

AI-Powered Public Speaking Training Platform: Game Implementation

Game 1: Rapid Fire Analogies – Workflow Steps

Round Selection & Timer Setup User selects the round (1, 2, 3, or Mastery) and sets time per prompt.

Prompt Generation Frontend requests a prompt from the backend. Backend randomly selects a prompt from a predefined list.

Display Prompt & Start Recording Prompt is displayed on the screen. Countdown timer begins simultaneously with audio recording in the browser.

User Response Capture User speaks their analogy response during the timer duration. Audio is automatically stopped when the timer ends.

Audio Transcription Recorded audio is sent to the backend. Whisper model converts audio to text transcript.

Evaluation of Response Transcript is evaluated using a language model for: - Response Rate - Speed of speech - Energy/Confidence Silent responses are marked as skipped (no evaluation scores).

Data Storage Transcript and evaluation scores are stored in the session for the round.

Next Prompt Cycle Steps 2–7 repeat until all prompts for the round are completed.

Round Completion & Feedback Generation All collected data is submitted. Feedback page displays: - Number of answered vs silent prompts - Average scores per metric - Round-specific statistics

Game 2: The Conductor – Features

Custom Topic Input – User can choose a default topic or enter their own.

Energy & Breathe Cues – Random energy levels (1–9) and breathe prompts guide speech delivery in real-time.

Audio Recording & Playback – Records speech during the game; playback includes a moving marker on the energy chart.

Visual Feedback – Energy meter and BREATHE overlay; line chart shows energy over time with cue points highlighted.

Audio Analysis – Uses Faster Whisper for transcription and librosa to compute vocal energy per second.

Topic Adherence – Measures similarity between transcript and topic using Sentence Transformers.

Post-Game Feedback – Displays transcript, topic score, average energy, energy transition success, response time, audio, and energy chart.

Session Metrics – Tracks energy transition accuracy, response speed, vocal variety, and recovery during BREATHE moments.

Game 3: Triple Step – Integration Under Pressure

Objective: Integrate random words into speech naturally while maintaining topic coherence and smoothness.

Features Implemented:

- Random Word Integration: 8 words displayed sequentially with 12s per word.
- Topic Guidance: Main topic 'Innovation in Technology' tracked for coherence.
- Speech Recognition: Browser-based real-time transcription.

Evaluation Metrics:

- Word Integration: Counts only once per word.
- Topic Coherence: 90% if topic mentioned, else 50%.
- Smoothness: Penalizes long pauses >1.5s.
- Semantic Similarity: BERTScore used to measure alignment between spoken transcript and topic.

UI: Styled words, metrics, and buttons for clear feedback.

Game Flow: Next word shown automatically; final score displayed after all words.

Tools Used: Faster Whisper (optional), Flan-T5 (Small), SBERT, BERTScore, Browser SpeechRecognition, JavaScript for metrics.