

ASR Experience Report

In this lab, the words tested were primarily non-English place names and long or complex words (as this relates to my final project which focuses on geographical location names).

A few key examples are:

Kyiv

- First attempt: { confidence: 0.05958548, utterance: "Queue" }

Kieve

- First attempt: { confidence: 0.70735157, utterance: "Kiev" }

Thiruvananthapuram

- First attempt: { Confidence: 0.11452396, utterance: "Thiruvananthapuram" }

Agrabah

- First attempt: { confidence: 0.05840482, utterance: "AG Hubba" }
- Second attempt: { confidence: 0.23866056, utterance: "Agrabah" }

Oswiecim

- First attempt: { confidence: 0.16863412, utterance: "Or switching" }

Sweden:

- First attempt: { confidence: 0.6711968, utterance: "Sweden" }

Stockholm:

- First attempt: { confidence: 0.8494827, utterance: "Stockholm" }

As observed by the examples provided, well-known names such as Stockholm, and Sweden tend to have a higher confidence score in comparison to less familiar names. I believe this aspect of familiarity also affected the system's struggle between Kyiv and Kiev (probably due to impact of more common usage). Additionally, names which are longer and more complex in pronunciation have low confidence despite its correct transcription. Thus, indicating uncertainty within the ASR system. Lastly, places such as Malmö (Memo, Melmo), Oswiecim (Or switching), Mizo (Mizzoh, Mi Zho) were misinterpreted in transcription and low in confidence scores, most likely due to difficulties in processing unfamiliar phonetic structures.

Despite the remarkable advances within ASR technology, there remains the challenges to continue expanding it to recognise diverse and specialised vocabulary. Further expanding research, and testing collection within the field aids in improving accuracy across wide range of content.