Final Year Project Proposal

TU856

Mapping the Spoken Word: A Geo-Spatial Platform for Voice Recognition and Analysis

\<Cianán Finn\>
\<C22393223\>

School of Computer Science
TU Dublin – City Campus

\<12/10/2025\>

Table of Contents

Contents

Declaration

I hereby declare that the work described in this dissertation is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed:

Cianán Finn

<Cianán Finn>

<12/10/2025>

Summary

My idea is to create a web-based application that accepts voice data input, which then analyses the voice data to predict where the accent may originate from. This result will then be visualised on a digital map. Although other projects have focused on individual aspects of my idea, I have not been able to find one that combines machine learning, web mapping and linguistics together.

The project's usefulness lies in how it ties speech patterns to specific geographical locations, offering a unique approach to visualising how accents change across locations. With further research and tweaking I believe it could become useful to a broader audience, such as actors wishing to learn or test an accent for a role, or simply by someone learning English wanting to practice their accent.

Background (and References)

Fundamentally, my project can be broken down into two main areas: web mapping and accent recognition. Accent recognition is a fascinating field that combines linguistics, speech processing, and machine learning. It focuses on patterns in speech data to group and classify accents, often with the goal of linking them to specific geographic regions. While the underlying mathematics has been established for decades, the availability of large datasets in recent years has made it possible to train models that produce accurate and reliable results.

Accent recognition is mainly used by large tech companies for voice assistants (eg. Siri), but these systems are hidden to the public and the data isn't visualised geographically. There are open source projects, such as Mozilla's "Common Voice" which collect speech data but still do not visualise the data in a geographical manner.

The goal of this project is to bridge the gap between accent recognition and web mapping, creating an innovative application that allows users to upload voice data and visualise accent predictions on a map. This approach combines machine learning, data visualisation, and linguistics in a novel way, and offers the potential for further development beyond the scope of this project timeline.

Proposed Approach

Initially, I plan to start with a research phase. There are a few key questions I want to look into before I start building the project so I avoid any time-consuming mistakes. The first question is which machine learning approach to use. Right now, I'm considering three options: MFCCs (Mel-Frequency Cepstral Coefficients) with classical classifiers, neural networks, or pre-trained models like Wav2Vec2. Which approach I choose will depend on the data I have available, which is the second question I will research. I want to find out how much labelled data each model needs, and what the minimum amount of data I could use is without losing too much accuracy, and what is the minimum amount of data needed to be uploaded for an accent to be determined.

Next, I'll research potential datasets. This will help me decide whether to focus on Irish-specific accents or expand to general Anglophone accents. The dataset I choose will also guide how I visualise the results on a map. If suitable datasets aren't available, collecting data may become an important part of the project.

The data visualisation part shouldn't need much research, as I'm currently studying Advanced Web Mapping, which covers the technical skills I will need. Finally, I plan to review existing literature related to accent recognition and geo-spatial visualisation to learn from methods used by others and make sure my approach is informed by best practices.

The user group this project will focus on is the general public, with a possible focus on the Irish people, specifically those with an interest in linguistics and accents. Their requirements will be:
- The ability to upload a speech data sample.
- Receive coordinates (continuous data) back.
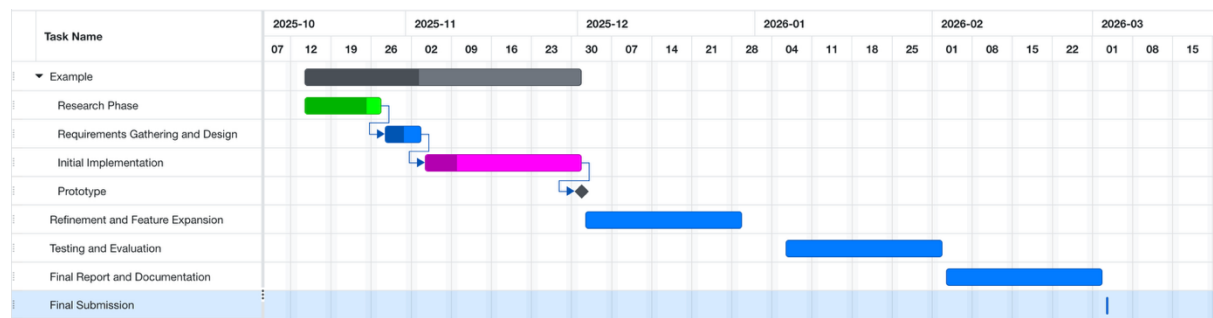- See the location-result on a map.

I plan to design this project with a Modular approach to make development and testing simpler and less time consuming. I will divide the project into the front end (using Leaflet and having a file upload method), the backend (using Django, to handle uploads, process the requests and pass data to and from the machine learning module), the database (Postgres and PostGIS) and the machine learning module. Input will be taken from the front end, passed to the back end which will then pass it to the machine learning module, which will return a predicted accent location, which the backend will then store in the database. The database will then be referenced to get the data to display is on the map.

The testing of the output data will be done using a dataset with known values that the machine has not analysed before. I will assess its accuracy, its speed and its consistency. I will be doing a lot of the User Interface testing myself, such as file uploads and the viewing of the output. After I am satisfied with the product, I will ask a random sample of people of different backgrounds(such as level of technical knowledge) to test it and to note any bugs or accessibility issues.

Deliverables
- User friendly working web application to upload input data and view output data.
- Machine learning model that can accurately predict accent origin and is able to be evaluated by speed, consistency and accuracy.
- Database for storing speech data and predicted outputs.
- Testing results.
- Final report

Project Schedule

Task	Start Date	End Date	Duration
Research Phase	14/10/2025	27/10/2025	13
Requirements gathering and Design	28/10/2025	03/11/2025	6
Initial Implementation	04/11/2025	01/12/2025	27
Prototype	01/12/2025	01/12/2025	0
Refinement and Feature Expansion	02/12/2025	29/12/2025	27
Testing and Evaluation	06/01/2026	02/02/2026	27
Final Report and Documentation	03/02/2026	02/03/2026	27
Final Submission	03/03/2026	03/03/2026	0

Technical Requirements

Frontend: Leaflet to create interactive maps to display the location data.

Backend: Django to handle requests and to facilitate communication between the user, the database and the machine.

Machine Learning Model: Python will be used. The libraries used will be determined by which machine learning approach I take.

Database: Postgres to handle the data, with PostGIS extension used to store the geographical data.

Others: I will be locally hosting it all on my Mac, and I will use Github for version control.

Conclusion

In this project, I plan to create a web-based application that accepts speech samples, which are then analysed by a trained machine learning model to predict the estimated geographic origin of the speaker. These results will be visualised on an interactive map, providing a clear connection between linguistic features and geographical data.

The project will be tough in two main areas: machine learning and linguistic analysis, and the data visualisation. By building off existing research carried out in accent recognition and

web mapping, I'll avoid previous issues and use the knowledge gained by other researchers to reach my project goal by March.

Given the timeline I have set out above, I believe the project is feasible with continuous thorough and hard work, and if all goes well, it will be a project to be proud of.

References

Hint:

Use Zotero to manage your references (see Brightspace resources).

Use the **Harvard** or **IEEE** numbering referencing style

- https://www.zotero.org/support/quick_start_guide

Mozilla "Common Voice"

https://commonvoice.mozilla.org

Open source speech data collection and classifying project.

Indian vs American English Accent Differentiation

https://github.com/lakshanakolur/Accent-Recognition-ML?utm_source=chatgpt.com

Supervised Machine Learning Model for Accent Recognition in English Speech using Sequential MFCC Features.

Leaflet.js

https://leafletjs.com

JavaScript Library for building interactive maps. To be used for data visualisation.

Title:
Investigating Issues with Machine Learning for Accent Classification

Student:
Zhewen Lou and Yitian Ren

Description (brief):
"Speech recognition has become a widely researched topic for decades, and there were already some successful products which has been put into commercial use, like Siri. However, sometimes it is hard to distinguish the word because words of different accents have different pronunciations. For instance, in Japanese English /r/ is usually pronounced as /l/. Therefore, it is natural to think that speech recognition could be divided into two parts. First attach a label to the audio about the accent, then recognize the contents based on the regular pattern of that accent. In this paper, we researched on several characteristics including voice onset region(VOR), vowels and formants to distinguish British English and American English. By applying both linear neural network and neural network with nonlinear classifications and two hidden layers(NN2HL), the accuracy rate reaches 86.67%, which is very satisfying."

What is complex in this project:

- Extracting meaningful features from speech, such as VOR, vowel formants, and other acoustic characteristics.
- Designing and training both linear and nonlinear neural networks, especially tuning a network with two hidden layers for non-linear classification.
- Getting a high accuracy despite variations in pronunciation across speakers.

Explain key strengths and weaknesses of this project, as you see it.

- High accuracy (86.67%)

- Using both linear and non-linear networks allows for comparison between the two.

Title:

English Accent Recognition Using Deep Machine Learning

Student:

Manokhin, Andryi V

Description (brief):

Its not available in person sadly, I will look for it this week.

What is complex in this project:

What technical architecture was used:

Explain key strengths and weaknesses of this project, as you see it.

Please check for spelling and grammar issues.

Appendix D: