IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 27, Issue 6, Ser. 1 (Nov. – Dec. 2025), PP 95-101
www.iosrjournals.org

An Experimental Study on Keystroke Dynamics for Caption Readability in Deaf and Hard-of-Hearing Users

Prashant Tiwari [1], Dr. Dhirendra Kumar Tripathi [2]
[1,2] Department of Computer Science, Mansarovar Global University, Sehore, M.P.

Abstract
There is a lack of objective ways for assessing caption readability, despite the fact that captioning is crucial in making multimedia material accessible to Deaf and Hard-of-Hearing (DHH) users. In order to objectively and quantitatively analyze the cognitive strain of reading captions, this study suggests keystroke dynamics. The experimental setup involved presenting DHH with a numeric key-input job while simultaneously exposing them to other forms of visual information, such as symbols, icons, closed captions, alphabetic text, and text shown centrally and peripherally. A variety of visual stimuli were tested for their effects on user performance by monitoring task processing time, keystroke frequency, and mistake rates. Processing durations rise and keystroke rates decrease substantially when exposed to linguistic information, especially continuous caption text, according to the results. Captioning led to the most noticeable decline in performance, but symbolic and icon-based stimuli were less disruptive than alphabetic information displayed on the periphery of the field of view. Keystroke dynamics offer an objective and dependable metric for assessing cognitive demand and caption readability, as shown by these results. In addition to lending credence to evidence-based caption optimization for DHH users, the study offers useful insights into the design of accessible captioning systems.
Keywords: *Caption Readability, Deaf, Keystroke, Peripheral, Symbol*

Date of Submission: 01-11-2025 Date of Acceptance: 13-11-2025

I. INTRODUCTION

Digital multimedia content's meteoric rise in popularity has revolutionized the way people all over the world access and consume information in fields such as education, entertainment, and professional training. In today's educational landscape, video-based learning platforms, MOOCs, taped lectures, webinars, and instructional videos play a crucial role. There has been an increase in accessibility for people with disabilities, especially those who are Deaf or Hard of Hearing (DHH), as a result of these advancements, which have increased reach and flexibility. Rather than being an afterthought, captions are the main means by which this demographic is able to access spoken material. Access to information, meaningful learning experiences, and social inclusion are all greatly impacted by how well captions are read and the overall quality of the captions.

The term "caption readability" describes how well the captioned text may be perceived, read, and digested by viewers in the allotted screen time. Captions, in contrast to regular text, have several unique requirements, such as being time-sensitive, adapting to different viewing rates, and maintaining semantic correctness. Captions are sometimes the only means of conveying auditory information for Deaf and Hard-of-Hearing users, therefore any problems with readability, such rapid speech, complicated sentence structures, unclear language, or inadequate segmentation, are very harmful. The accessibility aims that captions are meant to achieve can be undermined by poorly constructed captions, which can raise cognitive load, decrease understanding, tiredness, and disengagement.

Compliance is frequently concerned with whether or not captions are there, rather than the quality of the captions, even though worldwide accessibility standards like the Web Content Accessibility Guidelines (WCAG) stress their supply. When creating captions, many systems overlook important details like the DHH community's varying reading levels, the importance of keeping the text simple, and the importance of aligning the text with the time. The process by which deaf people acquire spoken and written language can vary greatly; for example, many may have grown up with sign language and only learned writing language later in life. Idiomatic idioms, fast speech patterns, and complicated grammar abound in spoken language, making it difficult for captions to accurately reflect it. This language barrier emphasizes the need to assess caption readability from the point of view of DHH users, instead than viewing captions as a literal copy of speech.

Concerns about readability have been heightened by the growing use of automated speech recognition (ASR) systems to generate captions. Despite the fact that ASR technologies have greatly enhanced the accuracy of transcription, they frequently put word-level accuracy ahead of user-centered readability concerns. Comprehension can be significantly impacted by automated captions due to issues such as erroneous punctuation, irregular pacing, absence of sentence boundaries, and lexical ambiguity. Users of DHH rely on real-time caption processing, therefore even little changes in timing or structure might impair comprehension. Consequently, readability is more than just correct grammar and spelling; it is a multi-faceted concept that includes cognitive, syntactic, temporal, and lexical aspects.

Caption readability is of utmost importance in educational settings. Students relying on captions for their education typically face time constraints as they read, analyze visual data, and incorporate new ideas. Appropriately constructed captions can improve DHH students' vocabulary learning, material retention, and conceptual comprehension, according to research. Captions that are excessively rapid, too thick, or too linguistically complicated, on the other hand, might hinder learning and make the gap between hearing and non-hearing students even wider. As the importance of inclusive education grows in both policy and practice, a crucial aspect of UDL is making sure that captions are of good quality and easy to read.

The assessment and improvement of caption readability has been made possible by new opportunities presented by recent advancements in machine learning and natural language processing. Machine learning algorithms can sift through mountains of caption data in search of trends indicating reading difficulties, ambiguity, and timing mismatches—unlike conventional rule-based methods. Computational modeling of features such as display time, word frequency, sentence length, words per second, and semantic ambiguity allows for more accurate prediction of reading levels. We may go beyond subjective evaluations and toward objective, repeatable metrics that are adapted to the demands of DHH users with the help of these tools, which allow for scalable, data-driven evaluation of captions across varied content areas.

Nevertheless, there are still a lot of unanswered questions when it comes to user-centered evaluation in the field of applying machine learning to caption readability. The specific challenges of captioned media for Deaf and Hard-of-Hearing audiences have not been adequately addressed in the majority of the available research, which mostly concentrates on voice recognition accuracy or generic text readability. The interplay between temporal and lexical variables and their impact on real-time understanding has also received little research attention. To fill these gaps, researchers in the fields of artificial intelligence, linguistics, cognitive psychology, and accessibility studies must work together to create solid frameworks for optimizing and evaluating captions.

The variety of the DHH population is another important factor in caption legibility. There is a large range in age, education, language ability, and communication style among those who are deaf or hard of hearing. Some depend mostly on sign language, while others are more proficient in spoken or written language. Some members of this demographic may feel excluded by caption systems that use a cookie-cutter approach. One encouraging step toward improving accessibility is the development of adaptive and personalized captioning systems guided by machine learning algorithms. These solutions can adjust to different reading speeds and language choices.

The study of caption readability is a social and ethical obligation in this environment, in addition to being a technological issue. Everyone has the right to access information, and the lack of captions that are easy to see can keep people from fully participating in society in areas such as work, education, and politics. Captions must be legible, understandable, and cognitively accessible for Deaf and Hard-of-Hearing users in order to promote digital equity, as digital platforms are becoming more influential in the diffusion of knowledge. Both technical progress and the larger aims of social fairness and inclusive development are aided by research in this field.

II. REVIEW OF LITERATURE

Oliver Alonzo et al., (2022) Software developed for Automatic Text Simplification (ATS) tries to simplify difficult text automatically. A number of user groups, including individuals who are deaf or hard of hearing (DHH), have benefited from access to ATS as a reading aid tool, according to previous studies. The needs and wants of some demographics who may make use of this technology, however, have received less attention. Our study aimed to examine the reading habits, interests, and opinions of DHH individuals with work experience in the computing industry regarding ATS-based reading assistance tools. We also wanted to know their thoughts on the social accessibility of these tools, taking into account previous research that found that computing professionals often need to read about new technologies to stay current in their profession. This group reads frequently, particularly for work-related purposes, and is interested in tools that can simplify complex texts; however, participants' ideal professional image may be at odds with the public's perception of them as a result of their public use of these tools. In light of these findings, future studies should focus on developing ATS-based reading assistance tools for DHH adults, taking into account factors like social accessibility, and determining which reading activities users would find most useful if this technology were to be implemented.

Kafle, Sushant et al., (2021) There are style rules for authors who highlight key terms in static text, such as bolding phrases in student textbooks. However, very little study has been done to explore highlighting in dynamic texts, such as captions during instructional films for users who are deaf or hard of hearing (DHH). In the experimental study that we conducted, participants from DHH were asked to make subjective comparisons between several design parameters for caption highlighting. These factors included decorating (underlining, italicizing, and boldfacing), granularity (sentence level versus word level), and whether or not to highlight just the initial instance of a term that appears multiple times. Despite the fact that previous research had not been based on experimental investigations with DHH users, we discovered that DHH participants favored boldface, word-level highlighting in captions. This somewhat contradicts the suggestions that had been made in earlier study. Our empirical findings offer direction for the development of keyword highlighting during captioned videos for DHH users, particularly in the context of instructional video genres.

Trokoz, Dmitry et al., (2021) Discussing the primary static and dynamic user identification methods via keyboard dynamics is the purpose of this article, which aims to explore these approaches. A generalized method of expressing the process of typing on the keyboard was proposed as part of the study based on the sequential change of the keyboard state. This method was proposed as part of the research. The dynamic identification technique is started by formulating the specification of the keyboard state context, which serves as the foundation for the method. Through the implementation of the suggested method, it will be feasible to use a wide range of static identification techniques, hence considerably increasing the collection of techniques that are utilized for dynamic user identification through keyboard dynamics.

Alsalamah, Anwar. (2020) When it comes to the utilization of captioning services to assist the academic performance of deaf and hard of hearing (DHH) students in higher education settings, the author conducts a comprehensive evaluation of studies that were published between the years 1989 and 2019. A total of seven papers were chosen for their analysis. Based on the outcomes of the studies, it was determined that the majority of DHH students profited from captioning services. After receiving captioned services, the students' performance on posttests that examined their comprehension of material linked to lectures improved. These tests were administered after the students had received captioning services. To determine whether or not captioning services are beneficial in assisting students with disabilities to achieve academic achievement in higher education, further study has to be carried out first.

Zhu, Guang et al., (2015) The computer keyboard is a prevalent, dependable, accessible, and efficient instrument for human-machine interaction and information transfer. Despite keyboards being utilized for centuries to advance human civilization, the analysis of human behavior through keystroke dynamics using smart keyboards continues to pose significant challenges. We present a self-powered, non-mechanical-punching keyboard that utilizes contact electrification between human fingers and keys, converting mechanical impulses into localized electrical signals without the need for external power. The intelligent keyboard (IKB) may sensitively activate a wireless alarm system upon mild finger tapping and also monitor and document written information by measuring the dynamic time intervals between keystrokes and the power used during each typing activity. These qualities offer potential for its application as an intelligent security system capable of detection, alerting, recording, and identification. Furthermore, the IKB can discern personal traits from many persons, aided by the behavioral biometric of keyboard dynamics. Moreover, the IKB can efficiently convert typing movements into power to charge commercial gadgets at any typing speed above 100 characters per minute. The aforementioned properties suggest that the IKB may be applicable not just to self-powered devices but also to artificial intelligence, cybersecurity, and computer or network access control.

Tyler, Michael et al., (2009) Television captioning possesses significant potential to provide deaf youngsters access to the aural content of programs. Nonetheless, the utilization of captions may be constrained by the comparatively lower English literacy levels among the deaf community relative to the general populace. This study examines the impact of caption delivery speed on the understanding of educational programs among deaf school pupils with varying reading abilities. Participants viewed three brief films, with subtitles shown at rates of 90, 120, or 180 words per minute (wpm). Comprehension was consistently superior at 90 and 120 wpm compared to 180 wpm across both reading levels. Regardless of caption rate, proficient readers had superior scores compared to less proficient readers. The findings indicate that a rate of 120 wpm can be utilized as the minimum pace for captions in children's television programs. Future research should seek to pinpoint the optimal rate, which appears to lie between 120 and 180 wpm.

III. RESEARCH METHODOLOGY

The purpose of this study is to investigate suitable captioning for persons with Down syndrome by measuring the speed of keystrokes while reading captions.

Viewing Information While Keying

We performed an experiment in which information is shown in the peripheral vision while keystrokes are being made to check if the speed of keystrokes can be used as an objective evaluation of the reading captions by persons who have Duchenne muscular dystrophy. Following the same experimental protocol as the previous study, in which participants entered keys while seeing symbols and letters in both their central and peripheral views, we compared the outcomes for individuals with and without hearing loss.

Method

Instead of using a head-mounted display (HMD) for augmented reality goggles, we set up a monitor in front of the researcher to serve as their experimental setting. The distance from the participants' line of sight to the screen was around 60 degrees. As a task, the participants had to use a keyboard to enter the numbers (0-3) that were shown in the middle of the screen. On the screen, five distinct kinds of material were displayed:

- **No Display:** No additional visual information was presented.
- **Symbol in Centre:** A geometric symbol (circle, cross, triangle, or rectangle) was displayed at the center of the screen.
- **Symbol in Peripheral Vision:** A geometric symbol was displayed at a peripheral viewing angle of approximately 28°.
- **Icon in Peripheral Vision:** An icon (e.g., mail symbol or music symbol) was displayed in peripheral vision.
- **Alphabet in Peripheral Vision:** Three randomly selected English alphabet letters were displayed in peripheral vision.

Every emblem or symbol was shown for three seconds at random intervals of two to three seconds. To verify seeing the symbol, each participant had to sign it or use their finger alphabet.

While working on the task of typing in the number displayed in the center of the screen, participants in this research viewed a backdrop image and its matching caption.

At the very bottom of the screen, you could see three lines of captioning, with no more than 23 characters per line. Twenty young adult DHH participants, all in their twenties, were shown the captions for around three minutes in one study.

IV. RESULTS AND DISCUSSION

Table 1: Average Task Processing Time According to Displayed Content

Display Condition	Average Task Processing Time (s)	Standard Deviation (s)	Error Rate (%)
No display	0.58	0.08	2.1
Symbol in centre	0.74	0.14	2.5
Symbol in peripheral vision	0.67	0.16	3.4
Icon in peripheral vision	0.65	0.15	3.2
Three-letter alphabet in peripheral vision	0.84	0.17	4.8

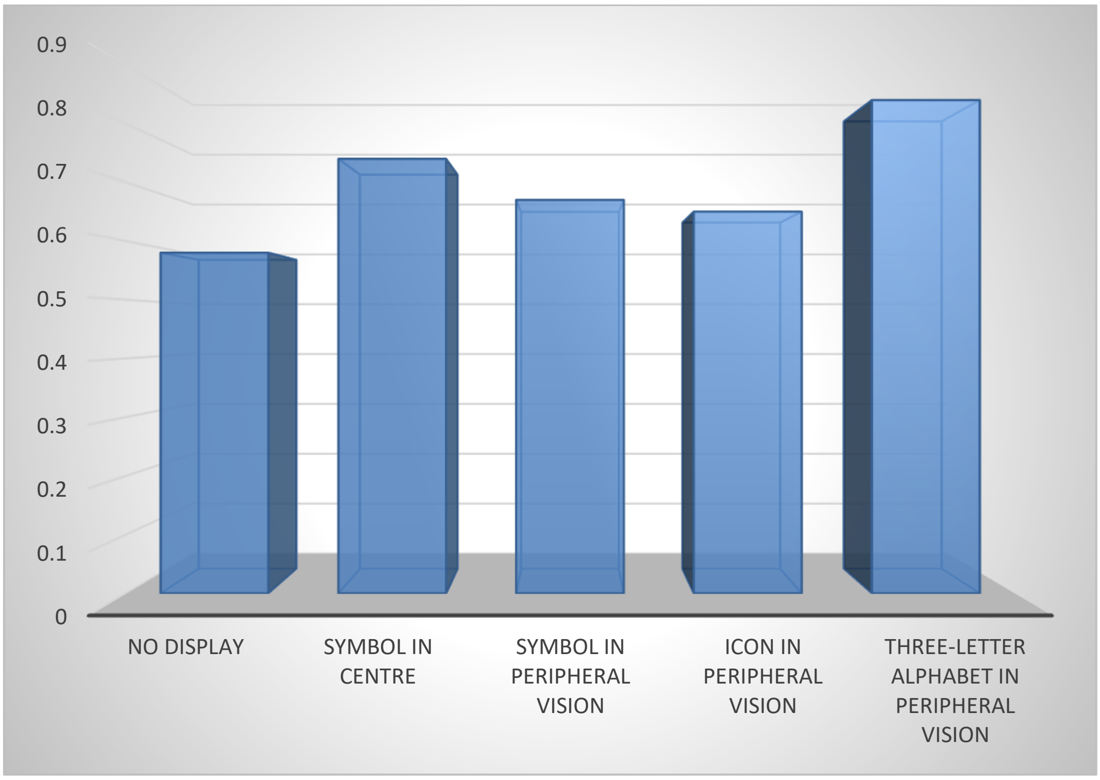

Figure 1: Average Task Processing Time According to Displayed Content

Table 1 shows that the visual information displayed has a noticeable effect on the execution of the activity. Average task processing time was 0.58 seconds and mistake rate was 2.1% in the absence of display, suggesting that participants' cognitive burden was negligible in this setting. The average processing time rose to

0.74 s when a symbol was shown in the center, indicating a substantial disruption of the main task caused by visual distraction in the center. This disparity stood out when contrasted with the no-display condition ($p < 0.01$), emphasizing the high level of attentional demand caused by material that is centrally shown.

Information displayed on the periphery of the field of view required different amounts of time to comprehend based on its kind. Seeing peripheral vision symbols and icons increased processing time by a substantial amount (0.67 s and 0.65 s, respectively), whereas the corresponding increases in mistake rates were rather minor. According to these results, there is a relatively low cognitive cost associated with processing non-linguistic visual information in peripheral vision. Alternatively, the most processing time (0.84 s) and mistake rate (4.8%), respectively, were caused by the display of a three-letter alphabet in peripheral vision. The findings of the t-test showed that task performance was much slower while seeing alphabetic information in peripheral vision, in comparison to both the no-display condition and icon-based peripheral displays ($p < 0.01$).

Table 2: Processing time for captioning presentation

Metric	Time
Average task processing time (s)	1.32
Standard deviation of task processing time (s)	0.41
Error rate (%)	3.6

The impact of closed captioning on task performance is seen in Table 2. Caption presentation averaged 1.32 seconds of processing time. Participants' cognitive load is increased when they read captions, as this difference was shown to be statistically significant ($p < 0.05$). Participants' varying degrees of reading ability and approaches to visual attention likely contributed to the somewhat larger standard deviation (0.41 s), which in turn indicated discernible individual heterogeneity in processing and reading rates. When participants are asked to process continuous linguistic information, the job becomes more challenging, as indicated by the observed mistake rate of 3.6%.

Table 3: Number of keystrokes per minute

Displayed Content	Average Number of Keystrokes per Minute	Percentage Compared to No Display (%)
No display	96.2	100
Symbol in centre	78.4	81
Symbol in peripheral vision	88.9	92
Icon in peripheral vision	86.7	90
Three-letter alphabet in peripheral vision	71.5	74
Captioning	45.6	47

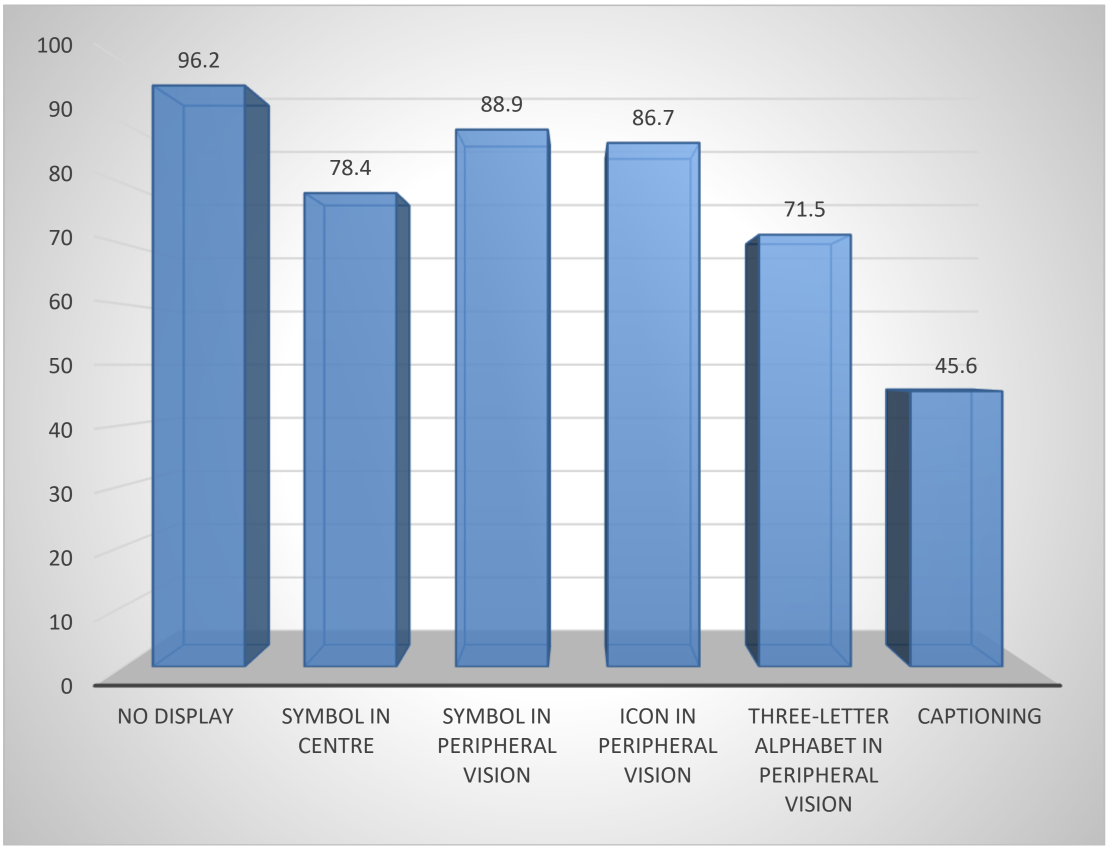

Figure 3: Number of keystrokes per minute

Table 3 demonstrates the effects of various display settings on keyboard performance, which may be indicative of changes in mental strain when tasks are being carried out. The lowest level of cognitive interference was indicated by the maximum keystroke rate (96.2 kbps) in the absence of display condition. Keyboard performance dropped to 78.4 kbps (81% reduction) when a symbol was shown in the middle, showing that having visual information in the center greatly interferes with focus and job processing.

Symbols and icons on peripheral displays reduced typing rates to a moderate degree (88.9 and 86.7 kbps, respectively), indicating that visual signals in peripheral vision that are not related to language do not place a heavy cognitive burden. On the other hand, a three-letter alphabet shown on the periphery of the field of view reduced the number of keystrokes per minute to 71.5 (74%), suggesting that processing linguistic stimuli still requires more effort even when they are not in the center of the screen.

The captioning process saw the greatest decrease, with the typing rate falling to 45.6 per minute, a 47% decrease. This significant decrease demonstrates the high cognitive and attentional demands of continuous caption reading, lending credence to keystroke rate as an objective and trustworthy metric for gauging caption readability and cognitive load in DHH users.

V. CONCLUSION

For the purpose of providing an objective assessment index for caption readability, this study empirically examined how well keyboard dynamics worked for users who are deaf or hard of hearing. The results show that various visual information kinds place distinct cognitive loads on the brain when performing tasks. The interference with keyboard performance induced by non-linguistic cues, such as symbols and icons, was rather low, especially when they were shown in the peripheral vision. Typic text and continuous captions are examples of linguistic stimuli that, on the other hand, raise cognitive strain by making task processing time greater while decreasing keystroke frequency. Confirming that persistent reading of textual material demands considerable attentional and cognitive resources, captioning generated the most pronounced impact. Keystroke frequency and speed are confirmed to be reliable, objective measures of caption readability and user workload, according to the results. The study shows that maximizing caption display and reducing language complexity are important ways to make DHH easier to utilize from a practical standpoint. Supporting the development of user-centered, accessible captioning systems, the suggested assessment approach is especially useful in developing scenarios like multimedia learning environments and augmented reality.

REFERENCES:

[1] S. Ahn, J. Kim, C. Shin, and J.-H. Hong, "Visualizing speech styles in captions for deaf and hard-of-hearing viewers," *International Journal of Human-Computer Studies*, vol. 194, no. 1, pp. 1–7, 2024.

[2] O. Alonzo, L. Elliot, B. Dingman, S. Lee, A. A. Amin, and M. Huenerfauth, "Reading-assistance tools among deaf and hard-of-hearing computing professionals in the U.S.: Their reading experiences, interests and perceptions of social accessibility," *ACM Transactions on Accessible Computing*, vol. 15, no. 1, pp. 1–10, 2022.

[3] S. Kafle, B. Dingman, and M. Huenerfauth, "Deaf and hard-of-hearing users evaluating designs for highlighting key words in educational lecture videos," *ACM Transactions on Accessible Computing*, vol. 14, no. 4, pp. 1–24, 2021.

[4] D. Trokoz, A. Martyshkin, E. Balzannikova, and I. Sergina, "Developing of keystroke dynamics analyzing systems using a presentation based on state contexts," *Nexo Revista Científica*, vol. 34, no. 4, pp. 1258–1271, 2021.

[5] A. Alsalamah, "Using captioning services with deaf and hard of hearing students in higher education: A systematic review," *American Annals of the Deaf*, vol. 165, no. 1, pp. 114–127, 2020.

[6] J. Butler, "The visual experience of accessing captioned television and digital videos," *Television & New Media*, vol. 21, no. 7, pp. 679–696, 2020.

[7] G. Zhu, J. Yang, Q. Jing, P. Bai, W. Yang, X. Qi, Y. Su, and Z. Wang, "Personalized keystroke dynamics for self-powered human–machine interfacing," *ACS Nano*, vol. 9, no. 1, pp. 105–116, 2015.

[8] R. S. Hallam and R. Corney, "Conversation tactics in persons with normal hearing and hearing-impairment," *International Journal of Audiology*, vol. 53, no. 3, pp. 174–181, 2014.

[9] W. S. Lasecki and J. P. Bigham, "Real-time captioning with the crowd," *interactions*, vol. 21, no. 3, pp. 50–55, 2014.

[10] M. F. Mueller, K. Meisenbacher, W. Lai, and N. Dillier, "Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization?" *Cochlear Implants International*, vol. 15, no. 1, pp. 36–42, 2014.

[11] E. Picou, E. Aspell, and T. Ricketts, "Potential benefits and limitations of three types of directional processing in hearing aids," *Ear and Hearing*, vol. 35, no. 3, pp. 339–352, 2014.

[12] C. Convertino, G. Borgna, M. Marschark, and A. Durkin, "Word and world knowledge among deaf learners with and without cochlear implants," *The Journal of Deaf Studies and Deaf Education*, vol. 19, no. 4, pp. 471–483, 2014.

[13] T. M. Grieco-Calub and R. Y. Litovsky, "Spatial acuity in 2-to-3-year-old children with normal acoustic hearing, unilateral cochlear implants, and bilateral cochlear implants," *Ear and Hearing*, vol. 33, no. 5, pp. 561–572, 2012.

[14] R. W. McCreery et al., "An evidence-based systematic review of directional microphones and digital noise reduction hearing aids in school-age children with hearing loss," *American Journal of Audiology*, vol. 21, no. 2, pp. 295–312, 2012.

[15] Y. Kaneko, I. Chung, and K. Suzuki, "Light-emitting device for supporting auditory awareness of hearing-impaired people during group conversations," *iSystems, Man, and Cybernetics*, pp. 3567–3572, 2013.

[16] A. Albertini and C. Mayer, "Using miscue analysis to assess comprehension in deaf college readers," *Journal of Deaf Studies and Deaf Education*, vol. 16, no. 1, pp. 35–46, 2011.

[17] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 1, pp. 2–16, 2010.

[18] M. Tyler, C. Jones, L. Grebennikov, G. Leigh, W. Noble, and D. Burnham, "Effect of caption rate on the comprehension of educational

television programmes by deaf school students," *Deafness & Education International*, vol. 11, no. 3, pp. 152–162, 2009.
[19] J. L. Luckner and C. M. Handley, "A summary of the reading comprehension research undertaken with students who are deaf or hard of hearing," *American Annals of the Deaf*, vol. 153, no. 1, pp. 6–36, 2008.
[20] T. Matthews, J. Fong, F. Ho-Ching, and J. Mankoff, "Evaluating non-speech sound visualizations for the deaf," *Behaviour & Information Technology*, vol. 25, no. 4, pp. 333–351, 2006.
[21] S. R. Gulliver and G. Ghinea, "How level and type of deafness affect user perception of multimedia video clips," *Universal Access in the Information Society*, vol. 2, no. 4, pp. 374–386, 2003.
[22] M. Lewis and D. Jackson, "Television literacy: Comprehension of program content using closed captions for the deaf," *Journal of Deaf Studies and Deaf Education*, vol. 6, no. 1, pp. 43–53, 2001.