Speech-to-Text APIs Comparison

1. Google Cloud Speech-to-Text

✔️Advantages

- Very high accuracy (90–95% for clear audio)
- Supports many languages
- Automatic language detection
- Strong noise-handling and Indian accents

✖️Disadvantages

- Free tier limited
- Requires Cloud setup + billing
- Implementation slightly complex

⬛Pricing (Approx)

- First 60 mins free / month
- After that: $0.006 – $0.012 per minute depending on model

🆎Accuracy

- English: 95%
- Hindi: 90–93%
- Mixed language: Medium accuracy

🗔How to Use

1. Create Google Cloud account
2. Enable Speech-to-Text API
3. Generate API Key / Service Account
4. Send audio file to `/speech:recognize` endpoint

---

2. AssemblyAI

✔️Advantages

- Automatic language detection
- High accuracy for long audios
- Easy API usage
- Free trial available

✖️Disadvantages

- Free tier limited
- Indian accent accuracy average

⬛ Pricing

- $0.012 per minute
- Free trial credits available

🅰️🅱️ Accuracy

- English: 92–95%
- Hindi: 85–90%
- Mixed language: Low–Medium

🪟 How to Use

1. Create account on assemblyai.com
2. Get API key
3. Upload audio to upload endpoint
4. Create transcript request with `{ language_detection: true }`

---

3. Microsoft Azure Speech Service

✔️ Advantages

- Very high accuracy
- Best in noisy environments
- Supports many languages
- Real-time recognition

✖️ Disadvantages

- Setup is complex
- Free tier limited

⬛ Pricing

- First 5 hours free
- After that: ~$1 per hour

🅰️🅱️ Accuracy

- English: 95%
- Hindi: 90–94%
- Mixed: Medium

🪟 How to Use

1. Create Azure account
2. Enable Speech resource
3. Get API key + region
4. Use `/speech/recognition` endpoint

---

4. OpenAI Whisper (Open Source)

✔️Advantages

- Fully free (self-host)
- Very high accuracy in Indian languages
- Works offline also

✖️Disadvantages

- Needs powerful server / GPU
- Setup required
- No language detection (model-specific)

◼️ Pricing

- Free if self-host
- Hosting cost depends on server

🆎Accuracy

- English: 95%
- Hindi: 94–97%
- Mixed Hinglish: 90–95% (best among all)

🗂️ How to Use

1. Install Whisper via Python
2. Run `whisper audio.mp3 --model large`
3. Get text output

---

5. LibreTranslate (NOT Speech-to-Text)

✔️Advantages

- Fully free if self-hosted
- Translation supported

✖️Disadvantages

- **Does NOT support speech-to-text**
- Only text translation
- Accuracy depends on model

◼️ Pricing

- Free (self-host)

Accuracy

- Depends on language pairs

How to Use

- Send text → receive translated text

🔲Final Recommendation (Based on Your Requirement)

If you want best accuracy for Hindi + English + Mixed languages:

➡ **OpenAI Whisper (Self-hosted)**

If you want easiest API integration:

➡ **AssemblyAI**

If you want highest accuracy but paid:

➡ **Google Cloud Speech-to-Text**

If you want real-time & accurate in noise:

➡ **Azure Speech Service**

Agar chaho to main aapke project ke liye **best API select karke full integration code** (React ya Node) likh kar de sakta hoon.