Lab 3: Vowel Space

Winda Hao, 260742006
Jiayao Zhang, 260767074
Joy Han, 260845974

0 Introduction

A vowel space chart is used to describe vowel locations in acoustic and articulatory space. In the start of this course, we were taught that the two-dimensional vowel space chart corresponds to a speaker's tongue-body position, with is indicative of the tongue's high versus low, and front versus back positions-- the articulatory space. Having said that, we want to explore the acoustic vowel space of an English-speaking female. The vowel diagram we will interpret examines the first two vowel formants, F1 and F2. The vertical axis of the vowel chart represents the frequency of formant F1, i.e. the height dimension, and the horizontal axis represents the frequency of formant F2, i.e. the front-back dimension.

In general, we expect our speaker's vowel plot will show similar patterns to that of the standard IPA chart (**Figure 3**). However, we do not expect our speaker's vowel space to perfectly match **Figure 3**, due to English dialectal variation. In particular, we expect our speaker's vowel chart to show a presence of Canadian raising for vowels, i.e [aɪ], before voiceless consonants, since our speaker is Canadian and was raised in Toronto for most of her childhood. So, in the case of [aɪ], its position on the vowel diagram should exhibit lower formants F1 and F2. This may be because the nucleus of the diphthong [a] drifts to the standard position of /ʌ/. Also, we expect to see an overlap in the vowel space shared by low back vowels [ɑ] and [ɔ] in the vowel diagram. Since [ɑ] and [ɔ] are mergers in North American English, they would have similar pronunciations. Lastly, we believe that [u]'s F2 formants could drift depending on its preceding and following consonants. It may move to a fronter position on, due to dialectal variation in North American English (**Figure 4**). These frontward drifts in the front-back dimension may be due to [u]'s environmental influences, i.e. [u]'s preceding and following sounds.

We run our experiment by obtaining a sound file from a young, Torontonian, female speaker. For monophthongs and the nucleus of diphthong tokens, we use the lab's provided script to measure the most stable region of F1 and F2 formants in the spectrogram.

1 Methods

To generate our vowel diagram, we obtained a 96-word sound file uttered by a native English speaker. The speaker is a twenty-one-year-old female born in China, but raised in Toronto, Canada. Although she learned Mandarin as her first language, we define her as a native English speaker since she started to learn and speak English at the critical age of four.

The 96 words were uttered in a relatively quiet room to reduce external noise. The words' vowel tokens were the main focus in our experiment, since we are interested in investigating the speaker's vowel space. After articulating each word and observing the vowels' spectrograms, we noticed the height of vowels correspond to formant F1, and the front and backness of vowels correspond to formant F2. For words with more than one vowel, the vowel token of interest was provided in the lab instructions. Lastly, for diphthongs, we focused on plotting their nucleus.

To create the vowel diagram, we paired up the sound file with the provided script in order for us to indicate F1 and F2 formants on the segment of the spectrogram which best represents the vowel token of interest. **Figure 1** and **Figure 2** are examples of how identified the vowel formants.

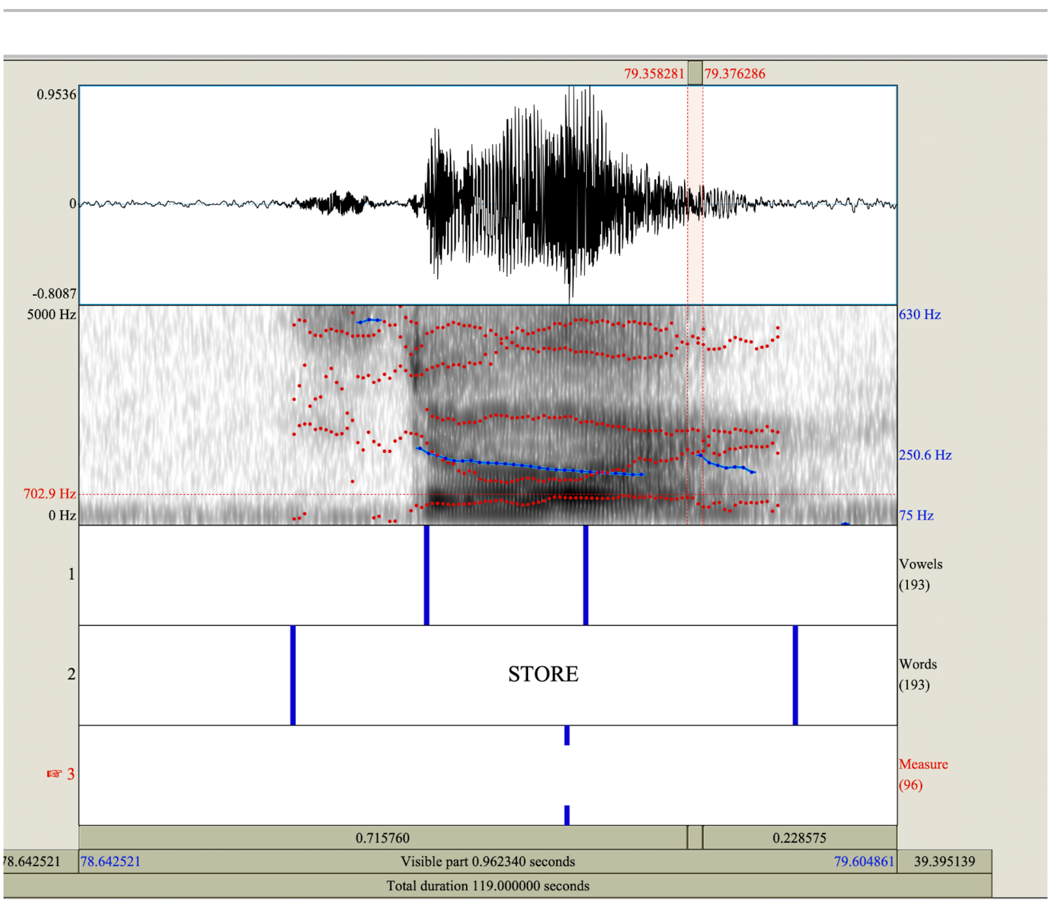

Figure 1

Figure 1 shows how we measured rhotacized vowels. It is hard to draw a clear boundary between a vowel its following rhotic; however, we identified the rhotic by examining its gradually increasing F2 and gradually decreasing F3. **Figure 1** shows how we marked [ɔ] in *store*; F2 and F3 gradually converge together as the vowel gets rhotacized. Also, there is greater energy associated with the vowel's acoustics, and this can be observed by its darker F1 and F2 formants, compared to the lighter F1 and F2 formants of the rhotic. Likewise, the vowel's formants are relatively stable.

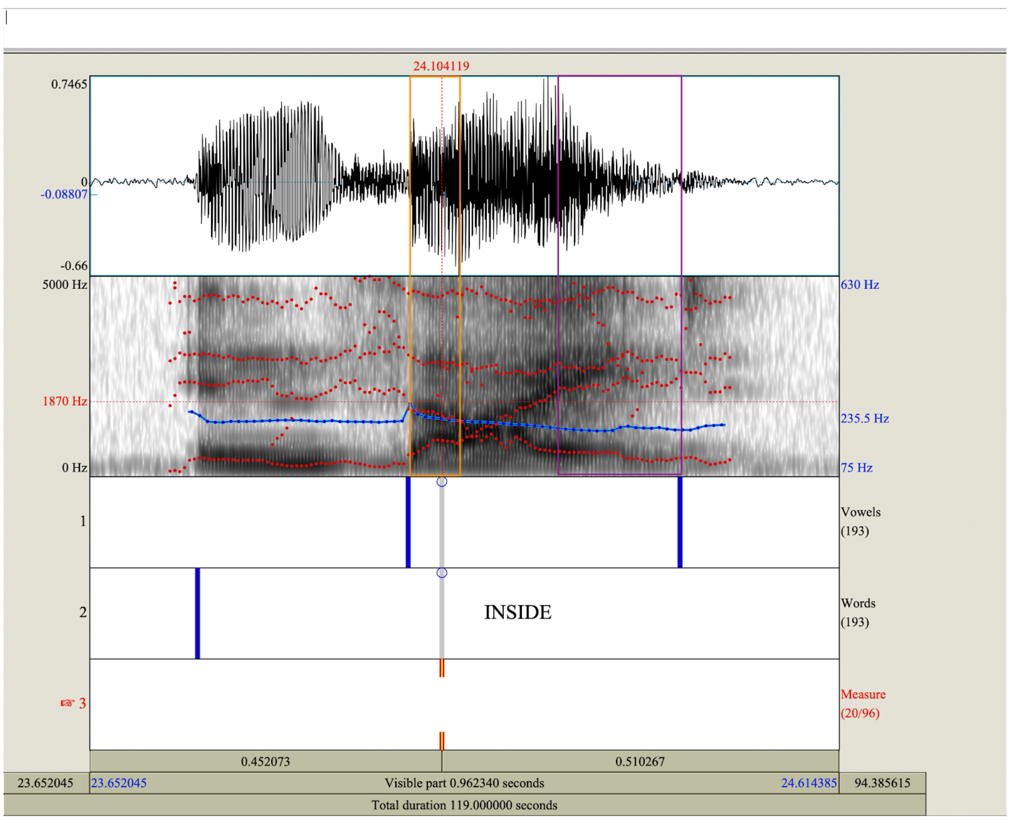

Figure 2

Figure 2 shows an example of how we distinguished a diphthong's nucleus. To optimally mark [aɪ]'s nucleus [a], we closely examined its F1. As learned, when our pronunciation shifts from [a] to [ɪ], the constriction moves from a lower to a higher position, which results in the lowering of F1 and a sloping upwards of F2. We think that the purple box is where the [ɪ] is pronounced, and the orange box may be where [a] is pronounced; however, it is hard to draw a clear line boundary between the two vowels.

2 Result

Upon examining our speaker's vowel diagram (**Figure 5**), we see that it is overall similar to the standard diagram (**Figure 3).** But after closely comparing the two figures, we observe some differences between certain vowels (i.e. [u], [ʊ], [aɪ], and [ɔ]) and their locations on our speaker's vowel diagram.

For high back vowel [u], its position on **Figure 3** is particularly back with very low F1 and F2 formants. In comparison, **Figure 5** shows that some of our speaker's [u] tokens have significantly higher F2 formants, distinguished by its considerable fronting. Additionally, it drifts slightly upwards, like for [u] in *renewed*, which makes it have a slightly lower F1 formant.

[ʊ] also differs in its vowel space positioning. Note, the standard space of [ʊ] is high and back. When we compared our [ʊ] to its standard version, [ʊ] drifted to a slightly lower and more central position. This means that its F2 formant is higher and moves to a more central position. Likewise, its height is near the middle of the diagram, with a higher F1 formant.

The third vowel, diphthong [aɪ], also drifts. The standard position of /aɪ/ shows it is low and relatively central. In comparison, [aɪ] in **Figure 5** drifts upwards on the vowel diagram. [aɪ] is observed to have a lowered F1.

The last vowel we observed which differed relative to the standard vowel diagram is [ɔ]. It concentrates at low back position, with a low F2 and relatively high F1. In our results, F1 of [ɔ] widely ranges from 550 to 900 Hz.

Having said that, we also noticed that there is no clear distinction between the front low vowels [ɛ], [æ] and [a], unlike how they are represented in **Figure 3**. **Figure 6** shows that the words uttered by our speaker with these three vowels highly overlap. For this lab, we were asked to explore whether there is evidence for Canadian raising in our speaker's speech. We hypothesized that our acoustic vowel diagram would show Canadian raising of [aɪ] before voiceless consonants. In particular, the values of both F1 and F2 formants in words that have voiceless consonants /t/, /s/, and /f/ following [aɪ] would be lower than words that have voiced consonants /d/, /z/ and /v/ following [aɪ]. However, as **Figure 5** shown, the nucleus of [aɪ] has high F2 formants, similar to that of the standard vowel diagram in **Figure 3**. This suggests that [a] does not shift to become more similar to [ʌ], due to the fact that F1 for [a] tend to become lower than

[a] in the IPA chart while formant F2 remains high. So, there is no obvious Canadian raising for [aɪ] in our experiment.

For our third main hypothesize, we predicted that [ɑ] and non-rhotic [ɔ] would merge together. By examining **Figure 4**, we see there is no clear distinction between [ɑ] and non-rhotic [ɔ]. They are both located in low and back positions (F1 = 800～950 Hz, F2 = 900～1200 Hz). Having said this, we were not able to clearly draw a boundary to distinguish these two token groups.

For the next question, we examined the distribution of [u]. According to **Figure 3**, [u] is a high back vowel. However, from what we discussed in class, [u] can shift to high front positions in American English dialects. We predicted that [u] would not always be pronounced as a high back vowel. From **Figure 6,** we see that formant F2 of [u] is distributed from ~2200 Hz to 700 Hz. When [u] is followed by [ɫ], it stays in its high back position. On the other hand, [ju] and [u] followed by a voiceless stop moves to a more front position.

Another observation is that the speaker of our dataset demonstrates a rhotic dialect when she utters words with vowel token [ʌ]. When [ʌ] is followed by [l], the speaker pronounces the word more back, with a lower F2 for [ʌ].

3 Discussion

Our results disprove our hypothesis that we would observe Canadian Raising in our speaker, according to **Figure 5-7**. As we learned, in Canadian English, diphthongs [aɪ] and [aʊ] begin with higher vowel [ʌ] instead of [a] before voiceless sounds. From our data, the nucleus of [aɪ], [a] remain at a low front position, with relatively high F1 and F2 values. Even though our speaker is from Toronto, she does not exhibit Canadian Raising when we look at her uttered words with [aɪ] in her vowel space. She does, however, have a Canadian accent, observed by the centralizing[1] of low back vowels. **Figure 7** indicates that the pronunciation of [ʌ] converges to [a] and [æ]. There might be a dialectical change in the Toronto area over time, due to the city's increasingly multi-ethnic population (similar to Cockney accent changes in London), which may have weakened Canadian raising amongst the younger populace.

We also predicted that [ɑ] and non-rhotic [ɔ] would merge, and our results support this. We observed that [ɑ] and non-rhotic [ɔ] are concentrated at low back positions, where F1 equals ~800 - 950 Hz, and F2 equals ~900 to 1200 Hz. The merge may be caused by the vowel shift and centralizing of low back vowels in Northern American English dialects[1]. [ɑ] and non-rhotic [ɔ] moved in the process of centralization to have a higher F2 and became indistinguishable. However, rhotacized [ɔ] tokens did follow this merging trend; their F1 formant lowered even more. According to spectrogram, we see that [ɹ] has a lower F1 formant in general. Therefore, we assume that during coarticulation, [ɔ˞] would be strongly influenced by [ɹ] and obtain a relatively low F1 value (F1 = 450 - 650Hz), i.e. [ɹ] would impose constraint around the alveolar ridge and lip which lowers the F1 of [ɔ].

Our last hypothesis which examines the pronunciation of [u], the results mirrored our prediction, which is the distribution of [u] depends on the environment of the [u] token, its front-back dimension may be affected by its preceding or following consonants. In particular, we found that if [u] is followed by a [ɫ] such as in *pool, tool, and school*, [u] is pronounced much more back, with a lowered F2 (~800 - 900Hz). The vocal tract would be more constricted at the velar position in order to prepare for the production of the [ɫ] sound. Also, [ɫ] does not change the lip roundness of [u]; therefore, it helps to maintain the low F2 of [u]. However, [u]'s F2 increases significantly when [u] follows /j/ (*renewed*). Onset [j] requires spread lips to be produced, and will affect [u]'s lip roundness due to coarticulation. Spread lips release the constriction on maximum pressure and raises F2 to 2400 Hz. When [u] is followed by voiceless stops (*shoot, cooped*), it moves front. We suggest that the fronting of [u] could be influenced by the front articulatory position involved in the production of [t] and [p]. Also, it may be due to the short vowel duration before voiceless stops, that [u] is not clearly pronounced. The pattern in our vowel diagram is similar to the patterns in **Figure 4** and may reflect dialectal variation in North American English.

In **Figure 5**, we see that [ʌ] is more centralized and clearly separate into two parts. We believe that this distribution might be affected by the coda of the syllable. When the coda is a voiceless consonant other than /l/ (based on data set we have), the vowel [ʌ] has a higher F2. As we discussed previously, there may be centralizing of [ʌ] due to our speaker's dialect. On the other hand, we observed that the F2 of *gulf*, *culvert*, *dull* are ~1200 Hz, whereas the F2 of the words *tuttle*, *dust* and *rushed* are ~1600 Hz. When the coda is a /l/, the vowel [ʌ] has a lower F2. Perhaps this has to do with coarticulation. In order to transition from [ʌ] to /l/, our speaker imposed a constriction around her velum, around the maximum velocity of F2, and produced [ʌ] with lower F2.

We also observed fronting and centralizing of [ʊ], which suggests that the words *could*, *would*, and *look*, were pronounced more like a [ə] by our speaker. According to the IPA vowel chart, /ʊ/ is a high back vowel with low F1 and F2. In our result (**Figure 5**) shows, [ʊ] has a higher F1 (~650 Hz) and F2 (~1450 Hz), and is a fairly central in its position, which is similar to where [ə] is located (**Figure 3**). The distribution of the words with vowel token [ʊ] is concentrated together relatively well in our speaker's vowel diagram, which may indicate that the shift of place and manner of articulation might not be a result of the vowel environments. This could be due to the speaker's accent. As we have stated, the speaker as a Torontonian might shift the back vowels to a more center position.

As we mentioned in Result, vowel tokens [ɛ], [æ] and [a] are not separable based on F1 and F2 formants. According to **Figure 6** and **Figure 7**, these vowels overlap. In the standard IPA chart, [a] has a higher F1 compared to [ɛ] and [æ]. In contrast, our speaker's F1 for [a] is similar to [ɛ] and [æ]'s F1. The lowered F1 of [a] may be because when producing the diphthong [aɪ], our speaker imposes a bigger constriction at the front of vocal track for [a] in order to easily produce the [ɪ] sound which is constricting at the alveolar ridge. We propose that the position of [ɛ] may be affected by its following consonant. For example, the reason why *sell* has a higher F1 (~ 900 Hz) may be because its coda consonant [ɫ] has a constriction around the velum, which raises F1.

For the unknown vowel *X*, we think the vowel could be [a]. As **Figure 5** shows, the *X* vowel has ~900 Hz for F1 and ~1500 Hz for F2. When we examine these frequencies in **Figure 7**, we observe that the words with vowel *X* mainly overlapping with the words contain [aɪ] which were measured based on [a], we believe that vowel *X* might [a].

4 Conclusion

To tie everything together, the purpose of our experiment was to investigate and interpret the English vowel space generated by a young, female, Chinese Canadian speaker. We were particularly interested in examining the effects of Canadian raising, low back vowels, and the pronunciation of [u].

From this lab, we gained a better understanding of how a speaker's individual vowel space may differ from the idealized IPA vowel chart, depending on which dialect of English they speak. We learned that there may be considerable overlap in terms of the height and front-back dimensions of certain vowels in terms of the space they take on the vowel chart, and this is most likely due to the wide range of dialectal variations existent in the English language.

5 Appendix

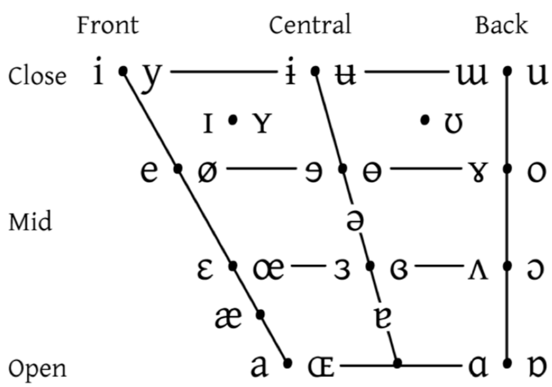

Figure 3

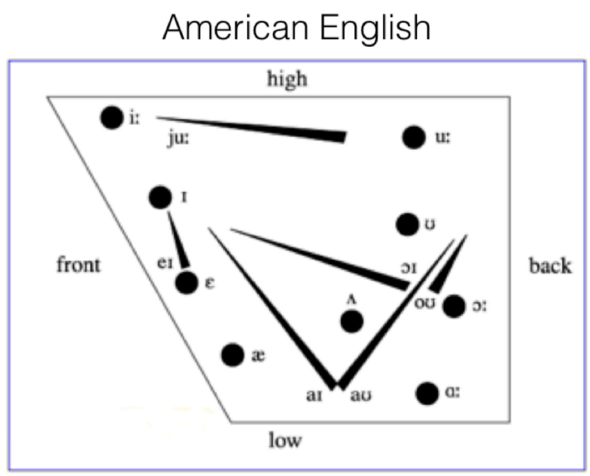

Figure 4

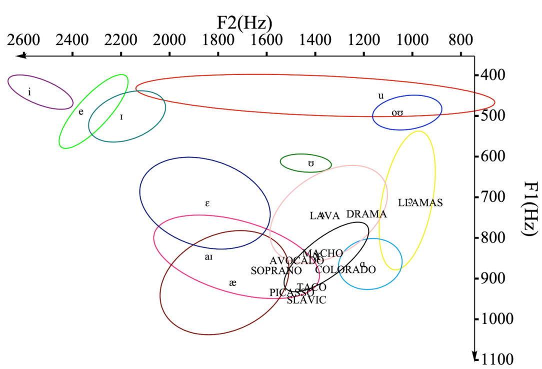

Figure 5

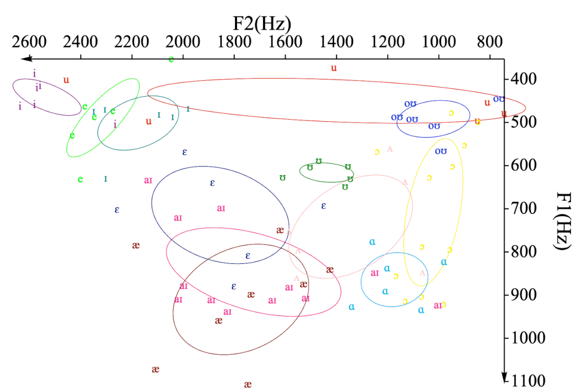

Figure 6

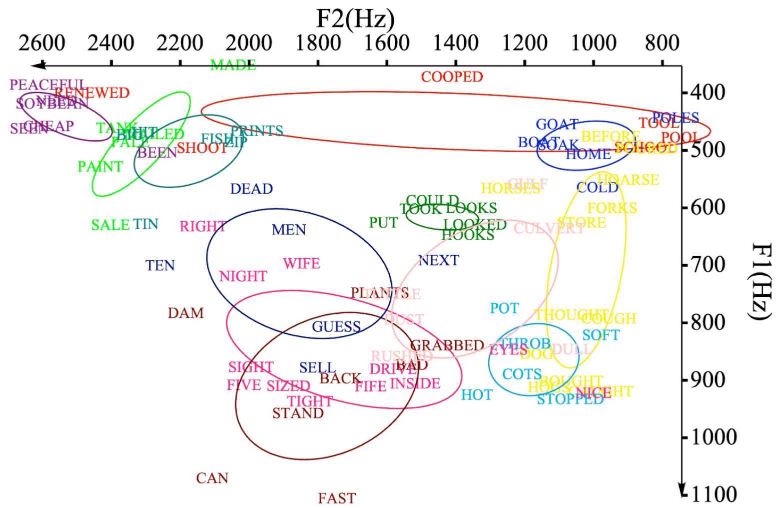

Figure 7

Word	Vowel	F1	F2	F3
COLD	oʊ	565	988	3095
SOAK	oʊ	492	1100	2416
POOL	u	479	744	2794
SHOOT	u	497	2133	2581
FIFE	aɪ	912	1646	2488
MADE	e	353	2045	2692
THROB	ɑ	837	1199	2389
GRABBED	æ	840	1424	2336
BEFORE	ɔ	477	948	2509
WIFE	aɪ	698	1846	2526
CAUGHT	ɔ	921	978	2541
NIGHT	aɪ	720	2015	2651

TOOK	ʊ	603	1500	2664
FAST	æ	1106	1745	2775
COULD	ʊ	588	1466	2626
HIT	ɪ	470	2302	3012
CULVERT	ʌ	637	1130	2875
DRIVE	aɪ	881	1580	2449
COOPED	u	373	1409	2580
INSIDE	aɪ	907	1515	2638
NICE	aɪ	923	998	2377
THOUGHT	ɔ	787	1063	2578
POLES	oʊ	444	760	2734
ZIP	ɪ	487	2037	2852
PALE	e	487	2345	2980
SCHOOL	u	496	847	2438

FIVE	aɪ	909	2013	2514
DAM	æ	784	2182	2707
SIGHT	aɪ	878	1993	2937
BEEN	i	505	2264	2988
DEAD	ɛ	568	1991	2645
FISH	ɪ	481	2091	2799
COLORADO	X	879	1274	2228
SEEN	i	463	2637	2789
MEN	ɛ	639	1883	2769
BOAT	oʊ	487	1159	2516
NEXT	ɛ	692	1449	2160
SOYBEAN	i	420	2571	2889
DUST	ʌ	795	1549	2423
BAD	æ	874	1527	2520
COUGH	ɔ	794	955	2529
PLANTS	æ	749	1618	2296
HOARSE	ɔ	552	897	2528
DOG	ɔ	855	1165	2405
PEACEFUL	i	387	2580	3101
HOGS	ɔ	913	1129	2500
HORSES	ɔ	567	1238	2236
LLAMAS	X	714	965	2532
GOAT	oʊ	456	1107	2462

TIN	ɪ	629	2299	3209
CAN	æ	1071	2105	2649
LOOKED	ʊ	630	1343	2671
BACK	æ	898	1732	2466
AVOCADO	X	857	1477	2519
FORKS	ɔ	601	944	2258
MAILED	e	473	2273	2965
RENEWED	u	401	2455	2812
STOPPED	ɑ	934	1066	2441
GULF	ʌ	560	1189	2486
TACO	X	922	1415	2434
STAND	æ	958	1858	2660
BOUGHT	ɔ	903	1064	2485
TANK	e	462	2383	2781
TOOL	u	453	809	2508
STORE	ɔ	627	1034	2468
SELL	ɛ	879	1800	2551
TEN	ɛ	701	2257	2853
CHEAP	i	459	2580	2970
SLAVIC	X	953	1438	3127
NEED	i	415	2558	2674
RIGHT	aɪ	633	2131	2480
HOOKS	ʊ	648	1363	2480

RUSHED	ʌ	859	1554	2418
COTS	ɑ	890	1205	2493
SALE	e	631	2400	2659
SOPRANO	X	881	1558	2660
PICASSO	X	935	1495	2485
PRINTS	ɪ	467	1977	2746
TIGHT	aɪ	938	1821	2638
SIZED	aɪ	911	1885	2592
HOME	oʊ	507	1014	2486
MACHO	X	839	1368	2649
PUT	ʊ	627	1608	2573
PAINT	e	529	2432	3188

LOOKS	ʊ	601	1353	2713
DULL	ʌ	847	1062	2577
SOFT	ɑ	822	976	2416
LAVA	X	746	1365	2546
HOT	ɑ	926	1337	2519
BIG	ɪ	474	2345	3072
POT	ɑ	776	1257	2529
DRAMA	X	742	1186	2278
EYES	aɪ	848	1245	2528
TUTTLE	ʌ	751	1584	2509
POURED	ɔ	498	844	2380
GUESS	ɛ	808	1745	2367

Figure 8

Reference :

1. Chambers, J. (1973). Canadian raising. Canadian Journal of Linguistics/Revue Canadienne De Linguistique, 18(2), 113-135. doi:10.1017/S0008413100007350