I have selected the following words of countries from the fictional narrative government game Suzerain, which was inspired from a variety of languages and geopolitical contexts.Below are the result of the speech recognition system:

Word: Kyrute (Turkish inspired-pronunciation, but vague in origin, IPA transcription of my pronunciation: /kaɪˈrɑːt/)

- First attempt: { confidence: 0.12123864, utterance: "Kairat" }
- Second attempt: { confidence: 0.15290421, utterance: "Kairat" }
- Third attempt: { confidence: 0.09283355, utterance: "Kairat" }

Word: Agnolia (Dutch inspired-pronunciation, IPA transcription of my pronunciation: /ɑˈnyːliə/)

- First attempt: { confidence: 0.053700436, utterance: "Aknulia" }
- Second attempt: { confidence: 0.05350034, utterance: "Ah, Nuliya" }
- Third attempt: { confidence: 0.05309956, utterance: "Aknulia" }

Word: Wehlen (German inspired-pronunciation, IPA transcription of my pronunciation /ˈveːlən/)

- First attempt: { confidence: 0.16626315, utterance: "Velun" }
- Second attempt: { confidence: 0.1689854, utterance: "Velun" }
- Third attempt: { confidence: 0.08822421, utterance: "Velun" }

The ASR's ability to recognize words with non-standard English pronunciations is quite low, as these pronunciations were not part of its training data. The confidence scores were generally in the range of 0.05 to 0.17. Despite these low confidence scores, we can definitely notice some phonetic similarity between the intended pronunciations and the ASR's transcriptions. I deliberately chose pronunciations inspired by Turkish, Dutch, and German for the purpose of our experiment. None of the attempts resulted in an accurate transcription of the words.The recognition failures are predominantly due to the unfamiliarity of these words to the system's vocabulary. There was notable vowel confusion. For example, when I pronounced "Wehlen" with a long 'e' vowel (/eː/), the system misinterpreted this, most likely as the short 'e' vowel (/e/). The final syllable was also misunderstood, likely because it contains a neutral vowel sound (/ə/). It was recognized as "un," a phonetically related but more stressed syllable. This suggests it is challenging for the ASR system to accurately map German,Dutch or Turkish phonemes to their closest English phonetic counterparts. The system attempts to approximate the input sounds using the closest phonemes it knows in English. It is also possible that variations in my own speech (clarity or speaking rate) have influenced the recognition, especially for words like 'Agnolia.' To achieve better sound recognition the ASR model would need to be fine-tuned on a dataset that includes these variations.