Multi agent-based framework for interactive learning system with conversational user interface for visually and speech impaired

V Ilango¹, Anand Shankar Raja M² and V R Uma³

¹CMR Institute of Technology, Bengaluru, India
²,³Christ University, Bengaluru, India
¹Email: researchguide8@gmail.com

Abstract. The purpose of this research article is to illustrate the role of Conversational User Interface to foster inclusive education among the Speech Impaired People (SIP) and the Visually Impaired People (VIP). The multiagent framework on "Conversational User Interface for interactive learning" will throw light on the role of the host and the learner in communicating and facilitating knowledge with each other. The article has been built after the in-depth exploration of various sources, which were handpicked for the study. This multiagent model can be put into implementation in the education industry. The scope of the study extends only to SIP's and VIP's because of their growing numbers across the globe, and yet they are deprived of formal education. The Conversational User interface can be used to bridge the gap, especially with the advancement of technologies such as VUI, MUI, WUI, BOTs, etc., for the VIPs and SIPs.

Keywords— Speech Impaired People, Visually Impaired People, Conversational User Interface, Inclusive education, Transcoders, BOTs.

1. Introduction

Visual and speech impairments are of several types. These impairments could have occurred due to several reasons. Further the intensity of the impairment can also differ from one individual to the other. Hence, the demand from the impaired community varies according to the acuity of the impairment. Many researchers and experts are actively working to cater to their needs [21]. Visual impairment, besides causing problems to conduct the daily routine activities, additionally hinders them from communicating with the external world [2]. Vision acuity levels vary from one individual to the other. To mention a few types of visual impairments, there is loss of central vision, blind vision, blurred vision, light sensitivity etc. [10], [11]. On the other hand, there are many speech-impaired people across the world [20]. There are several types of speech disabilities like stuttering, apraxia, dysarthria etc. [9]. In general, the majority of them who have been classified under SIP's and VIP's are deprived of normal education. However, information technology can aid such individuals in education. Many pieces of researches propose new models for the visually impaired and speech impaired users. The Automatic Speech Recognition (ASR) systems developed using Hidden Markov Model [23] Raspberry pi & OCR sensor [1] involves the development of wearable visual aid. It functions on voice commands [4] etc. There are many frameworks and models proposed to aid the VIP's and SIP's
however; there is a need to bring an inclusive framework to boost the quality. The VIPs and SIPs should be brought under the ambit of everyday learning. This research gap has been filled by making use of Conversational User Interface. This research accelerates the role of a conversational user interface in bridging the host and the learner to facilitate knowledge and learning.

2. Ease of Use

2.1 Empowering the Speech Impaired and visually impaired through Conversational User Interface and its sub tools (Study Setup)

The purpose of this research article is to enlighten the uses of Conversational User Interface (CUI) as a tool to empower the SIP's and VIP's (Speech Impaired People & Visually Impaired People). Inclusive education goal can be attained only when everyone in society are given equal opportunities to learn and contribute towards economic prosperity [17]. Moreover, the report by [8] mentions that it is essential to incorporate Industry 4.0 technologies that are robust. Blending robotics, 3D printing, and mixed reality capabilities into their teaching methods can enhance a better learning experience. Can the Ministries of Education across the globe with the help of 4.0 tools and technologies build holistic and inclusive educational institutions? If yes! The entire globe will be united in the thread of “Education without discrimination based on disabilities”. Disability cannot be avoided. The cause of disability could be natural or accidental, yet the impact is detrimental. However, the long-term impact can be reduced by providing them with education with the help of appropriate tools and technologies.

The study setup is clearly explained in Figure 1.

2.2 Conversational User Interface-An Overview

Virtual assistants are essential these days. Spurt in advancement in technology in the present era has lead to the use of virtual assistants in various fields. Conversational agents, when embedded in a dialogue system, can help a person to communicate in a natural language and in turn would also get a response in a similar language, thus facilitating seamless communication [22]. There are several task-oriented dialogue agents, the inbuilt software and database structure facilitated through Artificial Intelligence responds according to the query. To illustrate a task-oriented dialogue agent, let us consider the travel desk indicator software where a Traveler might ask a question to the dialogue system, and the latter would understand and give a proper feedback according to the query raised by the former. The second most important conversational tools are the chatbots that has the ability to strike a conversation between the humans and the Machines [7]. They are powered with artificial intelligence and also leverage the Natural Language Processing (NLP) to a larger extent.

The chatbox is a popular tool that has the ability to understand the request of humans, searches the best information as per the query placed, and it would also compose the response in the best possible manner according to the human expectation. Thus, AI has given opportunities for humans to make use of different machines, which are empowered to work for the benefit of the society and the economy itself [5]. Humans are able to enjoy these benefits because of artificial intelligence. AI is being used in various fields of computer science due to the incorporation of natural language processing NLP [12], [14]. Thus, the best for any conversational user interface is its dependence on the natural language processing technique, which helps in various purposes such as learning, conversion of text into audio or conversion of any information into useful and meaningful information as per the learner’s requirements. Also, it has the ability to convert unstructured data into a structured format, which can be used for better purposes [15], [3]. These conversational user interfaces are more flexible. They are devoid of emotions due to which a query can be raised multiple times, and the CUI would also respond without fatigue. Moreover, the ability to accept commands and respond in fraction of seconds facilitates quick decision making. Most importantly, CUIs have the ability to respond with clarity and accurate information. There, are different types of conversational user interface related tools like persona (A tool that is designed with artificial intelligence to convert the conversation into a written text, and text into conversation) [16]. Such type of tools would be very helpful for that community who cannot speak because of their disability.
Additionally, the visually impaired, who have the thirst to get educated would also benefit with the help of these advances tools and technologies [18]. On the other hand, there are other types of advanced tools and technologies such as the automatic speech recognition ASR, which has the ability to understand and extract the actual meaning to facilitate human understanding [6]. For example, there are many students who face issues because of the speech impairment. With the help of automatic speech recognition tools, the audio that is inputted by the speech impaired gets converted into an output that is understandable by a normal teacher thus facilitating normal conversation between the former and the latter.

The conversational user interface also has the most advanced version such as the follow-up intents like the Alexa which has the ability to understand the input given by a human and to detect what is in need and also to produce the same. Thus, such type of new tools and technologies has made human life very simple and more authentic. There are different types of Conversational User Interfaces CUI like Voice User Interface (VUI), Web User Interface (WUI) [19], Multilingual User Interface (MUI) [24], and Graphical User Interface (GUI) [13] Chatbots etc. The CUI is bound with transcoders, which plays the role of converting the audio, video, text into a desired format. In this process, data is converted into a desired format of the learner. It also facilitates better understanding by the tutor. The tutor can understand the disable learner better when the learner’s converse through braille/cloud telephony/App/VoIP/PDA/chatbots. These interfaces can help the student community that is impaired, especially the speech and the visually impaired, to undertake education as normal persons. It is also beneficial to Parents and teachers. They can understand their ward’s needs and expectation better. In simple words, the role of Conversational User Interfaces can uplift the life of disabled people in a meaningful way.

![Figure 1. Multi Agent based Framework for interactive learning system with Conversational User Interface for Visually and Speech Impaired](image)

Figure 1. Multi Agent based Framework for interactive learning system with Conversational User Interface for Visually and Speech Impaired
The speech and the visually impaired people can lead a normal life in the society. Figure 1 mentions the three core dimensions of Conversational User Interface. The proposed framework uses transcoder. It matches the learner’s preference or requirement and accordingly converts the input into an appropriate output without altering the digital content. The details of the transcoder that will be useful for disabled learners are listed in Table 1.

Table 1. Transcoder for disabled learners

| Transcoder          | Description                                                                 |
|---------------------|-----------------------------------------------------------------------------|
| Text-to-speech      | It synthesises the speech by converting the text input into voice or speech output. This is of great help to the visually impaired as they will not be able to read the digital content. It allows the learner to interact with the machine as a normal person. |
| Text-to-image       | The text input is converted to image output. Except the visually challenged, it is helpful for all other forms of disabled learners. |
| Text-to-video       | Similar to the text to image, the text to video transcoder also helps all other forms of disabled learners except the visually challenged. The input is given in the form of text, which is converted into a video output. |
| Speech-to-text      | The auditory impaired learners will not be able to listen to the speech or audio sessions in the digital learning content. The speech to text transcoder converts the audio input to text thus facilitating visual access to read. |
| Speech-to-video     | To facilitate better learning experience, the speech or audio input is converted to video output. It is helpful for the learners who are auditory impaired. |
| Image-to-animation  | The static images are converted into animation by the transcoder for a better learning experience. |
| Image-to-video      | The input is the image and it is converted to video. |

3. Conclusion

Sustainable education can be achieved through various technologies and tools available in this world. Conversational User Interface and its types are used in our routine lives in various forms for various purposes. However, there is a need to use the same in the educational sector to harmonize the VIP’s and SIP’s to learn and benefit and to help them to communicate and connect with the outer world without any discrepancies. CUI a tool powered by AI can easily help in the communication process between the learner and the host. Moreover, the CUI depends on NLP (Natural Language Processing) to communicate between the computer and the humans. Educating the speech and visually impaired can lead to economic prosperity as they would also get employment opportunities.

Suggestions and recommendations

Governments across the globe must facilitate education with the help of technological laboratories for the betterment of the society. Also, with the help of advanced AI tools, new skills can be developed among the SIP’s and VIP’s.
References

[1] A Karthick, S Prabhakaran, Voice Assistance for Visually Impaired People, International Conference on Communication, Computing and Internet of Things (IC3IoT), IEEE, pp 465-468, 2018.

[2] J F Allen, Byron, D K, Dzikovska M, Ferguson, G, Galescu, L, and Stent A. Toward Conversational Human-Computer Interaction, Artificial Intelligence Magazine, 22, 4, 2001.

[3] K R R and C Chitra, A Smart Shopping System for Visually Impaired,” International Journal of Innovative Technology and Exploring Engineering (IJITEE),8, Issue 11, pp 315–318, 2019.

[4] D. S. Sulamoyo, Preparing for the future, Community Action Leaders Rooting Out Poverty Local Lev, October, pp. 211–236, 2016.

[5] C C Chiu et al., State-of-the-Art Speech Recognition with Sequence-to-Sequence Models, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2018.

[6] N. Gerfelder, Novel user interface technologies and conversational user interfaces for information appliances, ACM Digital Library, 2000. [Online]. Available: https://dl.acm.org/doi/10.1145/633292.633320.

[7] V. Gupta, A. Kambil, S. Ghosh, and S. Mondal, Breaking the mold: The future of Indian Educators, 2019.

[8] A. K. Namasivayam et al., Treatment intensity and childhood apraxia of speech, Pediatrics-Official Journal of the American Academy of Pediatrics, 50, Issue No: 4, pp 529–546, 2015.

[9] B. D. Nuertey et al., Prevalence, Causes, and Factors Associated with Visual Impairment and Blindness among Registered Pensioners in Ghana, Journal of Ophthalmology, Article ID 1717464, 10 pageshttps://doi.org/10.1155/2019/1717464, 2019.

[10] Dr. Ananya Mandal, Types of visual impairment,[Online]. News Medical Life Science, 2019.

[11] A. R. Martinez, Natural language processing, Wiley Online Library, 2, Issue No: 3, May/June 2010.

[12] MathWorks Inc., Creating Graphical User Interfaces, MATLAB User Guide, pp 1-502, 2015.

[13] D. Maynard, K. Bontcheva, and I. Augenstein, “Natural Language Processing for the Semantic Web. Morgan and Claypool”, ISBN: 9781627059091, December 2016

[14] McTear, Michael, Callejas, Zorida, Griol Barres, David, The conversational interface: Past and Conversational Interface Talk to Smart Devices”, Springer, pp 1–422, 2016.

[15] Marta Medina-García. Equal opportunities in an inclusive and sustainable education system: An explanatory model. Sustainability, Sustain, 12, Issue No: 11, 2020.

[16] Narayanan, Potamianos, Creating conversational interfaces for children, IEEE Transactions on Speech and Audio Processing, Page: 65-78, ISSN: 1063-6676, Electronic ISSN: 1558-2353, February 2002.

[17] Xuedong Huang, An overview of modern speech recognition, Handbook of Natural Language Processing, Second Edition, 2010.

[18] M. C. Roy, O. Dewit, and B. A. Aubert, The impact of interface usability on trust in Web retailers,Internet Research, Submitted to Information Technology Department, University of Montreal Quebec, pp 1-20, 2001.

[19] Gemma Sellars, Learning to communicate with children with disabilities, National Library of Medicine, National Center for Biotechnology Information, 18, Issue No: 9, pp 26–28, 2006.
[20] Mick Healey, Alan Jenkins, Jonathan Leach, Carolyn Roberts, Issues in Providing Learning Support for Disabled Students Undertaking Fieldwork and Related Activities, *Geography Discipline Network (GDN)*, pp 1-65, 2010.

[21] Azeta, Ambrose A, Inam, Itorobong A, Daramola, Olawande, Voice-based e-examination framework for visually impaired students in open and distance learning, *Turkish Online Journal of Distance Education, 19*, Issue No: 2, Page: 34–46, 2018.

[22] Y. C. Huang and C. H. Tsai, Speech-Based Interface for Visually Impaired Users, *Proc. - 20th Int. Conf. High Perform. Comput. Commun. 16th Int. Conf. Smart City 4th Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2018*, pp. 1223–1228, 2019.

[23] Xiaozhong Liu, Full-Text Citation Analysis: A New Method to Enhance, *Journal of the American Society for Information Science and Technology, 64*, Page:1852–1863, 2013.