Virtual Reality Technology and Speech Analysis for People Who Stutter

Abeer Al-Nafjan[1], Najwa Alghamdi[2], Abdulaziz Almudhi[3, 4]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Imam Muhammad bin Saud University, Riyadh, Saudi Arabia
[2]Department of Information Technology, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[3]Department of Medical Rehabilitation Sciences, College of Applied Medical Sciences, King Khalid University, Abha, Saudi Arabia
[4]Speech Language Pathology Unit, King Khalid University, Abha, Saudi Arabia
Correspondence Author: annafjan@imamu.edu.sa

Abstract

Virtual reality (VR) technology provides an interactive computer-generated experience that artificially simulates real-life situations by creating a virtual environment that looks real and stimulates the user's feelings. During the past few years, the use of VR technology in clinical interventions for assessment, rehabilitation and treatment have received increased attention. Accordingly, many clinical studies and applications have been proposed in the field of mental health, including anxiety disorders. Stuttering is a speech disorder in which affected individuals have a problem with the flow of speech. This can manifest in the repetition and prolongation of words or phrases, as well as in involuntary silent pauses or blocks during which the individual is unable to produce sounds. Stuttering is often accompanied by a social anxiety disorder as a secondary symptom, which requires separate treatment. In this study, we evaluated the effectiveness of using a VR environment as a medium for presenting speech training tasks. In addition, we evaluated the accuracy of a speech analyzer module in detecting stuttering events.

Keywords: virtual reality, stutter, speech analysis, rehabilitation.

1. INTRODUCTION

Virtual reality (VR) is described as a form of human–computer interaction in which users are active participants in a computer-generated three-dimensional (3D) world [1]. In other words, a parallel 3D world is created for users that enables them to engage in such a world. VR can be classified according to its environment and the equipment used into the following categories: a head-mounted display, the Binocular Omni-Orientation Monitor, and the Cave Automatic Virtual Environment (CAVE). All VR systems contain hardware and software components. The hardware

components include a computer workstation, sensory displays, process acceleration cards, a tracking system, and input devices. The software components include 3D modeling software, two-dimensional graphics software, VR simulation software, and digital sound editing software [2].

VR technology has been identified as an emerging technology in many technology trend reports, and many VR systems and platforms have been developed and deployed during the past decade [3][4]. At present, VR technology has garnered increased attention in both research and applied settings. VR is used in various application areas, including medical and non-medical fields. In addition, it is relevant in a number of disciplines, such as rehabilitative therapy, electronic engineering, psychology, computer science, education, and entertainment.

Several studies [5][6] have reported positive results from using VR for rehabilitation purposes. Specifically, VR has helped individuals to regain the various physical or cognitive abilities that they had previously lost. In addition, it has desensitized and reduced individuals' negative feelings related to the fear of situations without relying on internal imagery or the ability to visualize. For example, individuals may have a fear of public speaking, which may lead them to hesitate in the sounds they make when speaking, as well as to exhibit avoidance behaviors, such as being afraid to pronounce certain sounds or escaping conversations [7].

In [8], Guitar defined stuttering as a complex communication disorder involving affective, cognitive, and behavioral components, such as sound/syllable repetitions and avoidance behaviors. Stuttering is a speech disorder that is characterized by interruptions in the forward flow of speech. The interruptions can be operationally described as monosyllabic whole-word repetitions, part-word repetitions, audible sound prolongations, or silent fixations/blockages [9]. This definition may provide a broader perspective of the dysfluencies observed in persons who stutter (PWSs), as the disorder is multi-dimensional and transcends the dysfluencies reflected in the definition. The deficits in stuttering may be overlaid onto other aspects of speech, such as the speaker's rate, pitch, loudness, inflectional patterns, articulation, facial expressions, and postural adjustments, thereby making stuttering a multifactorial disorder. Stuttering also has an internal component that is related to avoidance, fear, and anticipation [10].

Stuttering has various causes; as a result, stuttering treatment is an individualized process. The causes of stuttering can be categorized as follows: (i) developmental causes; (ii) psychological reasons, such as the fear of public speaking, anxiety, a lack of self-confidence, severe emotional trauma, and verbal or physical abuse; and (iii) physical and brain injuries. Stuttering can negatively affect one's quality of life. Specifically, stuttering may hamper self-respect in PWSs and may induce feelings of inferiority. In addition, stuttering can induce anxiety; PWSs may have difficulty with confronting different communicative situations. Stuttering can also affect

individuals' overall self-esteem; for example, talking in front of an audience or public speaking is a difficult task for PWSs [9].

A study [11] demonstrated a strong relationship between anxiety and stuttering; namely, anxiety can cause stuttering, and vice-versa. The author of [12] discovered through structured interviews that 44% of individuals undergoing treatment for stuttering each also had a diagnosis of social phobia. Other studies confirmed that PWSs also have some level of social anxiety similar to those of individuals with social phobias [7][13][14][15]. In addition, stuttering notably limits the achievements of individuals whose quality of life could be improved through appropriate assessment and treatment [1].

The treatment and assessment of stuttering cases present challenges for speech-language pathologists (SLPs). Several studies have reported that these challenges can be addressed with VR technology [6] [12]. A key challenge is the manual assessment of stuttering patients' speech in a clinical setting. The process of manually counting stuttering events is a laborious and subjective task. These challenges can be resolved with the intelligent processing of stuttered speech, which can provide accurate, simultaneous, and objective speech assessment [16–18]

In view of the above challenges, we aimed to develop a VR system to assist SLPs in offering better management sessions and to help PWSs to desensitize their stuttering and associated anxiety. PWSs may confront various difficulties in their routines. The primary difficulty is the fear of encountering certain communication situations. PWSs may encounter familiar and unfamiliar individuals, and they may be required to speak to individuals of various ages, genders, and ethnicities. This has been an ongoing challenge for health care professionals treating PWSs. Various communicative situations have been used in various stages of therapy to enable PWSs to overcome their fear and anxiety of confronting difficult communication situations.

In this study, we investigated how speech therapy can benefit from innovative VR and automatic speech recognition techniques. In addition, we investigated how SLPs can use VR combined with speech analysis as an advanced visualization tool for the assessment and treatment of stuttering, especially in the transfer and maintenance phases. This study's findings can help with providing a direction for evidence-based practice pertaining to VR technology for Arabic-speaking PWSs), as the technology may help these individuals to confront communicative situations more confidently.

The remainder of this paper is organized as follows. Section 2 summarizes related work, Section 3 presents the study contribution and originality, Section 4 describes the proposed system's design, Section 5 presents the experiment and analysis, and Section 6 presents the conclusion.

2. RELATED WORKS

Technology has been used in stuttering therapy and has revolutionized stuttering treatment [19][20]. Technology has also been used to provide feedback for PWSs and to alter the way in which feedback is delivered to them. Specifically, VR has been successfully used for PWSs in recent years, as it offers an enhanced solution for their assessment and treatment. VR provides a realistic, controllable environment that can benefit PWSs by allowing them to be in a situation that is similar to real life. Furthermore, VR benefits therapists by allowing them to control this environment and to achieve critical therapeutic goals at appropriate times during treatment [7] [14]. In addition, because VR situations are programmed with professional supervision, they are unpredictable only for PWSs. Thus, early failures that may interfere with treatment progress can be avoided. Moreover, the ability to control features signifies that environments, interactions, and reactions can be introduced at suitable times to maximize therapeutic session success. VR applications thus have the potential to enable functional assessment in a clinic setting. In the following, we present experimental studies that have used VR in the treatment of PWSs.

In [14], Walkom proposed using the VR exposure therapy (VRET) method for PWSs. In the study, a total of six participants with different types of stutters were introduced to a lecture theater environment through a VR headset, where they each had to give a speech for five minutes to an animated audience on a particular topic. The data gathered for this experiment included the participant's anxiety level, skin, and body temperature, and comments related to the participant's speech. The results indicated that VRET significantly reduced the participants' anxiety levels and likelihood of stuttering, with their speech exhibiting levels of improvement that they could then transfer to real-life situations.

In [12], Brundage et al. created a VR job interview environment under two conditions: challenging and supportive VR job interviews. A total of 23 participants with different types of stutters experienced both interviews' conditions. For each participant, the percentage of stuttered syllables was calculated manually for both interviews. Subjective and self-reported questionnaires were also collected to measure the sense of presence, pictorial realism, and environment's usability. The results revealed that the interview condition affected the amount of stuttering that the participants produced. For example, more stuttering were observed in challenging virtual interviews. Participants' subjective reports of the VR experience revealed reactions similar to those they reported experiencing in the real world.

In [7], Brundage and Hancock aimed to test the correlation between real-world and VR exposure for public speaking experience. To this end, they measured the cognitive, affective, and behavioral attitudes of 10 PWSs during five-minute speeches under two conditions: in front of virtual audiences and in front of a live audience. The frequency of stuttering was also measured during each speech. Their findings revealed that the affective, behavioral, and

cognitive experiences of PWSs during public speaking in VR environments were significantly correlated to those seen in the real-world environment. Therefore, the authors reported that the authentic, safe, and controlled environments provided via VR have utility in stuttering assessment and treatment.

In [13], Brundage et al. investigated the utility of VR environments in examining subjective ratings of distress in PWSs' physiological reactivity by means of their heart rates and skin conductance levels. Ten participants participated in the experiment. They collected physiological and subjective measures of arousal from the participants during four-minute speeches in two conditions: virtual audience and virtual empty room. Their findings revealed that stuttering frequency and physiological measures did not differ across speaking conditions; however, subjective ratings of distress were significantly higher in the virtual audience condition than in the virtual empty room.

3. ORIGINALITY

The study's contributions are as follows: (i) The study provides a framework that categorizes various ways in which VR can improve and augment speech and language pathology treatment. The proposed framework presents a methodology for augmenting stuttering therapy with VR systems during the transfer and maintenance phase. The principle of this methodology is to create a three-dimensional (3D) scene similar to a real-world scene and to subsequently record speech to test the fluency level. (ii) A speech processing framework is proposed that detects stuttering events using an off-the-shelf speech application programming interface (API). The study was aimed at providing a direction for evidence-based practice pertaining to VR technology for Arabic-speaking PWSs. We mainly aimed to investigate the effects and feedback of using VR as well as this system's effects on task speed, task performance, and subjective measures, such as comfort and ease of use. We were interested in exploring clinicians' acceptance and interest in using VR technology as unobtrusive devices used during speech-language therapy sessions.

4. SYSTEM DESIGN

The proposed speech-based VR system has two main components: the VR and the speech analyzer components as illustrated in Figure 1.
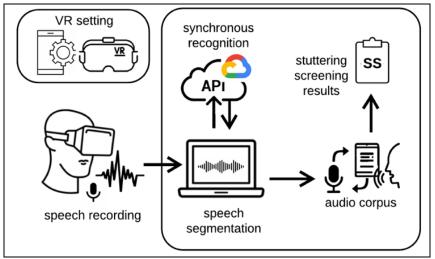
Figure 1. General structure of proposed speech-based VR system.

The first component is responsible for rendering the VR scene according to the therapist's settings, which control the number of audience members in the scene. As the test level progresses, the number of audience members increases (level 1, audience size = 5; level 2, audience size = 8; level 3, audience size = 11).

To create a VR environment, we used the Blender 3D modeling tool to create scene characters. We then used the Mixamo and Unity 3D tools to design the scene, animate the characters, and place the characters in proper seating positions as illustrated in Figure 2. When wearing the VR headset, PWSs find themselves situated in a simulated pitch presentation, playing the role of a presenter. In this simulated scenario, the PWSs' task is to read a script from a paper placed on a podium while interacting with the audience.

Figure 2. VR scene. When wearing the VR headset, PWSs can see the participants along with a paper containing the script placed on a podium.
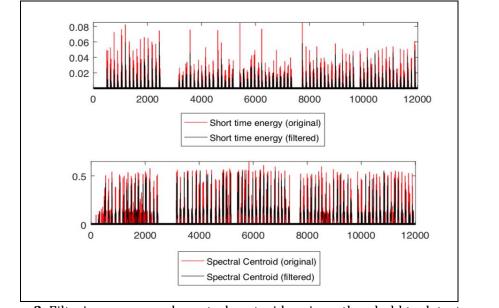
The second component is the speech analyzer. The main role of this module is to record PWS' reading tasks while they are immersed in the VR

scene and then analyze their recorded speech utterances to detect stuttering events. The detection framework uses an off-the-shelf speech API to compute reading accuracy (i.e., word error rate [WER]) and detect any prolongation (i.e., word duration abnormalities).

To detect the WER, the framework parses the automatic transcript of PAW reading generated via the API, and it compares it with the original script to compute errors (missing, repeated, or misread words). With respect to prolongation, the duration of each word read via PAW is compared with a predefined threshold computed using utterances from fluent speakers. A user utterance unit (word) is therefore labeled as *disfluent* in the following events:

1. If the duration of the user's utterance unit exceeds a predefined threshold computed using a fluent speech reference (prolongations)
2. If the automatic speech utterance recognition detects non-speech human vocal sounds (i.e., sounds that individuals produce when experiencing blockage)
3. When repetition in words is detected

The speech analyzer was initiated when the speech therapist asked to record the user's utterances while reading the script. When the user completed the task, the speech analyzer segmented the continuously read speech in the recording into utterances to facilitate analysis. Segmentation was achieved by thresholding the signal energy and spectral centroid extracted from continuous speech to detect speech segments. The output from this process defined the onset time for each utterance presented in the input audio track. Figure 3 illustrates the segmentation process.

Figure 3. Filtering energy and spectral centroids using a threshold to detect speech segments.

Every segmented utterance was then transcribed using the Google Cloud Text-to-Speech Python client library1 interface to a large vocabulary speech recognizer that uses pre-trained deep neural network models. Google Cloud Text-to-Speech was selected due to its high accuracy in under-resourced languages compared with other commercial speech-to-text APIs [21], as well as its support for Arabic dialects. The API service has three main recognition methods: synchronous recognition, asynchronous recognition, and streaming recognition. Synchronous recognition was selected to return end-pointed transcriptions of users' utterances (i.e., each word in the utterances was delimited in time, which allowed the word duration to be computed).

Prior to the recording experiment, a corpus of audio recordings was collected from three fluent female speakers reciting the same speech (number of words = 74) that was assigned to the participants in a quiet, isolated room. The recording for each speaker (SP) was segmented into extract utterances. Each utterance was then transcribed using Google Cloud Text-to-Speech, and the duration for the ith word in the speech task transcript was computed and recorded (duration_i). The threshold duration (th_i) for the ith word in the speech task transcript was then computed by averaging the word duration when the three speakers articulated it as follows:

$$th_i = \frac{\sum_{x=1}^{3} SP_x duration_i}{3} \tag{1}$$

where $SP_x duration_i$ is the ith word duration when speaker x articulates it ($SP_x, 1 \leq x \leq 3$). Then, th_i was used as the fluent speech reference for detecting prolonged speech segments as follows:

$$Prolongation = \begin{cases} True; p\_duration_i > th_i \\ False; otherwise \end{cases} \tag{2}$$

where p_$duration_i$ is the duration of the ith word when the participant recites it.

When experiencing blockage, individuals tend to produce non-speech vocal sounds, which the speech API fails to recognize; as a result, the API returns null values. Thus, a blockage was detected as follows:

$$Blockage = \begin{cases} True; p\_word_i = null \\ False; p\_word_i = word_i \end{cases} \tag{3}$$

where p_$word_i$ is the transcription of the speech segment that the participant recites such that the expected transcription is $word_i$.

---

1 https://cloud.google.com/speech-to-text/

Repetition was also recorded. The frequency of p_$word_i$ was compared with the frequency of $word_i$, and a repetition event was recorded when $count(p\_word_i) > count(word_i)$. The stuttering screening (*SS*) results were computed as follows:

$$SS = count \text{ (Prolongation) + count (blockage) + count (repetition)} \qquad (4)$$

5. EXPERIMENT AND ANALYSIS

An evaluation study was conducted to validate the use of the proposed system to support the clinical assessments of PWSs. The study's aim was twofold: (i) to evaluate the effectiveness of using a VR environment as a medium for presenting speech training tasks; and (ii) to evaluate the accuracy of the speech analyzer module in detecting stuttering events.

Three participants were recruited: two female participants and one male participant with ages ranging between 30 and 34 years (mean = 32 years, standard deviation = 1.6). All participants were healthy, had normal eyesight, and did not have prior experience with VR technology. More information about the participants is presented in Table 1.

Table 1. Participant information. Stuttering severity scores were provided by a speech therapist.

Participant	Age	Stuttering severity score
1	32	Moderate
2	34	Mild
3	30	Severe

The experiment was conducted in one session in an isolated room under the supervision of an expert SLP. The participants were the SLP's patients. The participants' main task was to read an Arabic script from a paper placed on a virtual podium facing the virtual audience (as in the VR scene). The number of words in the speech task was 74, and fluent speakers could read them in 44.7 ± 2.4 s.

The experiment started with a short orientation to ensure that the participants had a clear understanding of the required tasks. Thereafter, the participants put on a lavalier microphone digital Sony IC-Recorder (ICD-AX412F) and Samsung Gear VR (Oculus), lightweight VR glasses that are compatible with Android devices. They are required for displaying and rendering high-quality VR.

The participants were instructed to adjust their positions until they could see a clear VR scene, and to confirm the clarity of the text of the paper in the scene. The Audacity software was used to collect the recording, which was then fed into the speech analyzer to process the segmentation and to detect stuttering. After the experiment, the participants were interviewed to provide feedback on the use of VR in rehabilitation.

The results are presented in Table 2. The duration of the setup and preparation was between two and three minutes, whereas the duration of the observation sessions (reading sessions) ranged between 1:40 and 2:25 minutes. On average, the participants exceeded the mean fluent recitation duration of the script by 1:15 min.

Table 2 also displays the stuttering event percentage for the collected utterance for each participant. A strong positive correlation between the session length and the automatically detected stuttering event was found (R = 0.95), suggesting an acceptable performance of the speech analyzer in detecting stuttering events, especially prolongation instances.

Table 2. Results obtained from experimental study

Participant	Preparation and step up (min)	Session length (min)	Percentage of automatically detected stutter speech events (%)
1	2	1:40	8.6
2	2	2:25	20.8
3	3	1:47	4.8

Through subjective observation, the SLP did not observe a significant difference in the prosody of participants' speech when using VR compared with their speech outside of the VR environment. An interesting observation is that the participant with a mild stuttering severity exhibited a higher percentage of stuttering events. This observation suggests that VR may suit only individuals with higher stuttering severity. Additional data are required to validate this theory.

During the post-test interviews, the participants positively rated their VR experiences in different aspects. In terms of aesthetic design, the participants reported an acceptable resemblance between the VR scene and a real-word conference room. In terms of character design, the participants reported a mild uncanny valley effect (i.e., experiencing a lower level of human resemblance in a synthetically designed character). In terms of the immersion experience, the participants reported that using VR triggered a similar pitching experience in a VR scene as that of a real-life situation. The participants also stated that their feelings of fear were similar to those that they felt in real-world speaking-related activities. These comments indicated significant levels of presence and immersion in the environment.

VR offers benefits in the treatment of several phobias and fears. First, VR practice is safe; for example, it is much safer to desensitize oneself to a fear of heights on a virtual high building compared with a real building. Various situations evoking fear may be emulated or simulated by employing a VR system. PWSs may become desensitized to a situation by encountering it multiple times. Second, VR technology can enable PWSs to interact with multiple individuals in their vicinities, and they can regulate the way in which they communicate. In addition, based on individual needs, these interactions

are repeatable and can be arranged hierarchically, allowing for systematic practice and feedback from the treating professional. Third, the confidence with which PWSs interact can be increased, which, in turn, can allow them to confront routine situations more effectively.

The results shown are only preliminary but demonstrate the potential of embedding VR technology integration with speech analysis in speech-language therapy sessions. Our proposed system has the potential to operate as an adjunct technology in stuttering therapy, as it can be used in various phases of goal setting and task/situation design.

The results indicated that our system could be used as a supervised application for improving the quality of therapy. However, some issues need to be addressed to improve the system's effectiveness. These issues include system-related issues, where more consideration should be given to addressing adaptability, reliability, and safety issues associated with this technology. They also include user-related issues, where therapist training, user acceptance, cost, and time should be taken into consideration. A wider study with a larger number of participants should be conducted to validate these findings.

6. CONCLUSION

Due to the numerous advantages of VR, explosive growth has occurred in VR-based system research, publications, and applications in the management field. VR devices' portability, usability, reliability, and cost-effectiveness have all been cited as factors influencing the increased use of this method in applied research contexts. Moreover, advances in VR technologies have attracted the interest of an increasing number of scientists due to the promising practical applications of VR technologies in a human–machine relationship.

The assessment and treatment of PWSs present challenges for SLPs, such as developing a reliable method of measuring, determining which types of treatment programs are most effective for a specific individual, and measuring treatment processes. VR and speech analysis applications have the potential to improve speech therapy sessions by providing real-time scenes along with speech fluency measurement and the analysis of therapy sessions.

In this study, we examined VR application and speech analysis during a speech-language assessment session. Our aim was to use VR technology and speech analysis in a clinical setting to augment therapy. To this end, we conducted a case study that involved recording speech from three PWSs while they used the VR application. This work provides a framework and comprehensive background for VR. We conducted speech analysis experiments as an assessment and therapeutic tool in speech therapy. In addition, we examined the feasibility and potential benefits of applying VR technology in the field of speech and language pathology. Our preliminary results revealed the potential to incorporate such VR frameworks in clinical

settings, and further studies should be conducted to investigate additional potential benefits.

Future directions of this research include clinical trials that can be conducted for assessing the effectiveness of our proposed system to augment speech therapy for stuttering. In addition, studies can be conducted to build and validate alternative speech analysis approaches, as well as to extend design considerations to produce online analysis and visualization. Furthermore, studies can be conducted to integrate the experimental protocol into a speech therapy tutorial. A different protocol with a speech training task and challenging virtual VR environments that include interactive question-and-answer sessions could be considered. We also aim to increase our proposed system's performance and robustness. Other modalities can additionally be considered, such as the measurement of skin conductance to detect user emotions.

Acknowledgements

REFERENCES
[1] Schultheis, M.T., Himelstein, J.M., Rizzo, A.A., **Virtual reality and neuropsychology: upgrading the current tools**, *The Journal of head trauma rehabilitation*, Vol.17, pp.378–394, 2002.
[2] Steuer, J., **Defining Virtual Reality: Dimensions Determining Telepresence**, *Journal of Communication*, Vol.42, pp.73–93, 1992.
[3] Jackie, F., LeHong, H., **Hype cycle for emerging technologies**, 2011.
[4] Panetta, K.: **Top trends in the gartner hype cycle for emerging technologies**, 2017.
[5] Caughter, S., Dunsmuir, S., **An exploration of the mechanisms of change following an integrated group intervention for stuttering, as perceived by school-aged children who stutter (CWS)**, Journal of Fluency Disorders, Vol.51, pp.8–23, 2017.
[6] Brundage, S.B., **Virtual reality augmentation for functional assessment and treatment of stuttering**, *Topics in Language Disorders*. Vol.27, pp.254–271, 2007.
[7] Brundage, S.B., Hancock, A.B., **Real Enough: Using Virtual Public Speaking Environments to Evoke Feelings and Behaviors Targeted in Stuttering Assessment and Treatment**, *American Journal of Speech-Language Pathology*. Vol.24, pp.139–149, 2015.
[8] Guitar, B., **Stuttering: An integrated approach to its nature and treatment**. *Lippincott Williams & Wilkins*, 2013.
[9] Louis, K.O., **Measurement issues in fluency disorders**, *Current issues in stuttering research and practice*, pp.69–94, 2014.

[10] Bloodstein, O., & Ratner, N., **A handbook on stuttering**, S, San Diego, 1995.

[11] Mulcahy, K., Hennessey, N., Beilby, J., Byrnes, M., **Social anxiety and the severity and typography of stuttering in adolescents**. *Journal of Fluency Disorders*, Vol.33, pp.306–319, 2008.

[12] Brundage, S.B., Graap, K., Gibbons, K.F., Ferrer, M., Brooks, J., **Frequency of stuttering during challenging and supportive virtual reality job interviews**, *Journal of Fluency Disorders*, Vol.31, pp.325–339, 2006.

[13] Brundage, S.B., Brinton, J.M., Hancock, A.B., **Utility of virtual reality environments to examine physiological reactivity and subjective distress in adults who stutter**, *Journal of Fluency Disorders*, Vol.50, pp.85–95, 2016.

[14] Walkom, G., **Virtual reality exposure therapy: To benefit those who stutter and treat social anxiety**, *International Conference on Interactive Technologies and Games: EduRob in Conjunction with iTAG,* pp. 36–41, IEEE, 2016.

[15] Scheurich, J.A., Beidel, D.C., Vanryckeghem, M., **Exposure therapy for social anxiety disorder in people who stutter: An exploratory multiple baseline design**. *Journal of Fluency Disorders*, Vol.59, pp.21–32, 2019.

[16] Kourkounakis, T., Hajavi, A., Etemad, A., **Detecting Multiple Speech Disfluencies Using a Deep Residual Network with Bidirectional Long Short-Term Memory**, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 6089–6093, 2020.

[17] Lowder, M.W., Maxfield, N.D., Ferreira, F., **Processing of self-repairs in stuttered and non-stuttered speech. Language**, *Cognition and Neuroscience*, Vol.35, pp.93–105, 2019.

[18] Kostek, B., Czyzewski, A., Kaczmarek, A., **Intelligent Processing of Stuttered Speech**, *Journal of Intelligent Information Systems*, Vol.21, pp.143–171, 2003.

[19] Al-Nafjan, A., Al-Wabil, A., AlMudhi, A., Hosny, M., **Measuring and monitoring emotional changes in children who stutter**. *Computers in Biology and Medicine*. Vol.102, pp.138–150, 2018.

[20] Packman, A., Meredith, G., **Technology and the evolution of clinical methods for stuttering**, *Journal of Fluency Disorders*, Vol.36, pp.75–85, 2011.

[21] Herchonvicz, A.L., Franco, C.R., Jasinski, M.G., **A comparison of cloud-based speech recognition engines**. *Anais do Computer on the Beach*, pp. 366–375, 2019.