Endpoint-ID: deitiestest1

I tested Azure's ASR with names of fictional deities and places from the Dungeons and Dragons setting 'Forgotten Realms' (https://forgottenrealms.fandom.com/wiki/Seldarine).
As expected most of the terms were never correctly recognised in the baseline model, when I tested using full sentences as well as isolated utterances. Confidence for the incorrectly classified terms varied quite a bit, from unlikely readings (<0.01) to rather confidently incorrect assumptions (>0.6). Interesting was for one, that (assuming my pronunciation is not to blame entirely) the system also had trouble with some (less) common terms (I.e. 'elven', 'falchion', 'melee'). Also, the baseline system would occasionally correctly recognise some terms form which I would have expected that they hadn't been seen before ('Corellon' [0.083], 'Cormanthor' [0.303]).

Testing on the custom trained model was moderately successful at best: Regarding the lack of significant improvement on many words, I assume a major component may be the opaque orthography. Hardly any of the words below have only one legitimate possible pronunciation in English. Therefore, I assume that the presence of unusual to non-existent grapheme sequences and combinations in 'normal' English, might be the reason for the results. For these, it seems the system has trouble mapping sounds to the desired spelling, as there can be no exact prediction for the sound form of these segments (ie. <dh> in Angharradh, <vh> in Vhaeraun). And in fact, some of these segments are simply not present in the pronunciation (<tanari> vs. <tanar'ri>, <r'r> is realised as a single <r>, according to the https://forgottenrealms.fandom.com/wiki/Tanar'ri and reflected in my laptop's built-in TTS service). In such cases the ASR system would need to strongly favour the desired 'complicated' spelling over a 'simple' variant ('Vhaeraun' /veɪˈrɔːn/: <Veyron> [0.226]).
Similarly, unstressed vowel which are typically reduced to schwa can not be reliable reconstructed and lead to similar word forms competing ('Larethian': <Loretheon> [0.187], <Larathian> [0.305]).
Sometimes using the built-in TTS function helped me to gauge what sound form might be expected from spelling alone. In some cases this did lead me to get correct results from the ASR model, although the pronunciation might deviated from the canon in fiction.

Nevertheless, there were some terms that were recognised (more) reliably in the custom model, those terms and their maximum observed confidence score in the baseline -> the custom model being being: Seldarine (n.A. -> 0.291), Corellon (0.083 -> 0.334), Larethian (n.A. -> 0.280), Cormanthor (0.303 -> 0.345) and elven (0.064 -> 0.444). Of the 19 terms I gave the system 10 remained unrecognised, at best mapping to a single word ('Angharradh': <Angarath>), at worst only being transcribed as multiword expressions ('Mordenkainen': <More than Canaan>).

Tested expressions

Eventually recognised:
Seldarine
Elven
Arvandor
Corellon
Larethian
Cormanthor
Moonbow
Mordenkainen
Hanali

No successful recognition:
Iallanis
Araushnee
Vhaeraun
Ghaunadaur
Eilistraee
Angharradh
Aerdrie
Sehanine
Tanar'ri
Gruumsh