

## PROBLEM STATEMENT 1: Kannada Voice-to-Text Correction Challenge

### The Challenge

Create an interactive system that takes **your voice in Kannada**, converts it to text, and helps identify confusing characters that might have been misrecognized—giving you options to choose the correct one!

Think of it as building a "smart autocorrect" for Kannada that asks for your help when it's unsure.

### What You'll Build

A simple pipeline:

You speak → System writes it down → System asks "Did I hear ಓ or ಔ?" → You pick → Clean text!

**That's it!** No complex algorithms needed—just connect some existing tools and build a user-friendly way to fix mistakes.

### What You Need to Do

#### STEP 1: Record Yourself Speaking (Most Important!)

Pick **3 different paragraphs** and read them aloud. Each recording should be about **3 minutes long**.

What to Record	Type of Paragraph	How Long
Recording 1	Easy paragraph (simple, everyday Kannada)	3-4 minutes
Recording 2	Medium paragraph (newspaper-style writing)	3-4 minutes
Recording 3	Hard paragraph (bookish, formal Kannada)	3-4 minutes

**Total time you'll spend recording:** About 9 minutes of your voice!

**Save each recording as a .mp4 file** (your phone camera app works perfectly—just point it down at the table and speak!)

#### STEP 2: Convert Voice to Text

Use any free tool to turn your recordings into written Kannada:

- Google Speech-to-Text API (free tier available)
- Azure Speech Services (student account)
- Any other tool you find online

**Save the text output** (even if it has mistakes—that's expected!)

#### STEP 3: Build a "Which Letter?" Checker

Your system should:

**a) Find confusing letters like:**

- ವ್ಯ or ಮ್ಯ? (look similar)
- ವ್ಯ or ಫ್ಯ? (just a dot difference)
- ಒ or ಪ್ಯ? (easy to mix up)
- ಓ or ಓ? (curved shapes)

**b) Ask the user:** "Hey, I wrote ವ್ಯ here. Did you mean ವ್ಯ or ಮ್ಯ?"

**c) Let them click/select the right one**

**d) Move to the next confusing letter**

**e) Show the final corrected text**

That's your whole system!

**STEP 4: Do the Same for Joined Letters (Ottaksharas)**

Some letters stick together in Kannada (like ಕ್ಕಾ, ತ್ತಾ, ನ್ನಾ).

Your system should:

- Spot these joined letters
- Ask: "Is this ಕ್ಕಾ correct, or did it split weirdly into ಕ್ಕ + ಾ?"
- Let the user fix it

**STEP 5: Package Everything**

Save all your files properly (details below in "What to Submit").

**Where to Find Paragraphs?**

**Easy Paragraphs (12-15 sentences):**

- Class 5-7 Kannada textbooks
- Simple Wikipedia articles in Kannada
- Children's story websites

**Example topic:** My daily routine, My school, My city

**Medium Paragraphs (12-15 sentences):**

- Newspaper editorials (Prajavani, Vijaya Karnataka)
- Government announcements in Kannada
- Wikipedia articles on historical topics

**Example topic:** Bengaluru's traffic problem, Festivals of Karnataka

**Hard Paragraphs (12-15 sentences):**

- Kannada literature excerpts (Kuvempu, Bendre poems)
- Legal/technical documents in Kannada
- Academic articles

**Example topic:** Karnataka's cultural heritage, Scientific explanations in Kannada

**BONUS: Add Your Dialect! (Optional but Rewarded)**

Do you speak Kannada differently at home than in textbooks? Maybe you say "ಹೊಗ್ಗಿನಿ" instead of "ಹೋಗುತ್ತೇನೆ"?

If you want extra points, record the SAME paragraphs in your home dialect!

**Common Kannada Dialects:**

- **Bangalore Kannada:** Casual, mixed with English, shortened words
- **Mysore Kannada:** Traditional, uses more Sanskrit words
- **Coastal Kannada** (Mangalore area): Different pronunciations
- **North Karnataka** (Dharwad, Hubli): Influenced by Marathi
- **Havyaka Kannada:** Unique to certain communities

**What to do:**

1. Pick your dialect
2. Record the same 3 paragraphs in that dialect (speak naturally, like talking to family)
3. Note down what's different (e.g., "I said ಹೊಗ್ಗಿನಿ instead of ಹೋಗುತ್ತೇನೆ")

**Extra points:** +10 to +50 points depending on how well you document it!

**What to Submit**

Create one folder with your team's name. Inside it, organize like this:

YourTeamName\_VoiceChallenge/

```
|  
|   └── recordings/  
|       |   └── easy_paragraph.mp4      ← Your voice (3 min)  
|       |   └── medium_paragraph.mp4    ← Your voice (3 min)  
|       |   └── hard_paragraph.mp4     ← Your voice (3 min)
```

```
|   └── dialect_easy.mp4      ← OPTIONAL bonus  
|  
|  
└── text_outputs/  
    |   ├── easy_original.txt    ← What the speech tool wrote  
    |   ├── easy_corrected.txt   ← After you fixed confusing letters  
    |   ├── medium_original.txt  
    |   ├── medium_corrected.txt  
    |   ├── hard_original.txt  
    |   └── hard_corrected.txt  
|  
|  
└── confusion_fixes/  
    |   ├── easy_fixes.csv      ← Log of what letters you fixed  
    |   ├── medium_fixes.csv  
    |   └── hard_fixes.csv  
|  
|  
└── speaker_info.json      ← Your details (see format below)  
└── README.md                ← How you did it (tools used, steps)  
└── code/                    ← Your program files
```

### Format for speaker\_info.json

Copy this template and fill it in:

```
{  
  "team_name": "XYZ",  
  "student_name": "ABCD",  
  "age": 21,  
  "gender": "Female",  
  "native_kannada_speaker": true,  
  "home_region": "Bangalore",  
  "dialect_used": "Bangalore Kannada",
```

```

"is_dialect_recording_included": false,
"recording_device": "Samsung Galaxy M31 (phone)",
"recording_environment": "Quiet room at home",
"speech_to_text_tool_used": "Google Cloud Speech-to-Text API",
"date_of_recording": "2024-11-16"
}

```

If you did dialect recordings, change "is\_dialect\_recording\_included": true and add:

"dialect\_notes": "Used casual Bangalore style - said ಹೋಗೀನಿ instead of ಹೋಗುತ್ತೇನೆ, dropped some ನು/ಅನ್ನು endings"

### **Format for confusion\_fixes.csv**

Every time your system asks "ಎ or ಏ?" and the user picks, log it like this:

position,word,detected\_letter,alternatives,user\_picked,was\_correct

5,ಮನೆ,ಮು,"ಮ|ಎ",ಮು,yes

18,ಹೋಗುತ್ತಿದ್ದೇನೆ,ತ್ತೆ,"ತ್ತೆ|ತ್ತಾತೆ",ತ್ತೆ,yes

32,ಒರುವ,ಎ,"ಎ|ಏ",ಎ,yes

### **Columns explained:**

- position: Character number in the sentence
- word: The word it appears in
- detected\_letter: What the system thought it heard
- alternatives: Options you showed ("A|B|C")
- user\_picked: What the user chose
- was\_correct: Did user agree with original detection? (yes/no)

### **How You'll Be Judged**

What We're Checking	Points	What This Means
<b>Voice Quality</b>	20 pts	Clear audio, no noise, proper .mp4 format
<b>All Files Present</b>	15 pts	Did you submit everything in the right format?
<b>System Works</b>	30 pts	Does your letter-checker actually run and show options?
<b>Correction Accuracy</b>	20 pts	Did you correctly identify confusing letters?

<b>Documentation</b>	15 pts	Can we understand what you did from your README?
<b>BONUS: Dialect Work</b>	+50 pts	Extra credit if you documented your dialect!

**Total possible:** 100 points (150 with dialect bonus!)

### Which Letters Are "Confusing"?

Your system should check for the **common mix-ups**:

Looks Similar	Why They're Confusing
ಎ ↔ ಏ	Curved shapes are very similar
ಷ ↔ ಷ್ಟ	Just a tiny dot difference
ಒ ↔ ಒ೦	Dot vs small line
ಠ ↔ ಠ್	Both have loops
ಡ ↔ ಡ್	Dot position
ಕ್ಕ ↔ ಕ್ಳ	Small stroke difference
ಇ ↔ ರ್ಯಿ	Dot placement
ತ್ತ ↔ ತ್ಳ	Dot above or not
ಸ್ಸ ↔ ಷ್ಟ	Subtle curve difference
ನ್ನ ↔ ಣ	Bottom part slightly different

and others too!

### For joined letters (ottaksharas):

- ಕ್ಕು vs ಕ್ಕು + ಕು (should be joined!)
- ತ್ತು vs ತ್ತು + ತು
- ನ್ನು vs ನ್ನು + ನು
- ಷ್ಟು vs ಷ್ಟು + ಷು
- **Python + Tkinter** (desktop app)

### For Recording:

- Your phone's camera app (save as .mp4)
- Laptop webcam software

- Free apps like "Voice Recorder" (then convert to .mp4)

### Evaluation metrics

#### VOICE RECORDING QUALITY (20 points)

Criterion	Points	Requirements
<b>Audio Clarity</b>	8 pts	Clear speech, minimal background noise, audible throughout
<b>Recording Format</b>	4 pts	Proper .mp4 format, 44.1kHz+ sample rate
<b>Duration Compliance</b>	4 pts	Each recording 3-4 minutes (Easy, Medium, Hard)
<b>Content Variety</b>	4 pts	Three different difficulty levels properly represented

#### DATASET COMPLETENESS (25 points)

Criterion	Points	Requirements
<b>All Files Present</b>	7 pts	All required folders and files submitted
<b>Text Outputs</b>	6 pts	Original + corrected transcripts for all 3 recordings
<b>Confusion Logs</b>	7 pts	Properly formatted CSV with all corrections logged
<b>Metadata</b>	5 pts	Speaker info JSON complete and accurate

#### SYSTEM FUNCTIONALITY (25 points)

Criterion	Points	Requirements
<b>Speech-to-Text Works</b>	6 pts	Successfully converts voice to text
<b>Confusion Detection</b>	9 pts	Accurately identifies similar-looking characters (ಎ/ಎ್, ಒ/ಒ್, etc.)
<b>Interactive Interface</b>	6 pts	User-friendly UI for selecting correct characters
<b>Correction Application</b>	4 pts	System properly applies user-selected corrections

#### CORRECTION ACCURACY (15 points)

Criterion	Points	Requirements
<b>Character Detection Rate</b>	9 pts	% of confusing characters correctly identified (>85% = full points)
<b>Ottakshara Handling</b>	4 pts	Joined letters detected and corrected

<b>False Positives</b>	2 pts	Minimal incorrect flagging of correct characters
------------------------	-------	--

### DOCUMENTATION QUALITY (10 points)

Criterion	Points	Requirements
<b>README Completeness</b>	5 pts	Clear instructions, tools used, setup steps
<b>Metadata Accuracy</b>	3 pts	Speaker info properly filled, accurate details
<b>Code Comments</b>	2 pts	Well-commented, understandable code

### BONUS: DIALECT WORK (+5 points)

Criterion	Points	Requirements
<b>Dialect Recording</b>	2 pts	Same paragraphs recorded in home dialect
<b>Dialect Documentation</b>	2 pts	Clear notes on differences from standard Kannada
<b>Dialect Examples</b>	1 pt	Specific word/phrase differences identified

### PPT SUBMISSION (30 points)

Criterion	Points	Requirements
<b>Presentation Structure</b>	8 pts	Clear flow: Problem → Approach → System → Results (8-10 slides)
<b>Visual Quality</b>	7 pts	Professional design, proper formatting, readable fonts
<b>Content Clarity</b>	8 pts	Clear explanations, diagrams, screenshots of system
<b>Results Documentation</b>	7 pts	Accuracy metrics, before/after examples, confusion logs summary

#### PPT Must Include:

- Title slide with team details
- Problem understanding (1 slide)
- Approach & pipeline diagram (1-2 slides)
- System screenshots/interface (2-3 slides)
- Results with metrics (2 slides)
- Examples of corrections (1 slide)
- Challenges & learnings (1 slide)
- Conclusion (1 slide)

References -

**Kannada Similar-Sounding & Similar-Looking Letters Table**

**Legend**

- **Similar Sounding** → Letters whose pronunciation is close
- **Similar Geometrical / Similar Looking** → Letters whose shapes are visually confusing

**1. Vowels (ಸ್ವರಗಳು)**

Letter	Similar Sounding	Similar Looking
ಅ	—	ଓ (very close visually sometimes)
ಆ	—	ଓ
ಇ	ಈ	—
ಈ	ಇ	—
ಉ	ಊ	—
ಊ	ಉ	—
ಎ	ಏ	—
ಏ	ಎ	—
ಒ	—	—
ಒ	ಔ	—
ಔ	ಒ	—
ಔ	—	—

**2. Consonants – High Confusion Sets**

**(A) ಕ್ರ Group**

Letter	Similar Sounding	Similar Looking
ಕ್ರ	ಖ್	ಗ್ (shape somewhat similar)
ಖ್	ಕ್ರ	ಗ್ (sometimes)
ಗ್	ಕ್ರ, ಖ್	ಝ್
ಝ್	ಗ್	—
ಜ್	—	—

**(B) ಞ್ ಶ್ Group**

Letter	Similar Sounding	Similar Looking
ಶ್	ಞ್	—

ಂ	ಃ	ಂ (visually similar)
ಃ	ಂ	ಂ
ಜ	ಂ, ಃ	ಂ
ರ್ಯ	ಂ	—
ಞ	—	—

#### (C) ಂ Group

Letter	Similar Sounding	Similar Looking
ಟ	ಠ	ಡ (shape close)
ಠ	ಟ	ಡ
ಡ	ಟ, ಠ	ಡ
ಡ	ಡ	—
ಣ	ಣ	—

#### (D) ಂ Group

Letter	Similar Sounding	Similar Looking
ತ	ಂ	—
ಂ	ತ	—
ಡ	ಂ	—
ಡ	ಡ	—
ಣ	ಣ	—

#### (E) ಂ Group

Letter	Similar Sounding	Similar Looking
ವೆ	ಂ	ಂ (shape similar)
ಂ	ವೆ	ಂ
ಂ	ಂ, ಬೆ	ಂ
ಬೆ	ಬೆ	—
ಮೆ	—	—

#### (F) ಯ, ರ, ಲ, ವ Family

Letter	Similar Sounding	Similar Looking
ಯು	—	ꝝ (slightly similar in handwriting)
ರ	ꝝ (rare)	—
ಳ	ಽ	—
ಽ	ಳ	—
ವೆ	ಒ	—

#### (G) ಗ Family

Letter	Similar Sounding	Similar Looking
ಅ	ಷ್ಟ	—
ಷ್ಟ	ಅ	ಸ್ತ
ಸ್ತ	ಅ, ಷ್ಟ	—
ಹೆ	—	—
ಽ	ಳ	—

### 3. Comprehensive Confusion Pairs List (Quick Reference)

#### Similar Sounding

- ಕೆ ↔ ಒ
- ಗೆ ↔ ಫೆ
- ಚೆ ↔ ಟೆ
- ಚೆ ↔ ಜೆ
- ಟೆ ↔ ರೆ
- ಡೆ ↔ ಡೆ
- ತೆ ↔ ಧೆ
- ದೆ ↔ ಧೆ
- ಬೆ ↔ ಫೆ
- ಒ ↔ ಜೆ
- ಉ ↔ ಽ
- ಅ ↔ ಷ್ಟ ↔ ಸ್ತ

#### Similar Looking (Geometrical)

- ಅ ↔ ಏ
- ಕೆ ↔ ಗೆ

- ಜೆ ↔ ಜ
- ಟು ↔ ಡೆ
- ಪೆ ↔ ಬೆ
- ಫೆ ↔ ಬೆ
- ಷೆ ↔ ಸೆ

Links -

1. [Dictionary](#)