



Exploring Virtual Reality to Enable Deaf or Hard of Hearing Accessibility in Live Theaters: A Case Study

Mauro Teófilo^(✉), Alvaro Lourenço, Juliana Postal, and Vicente F. Lucena Jr.

SIDIA - Samsung Research Institute, Federal University of Amazonas, Manaus, Brazil
{mauro.t,alvaro.abl,j.postal}@samsung.com, vicente@ufam.edu.br

Abstract. Recent advancements in Virtual reality (VR) have made them a potential technology to improve speech understanding for deaf or hard of hearing (DHH). We want to extend this commodity to live communication in theater plays, which has not been investigated so far. For this, present study evaluates the efficiency of a language processing system, which makes use of VR technology combined with AI implementations for automatic speech recognition (ASR), sentence prediction and spelling correction. A quantitative and qualitative study was performed and demonstrated good overall results of the system regarding DHH understanding and satisfaction along entire play sessions. *abstract* environment.

Keywords: Virtual reality · Accessibility
Deaf or hard-of-hearing people · Captioning system
Subtitle display · Speech transcription

1 Introduction

In the last years, a rising number of researches around virtual environments (VE) has been scoping its possible benefits to compensate people for hearing disadvantages by improving their autonomy. In a previous study, VR technology has proved efficiency in helping impaired people [27], and with increased availability and public interest around virtual/augmented reality topic (VR/AR) [19,20], a number of authors started working on automated transcriptions for deaf or hard-of-hearing (DHH) people, attacking the matter from a number of models with different levels of success [3,22,24]. The overall approach leverages VR and AR capabilities of attaching message to objects in real-time [3] to improve usability within many specific contexts: from silent to noisy environments [19,20], single educational material [24] to entire library spaces and other indoor positioning systems [18].

When put in perspective, these works reveal a common consensus about the need for accuracy and the stratification of concerns that potentially solves the whole problem. Some argue of error suppression and its relevant impacts over

understandability [10]; others of the composition with further technology such as text-to-speech (TTS), automatic speech recognition (ASR) and audio-visual speech recognition (AVSR) [19,20]; all trying to reach a better transcription, caption, and signal language performance across many scenarios. Approaches measured confidence and reliability of display styles, output latency and precision of on-screen subtitles; observed and interviewed DHH users; and also inspected final levels of understanding with results ranging around 60 and 80%, without however declaring a gold standard to address one mainstream user experience [3,10,20,22,24].

1.1 VR and the Accessibility Research Field

The use Virtual Reality (VR) for DHH and other recent accessibility research for the inclusion of impaired people comes from a resourceful research sprint that ranges at least two strong decades. Within this time, plentiful reports applied the technology to a wide range of accessibility topics, targeting from those with learning disabilities to others with anxiety and phobias, other times leveraging mobility and life quality for spinal cord injured patients and those fighting along post strokes. Just recently, as more authors are following the fresh possibilities of the renewed generation of pervasive and popular technologies, a newer thread have been targeting the benefits of virtual accessibility to socialize the visual and auditory impaired people.

Before hitting the inclusion topic, however, this ever growing legacy for accessibility fields ranged at least three major research groups: Autism Spectrum Disorders (ASDs) [7], happening around 1996, continuing across 2005 with Physical Rehabilitation, and in later years with reports about VR therapies to ease Parkinson Disease (PD) and Traumatic Brain Injury (TBI) symptoms. Along the years, analysis of the VR benefits often included the improved safety and adherence to programs, the documentation and individualization capabilities, and of course some hard and better results.

Autism Spectrum Disorders. In autism, psychology authors considered that VR may provide improvements to the treatment of ASDs (Autism Spectrum Disorders) by increasing frequency and customization of role-playing tests while demanding less resources (e.g. teachers, parents, schools) [23]. Although theorized to offer enhanced results for ASD therapies, these methods were not tested so far [21]. Proponents of methodologies such as the ‘Behavioral’ and ‘Theory of Mind’ (ToM) discuss the different approaches and results for the topic, mainly arguing around the cost and outcomes in terms of learning and generalization. While the first offers better promises, it does so at more expensive school and structure requirements. ToM in the other hand focuses on teaching straightforward behaviors with the expense of less resources, with unfortunately lesser ambitions in terms of recovering children from their generalization handicaps. Acting in a position to overcome structure investments, some authors consider VR technology capabilities to automate and personalize therapies, and therefore improve final results.

Physical Rehabilitation. For motor rehab, compared studies between traditional and VR based therapies had consistently reported the VR superiority in a number of programs ranging from bicycle exercises to occupational therapy [28]. The studies generally presents the opportunity for automatic movement analysis and progressive challenge updates, while reducing boredom, fatigue and lack of cooperation with an acute increment of patients motivation [1, 4]. So far, these studies managed to improve posture, balance control, velocity and path deviation on bicycle exercises, inspecting increased pain tolerance levels, distance, duration, and energy consumption [12]. Similarly, studies on the field of occupational therapy also managed to inspect greater gains of dynamic standing tolerance in geriatric patients by using the VR technology. Studies of virtual therapies for traumatic brain injury and orthopedic appliances stated substantial improvements of patient enthusiasm [9, 17, 25, 26], confidence and motivation [5, 13], even when in absence of better hard performance indicators.

Parkinson Disease. Improvements on the negative impacts of Parkinson disease (PD) over patients cognitive and motor functions were also reported in 2010, succeeding to apply VR to exceed traditional treadmill training (TT) outcomes in aspects such as attention management, gait speeds, stride lengths, with additional development of cognitive strategies to navigate throughout virtual obstacles [14]. Results were inspected and proven to be retained after therapy sessions, being thoroughly scrutinized during and after program administration, confirming the retention and the occasional improvement of therapy gains in the months after. Authors assumes that TT with VR technology can actually relief cognitive and motor malfunctions of patients with PD, delivering better attention levels with the development of new strategies to overcome virtual obstacles.

1.2 Research Context

With a DHH population of 9 million people and a rich cultural environment, Brazil presents a challenging yet resourceful landscape for accessibility systems to be introduced and matured. Majority of impaired people in the country are unable to attend at theaters due to the lack of even basic accessibility services such as stage live-action subtitles and professional interpreters. All this poses Brazil as a good research lab to start and scale production of VE methods abroad, later hitting the impressive global DHH population of 5% [10].

From this scenario, our study takes advantage of Samsung Research efforts in Brazil to add more findings on this research thread. With first results publicized through “Theater for All Ears” campaign¹, the research underneath attempted to spot improvement venues for HCI in VE, inspecting challenges of DHH theater accessibility to propose a new concept that exceeds the efficiency and scalability results of traditional interpreters and stage captioning systems. We started from basic speech transcription models to inquire more than 40 users across 10 semi-controlled experiments that happened between May and July 2017, on

¹ <https://vimeo.com/217227242>.

weekly theater sessions in São Paulo, Brazil. Then we assessed new opportunities based on collected results and explored progressive models to improve accuracy of captioning systems using ASR, natural language processing (NLP), and artificial intelligence (AI); tame key usability factors; and achieve new state of art technology for better accessibility in live theaters.

In Sect. 2, we present related works that also scoped the automatic generation of subtitles for DHH people, followed by the proposal of a new IA based approach for language processing in Sect. 3. Our experiment methodology and results are discussed in Sect. 4, with conclusions further planned steps in Sect. 5.

2 Related Work

The use of Augmented Reality to improve communication with DHH people in real-time is a chance for research with strong social impacts as it enables the social inclusion of impaired people in theater entertainment, conferences and all sorts of live presentations. In this section we highlight related works that followed this important thread.

Mirzaei et al. [19] solution improves live communication between deaf and ordinary people by turning ordinary people's speech into text in a device used by the deaf communicator. In this situation, the deaf also can write texts to be turned back into speech, so the ordinary people can understand. The solution is composed by a device and a software in which narrator (ordinary person) speech is captured by ASR or AVSR (Audio-Visual Speech Recognition) and turned into text, the Joiner Algorithm uses the text generated by ASR or AVSR and creates an AR environment with the image of narrator and text boxes of his speech. TTS engines are used to convert texts written by deaf people into speech for the narrator, making possible a two-way conversation. The results pointed that the average processing time for word recognition and AR displaying is less than 3 s using ASR mode, and less than 10 s for AVSR mode. To evaluate the solution, they conducted a survey with 100 deaf people and 100 ordinary people to measure the interest rate of using technological communication methods between them and 90% of participants agreed that the system is really useful, but there's still opportunity for improvements with AVSR mode, which is more accurate in noisy environments.

Berke [3] believes that providing word and its confidence score in a subtitle using AR environment, in order to give more information about the narrators speech, will improve deaf people understanding of a conversation. The method proposal consists in a captioning which words generated by speech to text are displayed with its score of confidence in the subtitle and different colors are given for more confident and less confident words. These scores are calculated based on how sure speech to text algorithm are about the match of voice captured and the acoustic model of a word. The author also wants to study a way to present these information without confuse or make more difficult for the deaf to read the subtitle and pay attention on the narrator.

Piquard-Kipffer et al. [22] made a study to evaluate the best way to present the text generated by speech to text algorithm in French language. The study

covered 3 display modes: Orthographic, where recognized words are written into orthographical form; International Phonetic Alphabet (IPA), which writes all the recognized words and syllables in phonetic form by using the International Phonetic Alphabet, and lastly a Pseudo-phonetic where recognized words and syllables are written into a pseudo-phonetic alphabet. Some problems that challenge automatic subtitle systems such as noises captured by the device's unsophisticated microphones, implied in an incorrect word generation by ASR as reported in [19] thus flawed message understanding in deaf people's side. To minimize this negative, they included additional information about converted text within the subtitle for all display modes - like a confidence value for each word as proposed in [3, 22]. Experiments with 10 deaf persons found best reviews when using a confidence score to format correct words in bold while presenting the incorrectly recognized ones in pseudo-phonetic mode; and suggested that preceding training phase for the experiment would be necessary to make participants more familiar with pseudo-phonetic reading. All participants manifested interest for such a system and thought that it could be helpful.

Hong et al. [6] propose a scheme to improve the experience of DHH people with video captions, called Dynamic Captioning. It involves facial recognition, visual saliency analysis, text-speech alignment and other techniques. First, a script-face matching is done to identify which people the subtitles belong to in the scenes, this is based on face recognition, then a non-intrusive area is chosen in the video so that the caption can be positioned to avoid occlusion of important parts of the video and compromise the understanding of its content, the display of the caption emphasizing word for word is done through script-speech alignment and finally a voice volume estimation is done to display the magnitude indicator of the character's voice in the video. In order to validate the solution, the authors invited 60 hearing impaired people to an experiment that consists of watching 20 videos where some metrics such as comprehension and impression about the videos would be evaluated, in this experiment 3 captioning paradigms were tested: No Captioning, Static Captioning and Dynamic Captioning. The results showed that the No Captioning paradigm presented a poor experience for users, Static Captioning contributed to user distraction and 93.3% of users preferred Dynamic Captioning.

Beadles et al. [2] patent propose an apparatus for providing closed captioning at a performance comprise means for encoding a signal representing the written equivalent of spoken dialogue. The signal is synchronized with spoken dialog and transmitted to wearable glasses of a person watching the performance. The glasses include receiving and decoding circuits and means for projecting a display image into the field of view of the person watching the performance representing at least one line of captioning. The field of view of the displayed image is equivalent to the field of view of the performance. A related method for providing closed captioning further includes the step of accommodating for different interpupillary distances of the person wearing the glasses.

Luo et al. [15] designed and implemented a Mixed Reality application which simulates in-class assistive learning and tested at China's largest DHH education

institute. The experiments consisted in let these DHH children study a subject that is not in their regular curriculum and verify if the solution can improve the learning process. The solution has two main components, one component is the assisting console controlled by a hearing student, the other component is the virtual character displaying viewport which fulfills assistance. Both components use a dual-screen setup, within each component, one of the screens displays lecture video and the other screen displays mixed reality user interaction or rendering content. Videos on the screens of both components are synchronized in time. The hearing impaired student side of the system has a virtual character shown on the user content screen which can take predefined actions, while the hearing student side of the system has a control UI shown on the user content screen to manipulate virtual character at the other end to perform such actions. Results showed that the experience of being assisted by a virtual character were mostly very positive. Students rated this approach as novel, interesting and fun. 86,7% of them felt that with such help, it was easier to catch the pace of the lecture, understand the importance of knowledge points, and keep focused across the entire learning session.

Kercher and Rowe [11] propose a design, prototyping and usability testing of an AR head-mounted display system designed to improve the learning experience for the deaf, avoiding the attention split problem common among DHH people in learning process. The solution is focused in child's experience in a planetarium show. Their solution consists in three parts: Filming of the interpreter in front of a green screen, use a server to communicate the interpreter video to the headset and user interface for testing headset projection manipulation and optimization, then the interpreter will be always in the field of view of DHH spectator as it can also look freely to all directions and enjoy the show. The authors expect in the end of 3 years of research to not only help young children to have better experience in planetarium show but contribute in major changes in the experiences of the deaf in a variety of environments including planetariums, classrooms, museums, sporting events, live theaters and cinemas.

In our solution we present different strategies to deal with these problems, as detailed in the following section.

3 System Overview

This section presents a solution which uses speech recognition and AI to retrieve the correct subtitle of live play scenes using text from play script. Figure 1 shows the concept model for system's architecture.

3.1 The Actor Module

In Actors device module, a speech-to-text algorithm for Portuguese language converts recorded voice into text, using the first few words to retrieve the correct play speech using a sentence prediction algorithm before communicating it to server. The sentence prediction workflow is presented in Fig. 2. When the device

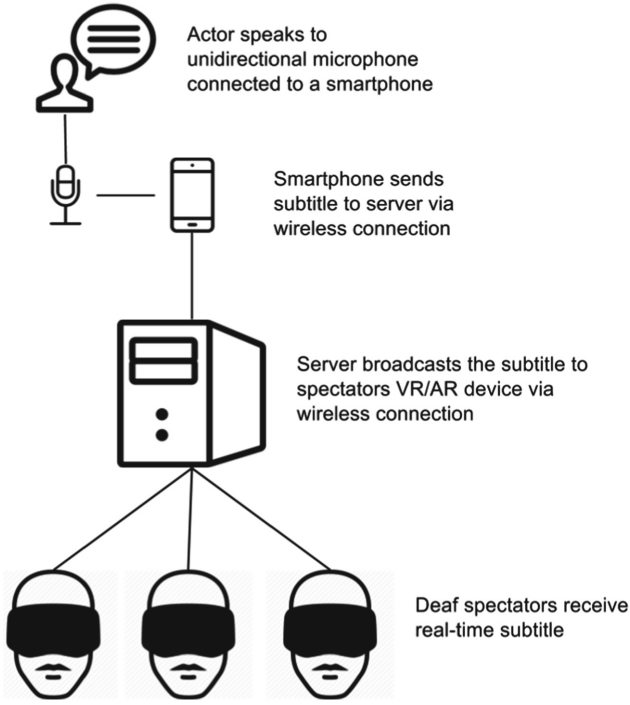


Fig. 1. System architecture.

application starts, play script is submitted to Ngram algorithm which counts and calculates probabilities for all sets of N sentences, building a data table that serves as language model which can be queried with first N words converted from speech-to-text as key, and returning the most probable sentence.

Sentence prediction will ensure the behavior of retrieving subtitles in real-time while scene performance is occurring. If a search for speech play retrieves no match, system realizes that there is an impromptu occurring, and if needed it passes raw converted words to a word correction algorithm, which then communicates with the server. As demonstrated by Fig. 3, word correction algorithm verifies if the words are present in Portuguese dictionary. If they are it skips correction, otherwise it searches for the most similar word in dictionary based on edit distance, and then overwrites those incorrect with similar words. Incorrect words are saved to be used as data for word correction retraining, as well as words captured from impromptu, which can be used in future sentence prediction retrain. Subtitles sent to server contain actor's character identification, line of speech identification and the subtitle itself. Server uses these line values to broadcast subtitles in the correct order. Lines for impromptu messages receives a special value.

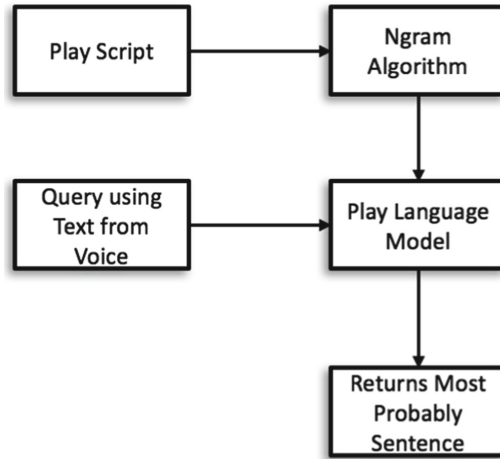


Fig. 2. System's sentence prediction.

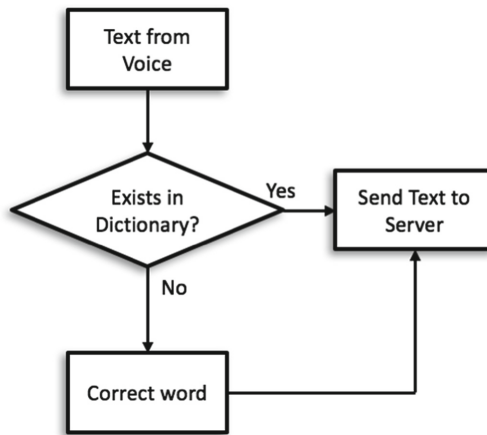


Fig. 3. System's word correction.

3.2 The Server Module

All performing actors in stage for the play carried a unidirectional microphone connected to a Samsung S8 device. Unidirectional microphones prevented undesired recording of noises from environment and voices from other actors. Each device had a copy of the play script to support the task of retrieving actor's speech as text. Actors voices were continuously recorded by their devices and no further interaction was required to operate the application.

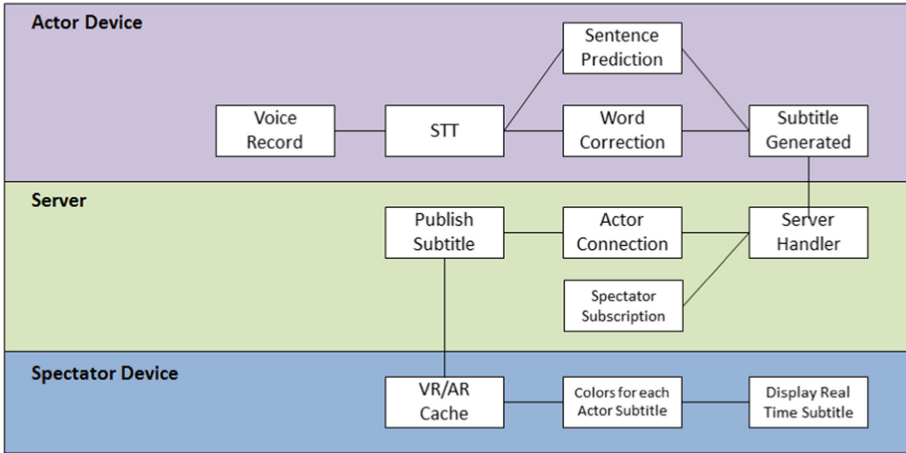


Fig. 4. System solution.

All actors and spectator’s device connections are managed by a single server, which broadcasts subtitles received from all actors to all spectators. The first few words converted into text are used to search for the exact play speech using IA before communicating it to server. If there is no match, the system understand it as an impromptu, and starts sending every word to server as soon as they are converted from actor’s voice. Figure 4 presents the solution and explain more specifically how system modules work.

3.3 The Spectator Module

Server have handlers for actor devices connections, disconnections and subtitle communication. Each actor’s device has its own connection which remains open until application closes. In the other side, spectator devices do not send messages to server, using only available subscription services. For each subscribed spectator, server broadcasts every received subtitle until unsubscription request happens. When subtitles are received by AR/VR application, they are added to a queue and wait to be displayed. The application renders each actor’s subtitles with a distinct color that is consistent during all play. When presenting these, a calculation is made to verify whether text fits in the UI space, and if they don’t, subtitle gets split and displayed by parts. Then a comfortable time estimation is given based in each subtitle’s size, defining when it will be overwritten by the next subtitle. When there is no subtitles in queue, application simply listens for upcoming server subtitles until the end of the play. Figure 5 shows an demonstration of how subtitle is in VR environment.



Fig. 5. Subtitle in gear VR.

4 Experiment and Results

In 2017, from May 12th to July 23rd, we conducted 10 user testing (UT) sessions in two mainstream theaters in São Paulo city, Brazil. Structured data and qualitative insights were collected from 43 DHH attendants over weekly performances of ‘O Pai’ play - ‘The Father’, from the original French ‘Le Père’.



Fig. 6. User's test participants and instructors.



Writer	Florian Zeller
Director	Léo Stefanini
Translators	Carolina Gonzalez and Lenita Aghetoni
Cast	Fulvio Stefanini, Carol Gonzalez, Lara Córdulla, Carol Mariottini, Paulo Emilio Lisboa and Wilson Gomes

Fig. 7. Cast actors and information about the play “O Pai”.

The Fig. 6 shows participants using Gear VR after the play in Fernando Torres theater.

Figure 7 shows the cast of play ‘O Pai’ and summarizes the technical information about play staged during tests.

4.1 User Test Method

Participants were selected with aid from regional Deaf association. Before play, participants were trained by supervisors about how to use the app on the VR device (Gear VR + Galaxy S7), eventual experiment issues and quick fixes, in case they occur. Figure 8 shows the DHH participants watching the play through Gear VR.

After play, participants answered 4 structured questions about image/display, subtitle, understanding and satisfaction using Likert-scale (1 poor to 5 best):



Fig. 8. DHH users watching the play.

- Image/display: whether they could see actors and stage with desired quality
- Subtitle: if they could read transcriptions with proper timing and readability
- Understanding: whether they could get the entire stream of speeches and emotions
- Satisfaction: how pleasant and rewarding it was to use the VR captioning system

The answers are summarized in the Boxplot chart showed in Fig. 9 and finally, participants went into an interview that collected qualitative insights about their experience along the play.

4.2 Results

Evaluation of Image/Display. Most participants had neutral to bad opinion about image provided by Gear VR + Galaxy S7, thus image/display was the worst factor of all analysis, mainly due to its inability to manage light intensity and to provide required definition. By sitting afar from stage, participant camera was unable to capture enough pixels to render details from actor’s facial expressions, suffering severe interference from stage lights, which sometimes added just too much brightness to final rendered image. Some users suggested the addition of features such as zoom, focal adjustments, and brightness control.

Evaluation of Subtitles. Subtitles had some dispersion on votes, but a consistent amount of these were around a good opinion, it means that subtitles may performed well, but there is still much room for improvements. There were

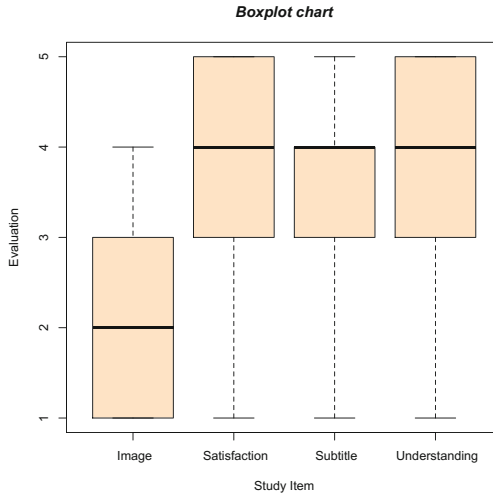


Fig. 9. Results from structured questions.

one relevant complaint about caption synchronization, and a minor critic about recognition errors. Users also suggested additional features to regulate caption size, adjust its placement and contrast on-screen.

Evaluation of Understanding and Satisfaction. Overall understanding and satisfaction were good, understanding were well voted by participants, that means they could follow all the play and be aware of surrounding spectators emotions, many participants freely stating that not only the technology helped on understanding but also that it was better than using professional interpreter services. Votes for satisfaction had similar distribution, this may suggest that understanding positively influences satisfaction.

Evaluation of Technical Setup. Technical setup, however, suffered from some issues during UT sessions. A third of users complained about head or eye strain during the play, most accusing solution from being too bright and Gear VR from being too heavy; Nearly 30% of users opted to remove the device at least once during play to take of the device to accommodate, clean their correction glasses or complained about image quality; It was necessary to replace devices some times along sessions due to overheating and some other unidentified application malfunction; And finally, 10% of users reported that technical breaks compromised the understanding at some extent.

4.3 Experiment Limitations

All participants were invited by Samsung (free of charge); We noticed some observer-expectancy effect; It is still important to test the device in different plays styles.

4.4 Improvement Chances

Great majority of issues seems to have an integrated solution with the adoption of lighter and unobtrusive AR glasses instead of Gear VR: from light intensity to image definition, facial expressions, inadequate rendering of stage lights, excessive brightness, head and eye strain, and excessive device weight. All seem to be easily solved by most of AR concepts available and probable to come.

Some other issues, however, can yet be discussed for further optimization: better software validation, overheating and simultaneous use of correction glasses.

5 Conclusion

Captioning system empowers DHH people by translating people's voice into text and then turning it into subtitles. This work extended this convenience to live theaters, by proposing a specific application that has not been investigated so far. Systems like this are far from optimal with many unsolved challenges for bullet-proof subtitle generation that ensures understanding and participation for DHH people in any live event. However, with the increasing interest of research community on the topic, authors and solutions tend to improve over time.

Current solutions for captioning in theaters which involves specialized sound equipment and professional typists raises two main problems:

- *Contract specialized sound and caption systems are expensive:*
There is a complex sound setup to support the typists because of special equipment to treat the sound (avoid noise in sound signal and enhance the voice of actor), the hiring of professional typist itself and also the softwares used to generate and send subtitles and to manage the connections of typists and Spectators Gear VR through the server. This condition may result in higher theater tickets prices that could discourage DHH people to attend the show or theater itself to no more provide this accessibility.
- *Typists cause delay in subtitles:*
Typists need to listen to all actors speech before typing the subtitle, so it means that subtitles will come always after the scene is already passed.

Despite the problems, our experiments to verify if DHH people found such system useful in this case of study of a live theater situation leveraged knowledge about usability of this special class of people and we were able to propose an improvement for current solution based on collecting their insights and improvement suggestions. Thus, knowing the problems and collecting suggestions we

proposed a new solution that uses a combination of AI implementations for language modeling and transcription for VE that exempts traditional human interpreter or typist work.

UT pointed that participants could follow the entire play with good understanding of both scene rationale and crowd emotion, with results pointing that these two components are crucial drivers for user overall satisfaction. Subtitles had good reviews with lesser complaints about subtitle synchronization, but it was expected once we knew that the typist would influence in this point and images had bad reviews mostly because of camera and hardware limitations.

In general, subtitle systems for theaters are well accepted by DHH spectators and we believe that a new version supported by our proposed technology, can improve DHH people experience watching the play. This points VR and AR devices as cost reduction alternatives for accessibility in theaters and possibly other live events.

5.1 Future Works

With results reveling opportunity for improvements in subtitles and image areas, we will conduct new UT sessions using proposed solution of Sect. 4 to test the main hypothesis of AR devices performing better and overcoming most of current image limitations, because usage of AR device will prevent any smartphones camera issues, eye strain and less concern about how stage lighting influences in camera, so DHH users will be able to see the show as it is, and the only virtual object will be the subtitle, this is a more natural and less tiring type of interaction, so in these new tests, participants will no longer watch the play having a device's camera projection as medium, but the real play itself.

Subtitles will become the sole virtual component in the scene, improved by yet another AI implementation that replaces the need for a typist. Delays and synchronization issues will be improved by an ASR word sequencing system that overcomes human listening and typing speeds. And finally, we will refine the N values when querying the language model to reach better accuracy, sentence prediction and spelling correction, reducing misinterpreted subtitles to a minimum and ensuring clear contexts for DHH spectators. Also, we want to build a more complete accessibility software that helps DHH people not only in theaters but in many other tasks of their daily lives and to improve their communication and interaction with hearing people.

Acknowledgments. The Authors would like to express their gratitude to Leo Burnett company staff and also Samsung Institute Research in São Paulo (SRBR) for the collaboration in some parts of the project.

References

1. Arthur, E., Hancock, P., Chrysler, S.T.: The perception of spatial layout in real and virtual worlds. *Ergonomics* **40**(1), 69–77 (1997)
2. Beadles, R.L., Ball, J.E.: Method and apparatus for closed captioning at a performance. US Patent 5,648,789, 15 July 1997
3. Berke, L.: Displaying confidence from imperfect automatic speech recognition for captioning. *ACM SIGACCESS Accessibility Comput.* **117**, 14–18 (2017)
4. Cromby, J., Standen, P., Brown, D.: The potentials of virtual environments in the education and training of people with learning disabilities. *J. Intellect. Disabil. Res.* **40**, 489–501 (1996)
5. Rivaetal, G.: Virtual reality in paraplegia: a VR-enhanced orthopedic appliance for walking and rehabilitation. *Stud. Health Technol. Inform.* **58**, 209–218 (1998)
6. Hong, R., Wang, M., Xu, M., Yan, S., Chua, T.S.: Dynamic captioning: video accessibility enhancement for hearing impairment. In: *Proceedings of the 18th ACM International Conference on Multimedia*, pp. 421–430. ACM (2010)
7. Howlin, P.: *Autism: Preparing for Adulthood* (1997)
8. Huang, X.D., Ariki, Y., Jack, M.A.: *Hidden Markov Models for Speech Recognition* (1990)
9. Inness, L., Howe, J.: The community balance and mobility scale (cbm) an overview of its development and measurement properties, vol. 22, pp. 2–6 (2002)
10. Kafle, S., Huenerfauth, M.: Effect of speech recognition errors on text understandability for people who are deaf or hard of hearing. In: *Proceedings of the 7th Workshop on Speech and Language Processing for Assistive Technologies, INTER-SPEECH* (2016)
11. Kercher, K., Rowe, D.C.: Improving the learning experience for the deaf through augment reality innovations. In: *18th International ICE Conference on Engineering, Technology and Innovation (ICE)*, 2012, pp. 1–11. IEEE (2012)
12. Kim, N., Yoo, C., Im, J.: A new rehabilitation training system for postural balance control using virtual reality technology. *IEEE Trans. Rehabil. Eng.* **7**, 482–485 (1999)
13. Kizony, R., Katz, N., et al.: Adapting an immersive virtual reality system for rehabilitation. *Comput. Anim. Virtual Worlds* **14**(5), 261–268 (2003)
14. Jones, L.E.: Does virtual reality have a place in the rehabilitation world? *Disabil. Rehabil.* **20**, 102–103 (1998)
15. Luo, X., Han, M., Liu, T., Chen, W., Bai, F.: Assistive learning for hearing impaired college students using mixed reality: a pilot study. In: *International Conference on Virtual Reality and Visualization (ICVRV)* 2012, pp. 74–81. IEEE (2012)
16. Mays, E., Damerau, F.J., Mercer, R.L.: Context based spelling correction. *Inf. Process. Manage.* **27**(5), 517–522 (1991)
17. Mc Comas, J., Sveistrup, H.: Virtual reality applications for prevention, disability awareness, and physical therapy rehabilitation in neurology: our recent work. *J. Neurol. Phys. Ther.* **26**, 55–61 (2002)
18. Meredith, T.R.: Using augmented reality tools to enhance childrens library services. *Technol. Knowl. Learn.* **20**(1), 71–77 (2015)
19. Mirzaei, M.R., Ghorshi, S., Mortazavi, M.: Combining augmented reality and speech technologies to help deaf and hard of hearing people. In: *14th Symposium on Virtual and Augmented Reality (SVR)* 2012, pp. 174–181. IEEE (2012)
20. Mirzaei, M.R., Ghorshi, S., Mortazavi, M.: Using augmented reality and automatic speech recognition techniques to help deaf and hard of hearing people. In: *Proceedings of the 2012 Virtual Reality International Conference*, p. 5. ACM (2012)

21. Scherer, M.J.: Virtual reality: consumer perspectives. *Disabil. Rehabil.* **20**, 108–110 (1998)
22. Piquard-Kipffer, A., Mella, O., Miranda, J., Jouvet, D., Orosanu, L.: Qualitative investigation of the display of speech recognition results for communication with deaf people. In: 6th Workshop on Speech and Language Processing for Assistive Technologies, p. 7 (2015)
23. Sanchez, J., Lumbreras, M.: Usability and Cognitive Impact of the Interaction With 3-D Virtual Interactive Acoustic Environments by Blind Children, pp. 67–73 (2000)
24. Sudana, A.K.O., Aristamy, I.G.A.A.M., Wirdiani, N.K.A.: Augmented reality application of sign language for deaf people in android based on smartphone. *Int. J. Softw. Eng. Appl.* **10**(8), 139–150 (2016)
25. Sveistrup, H., McComas, J., Thornton, M., Marshall, S., Finestone, H., Mc-Cormick, A., Babulic, K., Mayhew, A.: Experimental studies of virtual reality-delivered compared to conventional exercise programs for rehabilitation. *Cyber Psychol. Behav.* **6**, 243–249 (2003)
26. Sveistrup, H., Thornton, M., Bryanton, C., Mc Comas, J., Marshall, S., Finestone, H., McCormick, A., McLean, J., Brien, M., Lajoie, Y., Bisson, Y.: Outcomes of Intervention Programs Using Flat Screen Virtual Reality, pp. 4856–4858 (2004)
27. Teófilo, M., Vicente J., Lucena, F., Nascimento, J., Miyagawa, T., Maciel, F.: Evaluating accessibility features designed for virtual reality context. In: 2018 IEEE International Conference on Consumer Electronics (ICCE) (2018 ICCE), Las Vegas, USA, January 2018
28. Witmer, B., Bailey, J., Knerr, B.: Virtual spaces and real world places: transfer of route knowledge. *Int. J. Hum. Comput. Stud.* **45**, 413–428 (1996)