“Alexa, Can You Understand Me?” Ethnicity Based Bias in Voice Recognition Systems

Mitali Bhosekar
Clemson University

Pranita Camasamudram
Clemson University

Errol Francis II
Clemson University

Abstract

Over the last decade, the growing popularity of automated speech recognition platforms has led to increasing concerns regarding the security, privacy, and fairness of these devices. In this paper, we explore the accuracy and implicit biases of the Amazon Alexa voice recognition system. Using 1943 audio recordings, we measure the accuracy of Alexa’s responses to voice commands with different accents using a correct score calculations. Results reveal that while Alexa struggles to accurately complete voice commands when presented with a foreign accent, Serbian accents had the worst accuracy out of accents with 10 or more distinct recordings.

1 Introduction

In today’s world the use of automated speech recognition (ASR) platforms has become increasingly widespread. From speech to text recognition we have come a long way with technologies that even interpret our words, understand them, and reply accordingly. Platforms like Alexa, Siri, and Google Assistant have become a part of our daily life, from checking the weather to closing the doors. Due to such extensive use of these systems along with plethora of advantages there comes a few issues too.

With fairness and bias becoming a hot topic in machine learning these past few years, we are curious on how all these technologies with artificial intelligence implementations work against it. But when it comes to Artificial intelligence, it only works based on how it is trained. Which means that often these biases are a reflection of human bias which tend to get into the data that is fed to the algorithm. It reflects of what is included in the training data and what is excluded from it. Few of the data sets in the fields of medicine, crime, education, finance etc. are more prone to biases comparatively. Bias can exist in any form and can make their presence known during anytime of the model design. Although there is no definitive way of describing Fairness it is usually seen as the lack of prejudice or bias for any individual or group. The goal of fairness in such algorithms should be to make predictions that are not biased towards certain ethnic or demographic groups. In a lot of ways bias and fairness go hand in hand. The main bias that we see in this particular case is representation bias which is due to how we sample a population while collection of data. This gives rise to bias towards a certain culture. Another bias to be noted is data bias in which certain data is more represented than the others[5]. This formed as a major limitation for the accuracy of evaluation as explained more in the limitations section. The major question that we are asking is, “How fair are these speech recognition systems?”

In this paper we want to analyse how Alexa fares against bias and fairness problem. Often we see people struggling to get Alexa to understand what they’re trying to say and have to repeat the same statement over and over for Alexa to finally understand. Sometimes even that doesn’t work and they end up with silence as a reply to their question. This has lead to the frustration of the users and embarrassment of having to constantly make Alexa understand. Is it something in the internal workings of Alexa that’s causing this? Or is a problem with bias and fairness? Does the person’s accent play a part in this situation? Numerous surveys have found noticeable differences on how Alexa perceives a particular command in varying accents [2]. We have taken one such dataset, a collection of 2140 text samples read in English with distinct accents. While listening to the audio files Alexa gives a response and that is stored as an output in a csv file. We then performed the accuracy analysis by comparing each word of the response with each word of the expected response. This helped us assess the fairness of Alexa’s voice recognition system for each accent. The results were then tabulated and captured in the form of a graph.

1.1 Problem Statement

The problem that we are trying to implement is based on the concept of fairness in AI. Despite many useful applications, how do these systems stand against bias and unfairness is the question. Our project focuses on how the different accents affect Alexa's ability to understand and interpret the commands given. We drive our analysis on the following research question:

- **RQ1:** What is the accuracy level for different accents around the word when Alexa is asked to note down a command?

We feed the voice recognition system a dataset of audio files containing recordings of one command spoken in different accents. Based on that we then analyse and compare the expected result and what exactly Alexa hears with different accents and report the results.

1.2 Goals

Our main objective is to assess the fairness of the responses provided by the Amazon's ASR platform, Alexa. Using prerecorded audio files of speakers with varying regional accents reading the same text prompt, we will measure how accurately Alexa can capture the prompt. We calculate accuracy using our custom measurement then compare accuracy rates by regional groups. This analysis of response accuracy will be used to identify any existing ethnic disparities in the ASR platform.

2 Related Work

The study of racial bias in ASR platforms is a fairly young research topic and has yet to be explored in depth. Only a handful of research studies have been conducted in this area. In [4], researchers attempt to isolate the factors that cause ASR systems to exhibit bias against African American speakers. The researchers posited that African American Vernacular English (AAVE), specifically the use of the invariant form of "be" indicating a habitual aspect, would adversely affect the accuracy of ASR platforms. Two systems are used tested in this paper: DeepSpeech and Google Cloud Speech.

Both systems were fed over 105 hours of audio captured from the Corpus of Regional African American Language to analyze. Within this audio were 376 use cases of the habitual "be" and 2,974 use cases of the nonhabitual "be". Audio segments were classified into two types. Utterances were defined by single segments of speech delimited by pauses. A turn was said to be the entire set of a speaker's contiguous utterances prior to a new speaker beginning to talk. For DeepSpeech, the results showed that while the habitual "be" was more error prone than the nonhabitual "be" both use cases had a negative effect on accuracy. The results also suggested that speech rate had a negative effect on accuracy in utterances. Google Cloud Speech also displayed similar results, with the exception that accuracy was worse for the words preceding "be" than those that came after.

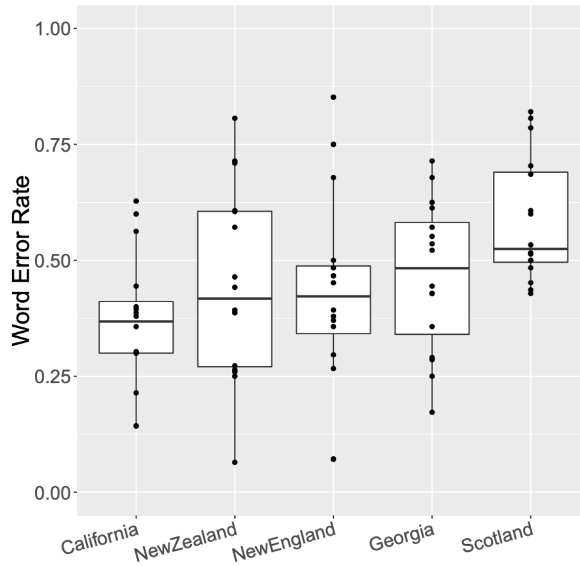

Figure 1: The results of the YouTube automatic caption word error rate experiment conducted by Tatman, grouped by dialect region. The points represent each individual speaker

In [7], YouTube's automated caption system was the focus of the research. Five different dialects of English were used across male and female participants. The data used was collected by hand checking the automatic captions provided by YouTube in selected "accent tag" or "accent challenge" videos. 80 speakers were used for analysis. The regional dialects selected were California, Georgia, New England (but only the states of Maine and New Hampshire), New Zealand, and Scotland. Eight men and eight women from each region were used in the study.

The results showed that the automated captions for speakers from Scotland were reliably less accurate than the captions for from the United States or New Zealand (Figure 1). There was also a significant difference in accuracy when comparing gender: automated captions were 13% more accurate for men than they were for women. The results also showed an interaction between gender and dialect region. The decrease in accuracy because of gender was not equal across all groups. The biggest discrepancy came from New Zealand where a

38% decrease in accuracy between men and women was recorded. Because the captions provided by YouTube are supported by Google speech recognition software, these results are comparable to those of other Google ASR platforms.

In [3], the authors have compared five ASR systems - Alexa, Siri, Google Home, IBM voice recognition system, and Microsoft Cortana - by conducting interviews with 42 white speakers and 73 African American speakers. The comparison comprised of different factors such as age, gender, location apart from Race. On average, the racial disparity, calculated by Word Error Rate (WER), was about 0.35 for African American speakers and 0.19 for white speakers. WER is calculated by $\frac{(S+D+I)}{N}$ where S stands for word substitutions, D for deletion, and I for insertions. N is the total number of words.

When mapped with gender, African American male speakers have a WER of about 0.41, whereas Females had 0.30. In comparison, white male and female speakers have about 0.21 and 0.17 WER rates. From this paper, we understand that many factors affect the WER rates. This also indirectly states that the sample population should have diversity in each factor considered—the paper talks about two races, i.e., African American and White speakers. In our research, we would like to select a few ethnicity's over which we can focus.

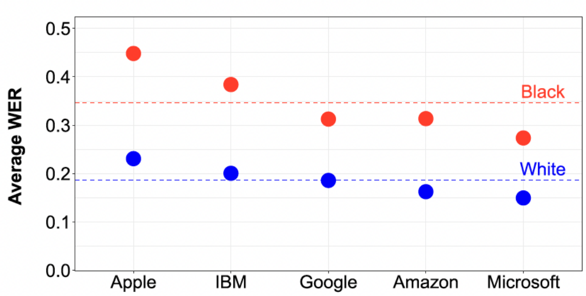

Figure 2: The average WER across ASR services is 0.35 for audio snippets of African American speakers, as opposed to 0.19 for snippets of white speakers.

In [2], a survey was conducted by The Washington Post in collaboration with research groups to study the Voice Assistants accent imbalance. A test with nonnative accents using Alexa showed 30 percent more inaccuracy. The article mentions that Google stated it needs to expand its datasets, and for Alexa, Vice President mentioned that as more people talk to Alexa with different dialects, Alexa will learn.The Article gives examples, if "China proposes removal of two-term limit, potentially paving way for Xi to remain president," is read to Alexa with a standard American accent, its reads "China pro-

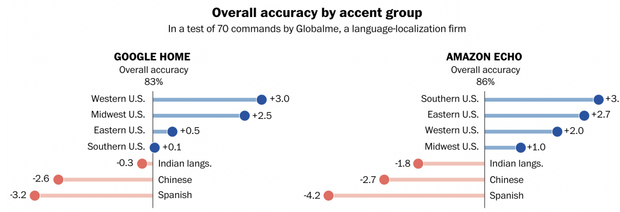

Figure 3: The chart above denotes Google home having 83 percent accuracy with worst inaccuracies for Spanish and Alexa Echo to have 83 percent accuracy with worst inaccuracies for Chinese

poses removal of to term limit potentially paving way for gh remain president". Suppose you read the same sentence in an Indian accent. In that case, Alexa perceives it as "China proposes removal of two-time limit for ten shirley pain way folks eight remind president."

The overall accuracy of nonnative accents of Chinese, Spanish and Indian is around 80 percent. The article states that nonnative accents are harder to train, which can be supported by research conducted on Google's speech recognition to auto-populate the YouTube subtitles with results showing the worst subtitles from women and southern and Scottish accents.

The problem of Alexa understanding accents has also been a huge issue in understanding the British accents as reported by Uswitch.com in article [1]. The data reveals that people have been increasingly upset about Alexa not able to interpret their commands given in regional accents. An analysis of 30 accents across UK has been done wherein people were suggested to ask their device the 10 most commonly asked questions. A score out of 10 was given for each accent for how well Alexa understood them with lesser the score more the errors. The article also states that the research went through how many searches of "Why can't Alexa understand me?" have been made on Google. The data was then summed up and then a final score was give. Based on the data and with 1550 google searches of "Why can't Alexa understand me?", the city where Alexa understood the commands the least was Cardiff, followed by Glasgow and Liverpool. On the contrary, the accent spoken in London was the easiest for the AI to understand with only 200 google searches monthly. Another independent experiment was done in 309 cities where the participants were told to ask their devices few common questions. This was then uploaded to Google Transcribe to see how many of those questions asked where not dictated properly. Google's Keyword Planner was then used to point out which if the the keywords were understood or not understood by Alexa.

3 Methodology

3.1 Generating the Output File

Generating the output file starts with selecting high quality audio files. Fortunately, we found a promising database that could be used in this endeavor [6]. The dataset was comprised of 2140 spoken text samples. Each sample comes from a different reader but all participants read the same sample. The readers hail from 177 countries using 214 distinct languages, but all text samples are read in English.

With the audio database selected, we employ a Python script to automate voice commands and record Alexa's responses. Every test start with an automated voice directing Alexa to "make a note." The program then uses the directory full of mp3 files as the input. The mp3 files are played through the computer speaker and Alexa to saves the played audio in a virtual note. After making the note, Alexa repeats it back to the program to ensure the note has been made correctly. The *speech_recognition* Python package enables the program to listen and record the responses Alexa returns and saves them to a csv file for later analysis. The test ends with the automated voice returning to tell Alexa, "Hey Alexa, exit."

Two Amazon devices were used in generating the output file. First, our team used an Echo Show 5 for the first phase of testing while still trying to refine the Python script. Once the script was complete, we used an Echo Dot (3rd Gen) for the final testing. The Python script - as well as all of the input audio - was run using a 2020 Macbook Pro with the Apple M1 chip.

3.2 Analysis

The output file generated by the Python script had various columns: audio file name, the Alexa generated output, elapsed time, and input text. Before running the Python script, the code was adjusted so that each recording only played for seven seconds. Since each accent has a different pace of speaking, the input provided to Alexa was not consistent. Our desired input was, "Please Call Stella. Ask her to bring these things with her from the store." In many cases the input played longer than the desired length. The output file consisted of 1943 rows of data after about 30 hours of running the testing script.

For our analysis we used the output file generated in CSV format. First, we transformed the file using Microsoft Excel so it could be used for our calculations. All of the file name had accents and numbers attached to it. Because of this, we had to separate the accent and it's accompanying numbers. We did this for 1943 inputs using excel formulas.

Next, we conducted a word-to-word comparison of the input against the noted Alexa output. Table 1 illustrates how accuracy was calculated. Each time a word was correctly noted, a positive score was calculated against it. If the word was missed, a zero score was calculated for it. The sum of positive score was divided against the total number of words to get an average score, as illustrated in (1). This process was conducted for all 1943 inputs. Therefore the max accuracy for an input that can be achieved is 1, which means that all 14 words were correctly noted by Alexa.

$$Score = \frac{Sum\ of\ positive\ score}{Total\ words} \tag{1}$$

The reason we did an exact word-to-word comparison was to measure how accurately Alexa can listen to a command. We made Alexa note exact what it could hear. Another method of interpretation of accuracy could be if Alexa could understand the commands in different accents rather than just note it down. This method was out of scope for our project since we could not get the right dataset for our it and would require to make Alexa perform an action.

4 Evaluation

The output file generated in methodology was used to create visualizations in tableau. The accents with 1943 inputs had a total of 190 accents. While creating visualizations, we added a parameter of count of accents from the 1943 inputs. Next we took the average of all the similar accents. With the kaggle data set we realised that each accent does not have a fixed number of inputs. For instance, there were 533 inputs for an English accents but only 92 input for an Arabic accents. As we aim our paper on fairness it would not be fair to compare the final output values of these two parameters together. For comparison, we club accents together with same input size due to the large nature of the dataset.

Input Count of 1

First we examine the scores for accents with only one recording in the sample. The Carolinian accent has the greatest accuracy, with a score of 0.92. The Temne accent placed second with a score of of 0.75. Around nine accents have an accuracy score 0.5, and almost 50 accents of the 190 in the collected have an accuracy score of 0.0. For more information, please refer to Appendix A.1.

Input Count of 2

Next we examine the scores for accents with only two recordings. When we compare accents with and input size of two, the Oriya accent has the maximum accuracy,

Example	Accent	Input	Output	Correct Words	Score
1	Russian	Please call Stella. Ask her to bring these things with her from the store	Please call Stella ask her to bring her things with her from the store	13	13/14 or 0.92
2	Kikongo	Please call Stella. Ask her to bring these things with her from the store	Please call Stella ask her to bring things with her from disturbed	7	7/14 or 0.50

Table 1: Two examples of how the accuracy is calculated using the original input and the recorded output from Alexa.

earning a score of 0.5. On the opposite end of the spectrum, there are 11 accents with accuracy score of 0.0. For more information, please refer to Appendix A.2.

Input Count of 3 and 4

The next grouping of accents have either three or four recordings in the sample. When comparing the accents with three recordings, we see the Pulaar - with a score of 0.59 - is the most accurate in the group. On the other hand, there are four accents that have an accuracy score of 0.0. In the accent group with four recordings, the kikuyu accent earns the highest score, with an accuracy of 0.5. The Uzbek accent is the least accurate with a score of 0.0. For more information, please refer to Appendix A.3.

Input Count Greater Than 5

When examining accents that have five or more recordings, we see a drastic decrease in accuracy scores. In the accent group with five to nine recordings, the Slovak accents - with five recordings - has the highest accuracy score of 0.28. The Ga accent, which has six recordings, is the least accurate, with a score of 0.02. In the accent group with 10 or more recordings, the Polish accent - with 23 recordings - has the highest accuracy score of 0.26 followed by czech and albanian with 0.0 accuracy with input size 7. In the accent group of more than 10 recordings, the Serbian accent, which had 15 recordings in the sample, is the least accurate, with a score of 0.07. Figure 4 illustrates the accuracy scores for accents with 30 or more recordings in the sample. For more information, please refer to Appendix A.4 and A.5, which contains the information on the accuracy scores for the accents in the sample.

From our analysis, even after considering the limitations, we can conclude that Amazon's Voice recognition system (Alexa) clearly produces a representation bias. With smallest input size one we can see that approximately 50 accents have accuracy 0. Similarities can be seen when the input size increases and accuracy drops further. Accuracy with 6 and 3 input size for ga and dari respectively drops to a lowest of 0.02 which in percentage is 2 % . Out of a total of 190 accents, 67 accents have 0.0 accuracy. This can clearly lead to frustration within users of Alexa and eventually stop them from using the device. We advice that Amazon use more accents in their training dataset of voice recognition systems. Furthermore, there has been further research on various voice recognition systems as illustrated in our background research which conveys word error rates. In general all Voice recognition systems infer representation bias. Efforts need to be made for inclusion of diversity in datasets and to improve accuracy further as voice recognition systems are a part of many technologies these days which can led to hazardous accidents, especially when it comes to automated car systems.

5 Limitations

While the analysis give us a great deal of information about the accuracy and fairness of the Alexa system, there are still few limitations that hinder the results from being a good percent of accurate. The dataset that we have has inconsistent recording inputs within the accents itself. For example the English accent has 533 recordings while Polish has only 23, and there are few others that have one recording. This kind of inconsistencies gives us inaccurate and unfair results. Even if we choose the the same number of recordings for all the accents in random and analyse, it will still be an unfair comparison as the selected accents may be either very easy to decipher or it can completely go the other way. Also another thing to be noted is that our analysis was based on comparing word to word and then calculating the average of true results. This may differ in reality for some cases as sometimes even if what Alexa hears is not the same as what is expected to hear, it may still understand the context and get the work done.

6 Future Work

The project due to it's limitations has a great potential

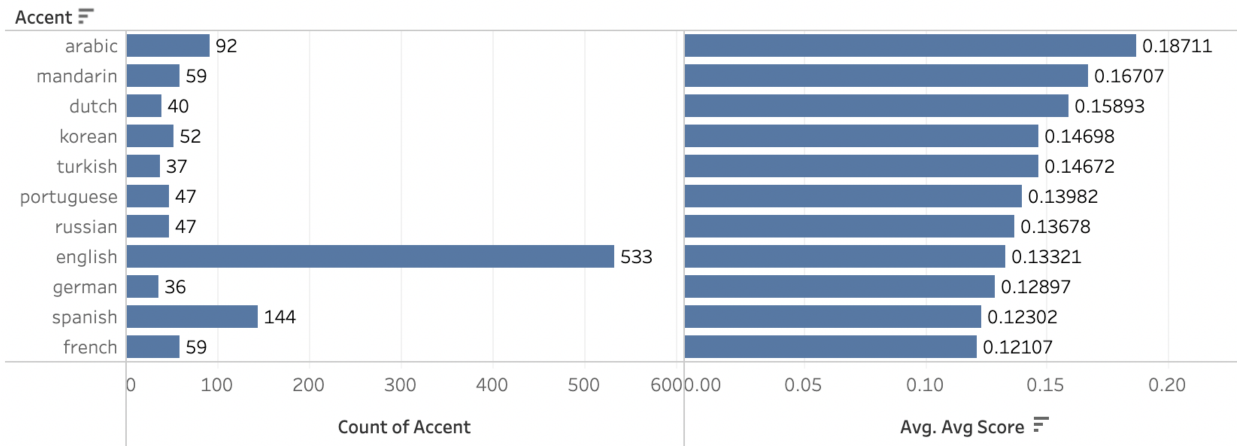

Figure 4: Captures the accuracy after of accents having input sizes 30 to 533. Maximum accuracy which can be see here is for Arabic language at 0.18 with a count of 92. The entire visualization can be seen in Appendix A.6 or using the following link: https://public.tableau.com/app/profile/mitalibhosekar/viz/Security_16378042735430/Story1

for future work. Basing on what was not possible for us to do with the dataset, good amount of work can be done with proper resources. One of the solutions can be to take a survey of people with the said accents and record their voices. We can then divide the people according to how comprehensible/not comprehensible their accents are. Then with that we can sort out equal number of recording for each accent and then analyse the accuracy of each one. This could give us a pretty good analysis with quite accurate results. A good research further on bias and fairness in Machine Learning would give us more insight and solutions on how to eradicate while it is in our hands.

7 Contributions

Contributions to the work are made equally by every group member. Our team meets biweekly to work collectively on the project and on upcoming presentations or reports. At the end of each meeting, action items are assigned to each group member to accomplish before the next gathering. GroupMe is used to communicate amongst team members to share results, ask questions, and plan outside of scheduled meetings.

The final phase of the project was split into two sections: testing and analysis. Errol took the lead on the testing portion of the final phase. He updated the Python code and ran the script along side the Alexa Voice Recognition System. Mitali and Pranita took the lead on the analysis portion of the final phase. They cleaned the csv data, conducted the accuracy calculations, and made informational visuals to accompany the data.

The final report was completed collaboratively. Errol drafted the Abstract. Pranita and Errol drafted the Introduction. Mitali, Errol and Pranita drafted the Related Works section while Mitali and Errol drafted the Methodology section. Mitali drafted the Evaluation Section. Pranita drafted the Limitations and Future Scope sections. Mitali and Errol drafted and formatted the Appendix. Once the drafts were complete, all team members helped in editing and revising the report for the final submission.

References

[1] FISH, T. Amazon echo: Alexa doesn't understand some regional uk accents. *Express* (2020).

[2] HARWELL, D. Why some accents don't work on alexa or google home. *The Washington Post* (2018).

[3] KOENECKE, A., NAM, A., LAKE, E., NUDELL, J., QUARTEY, M., MENGESHA, Z., TOUPS, C., RICKFORD, J. R., JURAFSKY, D., AND GOEL, S. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences 117*, 14 (2020), 7684–7689.

[4] MARTIN, J. L., AND TANG, K. Understanding racial disparities in automatic speech recognition: The case of habitual" be". In *INTERSPEECH* (2020), pp. 626–630.

[5] MEHRABI, N., MORSTATTER, F., SAXENA, N., LERMAN, K., AND GALSTYAN, A. A survey on bias and fairness in machine learning. *ACM Computiong Surveys* (2021).

[6] TATMAN, R. Speech accent archive. https://www.kaggle.com/rtatman/speech-accent-archive.

[7] TATMAN, R. Gender and dialect bias in youtube's automatic captions. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (2017), pp. 53–59.

A INPUT TABLES

A.1 Input Count of 1

Accent	Count	Accuracy
carolinian	1	0.928571429
temne	1	0.785714286
sesotho	1	0.5
sa'a	1	0.5
naxi	1	0.5
nama	1	0.5
hakka	1	0.5
gedeo	1	0.5
chittagonian	1	0.5
ashanti	1	0.5
agni	1	0.5
sinhala	1	0.428571429
bai	1	0.357142857
susu	1	0.285714286
sundanese	1	0.285714286
baga	1	0.285714286
yapese	1	0.214285714
teochew	1	0.214285714
nandi	1	0.214285714
mizo	1	0.214285714
mandingo	1	0.214285714
yakut	1	0.071428571
nuer	1	0.071428571
ilonggo	1	0.071428571
zulu	1	0
yupik	1	0
xasonga	1	0
turkmen	1	0
tatar	1	0
taishan	1	0
sylheti	1	0
sindhi	1	0
sicilian	1	0
shilluk	1	0
shan	1	0
rwanda	1	0
pohnpeian	1	0

Accent	Count	Accuracy
newari	1	0
mortlockese	1	0
moore	1	0
moba	1	0
mankanya	1	0
mandinka	1	0
maltese	1	0
malagasy	1	0
luxembourgeois	1	0
lingala	1	0
lamotrekese	1	0
lamaholot	1	0
kru	1	0
konkani	1	0
kirghiz	1	0
kanuri	1	0
kannada	1	0
kalanga	1	0
kabyle	1	0
jola	1	0
irish	1	0
ife	1	0
hindko	1	0
hainanese	1	0
gan	1	0
frisian	1	0
fataluku	1	0
faroese	1	0
fang	1	0
ewe	1	0
edo	1	0
dinka	1	0
chichewa	1	0
chamorro	1	0
cebuano	1	0
bamun	1	0

A.2 Input Count of 2

Accent	Count	Accuracy
oriya	2	0.5
chaldean	2	0.428571428
rotuman	2	0.357142857
bafang	2	0.321428571
slovenian	2	0.25
kikongo	2	0.25
gusii	2	0.25
burmese	2	0.25
bavarian	2	0.25
synthesized	2	0.107142857
papiamentu	2	0.107142857
ngemba	2	0.107142857
mauritian	2	0.107142857
filipino	2	0.107142857
amazigh	2	0.107142857

Accent	Count	Accuracy
tswana	2	0
telugu	2	0
shona	2	0
satawalese	2	0
quechua	2	0
luo	2	0
hmong	2	0
hadiyya	2	0
ganda	2	0
basque	2	0
bari	2	0

A.3 Input Count of 3 and 4

Accent	Count	Accuracy
pulaar	3	0.595238095
kikuyu	4	0.5
igbo	3	0.404761905
fijian	3	0.30952381
uyghur	3	0.261904762
marathi	3	0.238095238
latvian	3	0.238095238
vlaams	4	0.232142857
kambaata	3	0.19047619
malayalam	4	0.178571428
catalan	4	0.178571428
mende	3	0.166666667
icelandic	3	0.166666667
azerbaijani	3	0.166666667
wolof	3	0.142857143

Accent	Count	Accuracy
kazakh	3	0.142857143
fanti	3	0.142857143
lithuanian	4	0.125
estonian	3	0.071428571
xiang	4	0.053571428
krio	4	0.053571428
dari	3	0.023809524
tibetan	3	0
tajiki	3	0
oromo	3	0
belarusan	3	0
uzbek	4	0

A.4 Input Count Between 5 and 9

Accent	Count	Accuracy
slovak	5	0.285714286
danish	7	0.265306122
armenian	8	0.25
yoruba	5	0.228571429
hebrew	9	0.222222222
bosnian	9	0.206349206
pashto	7	0.193877551
georgian	5	0.185714286
twi	5	0.171428571
taiwanese	8	0.160714286
bambara	5	0.157142857
tamil	6	0.154761905
norwegian	6	0.154761905
gujarati	6	0.142857143
hausa	7	0.142857143
punjabi	8	0.142857143
somali	6	0.130952381

Accent	Count	Accuracy
afrikaans	5	0.128571428
indonesian	8	0.116071429
tigrigna	6	0.1
kiswahili	9	0.095238095
yiddish	5	0.085714286
malay	5	0.085714286
hungarian	6	0.083333333
khmer	7	0.081632653
mongolian	9	0.079365079
croatian	8	0.053571429
finnish	5	0.042857143
ukrainian	9	0.03968254
ga	6	0.023809524
czech	7	0
albanian	7	0

A.5 Input Count Greater Than 10

Accent	Count	Accuracy
polish	23	0.260869565
cantonese	20	0.214285714
swedish	20	0.207142857
farsi	20	0.203571429
urdu	13	0.203296703
bengali	17	0.197478992
miskito	11	0.188311688
arabic	92	0.187111801
nepali	13	0.181318681
thai	13	0.175824176
tagalog	12	0.172619048
italian	29	0.169950739
mandarin	59	0.167070218
greek	15	0.166666667
kurdish	10	0.164285714
japanese	24	0.160714286

Accent	Count	Accuracy
dutch	40	0.158928571
korean	52	0.146978022
turkish	37	0.146718147
portuguese	47	0.139817629
vietnamese	22	0.13961039
amharic	20	0.139285714
russian	47	0.136778116
english	533	0.133208255
macedonian	23	0.130434783
german	36	0.128968254
spanish	144	0.123015873
french	59	0.121065375
bulgarian	16	0.120535714
romanian	17	0.100840336
hindi	15	0.095238095
serbian	15	0.066666667

A.6 Tableau Table