AGAM AI LABS

Empowering Industry with AI Vision

Voice Conversion (Speaker A → Speaker B)

Preserve Linguistic Content, Convert Speaker Identity

Company:	Agam AI Labs
Assigned to:	Mohan K
Start Date:	16/09/2025
End Date:	22/09/2025

September 15, 2025

Contents

1 Introduction

The assigned task is to build a system that takes an audio clip of **Speaker A** and converts it so that it sounds like **Speaker B**, while preserving the original linguistic content (the words spoken). The solution must achieve high quality, naturalness, and intelligibility.

2 Task Scope

- Implement a voice conversion system that disentangles **content** and **speaker identity**.
- The system should generate speech that has the words of Speaker A, but the vocal characteristics of Speaker B.
- Participants may choose any models, architectures, or approaches they find suitable.
- Only conventional signal processing, machine learning, or deep learning methods are allowed. **No LLM usage is permitted for model design, training, or evaluation**.

3 Example Pipeline (for reference only)

Below are possible steps for inspiration. The final approach is up to the assignee.

1. **Content Extraction:** Separate linguistic features (e.g., PPGs, ASR embeddings).
2. **Speaker Embedding:** Obtain representation of Speaker Bs timbre, pitch, and style.
3. **Voice Conversion:** Combine content with target speaker embedding to generate new features.
4. **Vocoder:** Use a neural vocoder (e.g., HiFi-GAN, WaveGlow) to produce waveform.

4 Evaluation Metrics

- **Content Preservation:** Word Error Rate (WER) using ASR on converted speech.
- **Speaker Similarity:** Embedding similarity (cosine similarity) and subjective Mean Opinion Score (MOS).
- **Naturalness:** MOS rating for fluency, prosody, and quality.

5 Deliverables

Deliverable	Description	Due Date
System implementation	End-to-end code for training and inference.	22/09/2025
Documentation	Report covering chosen approach, architecture, and evaluation.	22/09/2025
Sample outputs	Converted audios (Speaker A → Speaker B).	22/09/2025
GitHub repository	Full codebase with README, requirements, and instructions.	22/09/2025

Table 1: Deliverables and Timeline

6 Contact / Questions

Project owner: **Naveen V**.

For clarifications, reach out to the project owner.

This document serves as the official specification and guideline for the assigned task. ***No LLM usage is permitted.***