Preliminary Paper

### Introduction

Treated as a social construct, self-identification along ethnic lines can be viewed through the lens of Social Identity Theory which argues that human categories are prototypes of integrated attributes (Hogg, 2016). In a nutshell, Social Identity Theory is defined as an individual’s knowledge that they belong to a social group and they know the emotional and social significance of the group, i.e. they know what it means to “be” a member of a group. This self identification forms an individual’s identity which is then mediated with the use of language particularly between two or more interlocutors (Ochs, 1993). This is all done by the interlocutors communicating by shared conventions that signify their identity and generate social meaning. One of the many forms of communicating conventions is through the production of specific sounds (Eckert & Labov, 2017) and stylistic choices in speech (Eckert & Podesva, 2011).

In the paper *Phonological Correlates of Ethnic Identity: Evidence of Divergence?* by Matthew J. Gordon, we find evidence of vowel production differences between four ethnic groups, White American, Mexican American, African American and biracial Americans. These differences are found along ethnic lines with regards to the Northern City Shift (NCS). African American and Mexican American speakers did not follow the general trend of /æ/ raising the same way that White Americans did. There were also fewer instances of “innovative” forms of production, i.e. White Americans led the change in vowel shifts while African and Mexican Americans tended to be more conservative in their speech. In Cananda, there has been evidence suggesting that Chinese Canadians had a vowel pattern that shifted differently when compared to White Canadians (Cardoso et al, 2016),

This previous research suggests to me that vowel differences will be present along ethnic lines in California. California is experiencing a shift known as the Low-Back Merger (LBM) which involves the merge of the /ɔ/ and /ɑ/ vowels (Doernberger & Cerny, 2008), the raising of /æ/, the fronting of /u/ and the lowering of /I/ and /ɛ/ (Sheydaei, 2024). This paper, by conducting a Linear Discriminat Analysis of ““White” vowels and the “East Asian” vowels from university students in California, will provide preliminary evidence of production differences in line with Gordon (2000), Cardoso et al. (2016), and Sheydaei (2024). A Signal Detection Theory analysis will also be conducted to check the reliability of the results of the LDA.

### Participants

The data for this study comes from 22 undergraduate students at the University of California, Davis. 12 self identified as East Asian (i.e. ethnically descendant of Japanese, Chinese, and Philippine peoples) and the remaining 10 self identified as White (i.e. ethnically descendant of European peoples). Each participant were recruited through the UC Davis SONA system.

### Data Collection

The procedure the participants followed was a straightforward one. After reading over information related to the study and signing a waiver allowing me to record them, they entered into a recording studio. After testing the sensitivity of the microphone I presented them with a Google Slides presentation containing the words I asked them to read. They were told that they would see a word on the screen and have 2 to 3 seconds to read the word before I switched to the next. They were told that not every word was a “real” word and to try their best.

Each word was typed in bright orange and presented against a white background. From my position outside of the recording booth, I clicked from slide to slide, word to word about every two seconds and made notes of words that the participant had difficulty with pronouncing. In a couple of instances, I asked participants to repeat a word after they had read through the entire list twice. For example, many participants would pronounce the word “head” in the same way the word “heed” is pronounced. Some words were unfamiliar to the participants even though they were real words. A couple were unsure about the pronunciation of the word “gin” and would produce the “g” with the voiced velar /g/ instead of the affricate /dz/. In total, each participant would repeat every word in the list twice. The list of words included HvD, BvD, BvT, pool/pull merger, pen/pin mergers, and two diphthongs the /oi/ before /l/ found in soil and the /ai/ found in sigh. The HvD words were to find neutral uses of the 10 English vowels (including /r/).

After the participants had completed their time in the recording booth, they were asked to answer a short questionnaire to gather specific information on their demographics. These included: what is your gender identity, what is your ethnic background, what area of California did you grow up in? These questions in an open-ended format because I wanted them to think about how they identified and I felt that yes or no questions were inadequate. I recognize my role in influencing the answers however as they could only answer the questions I asked them narrowing the ways in which they could identify. I asked these questions after they completed their time in the recording booth because I wanted them to read the provided stimuli as naturally as possible.

I recorded the participants using Praat and sampled at 44100 hertz. Each participant recording was done in a single wav file unless I had to re record an individual’s production. In that case, I created a separate wav file that contained only the repeated word or words. I then used the cut and paste feature in Praat to paste the repeated word into the original wav file. After the recording had finished, I took the wav file recordings and hand annotated them to align the spectrographs with the written words. Following this, I used DARLA (Dartmouth Linguistic Automation) to properly align the segments of the words to their respective slices in the spectrographs.

Once I was satisfied with the alignments, it was time to turn to vowel extraction and analysis. Using the FastTrack plugin for Praat (Barreda…) I extracted the vowels from each individual wav file using the aligned text grids and kept them stored as wav files for each production of a vowel and diphthong. Once completed, I used the TrackFolder command under the FastTrack plugin in Praat to run the LPC analysis to find the frequencies of the first three formants. I then normalized everyone’s formants following log-mean normalization (Barreda et Nearey, 2018).

### Data Analysis

This short paper will review the Linear Discriminant Analysis (LDA) of the “White” vowels and the “East Asian” vowels. Following that, a Signal Detection Theory (SDT) Analysis is conducted for each vowel. This steps are done for three reasons. First, if a model is able to correctly discriminate between the ethnicities using vowel formant data, its suggests that there exist differences in production. Second, If the sensitivity of the model (d’) is 1.5 or higher for each vowel, the model does a pretty good job of distinguishing between the two ethnicities. Third, if the c value is at or near zero for each vowel, then any confusion found by the model is likely due to phonetic overlap. The lda model (lda\_model) is lda\_model <- lda(vowel\_ethnicity ~ f11-f15+f21-f25+f31-f35).

In plain English, I expect the model to predict the relative location of each vowel for each ethnicity within the vowel space. I will make these predictions given the start and end of the first, second, and third formants of each vowel.

### LDA Predictions

|  |
| --- |
| LDA Prediction of Vowels by Ethnicity |

The plot above shows the location of each vowel for the two ethnicities based on what the model predicted. Before I report the results of the SDT analysis I will highlight a three generalizations that I note from the plot. First, between White speakers and East Asian speakers, the front vowels (front of ɝ) are predicted to be different from each other mainly in terms of height. The most dramatic difference appears to be /eI/ and /I/ while /ɛ/ is less dramatic but still noticeable.

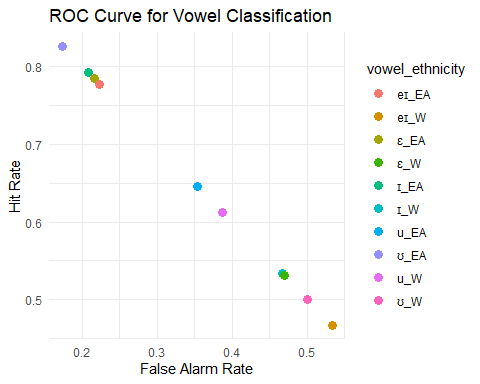
Second, the central vowels (between ɝ and ɔɪ) are not predicted to be different from each other. The model probably finds these vowels to be confusable. While it was generally clear that the East Asian speakers raised the front vowels more than the White speakers, I cannot find a generalizable pattern for the middle vowels.

Third, the back vowels (especially /u/) are predicted to be dramatically different.The most striking prediciton is the difference in the fronting of /u/. /u/ fronting, a relatively common feature in North American English (Havenhill, 2004)(Kataoka, 2010), appears to be much more pronounced in white speakers than in East Asian speakers. The fronting of /ʊ/ in the White population is very surprising. In contrast to the front vowels, the back vowels are predicted to shift more in the White population than the East Asian population.

With specific regards to the LBM, the fronting of /u/ fits the predicted pattern while the lowering of /ɪ/ and /ɛ/ don’t. I suspect that the White population is leading the shift more so than the East Asian population.

### SDT Analysis

The SDT Analysis is to test the reliability of the model. I specifically want to see the hit rate and miss rate for the five vowels mentioned above. Those vowels are, /eɪ/, /ɪ/, and /ɛ/ in the front and /u/ and /ʊ/ in the back. I will also report the sensitivity (d’) and the response bias in an appendix.



As can be seen in the plot above, most of the vowels had a hit rate higher than that of the false alarm rate. The /ʊ/ of the White population was split evenly while the /eɪ/ in that population had more false alarms than hits. Both vowels had a d’ and response bias at or just slightly below zero. This suggests that there exists confusion in the categories and that the model does not have any systemic bias.

In contrast the /eɪ/ and /ʊ/ in the East Asian population had a very high Hit Rate and a moderately reasonable d’ values (1.5 and 1.8 respectably). This could mean that the two populations have true differences in production though some overlap still exists. It could also mean that the White Population had more differences within itself which made the prediction more difficult.

The /ɛ/ vowel for both groups had high hit rates with d’ values of 1.57 for the East population and 1.52 for the White population. The response bias of the East Asian population is -2 suggesting that the model is more liberal in its predictions. However, given the false alarm rate is low, I think this suggests that while overlap of the categories still exists, the model’s bias doesn’t strongly sway the prediction.