Opportunities and Challenges of Designing Assistive Technologies for Aphasia Patients in Singapore: The Case of a Speech Evaluation Prototype

Andreea I. Niculescu, Jochen Ehnes, Minghui Dong, Changhuai You, Paul Chan Yaozhu
A*STAR Institute for Infocomm Research (I^2R)
1 Fusionopolis Way, Connexis South Tower, #21-01, Singapore 138632
{andreea-n, jwehnes, mhdong, echyou, ychan}@i2r.a-star.edu.sg

Abstract

Aphasia is a language disorder caused by brain damage, resulting in difficulties with speaking, understanding, reading, and writing. This study focuses on addressing challenges faced by local therapists in aphasia treatment. Through an ethnographic study involving observations and interviews, critical issues in current technological solutions for aphasia treatment were identified. These issues include the lack of feedback during patient training, limited localized content, repetitive materials, and a lack of options for conversational speech training. In our study, we tried to address one of these issues concerning automatic evaluation of patient speech pronunciation. By utilizing public datasets and local patient data, we developed a system that provides accurate pronunciation scores and personalized feedback, assisting therapists in guiding patient progress. The system supports customized pronunciations, including local accents and dialects. The system is designed for multiple platforms, ensuring accessibility, and can be extended to involve speech therapists to enhance its capabilities. This study emphasizes the importance of integrating research insights with clinical practice, empowering therapists, and enhancing the quality of aphasia treatment.

Introduction

Aphasia, derived from the Greek word *aphatos* meaning 'speechless,' is a condition that occurs due to damage to the left hemisphere of the brain (NIH, 2017). It is found in 21-38% of acute stroke patients and can also result from brain injuries, tumors, or progressive neurological disorders like Alzheimer's (Berthier, 2005). While aphasia causes difficulties in speaking, understanding speech, reading, and writing, it does not affect a person's intelligence (WebMD, 2023). However, the lack of speech production can have long-term effects on cognitive functions as speech and cognition are closely intertwined.

Currently, there is no medication or surgical intervention for aphasia and speech therapy is the only available treatment. Scientific evidence demonstrates that regular speech training can significantly improve the condition of aphasia patients (Roberts et al., 2022). However, therapy can be costly, with limited insurance coverage and additional expenses borne by patients. Financial burdens and a lack of caregivers often lead to patients discontinuing therapy shortly after leaving the hospital. Moreover, Singapore's healthcare system is currently facing a shortage of speech therapists in hospitals and care institutions (Guo, 2014). The therapy provided in the local context is often inadequate for most patients, falling below the necessary intensity for effectiveness.

Given the importance of speech therapy for enhancing patients' quality of life, it is crucial to ensure post-stroke accessibility to this treatment, despite the shortage of speech therapists in the local healthcare system. Possible solutions to these challenges include tele-rehabilitation (Towey, 2012) and interactive computer software for training.

Several software systems are available on the market that offer complementary training for aphasia patients, such as VASTtx (2017), Constant Therapy (2023), Bungalow Software (2021)., AphasiaScripts (2021), and Sentencesharper (2023). These programs provide training in pronunciation, writing, reading aloud, comprehension, memory and problem-solving tasks, voice therapy, word identification, and sentence memorization. Studies have shown that regular use of such programs can significantly aid in the rehabilitation process.

However, these programs have inherent limitations. Cognitive studies have indicated that training limited to a specific context makes it difficult for aphasia patients to generalize their skills to other conversational contexts (8). Although patients may experience improved speech production

proficiency, their communicative fluency often remains restricted to learned dialogue scripts, leading to ongoing difficulties in daily life activities. Additionally, many of these programs lack personalized guidance, localized context, and direct feedback on patients' input.

Field Study - The Human-Human Paradigm

To gain a deeper understanding of the needs within the treatment ecosystem for patients and therapists, we conducted ethnographic field studies that involved analyzing recorded data sessions from Open Aphasia data bank (AphasiaBank, 2017). Additionally, we interviewed speech therapists from Tan Tock Seng Hospital (TTSH) in Singapore.

Typically, therapy is provided by a speech therapist who designs a training program tailored to the patient's most pressing needs. A key element of this therapy involves engaging in conversations centered around specific topics to enhance overall communication skills in different settings. From our observations, we noticed that patients often struggled with organizing words in a proper sequence, constructing grammatically correct sentences, and with finding or pronouncing certain words. Training sessions often focused on storytelling and word category exercises, while therapist feedback regarding word choice and pronunciation proved to be crucial.

During our interviews with local speech therapists from TTSH, we learned that aphasia, a language disorder, is not only prevalent among elderly patients but can also affect younger individuals in their 30s and even 20s. Younger patients often seek therapy to regain their abilities and return to work, as they have a higher potential for recovery. In-patients receive intensive therapy once a day, whereas out-patients typically have sessions 2-3 times a week, although some patients desire additional sessions despite having to pay for them out of their own funds. Sustained improvement usually requires daily training for at least a year.

Local patients receiving therapy possess varying levels of education and proficiency in different languages. Some older individuals or those in middle age may struggle with literacy, encountering difficulties in reading and writing. Therapists conduct formal assessments to establish a baseline of the patients' functional abilities and identify areas of deficit. While there is usually a dominant language, such as Malay, English, Chinese or Indian sub-continent languages, therapists often need to adapt and determine which language works best for the patient. Typically, therapy focuses on one language, and both literacy level and patient proficiency in their dominant language play crucial roles in the treatment process.

Aphasia can impact verbal expression, comprehension, reading, and writing. Therapists primarily prioritizes verbal expression in their work, as it is often a top priority for patients. However, treatment approaches may vary depending on the patient's preferences. Therapy is highly personalized, considering the diverse needs of patients. While speech training is commonly prioritized, it is occasionally combined with writing exercises, including spelling training. However, training writing skills can be challenging if the patient also struggles with written language.

Aphasia patients often experience difficulties in auditory comprehension, struggling to follow conversations involving multiple people. Their dual-task processing abilities appear to be impaired. During therapy, it is crucial to conduct sessions in person, as therapists can slow down their speech, emphasize keywords, and use body language and facial expressions. Conducting such therapy over the phone for example would prove challenging.

Some patients also experience apraxia of speech. In these cases, therapy begins with basic exercises like counting, as counting involves automatic sequences. While swear words are also automatic, therapists avoid using them for treatment for obvious reasons. Instead, patients may practice singing songs that contain familiar phrases in their lyrics. Music therapy allows for slowing down or stopping the music, enabling patients to sing along. For patients without apraxia, therapy usually starts with naming objects relevant to the patient's daily life and gradually progresses to forming phrases or sentences. Mental connections between words and images are often used to trigger speaking. Establishing a core vocabulary for practice is essential, and therapists currently rely on the patient's family to provide a list of words the patient commonly uses as well as to understand patient’s environment.

Keeping patients motivated and engaged is crucial to ensure their commitment to therapy. Therapists aim to use words that patients are familiar with, based on their past activities and interests. Additionally, efforts are made to make therapy more enjoyable and engaging, even considering gamification techniques suitable for elderly patients. However, fatigue and limited cognitive stamina are common issues experienced by individuals with aphasia and therapy sessions need to take these into account.

Another challenge in therapy is the Westernized content used in available applications or training books for aphasia patients. For example, pictures depicting houses covered in snow may be unfamiliar to local individuals, making it difficult for them to relate. Additionally, certain apps like Constant Therapy require users to speak with an American accent, which can lead to inaccurate feedback for Singaporeans who may speak English with local accent.

While there are existing apps available for vocabulary and naming exercises, patients with milder forms of aphasia often desire to construct longer sentences, but they have limited options beyond training with a speech therapist. In such cases, therapists must devise their own resources to help patients determine the grammatical accuracy of sentences and effectively use prepositions and conjunctions.

As patients progress from severe to mild aphasia, their desire to communicate and participate in extended training

sessions intensifies. However, they continue to face difficulties with word retrieval and pronunciation, posing a challenge for therapists. This challenge becomes even more pronounced when therapists are not fluent in the patient's dominant language. A system trained on local languages that predicts patients' intended speech would serve as a bridge between their thoughts and verbal expression. This would not only enhance therapy outcomes but also facilitate everyday communication with others.

Currently, when patients train at home, there is no way to assess their progress, and they may unknowingly make numerous mistakes. Ideally, someone would be present to provide support during home training. Especially, an application that fulfills this role would be helpful, particularly for individuals who live alone. The lack of feedback is a current shortcoming for both patients and therapists as they do not currently have a standardized method to measure progress, relying instead on their notes and experience for regular patient evaluation check-ups.

Speech Evaluation Prototype – The Human – Agent Paradigm

As we embarked on our mission to enhance aphasia treatment, we thoroughly examined the challenges identified during our ethnographic study. Among these challenges, our team chose to focus on developing a technological solution for automatic evaluation of speech pronunciation of patients with mild aphasia who aim to train at home without direct therapist supervision.

To address the critical need for feedback during patients' training, which is currently lacking in existing technological solutions, we have created an initial prototype, as depicted in figure 1. The system uses different ways of prompting, such as text, images, audio and video clips with synchronized lip movement to guide the patients through the therapy

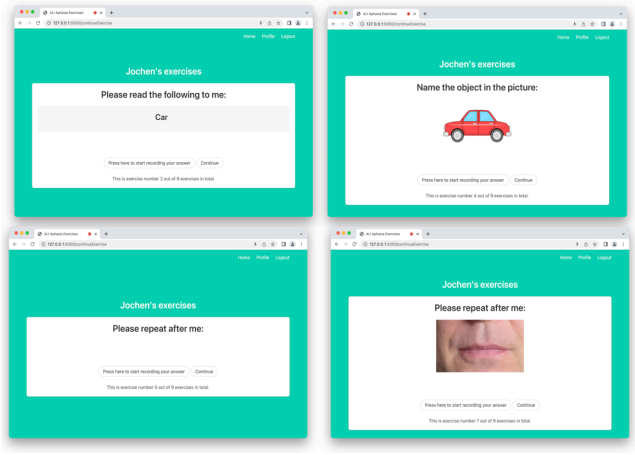

Figure 1: Speech evaluation prototype web implementation

session. This method would help patients with different degrees of aphasia to understand the questions and do their exercises.

When patients speak, the system carefully analyses the pronunciation, generating a corresponding score that allows assessing the accuracy of their speech. Moreover, the system provides specific pronunciation comments, highlighting areas that require improvement. It also accommodates customized pronunciations, taking into account local accents and dialects.

Speech Engine

While patients are guided through a therapy sequence, it remains crucial for a machine to understand their speech. To achieve this, we deployed a speech recognition engine trained with data from healthy speakers. The model used by this engine is derived from a large database of English speakers. The speech recognition model usually demonstrates good accuracy when substantial speech data for a given language is available.

To improve performance for local speakers, we adapt this model using a small set of Singaporean speech data. The speech engine converts speech into text, albeit with some possible errors, repetitions, or pauses as typically encountered in aphasia affected speech. This initial conversion provides a general understanding of what patients are saying, enabling a basic evaluation score by comparing the content with the target answer. We have established rules to identify various types of errors and convert them into scores.

Speech Evaluation

For a more detailed evaluation of pronunciation accuracy, we employ a pattern-based evaluation method called Dynamic Time Warping (DTW) (Senin, 2008). This method involves comparing the patient's speech with existing templates, calculating a score. Multiple templates of the same text are utilized to encompass enough variations. Notably, this method applies to any dialect or language, making it particularly suitable for low-resource languages. This proved to be beneficial for our study as collecting data from local aphasia patients was sparse and relatively challenging to obtain.

Content Generalization

While our current solution effectively handles exact phrase pronunciation, challenges arise when patients need to name objects based on pictures, where the precise utterance may not always match the expected answer. For example, if the target word is "Car," and a patient describes the picture as "red car" or "cartoon car," both of which are correct descriptions, they would be marked as incorrect due to deviation from the target response.

To address this issue, we are currently developing a solution that offers multiple possible descriptions of the image

as evaluation options. The system then compares the recognized speech to all the provided descriptions, providing an evaluation based on the best match. This approach significantly reduces the number of erroneous evaluations without resorting to more complex methods like language models.

In our pursuit of enhancing treatment options for aphasia patients, we recognize the potential of large language models (LLM) as a promising area for further research. These models offer greater flexibility, enabling patients to express themselves more freely. We are exploring options to incorporate and test LLMs in the next version of our prototype.

Our system is designed to be platform-agnostic, supporting iOS, Android, Windows, and Mac, ensuring accessibility across a wide range of devices. It also caters to multiple patients and allows for future extensibility, including the involvement of speech therapists to enhance its capabilities.

Conclusions

Our study aimed to address the challenges faced in aphasia treatment by therapists in Singapore. Through an in-depth ethnographic study, we identified the need for feedback, personalized guidance, and localized content to enhance therapy outcomes.

Our first step in trying to develop a technological solution targeted the development of an automatic evaluation system for speech pronunciation.

The developed prototype successfully analyzed pronunciation, providing accurate scores and targeted feedback to patients, while accommodating diverse accents and dialects. The inclusion of data collected from local speakers allowed for model fine-tuning, ensuring the system's effectiveness and applicability to the local context while the pattern-based evaluation method DTW proved to be beneficial for calculating the pronunciation evaluation scores.

The system's design as a web service will support scalability and allow access from multi-platforms. Furthermore, this architecture will enable therapists to supervise patients remotely and follow their progress after they leave the hospital.

Moving forward, future research will explore the integration of language models to enable more flexible and natural communication for patients. Additionally, ongoing efforts will focus on expanding the system's functionalities, incorporating additional therapeutic components mentioned by therapists during the interviews, and conducting rigorous clinical trials to validate its efficacy in improving aphasia treatment outcomes.

Ultimately, our work strives to empower individuals with aphasia, providing them with accessible and effective tools to regain their language abilities, enhance their quality of life, and promote social inclusion. By merging technology and ethnographic insights, we can make significant strides towards transforming aphasia treatment and improving the overall well-being of affected individuals.

Acknowledgments

We would like to express our gratitude to our clinical partners Kwong Seh Ling and Michelles Boo, as well as to our students Victoria Syn Hui En, Jodie Julian Liaw Kai for their help during our field studies

References

AphasiaBank.https://aphasia.talkbank.org/. Accessed: 2023-06-16

AphasiaScripts. https://www.sralab.org/aphasiascripts. Accessed: 2023-06-16

Berthier, M.L. 2005. Post-stroke aphasia; Epidemiology, pathophysiology and treatment. *Drugs and Ageing* 22(2): 163-182. DOI 02/651000025t2-200522020-00006.

Bungalow Software. Recover Speech and Language. https://www.bungalowsoftware.com. Accessed: 2023-06-16

Constant Therapy. https://constanttherapyhealth.com/constant-therapy/. Accessed: 2023-06-16

Guo, Y.E., Togher, L., Power, E. 2014. Speech pathology services for people with aphasia: what is the current practice in Singapore? *Disability and Rehabilitation* 36(8): 691-704. DOI: 10.3109/09638288.2013.804597

NIH. National Institute on Deafness and other communication disorders. Aphasia. https://www.nidcd.nih.gov/health/aphasia. Accessed: 2023-06-16

Roberts, S., Bruce, R.M., Lim, L., Woodgate, H., Ledingham, K., Anderson, S., Lorca-Puls, D.L., Gajardo-Vidal, A., Leff, A.P., Hope, T.M.H., Green, D.W., Crinion, J.T., Price, C.J. 2022. Better long-term speech outcomes in stroke survivors who received early clinical speech and language therapy: What's driving recovery? *Neuropsychological Rehabilitation* 32(9): 2319–2341. DOI: 10.1080/09602011.2021. 1944883.

Sentence Shaper. https://sentenceshaper.com. Accessed 2023-06-16

Senin, P. 2008. Dynamic time warping algorithm review. CSDL Technical report.

Towey, M.P. 2012. Speech Therapy Telepractice for Vocal Cord Dysfunction (VCD): MaineCare (Medicaid) Cost Savings. *International Journal Telerehabilitation* 4(1): 33-36. doi: 10.5195/ijt.2012.6095

WebMD. What is aphasia? https://www.webmd.com/brain/aphasia-causes-symptoms-types-treatments. Accessed 2023-06-16

VAStxt Sample. Aphasia Software Finder. https://www.aphasiasoftwarefinder.org/vasttx-therapy-samples. Accessed 2023-06-16

Sentence Sharper. An Innovative approach to language software for aphasia. https://sentenceshaper.com/. Accessed 2023-06-16