Voice Translation App: Functional Documentation

# Overview

The Voice Translation App is a mobile-responsive web application designed to enable real-time voice translation. It provides speech-to-text transcription, language translation, and text-to-speech synthesis to facilitate communication across different languages. Built using ReactPy, the app features an intuitive, mobile-first interface with support for selecting predefined or custom languages, displaying transcriptions and translations, and playing translated audio. This documentation details the app's core functionalities and UI elements for developers, stakeholders, and technical users, focusing on its capabilities without including a user guide.

# Key Functionalities

## Language Selection

- \*\*Input Language\*\*: Users can select the language of the spoken input from a predefined list (e.g., English, Spanish, French, German, Chinese) via a dropdown menu. Alternatively, users can type a custom language to accommodate unsupported languages or dialects.

- \*\*Output Language\*\*: Users choose or type the target language for translation, using the same predefined list or custom input option.

- \*\*Implementation\*\*: Two dropdown selectors in a top bar allow language selection. The UI supports custom input through text fields within the dropdown, with responsive design for mobile (vertical stacking) and larger screens (horizontal alignment).

## Audio Recording and Transcription Display

- \*\*Recording Mechanism\*\*: A floating action button (FAB) toggles audio recording. It displays a microphone icon (🎙️) when idle and a stop icon (⏹️) during recording, changing color from purple to red for visual feedback.

- \*\*Transcription Output\*\*: After stopping the recording, the app displays the transcribed text of the input audio in the selected input language. The transcription appears in a card with a "Transcript" header and a scrollable text area for lengthy content.

Backend Processing\*\*: Transcription leverages server-side processing with AI-enhanced accuracy, optimized for medical terminology.

## Translation Display

- \*\*Translation Process\*\*: A "Translate" button triggers translation of the transcribed text into the selected output language.

- \*\*Display Format\*\*: The translated text is shown in a separate card labeled "Translation," with a scrollable area for long text. The stacked card layout allows simultaneous viewing of the original transcript and translation.

- \*\*Real-Time Feedback\*\*: Translation results are updated dynamically upon user action, maintaining a clean and accessible presentation.

## Audio Playback

- \*\*Text-to-Speech\*\*: A "Play Translated Audio" button converts the translated text into audio for playback.

- \*\*Audio Player\*\*: An HTML5 audio player appears below the translation card, offering standard controls (play, pause, volume, seek).

- \*\*Seamless Integration\*\*: Audio is generated on-demand and played directly in the browser, ensuring a smooth user experience.

# Version Information

- \*\*Date\*\*: August 20, 2025

- \*\*Status\*\*: Based on the latest implementation with a mobile-first, responsive UI and server-side

This documentation focuses on the app's functionalities and design, omitting usage instructions as requested. For implementation details or troubleshooting, refer to the source code or contact the developer.