

# Pseudocode for Real-Time Audio Translation System

Xuhui Zhan, Siyu Yang

April 15, 2024

---

**Algorithm 1** Real-Time Audio Translation System

---

```
1: Initialize:
2: Load WhisperModel                                ▷ Speech to text model
3: Load HelsinkiNLPModel (zh-en or en-zh)            ▷ Translation model based on
   target language
4: Load XTTSV2Model                                  ▷ Text-to-speech model
5: Capture and Process Audio:
6: Capture complete audio from microphone              ▷ Record full session
7: Save audio as original_audio.wav                   ▷ Save original audio
8: transcribed_text ← WhisperModel.transcribe(original_audio.wav) ▷
   Transcribe audio
9: Save transcribed_text to file as transcription.txt
10: if target_language is English then
11:   translated_text ← HelsinkiNLPModelZH-
   toEN.translate(transcribed_text)                    ▷ Translate to
   English
12: else if target_language is Chinese then
13:   translated_text ← HelsinkiNLPModelEN-
   toZH.translate(transcribed_text)                    ▷ Translate to
   Chinese
14: end if
15: Save translated_text to file as translated.txt      ▷ Save translated text
16: speaker_embedding ← XTTSV2Model.create_speaker_embedding(original_audio.wav)
   ▷ Create voice profile
17: audio_output ← XTTSV2Model.synthesize(translated_text,
   speaker_embedding)                                  ▷ Generate speech
18: Save audio_output as output.wav                    ▷ Save synthesized speech
19: Play audio_output                                  ▷ Output translated audio
```

---