculagram Based Virtual Keyboard," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 8, pp. 2099-2108, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[5] Chi Zhang, Rui Yao, and Jinpeng Cai, "Efficient Eye Typing with 9-Direction Gaze Estimation," *Multimed Tools Applications*, vol. 77, pp. 19679-19696, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[6] Jang-Zern Tsai et al., "A Feasibility Study of an Eye-Writing System Based on Electro-Oculography," *Journal of Medical and Biological Engineering*, vol. 28, no. 1, pp. 39-46, 2008. [Google Scholar]

[7] Kingshuk Mukherjee, and Debdatta Chatterjee, "Augmentative and Alternative Communication Device Based on Eye-Blink Detection and Conversion to Morse-Code to Aid Paralyzed Individuals," *2015 International Conference on Communication, Information & Computing Technology (ICCICT)*, Mumbai, India, pp. 1-5, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[8] Antonia Pirrera et al., "Assistive Technologies and Quadriplegia: A Map Point on the Development and Spread of the Tongue Barbell Piercing," *Healthcare*, vol. 11, no. 1. pp. 1-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Sayyed G. Maisikeli, "Level of Assistive Technology Use and Awareness among People with Physical Disability in the United Arab Emirates," *Technology and Disability*, vol. 35, no. 1, pp. 11-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Mohd Rizwan Jafar, and D.S. Nagesh, "Literature Review on Assistive Devices Available for Quadriplegic People: Indian Context," *Disability and Rehabilitation: Assistive Technology*, vol. 18, no. 6, pp. 929-941, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Nordine Sebkhi et al., "Evaluation of a Head-Tongue Controller for Power Wheelchair Driving by People With Quadriplegia," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 4, pp. 1302-1309, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Petra Karlsson, Megan Shepherd, and Ingrid Honan, "Accommodations to Cognitive Assessment for a Child with Dyskinetic Cerebral Palsy: Case Study", *Disability and Rehabilitation: Assistive Technology*, vol. 19, no. 2, pp. 360-366, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Shengchang Lan, Linting Ye, and Kang Zhang, "Applying mmWave Radar Sensors to Vocabulary-Level Dynamic Chinese Sign Language Recognition for the Community With Deafness and Hearing Loss", *IEEE Sensors Journal*, vol. 23, no. 22, pp. 27273-27283, 2023. [CrossRef] [Google Scholar] [Publisher Link]