

The Role of Customization in Supporting Speech Therapy: the Case of e-SpeechT

Vita Santa Barletta¹, Miriana Calvano¹, Antonio Curci^{1,2}, Rosa Lanzilotti¹ and Antonio Piccinno¹

¹University of Bari "Aldo Moro", Via Edoardo Orabona 4, 70125, Bari, Italy

²University of Pisa, Largo B. Pontecorvo 3, 56127, Pisa, Italy

Abstract

Speech therapy is the medical field in which speech impairments are treated. It mainly concerns the inability of people to adequately enunciate words, to construct and elaborate appropriate sentences when speaking, and generally lack linguistic skills.

This research work presents a software system, called e-SpeechT, supporting speech therapies performed in presence and remotely from both physicians' side and other actors involved in the therapies, i.e., patients, caregivers, and also teachers, schools, etc. The study explores the requirements and functionalities with respect to customization elements and how Artificial Intelligence (AI) can support all parties involved in the process in performing their tasks. The system was designed and developed following the participatory design focusing on the different roles, experiences, and demands of the involved end users and inviting them to contribute from different perspectives.

Keywords

Participatory design, Speech Therapy, Customization, Software engineering, Artificial Intelligence (AI),

1. Introduction

Today, medicine is constantly being exposed to innovative technologies, and new products and services. Many professionals are beginning to trust the use of new instruments to perform their daily activities, such as administering therapies to their patients. This research work focuses on speech therapy and how technology can be a powerful instrument to support all parties involved. More specifically, this case study examines the roles of speech therapists, caregivers, and patients who are children aged 4 to 7 years. Speech therapy refers to the medical field that focuses on the study of linguistic impairments that can negatively influence the personal and professional sphere of an individual, affecting his ability to understand and use spoken language with others [1]. However, subjects who suffer from speech disorders can be rehabilitated by performing exercises that aim to strengthen the muscles of the face and throat, as well as their cognitive skills [2]. The traditional approach to speech therapy, without the aid of technology, is characterized by limitations in multiple aspects, such as accessibility,

Proceedings of the 8th International Workshop on Cultures of Participation in the Digital Age (CoPDA 2024): Differentiating and Deepening the Concept of "End User" in the Digital Age, June 2024, Arenzano, Italy

✉ vita.barletta@uniba.it (V. S. Barletta); miriana.calvano@uniba.it (M. Calvano); antonio.curci@uniba.it (A. Curci); rosa.lanzilotti@uniba.it (R. Lanzilotti); antonio.piccinno@uniba.it (A. Piccinno)

🆔 0000-0002-0163-6786 (V. S. Barletta); 0000-0002-9507-9940 (M. Calvano); 0000-0001-6863-872X (A. Curci); 0000-0002-2039-8162 (R. Lanzilotti); 0000-0003-1561-7073 (A. Piccinno)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

feasibility, and personalization. The introduction of e-health becomes crucial because it has the goal of overcoming the challenges of patients and physicians when dealing with medical issues, gaining benefits in terms of costs and time resources [3].

In the current scenario, there are systems and mobile applications developed to support patients to perform therapy's activities and therapists to administer the treatment. For instance, *Happi Scrive*¹ and *KidEWords*² are two crossword puzzle applications for children used for the development of writing skills through an attractive graphical interface. Another example is *Teach and Touch*³ which is an application designed to support the speech therapist when administering the therapy; the child is allowed to perform a personalized rehabilitation path on specific morphosyntactic difficulties.

In this context, a new web application integrating AI and customization functionalities, called *e-SpeechT*, can be introduced.

The integration of Artificial Intelligence (AI) in this scenario provides new solutions based on the needs and requirements of patients and therapists by not only providing functionalities that allow to automate tasks and reduce the cognitive demand of repetitive activities, but also by supporting the management and evolution of the system [4, 5]. In particular, patients can be able to perform therapies in the comfort of their homes, while allowing continuous monitoring and control from therapists to obtain a broader and more complete understanding of their patient's medical situation [6, 7].

Customization in e-health has the potential of revolutionizing speech therapies promoting inclusion and the establishment of a symbiosis between humans and machines.

This paper presents and describes a web-application, called the e-SpeechT, created to support the actors involved (e.g. patients, caregivers and therapists) while performing, administering, and managing therapy; through the integration of AI techniques these activities can be automated and the system can be further adaptable to the patient's performance.

This solution aims to improve the effectiveness and accessibility of speech rehabilitation, showing the requirements and architectural components of e-SpeechT and illustrating how artificial intelligence is integrated into it.

2. e-SpeechT Design

e-SpeechT was developed according to the SCRUM framework, as an agile project; the framework proved to be powerful and useful, especially due to its iterative nature and flexibility since it allows one to meet user needs efficiently and effectively [8]. Speech therapists were involved throughout the process, employing the Human-Centered Design (HCD), which allowed to obtain rich insights into their needs and preferences to create compliant solutions [9]. The ultimate objective is to create an interactive system that can be customizable and tailorable by end users as domain experts in the design team [10].

¹<https://apps.apple.com/it/app/happi-scrive/id464675842>

²<https://apps.apple.com/it/app/kidewords-by-chocolapps/id879490139>

³<http://www.teachandtouch.it/>

2.1. Requirements

User requirements are fundamental when designing or developing any kind of software; they describe the functionalities the final system will have and provide guidance regarding the priority of tasks and activities to carry out throughout the process.

For e-SpeechT, a number of actors were identified: speech therapist, caregiver, patient, school, teacher, and others to encompass the multi-faceted roles that end users can play, but for sake of simplicity, only the three main actors are here represented. The requirements were grouped in three different categories, one for each type of the main actors, and elicited through multiple interviews involving professionals. Through their collaboration in every sprint of the SCRUM process, it was possible to understand their preferences, priorities and needs, as well as those belonging to caregivers and patients. Some of the main requirements and the corresponding functionalities, that were identified for each role, are listed below and are represented in Figure 1.

Speech therapist Speech therapists are professionals who organize, manage and administer therapies to children; they also have the important duty of communicating and helping caregivers, such as parents, to help their children throughout the process [11]. In particular, the application must allow therapists to:

- Create and manage diagnoses for their patients.
- Share words and exercises with other professionals who use it.
- Manage appointments and schedule them, even in their office.
- Verify the results of a therapy performed by a patient.
- Be supported by an automatic correction feature to reduce their workload, while allowing them to change the decisions they made to keep professionals in control.
- Send notifications to patients about the status of their therapy.
- Automatically classify new patients according to the level of severity of their impairment.
- Change the criteria based on which the patients are classified by the system.

Caregiver Caregivers are the intermediaries between therapists and patients. This is a case of multi-tiered proxy design problems in which the end user, who must be able to tailor and customize the system, and the caregiver do not coincide with the actual end user of the system, i.e. the child [12]. Caregivers have to follow them and ensure that their children perform all assigned tasks, as well as accompanying them to appointments. In particular, the application must allow caregivers to:

- Start and stop the child's execution of a therapy.
- Manage the way the system is presented to the child, in terms of wallpapers, scenarios, and graphical elements.

Patient Patients are the target of therapies and are affected by speech impairments of different degrees of severity. Their age ranges from 4 to 7 years old. The requirements collected for their personal area of e-SpeechT state that the system must:

- Display playful and gaming aspects to avoid feelings of stress or frustration.
- Include gamification, encompassing game mechanics such as rewards and playful feedback.
- Allow the child to log in without the aid of a professional or caregiver.

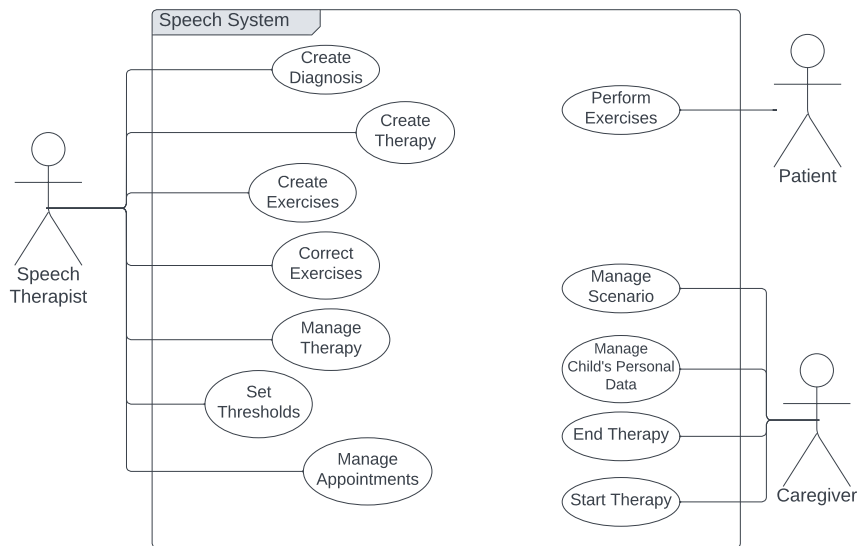


Figure 1: Use Cases Diagram representing the actors and the respective functionalities.

3. Customizing the e-SpeechT Behavior

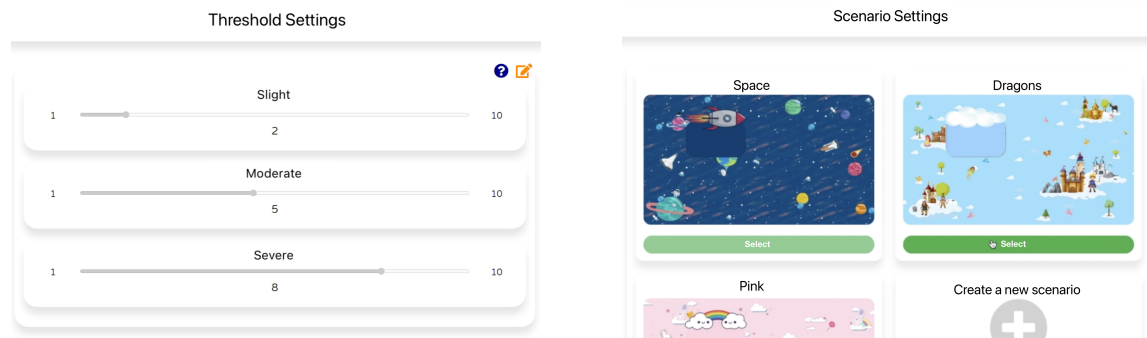
After eliciting the requirements, creating the use cases, and carrying out the preliminary phases of the Human-Centered Design (HCD), e-SpeechT was designed and developed. This section illustrates how customization plays a pivotal role in the web application and how it aims to guarantee the user-specific solutions to health problems, carry out the right diagnoses and avoid stress, boredom and frustration. e-SpeechT embodies End-User-Development (EUD) in multiple aspects and functionalities, which differ depending on the type of user involved.

3.1. Speech-Therapist's Participation

The e-SpeechT application allows different types of intervention, from customization to EUD activities, to personalize its behavior. From the Speech Therapist's side, the application allows to create new examples to evaluate their patients, since each word is characterized by peculiarities that result in different outcomes in exercises; it is also possible to assign images to words with no constraints, meaning that therapists can pick images that bring continuity to the work done during on-site appointments. Exercises can be created and designed based on patients' needs, diagnosis and level of impairment; they belong to one of the three categories - i.e., *naming images, minimum pair recognition and repetition of words*

In addition, therapists are free to create and modify therapies depending on the patient and their diagnosis by manipulating attributes like exercises to be administered, beginning and end date and date, time and place of the appointment. Therapies can contain one or multiple exercises, which can be administered individually or through series.

e-SpeechT offers an exercise automatic correction functionality, which exploits machine learning and speech recognition algorithms in order to facilitate the process of correcting exercises from the therapist’s standpoint. They are used to adapt how the system behaves according to specific groups of patients. Thresholds are used to adjust the level of strictness of the system when it comes to automatically correct exercises during the evaluation of the patient’s vocal interaction; three categories of thresholds are available, which depend on three levels of severity of impairments: *Moderate*, *Severe*, and *Slight*, as shown in Figure 2.



(a) Customization of the thresholds of the automatic correction

(b) Customization of the Patient’s area

Figure 2: Examples of functionalities that include user’s preferences and customization

3.2. Participation of the Caregiver

User-specific personalization is crucial to ensure that patients (children aged from 4 to 7 years old) feel welcomed, comfortable, and not under examination.

In e-SpeechT, scenarios are used to differentiate the style, aesthetic and atmosphere of the patient’s page during the therapy sessions, when carrying out exercises. Caregivers, through EUD activities, are able to modify the whole scenarios by choosing among the ones proposed by the system or creating a new one from scratch. Visual familiarity can prevent children from feeling examined or judged.

Another important aspect of the the system lies in setting a graphical password for the patient, which children are able to log in the application with.

4. Discussion and Future works

This research work presents a web-application, *e-SpeechT*, that aims at supporting therapists, patients, and caregivers during treatments. It describes the design and development process of the application in question, which was performed in accordance with the Human-Centered

Design highlighting its participatory process. In this way, the role of the end users can be further investigated to analyse how they are progressing into active participants during the design and development phases.

In this context, e-health systems that include AI can bring advantages automating tasks performed by therapists and providing patient-specific solutions. Additionally, through the improvement of speech recognition algorithms it would be possible to easily and automatically detect errors in enunciating/repeating words [13]. However, AI can present some limitations such as, for example, the limited context understanding and the risk of overreliance. For this reason, it is necessary to guarantee the right balance between the technology and the human intervention in clinical practice.

Another important aspect is to understand how to guarantee the correct balance between the end users' contribution and how to guarantee that they will be motivated to contribute to the system's performance over long periods of time. Therefore, as a starting point, usability studies were carried out to understand the weaknesses and strengths of the system; as a future work, it is intended to conduct a longitudinal study to test the medical feasibility by gathering information on how children respond to therapies when using the system in question and whether all the involved actors continue to be motivated to use e-SpeechT.

Future work also involve conducting tests to evaluate the application with respect to the WCAG (Web Content Accessibility Guidelines)⁴; this is an important aspect because this application is primarily aimed at individuals possibly affected by disabilities.

Another future work is to expand the application's environment by integrating smart assistants and Internet of Things (IoT) devices. This feature has the objective of automatizing the process of guiding children in their daily activities when having to carry out exercises at home. This project has already started, but is still in its early development stages.

In conclusion, the intersection between speech therapy and technology can revolutionize the approach to deal with speech disorders. However, it is important to underline that the role of AI in this scenario is not a panacea, but an instrument capable of helping users perform their tasks without replacing them.

References

- [1] P. Butcher, A. Elias, R. Raven, J. Yeatman, D. Littlejohns, Psychogenic voice disorder unresponsive to speech therapy: Psychological characteristics and cognitive-behaviour therapy, *International Journal of Language & Communication Disorders* 22 (1987) 81–92. doi:<https://doi.org/10.3109/13682828709088690>.
- [2] V. S. Barletta, F. Cassano, A. Pagano, A. Piccinno, A collaborative ai dataset creation for speech therapies, in: *CEUR Workshop Proceedings*, volume 3136, 2022, pp. 81–85. URL: <https://hdl.handle.net/11586/407013>.
- [3] E. Sillence, L. Little, P. Briggs, E-health, in: *British Computer Society Conference on Human-Computer Interaction*, 2008. URL: <https://api.semanticscholar.org/CorpusID:209022601>.
- [4] V. S. Barletta, D. Caivano, L. Colizzi, G. Dimauro, M. Piattini, Clinical-chatbot ahp evaluation based on “quality in use” of iso/iec 25010, *International Journal of Medical Informatics* 170 (2023) 104951. URL: <https://www.sciencedirect.com/science/article/pii/S1386505622002659>. doi:<https://doi.org/10.1016/j.ijmedinf.2022.104951>.

⁴<https://www.w3.org/WAI/WCAG22/Understanding/understanding-act-rules.html>

- [5] P. Tagde, S. Tagde, T. Bhattacharya, P. Tagde, H. Chopra, R. Akter, D. Kaushik, M. H. Rahman, Blockchain and artificial intelligence technology in e-Health, *Environmental Science and Pollution Research* 28 (2021) 52810–52831. URL: <https://link.springer.com/10.1007/s11356-021-16223-0>. doi:10.1007/s11356-021-16223-0.
- [6] F. Cassano, A. Pagano, A. Piccinno, Supporting speech therapies at (smart) home through voice assistance, volume 483, 2022, p. 105 – 113. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137986566&doi=10.1007%2f978-3-031-06894-2_10&partnerID=40&md5=de8c5b2f0a3ed3f354cb010281ad4004. doi:10.1007/978-3-031-06894-2_10.
- [7] V. Barletta, M. Calvano, A. Curci, A. Piccinno, A new interactive paradigm for speech therapy, in: J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno, M. Winckler (Eds.), *Human-computer interaction – INTERACT 2023*, Springer Nature Switzerland, Cham, 2023, pp. 380–385. doi:10.1007/978-3-031-42293-5_39.
- [8] T. Badriyah, E. N. Fadila, I. Syarif, N. Jauari Akhmad, Rapid development of m-health application with the sprint design approach and scrum process : Application development for e-prescribing, in: *2018 International Conference on Applied Science and Technology (iCAST)*, 2018, pp. 336–342. doi:10.1109/iCAST1.2018.8751603.
- [9] V. S. Barletta, M. Calvano, A. Curci, A. Piccinno, A. Pagano, Poster: Speech therapies and smart assistants: An interaction paradigm proposal, 2023. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85173634861&doi=10.1145%2f3605390.3610823&partnerID=40&md5=62e9ad3f964c7c3df3b04931fd4fd68a>. doi:10.1145/3605390.3610823, cited by: 0.
- [10] M. F. Costabile, D. Fogli, A. Marcante, P. Mussio, L. Parasiliti Provenza, A. Piccinno, Designing customized and tailorable visual interactive systems, *International Journal of Software Engineering and Knowledge Engineering* 18 (2008) 305–325. URL: <http://www.worldscientific.com/doi/abs/10.1142/S0218194008003702>. doi:10.1142/S0218194008003702.
- [11] T. Pring, E. Flood, B. Dodd, V. Joffe, The working practices and clinical experiences of paediatric speech and language therapists: a national uk survey, *International Journal of Language & Communication Disorders* 47 (2012) 696–708. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-6984.2012.00177.x>. doi:<https://doi.org/10.1111/j.1460-6984.2012.00177.x>. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-6984.2012.00177.x>.
- [12] D. Fogli, A. Piccinno, Co-evolution of End-User Developers and Systems in Multi-tiered Proxy Design Problems, volume 7897 of *Lecture Notes in Computer Science*, Springer, Berlin Heidelberg, 2013, pp. 153–168. URL: http://dx.doi.org/10.1007/978-3-642-38706-7_12. doi:10.1007/978-3-642-38706-7_12.
- [13] M. Calvano, A. Curci, A. Pagano, A. Piccinno, Speech therapy supported by ai and smart assistants, in: R. Kadgien, A. Jedlitschka, A. Janes, V. Lenarduzzi, X. Li (Eds.), *Product-Focused Software Process Improvement*, Springer Nature Switzerland, Cham, 2024, pp. 97–104.