Problem Statement and Goals
Software Engineering

Team 13, Speech Buddies
Mazen Youssef
Rawan Mahdi
Luna Aljammal
Kelvin Yu

Table 1: Revision History

Date	Developer(s)	Change
Sept 18	Mazen	Drafted & Reviewed the Problem Statement/Goals
Sept 22	Mazen	Made Changes based on team's Feedback and filled Appendix
Sept 28	Rawan	Added Citations to problem statement based on peer review

1 Problem Statement

1.1 Problem

Individuals with speech impairments, such as those caused by neurological conditions (e.g., Parkinson's disease, stroke, or cerebral palsy), often struggle to interact with modern digital devices [1]. Though there have been recent advancements in state-of-the-art Automatic Speech Recognition (ASR) models, widely used models like Whisper and wav2vec show significantly reduced performance on atypical speech [2]. The lack of accurate tools for this population limits their independence, restricts accessibility, and creates barriers to everyday tasks such as sending messages, browsing the internet, or operating applications. A solution is needed that can reliably interpret impaired speech and seamlessly integrate with existing devices to empower users with greater autonomy.

1.2 Inputs and Outputs

At a high level, the system will take speech audio input from users with impaired or slurred speech, captured through a microphone. The input is then processed by a custom-trained machine learning model designed to handle atypical speech patterns. The system will generate text output representing the user's intended words. This text is further mapped into commands or actions that can be executed on a phone or computer interface, allowing the user to control applications, send messages, or perform other digital tasks without requiring standard speech clarity. By bridging the gap between impaired speech and device interaction, the system transforms inaccessible voice commands into actionable outcomes.

1.3 Stakeholders

The primary stakeholders are individuals with speech impairments, such as those with Parkinson's disease, stroke-related conditions, cerebral palsy, or other disorders that affect articulation. These users directly benefit from improved accessibility and independence in controlling digital devices.

Secondary stakeholders include:

- Caregivers and family members, who gain relief knowing their loved ones can communicate and operate devices more independently.
- Healthcare professionals and speech therapists, who can recommend or integrate this technology into treatment plans to support rehabilitation and daily functioning.
- Accessibility advocates and organizations, who seek inclusive technology solutions to close the digital accessibility gap.
- Software developers and device manufacturers, who may integrate this system into their platforms to expand accessibility features.

1.4 Environment

The system is expected to operate on consumer-grade devices such as smartphones, tablets, or personal computers, equipped with a standard microphone for capturing speech input. It must function in everyday environments where background noise may be present, such as at home, in clinics, or in public settings.

The software environment will include:

- A machine learning backend trained on impaired speech datasets, capable of running locally or through cloud-based services.
- A user interface layer on the device to display the transcribed text and execute mapped commands.

- Integration with existing operating system accessibility frameworks to enable device-level control (e.g., opening apps, sending messages, browsing).

The solution must be designed to be lightweight and user-friendly, requiring minimal setup, so that individuals of varying technical ability can use it independently.

2 Goals

The following are goals that VoiceBridge must achieve in order to be considered acceptable and complete.

Table 2: Minimum Viable Product Goals

Goal	**Explanation**	**Reasoning**
Accurate Transcription	The system must reliably convert impaired or atypical speech audio into text output with measurable accuracy (70%+ translation accuracy).	Without accurate transcription, the product fails its core purpose of enabling users to communicate.
Real-Time Processing	The system should process speech input and generate text/commands with minimal delay (<1 second delay).	Enables natural interaction with devices, making the system practical for everyday use.
Usability of Input	The product must accept speech input through a standard microphone without requiring specialized equipment.	Keeps the solution accessible and cost-effective, lowering barriers for adoption.
Non-Disruptive Feedback	The transcribed text should be displayed clearly to the user for verification.	Transparency builds user trust and allows them to correct errors if necessary.

3 Stretch Goals

The following are goals that are not necessary for the product to be viable, but will provide added benefits should they be achieved.

Table 3: Stretch Goals

Goal	**Explanation**	**Reasoning**
Device Control Integration	Transcribed text should be mapped to commands that can interact with applications on a phone or computer.	Extends the system from communication-only to full device control, empowering users with independence.
Personalization	Allow users to customize command mappings (e.g., "open browser" or "send text").	Personalization increases usability and adapts the system to individual needs and preferences.
Robustness to Noise	System should handle moderate background noise in typical environments (home, clinic, public).	Ensures reliability in real-world conditions, not just in quiet test settings.
Autonomy	Once activated, the system should run with minimal user intervention.	Reduces reliance on caregivers and maximizes independence for the user.

4 Extras

- Literature Review Report
- Performance Report

Appendix — Reflection

The purpose of reflection questions is to give you a chance to assess your own learning and that of your group as a whole, and to find ways to improve in the future. Reflection is an important part of the learning process. Reflection is also an essential component of a successful software development process.

Reflections are most interesting and useful when they're honest, even if the stories they tell are imperfect. You will be marked based on your depth of thought and analysis, and not based on the content of the reflections themselves. Thus, for full marks we encourage you to answer openly and honestly and to avoid simply writing "what you think the evaluator wants to hear."

Please answer the following questions. Some questions can be answered on the team level, but where appropriate, each team member should write their own response:

1. What went well while writing this deliverable?

   We quickly aligned on the core idea and broke it down into a clear, consolidated problem statement with agreed-upon goals. Using the template kept us focused on abstraction (inputs/outputs, stakeholders, environment) and made it easy to divide.

2. What pain points did you experience during this deliverable, and how did you resolve them?

   No major pain-points as this deliverable had a small scope.

3. How did you and your team adjust the scope of your goals to ensure they are suitable for a Capstone project (not overly ambitious but also of appropriate complexity for a senior design project)?

   We narrowed the scope by defining a concrete MVP and separating stretch features. The MVP focuses on reliable transcription for impaired speech with basic command mapping and clear feedback, while stretch goals (e.g., broader device control, real-time latency targets, personalization, noise robustness) are aspirational, not required for success.

References

References

[1] National Institute on Deafness and Other Communication Disorders. (2024) Assistive devices for people with hearing, voice, speech, or language disorders. [Online]. Available: https://www.nidcd.nih.gov/health/assistive-devices-people-hearing-voice-speech-or-language-disorders

[2] N. Pokel, P. Moure, R. Boehringer, S.-C. Liu, and Y. Gao, "Variational low-rank adaptation for personalized impaired speech recognition," *arXiv preprint arXiv:2509.20397*, 2025. [Online]. Available: https://arxiv.org/abs/2509.20397