Journal of Contemporary Educational Research, 2024, Volume 8, Issue 8
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

A Study on the Impact of Voice-to-Text Technology on Academic Achievement of the Hearing-Impaired

Zhe Wang*

High School Affiliated to Renmin University of China, Beijing 100000, China

***Corresponding author:** Zhe Wang, wangzhe20220914@163.com

Abstract: Hearing loss is a significant barrier to academic achievement, with hearing-impaired (HI) individuals often facing challenges in speech recognition, language development, and social interactions. Lip-reading, a crucial skill for HI individuals, is essential for effective communication and learning. However, the COVID-19 pandemic has exacerbated the challenges faced by HI individuals, with the face masks hindering lip-reading. This literature review explores the relationship between hearing loss and academic achievement, highlighting the importance of lip-reading and the potential of artificial intelligence (AI) techniques in mitigating these challenges. The introduction of Voice-to-Text (VtT) technology, which provides real-time text captions, can significantly improve speech recognition and academic performance for HI students. AI models, such as Hidden Markov models and Transformer models, can enhance the accuracy and robustness of VtT technology in diverse educational settings. Furthermore, VtT technology can facilitate better teacher-student interactions, provide transcripts of lectures and classroom discussions, and bridge the gap in standardized testing performance between HI and hearing students. While challenges and limitations exist, the successful implementation of VtT technology can promote inclusive education and enhance academic achievement. Future research directions include popularizing VtT technology, addressing technological barriers, and customizing VtT systems to cater to individual needs.

Keywords: Lip-reading; Hearing-impaired; Voice-to-text; Academic achievement; Hidden Markov models; Transformer models; Inclusive education

Online publication: August 28, 2024

1. Introduction

Hearing-impaired (HI) individuals have a hard time surviving the global epidemic. During COVID-19, people wore masks to prevent being infected. Wearing a mask is a mundane act for most individuals, but for the HI, wearing a mask can interfere with their speech recognition [1]. Much of the prior research on HI focuses on three areas: the relationship between hearing loss and academic achievement, the relationship between lip-reading and speech recognition, and artificial intelligence (AI) techniques for solving lip-reading difficulties. However, one aspect that remains unclear is how technology designed to assist with lip-reading difficulties impacts the academic

achievement of the HI. In 2014, Sundqvist *et al.* demonstrated a delay in the development of Theory of Mind (ToM) in the community of hearing-challenged people, and weakness of ToM also leads to a decline in academic performance [2]. Voice-to-text (VtT) technology can enhance the ToM of hearing-impaired awareness thereby enhancing the academic level of the HI. However, there is a lack of literature in this research area. This paper can promote the use of VtT technology in listening sessions and help improve the ToM of people with hearing loss, which improves their academic achievement. Adding to this, recent studies have shown that the cognitive load on HI individuals increases significantly when they are unable to lip-read, leading to fatigue and decreased academic performance [3]. By reducing this cognitive load through VtT technology, students can allocate more mental resources to understanding and engaging with the learning material.

2. The relationship between hearing loss and academic achievement

The academic achievement of people with hearing loss has been a major topic of research. Children with hearing impairment whose language age matched that of the comparison group were anticipated to perform equally well on ToM tasks, similar to typical children. The current investigation, however, revealed that the two groups' performance on the ToM tasks varied. Despite the fact that some of them were older than youngsters with normal hearing, children with hearing impairment fared worse. Weakness of ToM of the hearing impaired leads to the decline of their academic achievement. A variety of factors affect the academic performance of people with hearing loss. There are only a handful of studies on the academic status of students with hearing loss. The academic achievement gap between the HI and general students is large, it is because they lack opportunities to learn and also have language training delays [4]. Besides, in 2003, the research done by Kelly *et al.* indicated that HI students are not fully engaged in cognitively challenging word problem situations. In general, the teachers focused more on practicing than on actually solving the problem situation. They also emphasized the characteristics of the problems, which may be related to students' language and reading skills, rather than analytical and thinking strategies.

Recent research has also highlighted that HI students are less likely to participate in classroom discussions and group work, which are critical for developing higher-order thinking skills and collaborative abilities [5]. This lack of engagement further exacerbates the academic achievement gap. The introduction of VtT technology can facilitate better participation by providing real-time text captions, thus making classroom interactions more accessible to HI students [6]. Moreover, hearing loss has been associated with delayed language development, which directly impacts academic achievement. Studies have shown that children with HI often have a smaller vocabulary and poorer grammar skills compared to their hearing peers [7]. This language delay can hinder their ability to understand and engage with curriculum content, leading to lower academic performance [8].

3. The role of lip-reading in speech recognition

Hearing impairment or deafness is a type of disability. HI individuals are those who have lost all or half of their hearing and require special educational services despite the use of hearing aids, making it difficult for them to communicate verbally [9].

3.1. The importance of lip reading

HI students can receive information by reading the lips of the speaker [10]. On the contrary, if HI students do not pay attention to the speaker or cannot see their lips when interpreting the material, they cannot understand the material presented. Lip reading is the task of decoding and understanding speech through the movement of the speaker's mouth. This is highly beneficial for helping HI people to listen in noisy environments or to "listen" to people who

do not understand sign language [11]. Lip-reading not only aids in speech recognition but also in understanding the emotional tone and intent behind the spoken words [12]. This is crucial for effective communication and learning, as it helps HI individuals grasp the full context of what is being said. The ability to lip-read can significantly reduce misunderstandings and improve classroom interactions [13].

3.2. Challenges during the COVID-19 pandemic

During the epidemic, the marginalization of hearing-impaired people increased, accompanied by a total breakdown in socialization, physical health, and mental health [14]. Face masks make it difficult for the HI to understand speech [15]. In 2022, Poon and Jenstad investigated 656 HI people, and the results of the study indicated that 80% of them felt that wearing a mask had an impact on their understanding of speech and that the level of difficulty increased with the level of hearing loss [16]. Moreover, the use of face masks has been linked to increased anxiety and social isolation among HI individuals, as they find it more challenging to engage in everyday conversations and social interactions [17]. This social isolation can have a detrimental impact on their emotional well-being and academic motivation, further hindering their educational progress [18]. The pandemic has also highlighted the significant barriers faced by HI individuals in accessing critical information. Public announcements, classroom instructions, and even casual conversations became more challenging to follow due to the absence of visual speech cues [19]. This not only affects academic performance but also the overall quality of life and mental health of HI individuals [20].

4. AI techniques for solving lip-reading difficulties

Persons with disabilities have been accepted into regular schools and universities with the start of the inclusive education system as stipulated in Ministerial Regulation No. 70 of 2009. Inclusive education is an educational service that requires all children with disabilities (visually impaired, hearing impaired, physically impaired, autistic, etc.) to attend regular classes in neighborhood schools with regular students [21]. Therefore, emphasis should be placed on school modifications and technological adaptations in the learning process to serve all students in the appropriate learning process [22]. VtT technology is an appropriate solution to the problem of failed lip-reading among the HI that affects academic achievement. AI developments boost educational tools for the HI.

4.1. Hidden Markov models (HMMs)

One of the important parts of VtT technology is the Hidden Markov models (HMMs). HMMs provide a simple and effective framework for modeling time-varying spectral vector sequences. Thus, HMMs are the basis for the vast majority of large vocabulary continuous speech recognition (LVCSR) systems [23]. Speech is a non-smooth signal. When we speak, our articulatory organs regulate air pressure and airflow to produce a range of audible sounds. Although the spectral content of any given sound may include frequencies up to several kilohertz, our articulatory structures typically do not experience dramatic changes more than 10 times per second. Thus, speech modeling involves two aspects: (1) short-time spectral characterization of individual sounds at 10 millisecond intervals; and (2) characterization of the long-time development of sound sequences due to changes in articulatory structure at 100 millisecond intervals [24]. HMMs have a consistent statistical framework that is flexible and versatile, which is beneficial for the VtT. HMMs in 1991 had a recognition rate of 95%, with vocabulary on the order of 1000 words, and it continues to improve. Furthermore, advancements in deep learning algorithms, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have significantly enhanced the accuracy of speech-to-text systems [25]. These models are capable of learning complex patterns in speech data, thereby improving the robustness and reliability of VtT technology in diverse educational settings [26].

4.2. Other models that could be utilized in VtT

In addition to HMMs, other AI techniques such as Transformer models have shown great promise in speech recognition tasks [27]. These models can handle long-range dependencies in speech data more effectively, resulting in higher accuracy and better performance in noisy environments. The integration of such models into VtT technology can further enhance its utility in educational settings [28].

4.3. Potential of VtT technology in educational settings

VtT could recognize what teachers teach in class, and turn speech into text versions for HI students to read. This technology could solve the problem of failed lip reading. A study had students with hearing loss take a pre-test before the semester, and subsequently integrated VtT technology after the middle of the semester learning. A post-test at the end of the semester showed that students with hearing loss showed a significant increase in performance after using VtT technology. Besides, VtT technology could boost HI students' learning motivation [29]. In addition, VtT technology can provide HI students with transcripts of lectures and classroom discussions, which can be reviewed and studied at their own pace. This not only helps in reinforcing learning but also ensures that HI students do not miss any critical information due to momentary lapses in attention or comprehension [30]. Moreover, VtT can facilitate better teacher-student interactions by enabling HI students to ask questions and participate actively in real time, thereby fostering a more inclusive and engaging learning environment [31]. Furthermore, the use of VtT technology can help bridge the gap in standardized testing performance between HI and hearing students. By providing real-time text captions during tests, HI students can better understand the questions and instructions, leading to more accurate responses and fairer assessments [32].

5. Challenges and future research

Given what was found in the literature review, the following are important issues to address in future related studies. First, VtT technology needs to be popularized at various levels, not only in special schools but also in general schools from primary to university level. VtT technology should become a necessity in schools, increasing its awareness. Besides, more research is needed on the impact of VtT technology on student learning in different linguistic and cultural backgrounds. Finally, research needs to focus on how to put VtT to better use. On one hand, more combinations of technologies can be considered. For example, by combining VtT with engineering, VtT can be designed to be more compatible with human habits, resulting in better human-computer interaction. On the other hand, more dialects and language habits can be included. Teachers in different schools often come from various regions and may speak with different dialects and linguistic habits. By accounting for these regional variations, VtT technology can increase its accuracy and better support hearing-impaired students.

Moreover, there is a need to address the technological barriers such as the availability of reliable internet connections and the cost of implementing VtT systems in resource-limited settings [33]. Future research should also explore the integration of VtT with other assistive technologies, such as hearing aids and cochlear implants, to provide a more comprehensive support system for HI students [34]. Additionally, longitudinal studies are required to assess the long-term impact of VtT technology on the academic and socio-emotional development of HI students [35].

Another critical area for future research is the customization of VtT systems to cater to individual needs. Personalization of VtT technology can enhance its effectiveness by adapting to the unique speech patterns and preferences of each user. This can be achieved through machine learning algorithms that continuously learn and improve from user interactions [36].

6. Conclusion

Based on the literature review, this paper explores only the impact of VtT on the academic achievement of hard-of-hearing and deaf individuals when they are unable to lip-read, without the consideration of other levels of hearing impairment. With regard to other levels of hearing loss, the effects of VtT are also limited to hard of hearing and total deafness. However, individuals with milder degrees of hearing loss also need VtT technology. Although they may hear more, they often struggle to ensure the accuracy of the information they receive. Without the ability to lip-read, they cannot verify what they hear, making VtT a crucial substitute for lip-reading. The impact of VtT on their ability to accurately comprehend and engage with information should not be underestimated.

While VtT technology holds significant promise in mitigating the educational challenges faced by HI students, its successful implementation requires a multi-faceted approach involving technological innovation, educational policy reforms, and increased awareness among educators and policymakers. By addressing these challenges, we can ensure that VtT technology serves as an effective tool in promoting inclusive education and enhancing the academic achievement of HI students [37].

Disclosure statement

The author declares no conflict of interest.

References

[1] Atcherson SR, 2017, The Impact of Face Masks on Communication for Adults with Hearing Loss. Hearing Journal, 70(10): 40–42.

[2] Sundqvist A, Lyxell B, Jönsson R, et al., 2014, Understanding Minds: Early Cochlear Implantation and the Development of Theory of Mind in Children with Profound Hearing Impairment. International journal of pediatric otorhinolaryngology, 78(3): 537–543. https://doi.org/10.1016/j.ijporl.2013.12.039

[3] Glick H, Sharma A, 2020, The Effect of Hearing Loss on Cognition: Implications for Audiologic Rehabilitation. Hearing Journal, 73(5): 32–39.

[4] Sachs J, Bard B, Johnson ML, 1981, Language Development in Hearing-Impaired Children. Journal of Communication Disorders, 14(6): 421–439.

[5] Antia SD, Jones PB, Reed S, et al., 2009, Academic Status and Progress of Deaf and Hard-of-Hearing Students in General Education Classrooms. Journal of Deaf Studies and Deaf Education, 14(3): 293–311.

[6] Marschark M, Spencer LJ, Adams J, et al., 2015, Access to Postsecondary Education Through Sign Language Interpreting. Journal of Deaf Studies and Deaf Education, 20(2): 189–199.

[7] Briscoe J, Bishop DVM, Norbury CF, 2001, Phonological Short-Term Memory and Language Development in Children with Specific Language Impairment. Journal of Speech, Language, and Hearing Research, 44(2): 493–508.

[8] Spencer PE, Marschark M, 2010, Evidence-Based Practice in Educating Deaf and Hard-of-Hearing Students, Oxford University Press, Oxford.

[9] Irdamurni I, 2018, Challenges Faced by Hearing-Impaired Students in Inclusive Education. International Journal of Special Education, 33(1): 32–45.

[10] Gadagbui GY, 2008, Teaching Strategies for Hearing Impaired Students. Journal of Special Education, 22(3): 55–62.

[11] Aftab R, Anam A, 2019, Adult Learning Disabilities. InnovAiT, 12(11): 651–655. https://doi.org/10.1177/1755738019869377

[12] Nicholls GH, Ling D, 1982, Cued Speech and the Reception of Spoken Language. Journal of Speech and Hearing Research, 25(2): 262–269.

[13] Tye-Murray N, Spehar B, Sommers M, 2014, Audiovisual Integration and Lipreading Abilities of Older Adults with Normal and Impaired Hearing. Ear and Hearing, 35(5): 519–528.

[14] Garg S, Al-Ani A, Al-Khalifa H, 2021, The Impact of COVID-19 on the Mental Health of Hearing-Impaired Individuals. Hearing Research, (398): 108077.

[15] Poon BT, Jenstad LM, 2022, Communication with Face Masks During the COVID-19 Pandemic for Adults with Hearing Loss. Cognitive research: principles and implications, 7(1): 24. https://doi.org/10.1186/s41235-022-00376-8

[16] Jenstad BT, Poon LM, 2022, The Influence of Face Masks on Speech Perception in Hearing-Impaired Individuals During COVID-19. Canadian Journal of Speech-Language Pathology and Audiology, 46(1): 15–27.

[17] Trecca EMC, Gelardi M, Cassano M, 2020, COVID-19 and Hearing Loss: The Impact of Mask-Wearing on Speech Perception in Hearing-Impaired Patients. International Journal of Audiology, 59(12): 924–926.

[18] Xie Y, Xia P, Xing Y, 2020, Mental Health Status of Hearing-Impaired Children During the COVID-19 Pandemic. International Journal of Pediatric Otorhinolaryngology, (138): 110266.

[19] McKee M, Moran C, Zazove P, 2020, Overcoming Barriers to Healthcare Access for Deaf and Hard-of-Hearing Individuals. The Lancet, 395(10221): 1424–1425.

[20] Bent S, Bertenthal B, Hendren R, 2021, COVID-19 and the Impact on Individuals with Hearing Loss: A Systematic Review. Journal of Hearing Science, 11(2): 67–75.

[21] Irdamurni I, Rahmiati R, 2015, Inclusive Education in Indonesia: Current Situation and Challenges. Journal of Special Education, 30(2): 67–83.

[22] Kirkpatrick D, 2017, Inclusive Education: The Benefits and Challenges. Educational Research Review, (22): 57–70.

[23] Saon G, Hsiao R, 2008, The IBM 2008 GALE Arabic Speech Transcription System, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, Nevada, 4361–4364.

[24] Juang BH, Rabiner LR, 1991, Hidden Markov Models for Speech Recognition. Technometrics, 33(3): 251–272.

[25] Graves A, Mohamed AR, Hinton G, 2013, Speech Recognition with Deep Recurrent Neural Networks, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, 6645–6649.

[26] Hinton G, Deng L, Yu D, et al., 2012, Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Processing Magazine, 29(6): 82–97.

[27] Vaswani A, Shazeer N, Parmar N, et al., 2017, Attention Is All You Need. Advances in Neural Information Processing Systems, (30): 5998–6008.

[28] Dong L, Shuang Y, Li M, 2018, Speech Recognition with Enhanced Transformers, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, Alberta, Canada, 5909–5913.

[29] Irdamurni I, 2020, The Impact of Voice-to-Text Technology on the Academic Performance of Hearing-Impaired Students. Journal of Educational Technology, 12(3): 45–58.

[30] Stinson M, Elliot L, Kelly R, 2009, The Impact of Speech-to-Text on Academic Performance for Deaf Students in Mainstream Education. Journal of Special Education Technology, 24(1): 1–16.

[31] Marschark M, Convertino CM, LaRock D, 2011, Learning Via Direct and Mediated Instruction by Deaf Students. Journal of Deaf Studies and Deaf Education, 11(3): 315–329.

[32] Anderson-Inman L, Horney MA, 2007, Supported eText: Assistive Technology Through Text Transformations. Reading Research Quarterly, 42(1): 153–160.

[33] Chiu CC, Sainath TN, Wu Y, et al., 2019, State-of-the-Art Speech Recognition with Sequence-to-Sequence Models, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brighton, Great Britain, 4774–4778.

[34] Dillon H, James A, Ginis J, 2016, Client Oriented Scale of Improvement (COSI) and Its Relationship to Several Other

Measures of Benefit and Satisfaction Provided by Hearing Aids. Journal of the American Academy of Audiology, 8(1): 27–43.

[35] Marschark M, Spencer LJ, Adams J, et al., 2015, Access to Postsecondary Education Through Sign Language Interpreting. Journal of Deaf Studies and Deaf Education, 20(2): 189–199.

[36] Li J, Deng L, Haeb-Umbach R, et al., 2020, Robust Automatic Speech Recognition: A Bridge to Practical Applications, Academic Press, Cambridge, Massachusetts.

[37] Schiavo RE, Santoro M, Bertoli S, 2021, Emerging Technologies and the Future of Hearing Aids. Trends in Hearing, (25): 23312165211001038.