Li-Fi Technology to Improve Communication Accessibility for Dysarthric Speakers

Selva Nidhyananthan.S *,[a], Bhavasri.V[b], Shantha Selvakumari.R[c] and Suganya.D[d]
[a]Senior Associate Professor, [b]PG Scholar, [c]Senior Professor & Head, [d]Research Scholar
[a-d] Department of Electronics and Communication Engineering,
Mepco Schlenk Engineering College, Sivakasi, TamilNadu, India

*Corresponding Author

Abstract— An approach for communication enhancement for individuals with Dysarthria, a speech disorder characterized by difficulty articulating words is proposed. The proposed system leverages Li-Fi (Light Fidelity) technology to transmit Dysarthria-corrected audio samples. The system takes a Dysarthria-affected audio sample as input, processes it using a speech recognition engine to convert it to natural-sounding text, and then transmits the recognized audio through an audio jack to a Li-Fi transmitter where an LED acts as the transmitter. The transmitted signal is received by a photodiode integrated into a solar panel, amplified using an integrated circuit for clear playback. Dysarthria corrected speech output is subjected to qualitative analysis of prosodic features, Mel Cepstral Distortion (MCD) measure and subjective evaluation measure. This system offers a promising solution for improving communication accessibility and clarity for people with Dysarthria.

Keywords: Dysarthria, Speech Recognition, Li-Fi, Communication Enhancement, Mel Cepstral Distortion

1. INTRODUCTION

Dysarthria significantly impacts communication for individuals by affecting their ability to control the muscles used for speaking [1]. This result in a range of speech impairments, such as slurring, distorted sounds, and problems with volume or pitch control [2]. The severity of Dysarthria can vary depending on the underlying neurological condition. The frustration of not being understood and the social isolation arising from Dysarthria highlight the need for innovative assistive technologies to empower individuals with this speech disorder [3]. Speech recognition technology has emerged as a promising approach to enhance communication for people with dysarthria [4]. These systems convert spoken language into text, allowing individuals with Dysarthria to communicate more effectively through alternative channels, such as text-to-speech synthesis [5]. However, accurately capturing and interpreting speech affected by Dysarthria remains a challenge. While advancements have been made, research continues to explore alternative methods and communication enhancement techniques to better address the unique needs of individuals with Dysarthria.

Li-Fi (Light Fidelity) technology emerges as a potential innovation in communication, offering a compelling alternative to traditional radio frequency (RF) communication [6]. It utilizes light-emitting diodes (LEDs) to transmit data by modulating their light intensity. This approach presents several advantages over RF-based communication such as higher data rates as Li-Fi has the potential to achieve higher data rates compared to RF, especially in congested wireless environments [7]. This increased bandwidth can be crucial for future communication demands. Li-Fi signals are not susceptible to electromagnetic interference which can disrupt RF communication in settings with numerous electronic devices [8-10]. This immunity makes Li-Fi ideal for hospitals, airplanes, and other sensitive environments. Due to the limited range of Li-Fi signals, data transmission is less likely to be intercepted unintentionally [11]. These characteristics of Li-Fi technology suggest its potential for various applications beyond simply replacing existing wireless communication methods.

This paper proposes a novel approach that utilizes Li-Fi technology to address communication challenges faced by individuals with Dysarthria. While various technologies and methodologies have been explored to aid individuals with speech disorders, the integration of Li-Fi technology for this purpose is a novel contribution. Traditionally, communication aids for Dysarthria have predominantly relied on techniques such as speech therapy, augmentative and alternative communication (AAC) devices, and signal processing algorithms for speech enhancement [12-14]. However, these approaches often have limitations in terms of accurate recognition quality and fast transmission. The proposed Li-Fi-based system has the potential to significantly improve communication accessibility and empower individuals to express themselves more effectively.

The paper is organized as follows. In section 2, the system design of Speech recognition module and Li-Fi transceiver is described. Then, in section 3, results of the quantitative and qualitative measures carried out for the corrected speech are discussed.

2. SYSTEM DESIGN

The system design incorporates a speech recognition module as shown in Figure 1 to accurately transcribe Dysarthric speech input into text format. Through robust feature extraction, the phonetic and linguistic cues embedded within the dysarthric speech are extracted as text. Following speech recognition, the system integrates a text-to-speech (TTS) engine to transform the transcribed text into clear and natural-sounding speech output.

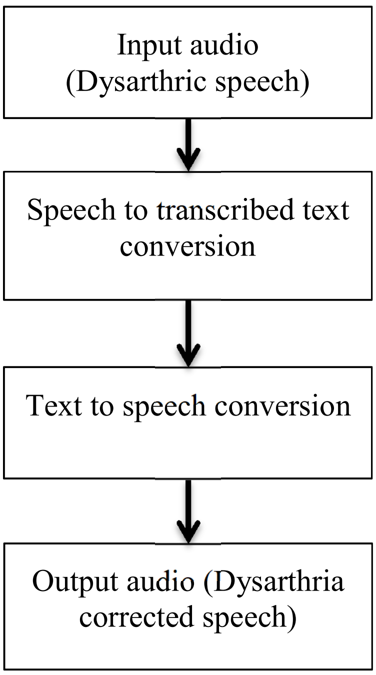


Figure 1 Speech recognition module

The system uses Li-Fi technology for the transmission of corrected Dysarthria speech generated by the recognition module. The Li-Fi transceiver module is designed as separate transmitter and receiver modules. Figure 2 shows the two primary modules of Li-Fi transmission with key components. The Li-Fi transmitter module acts as the information source, converting the Dysarthria-corrected audio signal into a data (light) stream suitable for Li-Fi transmission. The custom-designed unit is built around an LED driver circuit. LED serves as the light source for Li-Fi transmission. The LED's intensity is modulated according to the data stream from the Li-Fi transmitter, effectively encoding the Dysarthria-corrected audio information onto the light signal. 3.5 mm Audio Jack serves as the audio input for the system. The user with Dysarthria will connect a microphone through this jack to capture their speech.

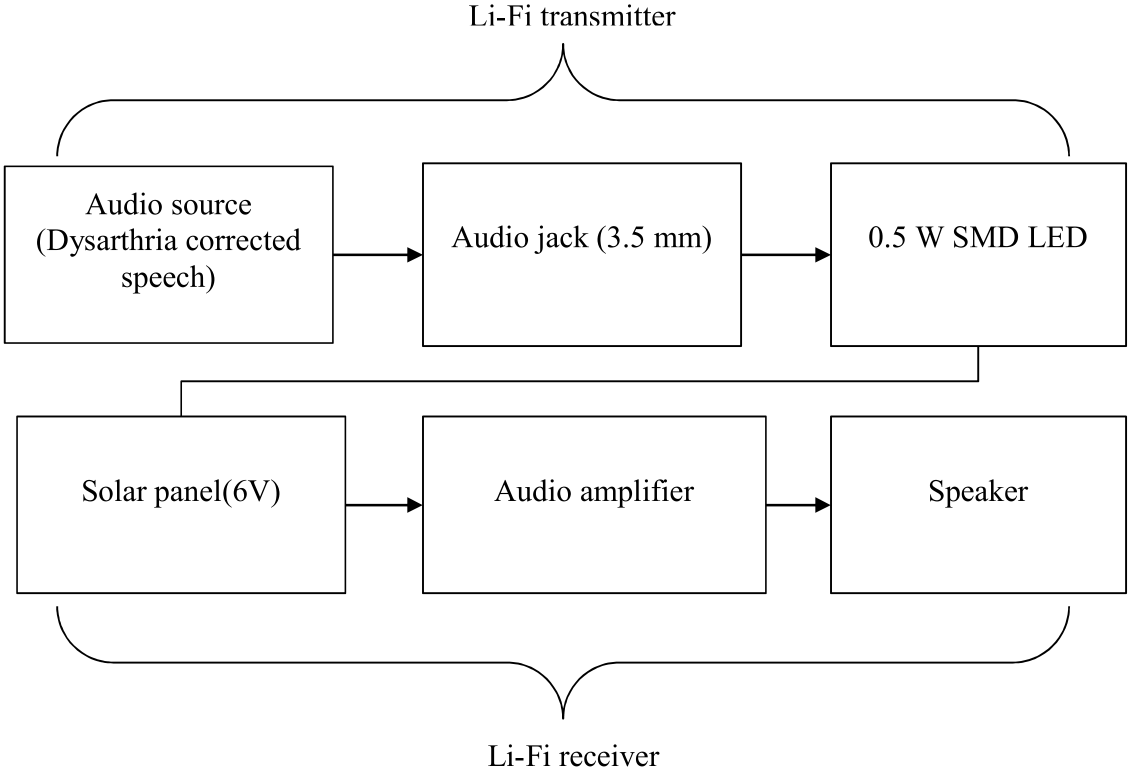
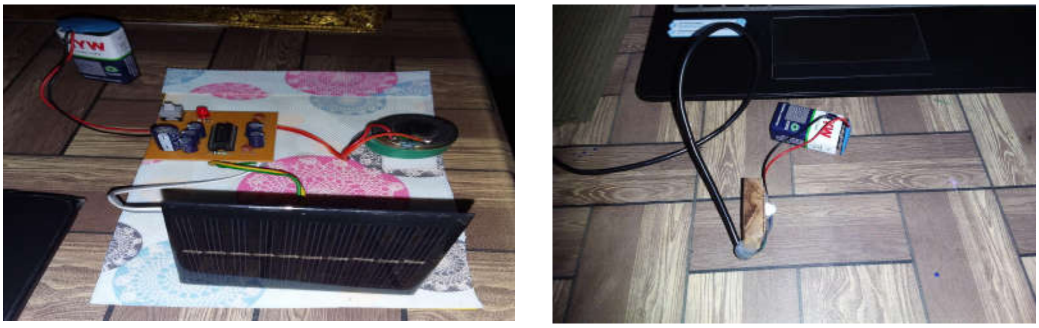


Figure 2 Li-Fi transceiver

The Li-Fi Receiver module receives the Li-Fi signal transmitted by the LED and converts it back into an electrical data stream. The Solar Panel acts as the photodiode at the receiver. The 5V output voltage aligns with the typical operating voltage range of many Li-Fi receivers. Audio Amplifier circuit increases the amplitude of the demodulated audio signal from the Li-Fi receiver, making it powerful enough to drive the speaker. Amplifier Circuit utilizes the IC 2025 to amplify Dysarthria corrected audio signals. Speaker reproduces the Dysarthria corrected audio received through the Li-Fi channel, allowing the listener to hear a clear and understandable version of the speaker's message. The Li-Fi transceiver is shown in Figure 3.



(a) (b)

Figure 3 Hardware circuit of Li-Fi a) Receiver b) Transmitter

3. RESULTS AND DISCUSSION

Dysarthric speech samples were obtained from the Kaggle Dysarthria detection dataset. This dataset contains 2000 samples for dysarthric males, dysarthric females, non-dysarthric males, and non-dysarthric females. Mel-Cepstral Distortion (MCD) serves as a crucial metric for evaluating the dissimilarity between clean speech and speech samples affected by Dysarthria [15]. It measures the spectral distortion between the original speech signal and the processed or corrected version, providing insights into the effectiveness of speech enhancement techniques [16]. The MCD values obtained as shown in Table 1 reveals a clear distinction between the original dysarthric speech samples and the dysarthria-corrected versions.

Table 1. MCD comparison of Dysarthric speech and corrected speech

Dysarthric sample from Dysarthria detection dataset	MCD value of Dysarthric speech	MCD value of Dysarthria corrected speech
Sample 1	845.25	412.78
Sample 2	825.17	415.65
Sample 3	840.15	410.35
Sample 4	835.20	417.14
Sample 5	800.00	409.92
Sample 6	810.23	418.22
Sample 7	815.12	416.55
Sample 8	830.00	412.25
Sample 9	820.83	414.62
Sample 10	805.55	409.01
Average	823.45	412.78

The average MCD for the Dysarthric speech samples was 823.45. This relatively high value indicates a significant level of Mel-Cepstral distortion, reflecting the speech impairments characteristic of Dysarthria. The average MCD for the Dysarthria-corrected speech generated by the system was 412.78. This value is considerably lower compared to the original samples, suggesting a successful reduction in Mel-Cepstral distortion through the speech recognition software's processing. The observed reduction in Mel-Cepstral Distortion (MCD) holds promising implications for the field of assistive communication technology and speech therapy for individuals with Dysarthria. By quantitatively measuring the improvement in speech clarity and intelligibility, MCD provides objective evidence of the efficacy of the proposed system in addressing the communication challenges posed by Dysarthria. This validation is crucial for building confidence in the system's effectiveness and fostering wider adoption among clinicians, speech therapists, and individuals with Dysarthria.

The comparison of prosodic features of a single sample from female dysarthric speech samples is as shown in the Figure 4 between dysarthria-affected speech and speech produced by the designed speech recognition module provides valuable insights into the differences in speech production and recognition dynamics .In terms of pitch, corrected speech produced by the designed speech recognition module typically exhibits frequency ranges of 150-300 Hz, reflecting a typical vocal range of an adult female [17]. This consistency in pitch characteristics highlights the effectiveness of the designed speech recognition module in producing clear and consistent speech outputs.

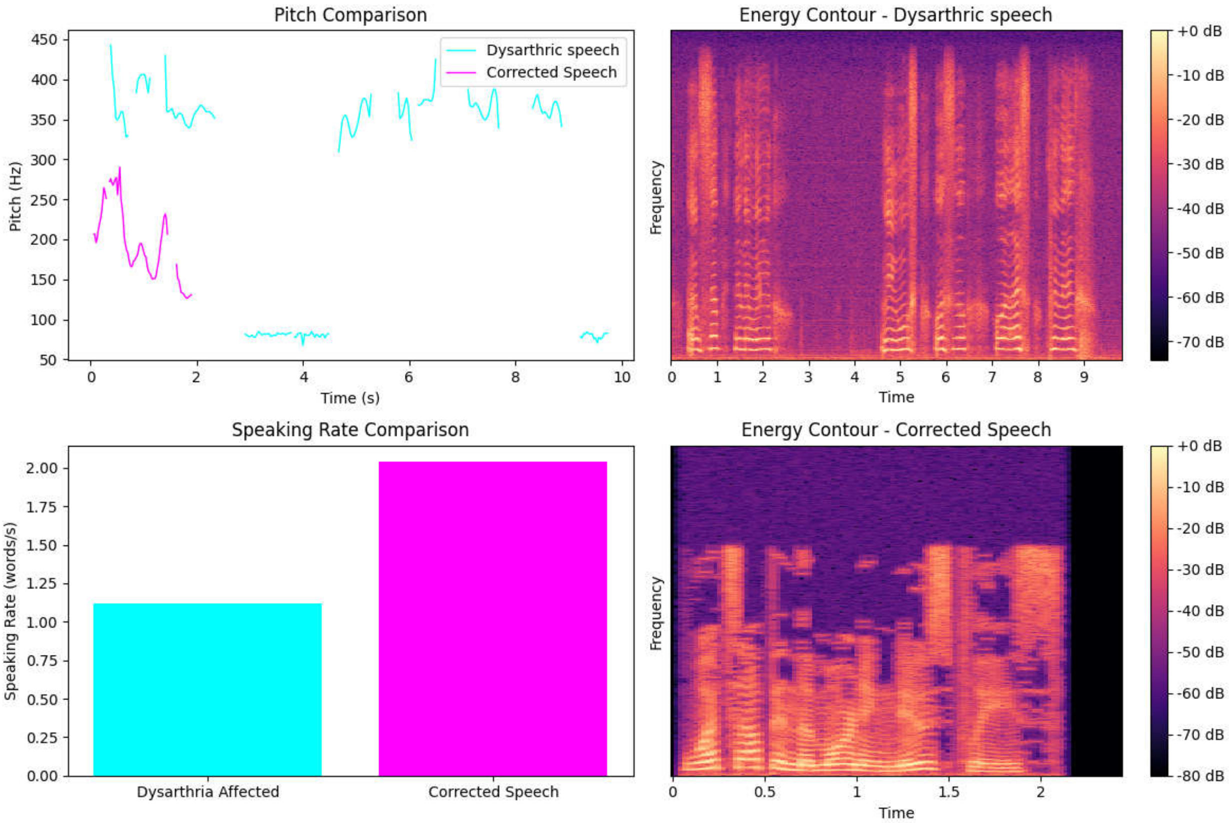


Figure 4 Comparison of prosodic features of female speech sample

Conversely, dysarthric speech demonstrates notable variations in pitch, with frequency ranges extending from 350-450 Hz in some instances. Additionally, instances of near-zero frequency suggest potential breaks or interruptions in vocalization, which may be attributed to dysarthria-related motor control issues [18-19]. These deviations from typical pitch patterns underscore the challenges in maintaining consistent vocal production in individuals with Dysarthria. Regarding speaking rate, dysarthric speech shows a significantly lower rate, averaging around one word per second, compared to the approximately double the rate observed in corrected speech produced by the designed speech recognition module. This substantial difference in speaking rate highlights the impact of dysarthria on speech fluency, articulatory speed, and motor control. The energy contour of Dysarthric speech exhibits irregular phonation characterized by glottal intervals [20] while the spectrogram of the corrected speech resembles that of a normal female speech indicating the improvement in quality through the speech synthesis module. Ten listeners participated in a comparison Mean Opinion Score (MOS) test for subjective evaluation [21-22]. Participants were presented with audio samples of both dysarthric and corrected speech and were asked to rate the quality of corrected speech over the Dysarthric speech. The results in Figure 5 show that the quality of the corrected speech has increased for the subjective listener's perception as well. The category ratings of the test is shown in the Table 2.

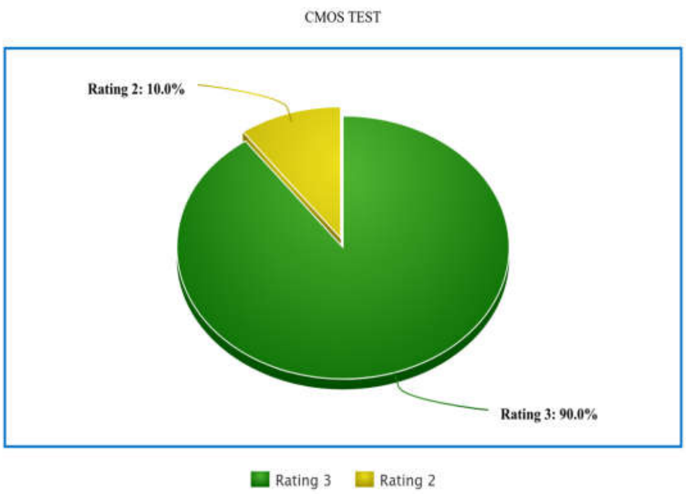


Figure 5 CMOS Test for subjective evaluation of speech quality

Table 2 Comparison category ratings used in the comparison mean opinion score (CMOS) test

Rating	Quality of corrected speech compared to the Dysarthric speech is
3	Much better
2	Better
1	Slightly better
0	About the same
-1	Slightly worse
-2	Worse
-3	Much worse

Future research directions may focus on refining the speech processing algorithms and techniques to further reduce Mel-Cepstral Distortion and enhance speech clarity. Exploring advanced machine learning and deep learning approaches tailored specifically for Dysarthria holds promise for achieving even greater improvements in communication accessibility. Additionally, integrating the Li-Fi communication system with existing speech therapy tools and platforms could provide valuable real-time feedback on articulation and pronunciation, thereby enhancing the efficacy of speech therapy interventions for individuals with Dysarthria.

4. CONCLUSION

The proposed system with the facility for recognizing Dysarthric speech and transmission of the corrected speech signal through a Li-Fi transceiver paves way for improvement in communication accessibility for Dysarthric speakers. The quality of the corrected speech is assessed through MCD and various prosodic measures followed by subjective qualitative measure which shows a prominent enhancement. Use of Li-Fi also eliminates the need for separate RF broadband allocation ensuring faster communication. Hence, the integration of Li-Fi with various communication assistive devices holds a promising future.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Yorkston, Kathryn M., and Andrew R. Strand. "Management of Dysarthria in Adults." Physical Medicine and Rehabilitation Clinics of North America, vol. 21, no. 4, 2020, pp. 789–807.: https://www.ncbi.nlm.nih.gov/books/NBK592453
2. Duffy, Janet R. Motor Speech Disorders. 3rd ed., Mosby Elsevier, 2019.: https://shop.elsevier.com/books/motor-speech-disorders/duffy/978-0-323-53054-5
3. Beukelman, David R., and Katherine M. Yorkston. Communication Enhancement for Individuals with Dysarthria. 2nd ed., Paul H. Brookes Publishing Co., 2021.: https://journals.sagepub.com/doi/10.2466/pms.110.3.965-982
4. Pathak, Prathap H., et al. "Visible Light Communication, Networking, and Applications: A Survey." IEEE Communications Surveys & Tutorials, vol. 15, no. 2, 2023, pp. 490–505.: https://ieeexplore.ieee.org/document/9780735/
5. Karunaratne, Chathurika, et al. "Li-Fi Security: A Comprehensive Review." Journal of Network and Computer Applications, vol. 161, 2020,
6. Ali, Akbar, et al. "Audio Streaming Using Li-FI Communication." Irish Interdisciplinary Journal of Science & Research 7.1 (2023): 1-7.
7. A. Devrani, D. Darshan, A. Vidyarthi, R. Prakash and R. Gowri, "Design of a LiFi System for Short-Distance Communications," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023
8. Saranya, S., et al. "Audio transmission using visible light communication and Li-Fi technology." 2021 6th International Conference on Inventive Computation Technologies (ICICT). IEEE, 2021.
9. Kent, R. D., Weismer, G., Kent, J. F., Vorperian, H. K., & Duffy, J. R. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication*

Disorders, *32*(3), 141-186. https://doi.org/10.1016/S0021-9924(99)00004-0
10. Bőhm, Tamás & Both, Zoltán & Németh, Géza. (2019). Automatic Classification of Regular vs. Irregular Phonation Types. 5933. 43-50. 10.1007/978-3-642-11509-7_6.
11. Loizou, P.C. (2021). Speech Quality Assessment. In: Lin, W., Tao, D., Kacprzyk, J., Li, Z., Izquierdo, E., Wang, H. (eds) Multimedia Analysis, Processing and Communications. Studies in Computational Intelligence, vol 346. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-19551-8_23
12. T.D. Subha, T.D. Subash, N. Elezabeth Rani, P. Janani,Li-Fi: A Revolution in Wireless Networking,Materials Today: Proceedings,Volume 24, Part 4,2020,Pages 2403-2413,ISSN 2214-7853,https://doi.org/10.1016/j.matpr.2020.03.770.
13. Qian, Z.; Xiao, K. A Survey of Automatic Speech Recognition for Dysarthric Speech. *Electronics* **2023**, *12*, 4278. https://doi.org/10.3390/electronics12204278
14. D. Shah, V. Lal, Z. Zhong, Q. Wang and S. R. Shahamiri, "Dysarthric Speech Recognition: A Comparative Study," *2023 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, Bucharest, Romania, 2023, pp. 89-94, doi: 10.1109/SpeD59241.2023.10314881.
15. Brandt, Erika & Zimmerer, Frank & Andreeva, Bistra & Möbius, Bernd. (2022). Mel-Cepstral Distortion of German Vowels in Different Information Density Contexts. 2993-2997. 10.21437/Interspeech.2017-838.
16. Vasilijević, A., & Petrinović, D. (2021). Perceptual Significance of Cepstral Distortion Measures in Digital Speech Processing. *Automatika*, *52*(2), 132–146. https://doi.org/10.1080/00051144.2011.11828412
17. Weijie Wang, Gaopeng Zhang, Luming Yang, V.S. Balaji, V. Elamaran, N. Arunkumar,Revisiting signal processing with spectrogram analysis on EEG, ECG and speech signals,Future Generation Computer Systems,Volume 98,2023,Pages 227-232,ISSN 0167-739X,https://doi.org/10.1016/j.future.2018.12.060.
18. Badeel, R.; Subramaniam, S.K.; Hanapi, Z.M.; Muhammed, A. A Review on LiFi Network Research: Open Issues, Applications and Future Directions. *Appl. Sci.* **2021**, *11*, 11118. https://doi.org/10.3390/app112311118
19. Mohammad Reza Ghaderi,LiFi and Hybrid WiFi/LiFi indoor networking: From theory to practice,Optical Switching and Networking,Volume 47,2023,100699,ISSN 1573-4277, https://doi.org/10.1016/j.osn.2022.100699.
20. S. Alharbi *et al*., "Automatic Speech Recognition: Systematic Literature Review," in *IEEE Access*, vol. 9, pp. 131858-131876, 2021, doi: 10.1109/ACCESS.2021.311253.
21. Hosier, Jordan, et al. "A systematic study of open source and commercial text-to-speech (TTS) engines." Text, Speech, and Dialogue: 23rd International Conference, TSD 2020, Brno, Czech Republic, September 8–11, 2020, Proceedings 23. Springer International Publishing, 2020.
22. SpeechTokenizer:Unified Speech Tokenizer for Speech Large Language Models – Scientific Figure on ResearchGate. Available at https://www.researchgate.net/figure/Melspectorgram-of-German-speech-and-resynthesized-speech-of-SpeechTokenizer-RVQ-1-tokens_fig3_373552537 [accessed 15 Apr, 2024]