Artificial Intelligence for Severe Speech Impairment: Innovative approaches to AAC and Communication

Monica Murero[1, 2], Salvatore Vita[1], Andrea Mennitto[1], Giuseppe D'Ancona[3]

[1] University Federico II, (Italy)
[2] Distributed Artificial Intelligence Laboratory, Technische Universität, Berlin (Germany)
[3] Vivantes Klinikum Group, Berlin (Germany)
monica.murero@unina.it

Abstract. This paper aims to analyze how innovative Artificial Intelligence (AI) systems (Voiceitt®) for non-standard speech recognition may revolutionize Augmentative Alternative Communication (AAC) technology for people with severe speech impairments. By using built-in capabilities of portable devices, the AI-based algorithm may "understand" dysarthric speech and "translate" it into a fluid real-time user communication, thanks to a "voice donor" outcome system. The pattern classification algorithm is customized for non-standard speech recognition. The AI based system is personalized for each person unique language production and offers a real step forward in AAC efficiency. Earlier empirical findings show limitations in analogic assistive tools addressing Speech, Language, and Communication Needs (SLCNs). Recently, Speech-Generating Devices (SGDs) have been successfully used to support communication in patients with Autism and Dysarthria.

With impressive improvements in recognizing non-standard natural language, AI-based technology (supported by deep learning, big data, and cloud processing) is offering a turning point for personalized Augmentative Alternative Communication (AAC). Upcoming AI-based innovations promise to generate an immense transformative effect on the everyday life of the speech impaired people, their caregivers, significant others, and the entire society.

Keywords: Communication Disorders, Artificial Intelligence, non-standard speech recognition system.

1   Introduction

Speech, language, and communication needs (SLCNs) affect up to 1% of the world population [1]. Recent empirical findings have shown that about 8% of children between 3 and 17 years of age are affected by a communication disorder [2], defined as a deficit of the speech, language, voice quality, or a swallowing problem. A deficit in learning, using, and understanding words may also result in a communication problem. Reducing communication difficulties is fundamental because when a child has a linguistic and communication deficit, the ability to obtain information from the envi-

ronment, to develop cognitive potential, and to interact socially is significantly compromised, with negative consequences on the child's development and behavior [3].

In the era of Information and Communication Technologies (ICTs), Augmentative Alternative Communication (AAC) systems may offer new opportunities to address communication challenges. AAC's primary mission is to "compensate, temporarily or permanently, the patterns of disorder or disability of individuals with severe disorders of communication" (American Speech-Language-Hearing Association [ASHA], 1989). AAC systems provide effective means of communicating and substituting conventional language, representing a real option for people who do not learn the language normally and readily [1, 2]. A systematic literature review [4, 5] has shown the effects of augmented input techniques on communication production (reception, expression, pragmatics, and syntax) in people with developmental disabilities and verbal apraxia. AAC systems can improve child skills in producing single words and multiple symbols sentences. Among the AAC tools, we should mention communication boards [6], handbooks for signs, tangible objects stimulating child creativity and initiatives, and Speech-Generating Devices (SGDs).

2 SGD

Among the AACs, the SGDs have always played a leading role, thanks to their multiple advantages over other assistive communication tools. People with SLCNs often report difficulties in everyday interpersonal interactions [1]: SGDs capabilities may result in a reduction of social distancing. The use of SGDs may apply to various forms of speech disabilities that are secondary to chronic pathologies, such as dysarthria (difficulty in articulating words), apraxia (difficulty coordinating mouth and speech movements), or aphasia (partial or total loss of the ability to communicate verbally or using written words) [7]. Moreover, SGDs can be easily integrated and used in everyday life. Some studies underline that SGDs are preferred over Picture Exchange Communication Systems (PECSs) and sign manuals [4, 8].

The diffusion of new ICTs [9] has made SGDs increasingly efficient and ergonomic [1], primarily when supported by mobile devices such as tablets and smartphones [4]. Recent research shows that SGDs are not simple substitutes for language, but rather a framework supporting the entire communication process by increasing the request-response process [10]. At the present stage, SGDs are among the most effective and efficient AAC tools for addressing numerous diseases [4, 10]. For example, in the case of autism, recent literature has shown that the use of SGDs could have a significant impact on the participant's communication skills, resulting in an increased speech production [11, 12], increased vocabulary (even in non-vocal children) [11], and increased social-relational skills [13]. Furthermore, a significant advantage offered by SGD is that their current communication output mode consists of voice messages.

However, the efficiency and effectiveness of SGDs use remain related to and limited by the user's capabilities [14].

3   Voiceitt: Artificial Intelligence For Non-Standard Speech Recognition

Speech impairments can result in isolation, depression, frustration, sense of inadequacy, lack of self-esteem, and lower quality of life not only for the person living with the disability, but also for those interacting with the disabled person including her/his family, caregivers, friends, and colleagues [15, 16].

A real innovation for Speech-Generating Devices that may overcome the limitations imposed by the users' capabilities and the current obstacles to Speech, Language, and Communication Needs lies in the use of Artificial Intelligence (AI). The term "artificial intelligence" was officially coined by John McCarthy in 1956 and, since then, AI has grown into a structured field of science and engineering. According to Russell & Norvig [17], there are eight prominent AI definitions that can be applied according to the AI context, design, and application.

For example, Artificial Intelligence systems based on standard or non-standard voice recognition can be defined as "Computational intelligence", or "the study of the design of intelligent agents" [18]. In this context, the definition of AI reflects the "rational agent" approach – i.e., machines that can act and behave in "rational" ways and try to achieve the best possible outcome, or the best-expected result [17, 19].

In recent years, AI has been providing innovative approaches and disruptive promises to our society. Rational agents for standard speech-recognition software have become increasingly accurate. Interestingly, after the initial training of the input algorithms, no programming is done by humans to enable Machine Learning (ML) algorithms to perform their tasks [20]. Recent advances in Deep Learning and algorithms development have further improved standard speech recognition systems, thanks to the availability of large datasets, to the increased access to vast computational power, and to the reduced costs required for accessing and storing large amounts of data [20, 21]. Thanks to AI, SGD systems appear more and more performant and ergonomic [1, 11], creating a new phenotype of SGD: communicators who learn from the speaker.

AI system for standard speech recognition (Alexa, Siri, and thousands of voice-activated apps) do not haves non-standard speech recognition capabilities and, for this reason, they are not able to "understand" speech-impaired people.

Recently, Deep Learning speech recognition technology has been designed to understand unique speech impairments, disorders, or disabilities [22]. Voiceitt® is a smart device software application that translates unintelligible sounds into clear speech in real-time. It can foster inclusion and allow the disabled to be more independent by enabling those with a motor or cognitive disability communicate with caregivers, family members, health care professionals, and society as a whole. In this context, Voiceitt® represents an unprecedented AI-based system designed to understand non-standard communication speech (i.e., dysarthria) secondary to congenital or acquired conditions, including Cerebral Palsy, Autism, cerebrovascular accident, Parkinson disease, brain tumor, or Traumatic Brain Injury. Non-standard speech recognition uses different Deep Learning techniques, including pattern classification technology, that is personalized for each speaker. Unlike standard speech recognition, the system is not language-dependent but rather speaker-dependent. It is the only

Augmentative and Alternative Communication (AAC) device that allows people with speech disabilities to communicate with their own voice as input, and to speak through a donor voice technique as output.

Voiceitt® can combine unique statistical modeling and machine learning into its app. In this way speech-impaired people can be understood by anyone, not only caregivers. In fact, it has been observed that while people that do not routinely deal with the speech-impaired person may struggle understanding her/his voice, the family members, the friends, and the caregivers can often understand the disabled with ease, because they have learned how to adapt to her/his unique pattern of speech. From these observations, Artificial Intelligence has been able to construct algorithms replicating this pattern. The app offers a new algorithmic solution for recognizing impaired speech. Each individual user has a distinct phonetic inventory. Adapting to each individual's communication is similar to adjusting the system to a new language. Voiceitt® manages this in a two-stage solution:

- Collect recordings of known utterances: Upon launch of the application, an initialization requires the user to provide a tiny sample of recordings (five words, each repeated twice). The algorithm learns how the user says these specific words and recognizes them when spoken. As a single user continues to use the application in this form, the algorithm steadily learns from the new recordings of the user.
- Apply clustering techniques to the user's phonetic inventory: For example, by collecting a large number of recordings from any one user, the algorithm can run clustering methods on the phonetic data and identify units of sound in the user's unique speech style; this process allows the mapping of the units of sound in standard speech recognition and the application to a more extensive, unlimited vocabulary.
- When used in combination with the Internet of Things (IoT), the app has proven to be effective in increasing autonomous behavior [1]. Speech-impaired people may ultimately control the environment and internet-connected interdigital objects by clear voice-commanded applications [22].

4   Conclusion

Speech impairments can result in isolation, depression, frustration, sense of inadequacy, lack of self-esteem, and reduced quality of life [15, 16, 23].

AI may accelerate progress in serving people with complex communication needs, non-standard speech impairments, and multiple disabilities. Voiceitt® may offer voice-activated means of living. In this way, more autonomous perspectives are foreseen for the disabled. Moreover, upcoming AI-based technological systems combined with convergent applications (Internet of Things, IoT) will offer challenging opportunities and implementing capabilities. Future research should focus on improving the prototype features, increasing the non-standard speech datasets, enlarging the target of possible users, and expanding AI-based solutions converging into IoT. These challenges will lead to new opportunities for the speech-impaired people, including improved chances to communicate and participate within the society. In conclusion,

innovative AI systems are a turning point for the production of unprecedented Augmentative Alternative Communication (AAC) solutions addressing Speech, Language, and Communication Needs in people with speech impairment.

References

1. Elsahar, Y., Hu, S., Bouazza-Marouf, K., Kerr, D., Mansor, A.: Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability. Sensors. 19, 1911 (2019).
2. Black, L.I., Vahratian, A., Hoffman, H.J.: Communication Disorders and Use of Intervention Services among Children Aged 3-17 Years: United States, 2012. NCHS Data Brief. Number 205. Centers Dis. Control Prev. (2015).
3. Hirotomi, T.: An AAC system designed for improving behaviors and attitudes in communication between children with CCN and their peers. In: International Conference on Universal Access in Human-Computer Interaction. pp. 530–541. Springer (2018).
4. Rega, A., Mennitto, A., Iovino, L.: Liar (Language Interface For Autistic's Rehabilitation): Technological Aids For Specialists Supporting The Acquisition Of Verbal Behavior In Persons With Autism. In: 9th International Conference on Education and New Learning Technologies. IATED (2017).
5. Logan, K., Iacono, T., Trembath, D.: A systematic review of research into aided AAC to increase social-communication functions in children with autism spectrum disorder. Augment. Altern. Commun. 33, 51–64 (2017).
6. Bondy, A.S., Frost, L.A.: The picture exchange communication system. Focus autistic Behav. 9, 1–19 (1994).
7. Kerr, D., Bouazza-Marouf, K., Gaur, A., Sutton, A., Green, R.: A breath controlled AAC system. (2016).
8. Gevarter, C., O'Reilly, M.F., Rojeski, L., Sammarco, N., Lang, R., Lancioni, G.E., Sigafoos, J.: Comparisons of intervention components within augmentative and alternative communication systems for individuals with developmental disabilities: A review of the literature. Res. Dev. Disabil. 34, 4404–4414 (2013).
9. Murero, M., Rice, R.E.: The Internet and health care: theory, research, and practice. Routledge (2013).
10. Ricci, C., Miglino, O., Alberti, G., Perilli, V., Lancioni, G.E.: Speech generating technology to support request responses of persons with intellectual and multiple disabilities. Int. J. Dev. Disabil. 63, 238–245 (2017).
11. Wahl, B., Cossy-Gantner, A., Germann, S., Schwalbe, N.R.: Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Glob. Heal. 3, e000798 (2018).
12. Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., Murphy, S., Almirall, D.: Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. J. Am. Acad. Child Adolesc. Psychiatry. 53, 635–646 (2014).
13. Allen, A.A., Schlosser, R.W., Brock, K.L., Shane, H.C.: The effectiveness of

aided augmented input techniques for persons with developmental disabilities: A systematic review. Augment. Altern. Commun. 33, 149–159 (2017).

14. Bourque, K.S., Goldstein, H.: Expanding communication modalities and functions for preschoolers with autism spectrum disorder: Secondary analysis of a peer partner speech-generating device intervention. J. Speech, Lang. Hear. Res. 63, 190–205 (2020).

15. Feeney, R., Desha, L., Ziviani, J., Nicholson, J.M.: Health-related quality-of-life of children with speech and language difficulties: A review of the literature. Int. J. Speech. Lang. Pathol. 14, 59–72 (2012).

16. Zhong, B.-L., Luo, W., Xu, Y.-M., Li, W.-X., Chen, W.-C., Liu, L.-F.: Major depressive disorder in Chinese persons with speech disability: High rates of prevalence and perceived need for mental health care but extremely low rate of use of mental health services. J. Affect. Disord. 263, 25–30 (2020).

17. Intelligence, A.: Artificial Intelligence. Wikipedia Semi-Structured Resour. (2013).

18. Poole, D., Mackworth, A., Goebel, R.: Computational Intelligence. (1998).

19. Iovino, L., Vita, S., Mennitto, A.: " CareMe": a new way to face problem behaviors at school. In: PSYCHOBIT (2019).

20. Parloff, R.: Why deep learning is suddenly changing your life. Fortune. New York Time Inc. (2016).

21. Ark, T. Vander: The Rise of AI: What's Happening, What it Means, How to Prepare?, https://www.gettingsmart.com/2018/03/rise-ai-whats-happening-means-prepare/.

22. Murero, M.: Wearable internet for wellness and health. Geogr. Internet. (2020).

23. Scherbaum, R., Hartelt, E., Kinkel, M., Gold, R., Muhlack, S., Tönges, L.: Parkinson's Disease Multimodal Complex Treatment improves motor symptoms, depression and quality of life. J. Neurol. 267, 954–965 (2020).