Studia i Materiały Informatyki Stosowanej, Tom 16, Nr 2, 2024
str. 28-31

IT CHALLENGES OF SPEECH THERAPY AND PHYSIOTHERAPY OF THE ORAL-FACIAL TRACT

Emilia Mikołajewska

Ludwik Rydygier Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Toruń, Poland
Department of Physiotherapy, Faculty of Health Sciences
Jagiellońska 13-15, 85-087 Bydgoszcz

Abstract: *Speech therapy and physiotherapy of the oro-facial tract can already be effectively supported by IT solutions, both in the area of diagnosis and therapy. However, this requires an integrated approach based on interdisciplinary cooperation between speech therapists, physiotherapists and computer scientists. The aim of the article is to assess the extent to which current IT capabilities are used to support diagnostics and therapy in speech therapy and physiotherapy in the oro-facial tract and to determine the potential for their further stimulated development.*

Słowa kluczowe: *speech therapy, neurology, physiotherapy, dental physiotherapy, physiotherapy of the oral-facial tract, informatics, physiotherapy of the masticatory organ, resting position of the tongue, tongue*.

Informatyczne wyzwania logopedii i fizjoterapii traktu ustno-twarzowego

Streszczenie: *Logopedię i fizjoterapię układu ustno-twarzowego można już skutecznie wspierać rozwiązaniami informatycznymi, zarówno w obszarze diagnostyki, jak i terapii. Wymaga to jednak zintegrowanego podejścia, opartego na interdyscyplinarnej współpracy logopedów, fizjoterapeutów i informatyków. Celem artykułu jest ocena, w jakim stopniu obecne możliwości informatyki są wykorzystywan dla wsparcia diragnostyki i terapii w logopedii i fizjoterapii traktuv oro-facjalnego oraz określenie potencjału do ich dalszego stymulowanego rozwoju.*

Słowa kluczowe: *logopedia, neurologopedia, fizjoterapia, fizjoterapia stomatologiczna, fizjoterapia traktu ustno-twarzowego, informatyka, fizjoterapia narządu żucia, pozycja spoczynkowa języka, język.*

.

1. Introduction

There is great interest and enormous potential in the use of IT solutions (currently increasingly based on artificial intelligence - AI) in speech therapy and physiotherapy of the oro-facial tract. Most AI-based research currently focuses on natural speech and less on speech deficits [1-4]. Multicenter studies across diverse populations, conducted in collaboration between clinicians, scientists and engineers, can further develop this technology in a way that addresses the greatest societal needs in this area.
The aim of the article is to assess the extent to which current IT capabilities are used to support diagnostics and therapy in speech therapy and physiotherapy in the oro-facial tract and to determine the potential for their further stimulated development.

2. Results of the literature review

Current tongue tracking devices do not work well in everyday clinical practice. A certain solution is a larticulograph - an electromagnetic device that tracks the tongue. Continuous speech recognition of two people showed phoneme error rates of 73.92% and 66.73%, respectively, which are comparable results to commercially available alternatives. Further research may focus on visual feedback-based speech therapy and second language learning [1]. So far, ultrasound imaging of the tongue has often been used in linguistic research and speech recognition, also as visual feedback in the study of speech disorders and their treatment, research and analysis of articulation, swallowing and 3D modeling of the tongue. A comparison of traditional computer vision and image processing algorithms using machine learning (ML) and

deep learning (DL) for language segmentation showed the advantage of the latter in terms of performance and generalizability of the algorithm. Traditional techniques are better at supporting learning, so real-time hybrid approaches are recommended [2]. AI/ML-based biologically inspired speech processing can increase the accuracy of diagnostics and therapies, but requires further research [3]. Magnetic resonance imaging (MRI) is the leading standard in imaging speech and vocal tract function, but it is too slow and the quality of 3D images is insufficient. Cross-sectional consistency and image sharpness are improved by the use of visual stimuli and short text fragments, especially during natural speech [4]. 3D MRI of the entire vocal tract in high temporal (61 ms) and spatial resolution (2.4 × 2.4 × 5.8 mm3) during natural speech shows the dynamics of the vocal tract articulators, without the need for multiple repetitions. However, it is quite expensive and not widely available [5]. It is clear from this that AI and neuroscience are interconnected and it is worth using the mutual support of both. areas in both speech therapy and physiotherapy of the oro-facial tract. This applies to text processing, speech recognition, detection of dysfunctions and their classification/prediction, and validation of existing models, to learning complex therapeutic strategies without explicit instructions. above large-scale AI-based simulations help test hypotheses faster and more effectively, including those that are difficult to implement in real conditions. This will help reduce the workload of speech therapists and physiotherapists and support their decision-making processes regarding diagnosis and therapy. In a broader context, AI can be effectively used to predict and detect speech or muscle movement disorders, including neurological ones [6]. Thus, new diagnostic devices are emerging to accurately assess nasal expiratory flow, children's sensations and correlation with each child's profile [7]. Modern neuroscience and AI-based models can also provide insight into, for example, the mechanisms of disfluent speech in aphasia and the development of new techniques in the rehabilitation of disfluent aphasia [8]. In the case of hoarseness and shortness of breath, these possibilities are much lower, but automatic analysis based on AI can still be used to significantly objectively support perceptual analysis [9].

3. Discussion

3.1. Limitations of current studies

The main unresolved problems and limitations concern:

- Technological barriers and access to IT systems in areas distant from metropolises (Poland is to be fully covered by the 5G network only by 2027) may hinder access to IT-based speech therapy services, limiting their access to those who need them most;
- Furthermore, over-reliance on IT platforms for speech therapy may widen access gaps for groups who do not already have the necessary digital resources or skills to effectively use such technology (e.g. elderly people in rural areas who, in the next 10 years, may constitute 40% of Polish society);
- Lack of personalized feedback may hinder the progress of therapy;
- Limited standardization and regulation of speech therapy IT tools and applications can pose challenges in comparative research, ensuring quality control and effectiveness of different platforms and providers, and in transitioning patients from platform to platform;
- Technical problems (e.g. software compatibility or file formats) can disrupt therapy sessions, disrupting continuity of care and impairing patient motivation and engagement;
- Privacy issues related to the storage and security of sensitive patient data pose a challenge when implementing IT-based speech therapy solutions and require robust measures to ensure confidentiality.

3.2. Directions for further research

Essential research directions for IT support of speech therapy and physiotherapy of the oro-facial tract include:

- Imaging of tongue position during closed mouth - resting tongue position (other than x-ray, MRI, available to date);
- Assessment of tongue activity during speech and swallowing (other than x-ray, MRI, available to date);
- Personalised AI-based therapy to analyse speech patterns and tailor therapeutic exercises to individual needs, increasing the effectiveness and efficiency of therapy;
- Immersive virtual reality (VR) applications to simulate real communication scenarios to provide a more engaging and interactive platform for speech therapy sessions, potentially improving patient outcomes and motivation;
- Mobile apps and wearables focusing on the development of user-friendly mobile apps and wearables equipped with sensors to monitor speech production in real time,

enabling continuous feedback and tracking of progress beyond the clinical setting;
- Remote monitoring and testing the effectiveness of telepractice models for remote speech therapy, including the use of video conferencing tools, remote monitoring technologies and telehealth platforms to overcome geographical barriers and increase access to systematic professional care;
- Exploring the integration of gamification elements and digital incentives in speech therapy programmes to increase patient engagement, compliance and long-term adherence to treatment plans;
- Exploring the benefits of combining IT-based speech therapy with other modalities, such as oro-facial tract physiotherapy, to address complex communication disorders and improve overall treatment outcomes [10-14].

AI is increasingly integrated into the wider healthcare service area (including patient monitoring and secondary lesion prevention), and there is a need to assess the feasibility and effectiveness of AI support in various subspecialties of clinical care, including otolaryngology, speech therapy and physiotherapy. Already, reviews are underway of, for example, the diagnostic capabilities of ChatGPT, the ability to communicate descriptions of conditions in simple words, the accuracy in making treatment recommendations, and the appropriateness in follow-up and post-treatment recommendations for typical conditions, both in support of traditional diagnosis and therapy and telecare [11]. On the other hand, many scenarios have been developed so far using automatic speech recognition (ASR). e.g. in the treatment of dysarthria (to support articulation). However, the variability of speech in the course of communication, its disorders and the lack of corpora of disordered speech in languages other than English mean that there is a lack of highly reliable results for voice recognition using disordered speech models. A convolutional neural network that detects a small number of key words (also in atypical speech dependent on the speaker's deficit) has been developed for articulatory studies for speech therapy purposes. This will allow us to consider the introduction of more accurate remote diagnostics and monitoring pa supervised by remote speech-language pathologists [13].

4. Conclusions

Speech therapy and physiotherapy of the oro-facial tract can already be effectively supported by IT solutions, both in the area of diagnosis and therapy. However, this requires an integrated approach based on interdisciplinary cooperation between speech therapists, physiotherapists and computer scientists. It is also clear that due to low intra- and inter-personal reliability, perceptual voice assessment should be supported by objective, semi-automatic and automatic AI-based methods, especially ML. Examples include computer-assisted prosodic analysis and speech measurements.

References

1. Cao B., Ravi .S, Sebkhi N., Bhavsar A., Inan O.T., Xu W., Wang J., „MagTrack: A Wearable Tongue Motion Tracking System for Silent Speech Interfaces", J Speech Lang Hear Res. 2023, 66(8S), 3206-3221.
2. Al-Hammuri K., Gebali F., Thirumarai Chelvan I., Kanan A., „Tongue Contour Tracking and Segmentation in Lingual Ultrasound for Speech Recognition: A Review", Diagnostics, 2022, 12(11). 2811.
3. Wei H., Tao F., Huang Z., Long Y., „Bioinspired Artificial Intelligence Applications 2023", Biomimetics 2024, 9(2), 80.
4. Isaieva K., Odille F., Laprie Y., Drouot G., Felblinger J., Vuissoz P.A., „Super-Resolved Dynamic 3D Reconstruction of the Vocal Tract during Natural Speech", J Imaging. 2023, 9(10), 233.
5. Lim Y., Zhu Y., Lingala S.G., Byrd D., Narayanan S., Nayak K.S., „3D dynamic MRI of the vocal tract during natural speech", Magn Reson Med. 2019, 81(3), 1511-1520.
6. Surianarayanan C., Lawrence J.J., Chelliah P.R., Prakash E., Hewage C., „Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders-A Scoping Review", Sensors. 2023, 23(6), 3062.
7. Doi R., Akagami S., Kondo K., Yoshida Y., Chiuriki N., Ikuta M., Saiki K., Kataoka T., Narai T., Fujii N., Kawasaki M., Otsuki K., Kodani I., „Study on the Development of a New Device with Dual Cameras for Evaluating Expiratory Nasal Flow", Yonago Acta Med. 2020, 63(4), 255-265.
8. Bonilha L., Hillis A.E., Wilmskoetter J., Hickok G., Basilakos A., Munsell B., Rorden C., Fridriksson J., „Neural structures supporting spontaneous and assisted (entrained) speech fluency", Brain. 2019, 142(12), 3951-3962.
9. Haderlein T., Schwemmle C., Döllinger M., Matoušek V., Ptok M., Nöth E., „Automatic Evaluation of Voice Quality Using Text-Based Laryngograph Measurements and Prosodic Analysis", Comput Math Methods Med. 2015, 2015, 316325.
10. Duch W., Nowak W., Meller J., Osiński G., Dobosz K., Mikołajewski D., Wójcik G.M., „Consciousness and attention in autism spectrum disorders", Proceedings of Cracow Grid Workshop 2010, 202-211.
11. Langlie J., Kamrava B., Pasick L.J., Mei C., Hoffer M.E., „Artificial intelligence and ChatGPT: An otolaryngology patient's ally or foe?", Am J Otolaryngol. 2024, 45(3), 104220.
12. Prokopowicz P., Mikołajewski D., Mikołajewska, E., Kotlarz P., „Fuzzy system as an assessment tool for analysis of the health-related quality of life for the people after stroke".

Artificial Intelligence and Soft Computing: 16th International Conference 2017.
13. Mulfari D., La Placa D., Rovito C., Celesti A., Villari M., „Deep learning applications in telerehabilitation speech therapy scen arios", Comput Biol Med. 2022, 148, 105864.
14. Mikołajewska E., Mikołajewski D., „Neurorehabilitacja XXI wieku. Techniki teleinformatyczne", Impuls, Kraków 2011.