

SAMPLE PROVIDED TO THE A.I.	
CURRENT PROBLEMS	POTENTIAL IMPROVEMENTS
Had multiple languages, Hindi and English.	Should have single language, preferably English.
It was a natural talking tone.	We need a narration tone.
Ups and downs in the speed of speech.	Variation in speed is okay, but should not be too frequent and drastic.
Poor quality recording, with too much background noise.	Better quality recordings would improve the outputs.
Small sample size, only 4 to 5 minutes.	Need 30 minutes, at least, but 3 hours would be ideal.

Problems with outputs:

- We do not have absolute control over the modulation. There is only one “style exaggeration” slider, which can give unpredictable and unstable outputs. Learning to do all this manually would take days.
- Text-to-speech generations could feel lifeless, and we do not have anyone to match his style and tone to generate a speech-to-speech voice.

Currently, we have Eleven Lab’s Starter plan (\$5), but to create a “professional voice clone”, where we will provide a larger, clearer voice sample, we need their Creator plan (\$11 for first month, \$22 after that). I believe that going for the Creator plan, we can achieve better results. However, I cannot guarantee that this will be absolutely perfect.