

# AI-Powered Pronunciation Assessment Platform for English Teachers

AnJiang  
University of Edinburgh  
United Kingdom

## 1 MOTIVATION AND WHY I CHOSE THIS TOPIC

The Informatics Large Practical (ILP) course focuses on building **distributed, containerised, data-driven systems** that solve real-world operational problems using **REST microservices**. For CW3, I wanted to create something that not only extends the spirit of ILP, but also solves a problem that affects someone close to me: my girlfriend, an English teacher.

A typical class requires her to manually listen to **20–40 audio recordings per student** for pronunciation tasks. For a class of 30 students, this can exceed **900 recordings**. This takes hours and makes personalised feedback almost impossible within the constraints of a busy teaching schedule. Her challenge mirrors the core ILP principles: a heavy manual workload, limited resources, and the need for **automated decision-making**.

Just as ILP automates drone deliveries, my CW3 project automates pronunciation delivery: collecting student recordings, assessing pronunciation quality, and routing results back to teachers. Time saved for teachers directly translates to better learning outcomes for students. And practically speaking, “time is money.”

This use case is meaningful, relatable, and technically aligned with ILP’s focus on service orchestration, data processing, and microservice design.

## 2 PROBLEM DEFINITION

The key problems English teachers face are:

- (1) **High manual workload:** Teachers repeatedly listen to each recording to determine pronunciation correctness. This is time-consuming.
- (2) **No scalable feedback:** Manual marking limits how many assignments can be reviewed, which reduces student practice frequency.
- (3) **No structured workflows:** Assignments (20–40 word tasks) are managed in spreadsheets or group chats, making it difficult to track progress.

The system I built addresses all three problems with a structured, automated, ILP-style microservice solution.

## 3 TARGET AUDIENCE

The platform is designed for:

- **Primary audience:** English teachers in schools, training centres, and private tutoring.
- **Secondary audience:** Students aged 8–18 preparing for exams such as IELTS speaking exam.

The system optimises workflows for teachers while providing students with an intuitive, progress-based learning environment.

The conceptual alignment with ILP is strong and intentional.

## 4 FEATURES IMPLEMENTED

The system automates the full assignment lifecycle.

Table 1: Features for Teachers and Students

For Teachers	For Students
Browse pre-built databases (IELTS, Zhongkao, TOEFL)	View assigned tasks
Select 20–40 words to create an assignment	Record pronunciation directly in the browser
Assign to one or multiple students	Receive instant pronunciation scoring (0–100)
View student progress (completion %, average score)	Phoneme-level feedback
Review submitted recordings	Track progress visually with AI feedback attached

## 5 IMPLEMENTATION DETAILS

### 5.1 Architecture Overview

The system follows a three-tier microservice architecture with clear separation of concerns.

- **Backend Layer (FastAPI + Python 3.9):** Handles REST endpoints, authentication (JWT), and communication with the Assessment Engine.
- **Frontend Layer (React 18 + Vite):** A modern Single Page Application (SPA) for teacher assignment management and student practice/recording interface.
- **Database Layer (SQLite/PostgreSQL):** A normalized relational database storing users, assignments, and detailed assessment results (JSONB).

### 5.2 Pronunciation Assessment Engine

The system integrates with Azure Speech Service’s Pronunciation Assessment API.

- **Audio Processing Pipeline:** The Frontend uses the **MediaRecorder API** to capture audio, converting it to a **16-bit PCM WAV format (16kHz)** required by Azure.
- **Assessment Metrics:** The API provides Accuracy, Fluency, Completeness, and an overall Pronunciation Score (0–100).
- **Automated Feedback:** The system parses the JSON response to generate contextual, phoneme-specific guidance

### 5.3 Code Quality and Testing

The development prioritised the architectural quality and unit testing, crucial aspects of the ILP marking criteria.

- **Code Structure:** The backend Python code is organized into clear packages: routers (endpoints), services (business logic), and models (database and Pydantic schemas), ensuring **clarity and readability** for future maintenance.
- **Testing Strategy:** The system includes **unit tests** for core business logic (e.g., assignment validation, progress calculation) using **pytest**. Integration tests verify **API endpoints** return correct status codes. All services are protected by custom exception handlers to ensure the service remains robust and consistently returns appropriate HTTP status codes, preventing runtime failures in the container environment.

#### 5.4 Containerisation and Deployment

The backend service is **containerised** using a minimal Python **Dockerfile**. The system has been deployed to an **Azure VM** and is publicly accessible at [speakright.uk](http://speakright.uk). The production deployment architecture includes:

- Docker containers for backend services
- Nginx as a reverse proxy
- PostgreSQL for production data persistence
- Let's Encrypt SSL certificates for secure HTTPS

The platform has been tested with **2 teachers** and **3 students**, processing over **200 recordings** successfully in a pilot program.

#### 5.5 Known Limitations

To provide a complete and honest assessment of the solution, the following limitations are noted:

- **External API Dependency:** Reliance on the Azure API creates latency and a monetary cost per call.
- **Missing Features:** The student interface does not currently support an offline mode.
- **Scope:** The assessment is currently limited strictly to English pronunciation.

### 6 KEY INNOVATION / IDEA

- (1) **True Workload Reduction:** Teachers can now review 40 recordings in under 5 minutes, as AI generates detailed scoring and comments.
- (2) **Scalable Personalised Learning:** Large classes can receive personalised pronunciation feedback instantly.
- (3) **Phoneme-Level Insights:** Students receive actionable guidance such as “Your /r/ sound is too weak.”
- (4) **Alignment with ILP Mindset:** A clean microservice system that leverages AI for a complex, human-centric decision-making process.

### 7 CONCLUSION

This CW3 project is a practical, innovative extension of ILP’s core themes, automation, distributed design, REST architecture, and meaningful real-world impact. It solves a problem with a production-ready tool, demonstrates technical depth, and aligns with the marking rubric for innovation, implementation, completeness, and clarity.