LING 401 Final Project:

Physical and Perceptual Differences Between Allophones, Quasi-Phonemes, and True Phonemes in English, Turkish, and Vietnamese

Rosemary Lach

13 December 2021

I. Introduction

The purpose of this study is to analyze the physical and perceptual differences between phonemes that have different forms of phonological contrast in different languages. Particularly, we look at the voiceless palatal and velar stops [c] and [k] in the languages English, Turkish, and Vietnamese. These language and phoneme combinations are of interest, because in English, [c] and [k] are allophones, while they are quasi-phonemes in Turkish, and true phonemes in Vietnamese. Since they all have functional differences in how they are used in the different languages in terms of phonological contrast, we wanted to examine whether this led to any phonetic differences in how the sounds are produced and perceived by speakers of those languages. We tested this by running three tests on speakers of all three languages: spectrum analysis of the produced sounds, palatograms to determine articulation location, and a phoneme confusion test to measure perceptual distance. After running these tests, we were able to see a difference between all three languages in terms of distance between [k] and [c]. True phonemes showed the most physical and perceptual distance, followed by quasi-phonemes, and then allophones. Not all results from this study were significant due to the small sample size of speakers, but this primary finding is clear nonetheless.

All code for the project can be found at https://github.com/rosemarylach/palatal-stops.git. The repository also includes all data generated for this project. There was much more data collected than could be included in this report, so we encourage looking here if you are interested in seeing more intermediate data or data for individual speakers.

Lastly, I would like to thank all of the speakers who volunteered for this project. All of them were extremely cooperative and helpful, and I could not have explored this topic without their participation.

II. Background

The idea for this project stems from the phenomenon in Turkish phonology where the [k] and [g] velar stops are often palatalized to produce [c] and [ɟ] in front of front vowels.

These sounds appear in contrastive distribution, but since Turkish speakers recognize the velar and palatal versions as distinct sounds with “hard” and “soft” forms, the sounds qualify as quasi-phonemes [1]. Since I was interested in continuing this phonetic analysis, as well as linking it to phonology, I became curious about how the quasi-phoneme relationship drove the distance between the sounds because theoretically they could be pronounced very similarly and meaning in the language would not change.

As I searched for languages that had either [c] and [k] or [ɟ] and [g] as separate phonemes, I came across Vietnamese, which uses the voiceless sounds as two separate phonemes [2]. Lastly, I decided to analyze English along with Turkish and Vietnamese because for some speakers, the [k] and [c] sounds are allophones, which provides a third category of phonological contrast for comparison.

Going into this experiment, my predictions were that if there were a phonetic difference between sounds that are phonemes, quasi-phonemes, or allophones, then the true phoneme case would have the sounds be most phonetically distinct both physically and and perceptually. This is because an ability to produce sounds that are distinct is most important when distinguishing between the sounds has an impact on meaning, which is not the case for allophones or quasi-phonemes. I also predicted that allophones would have the least phonetic difference between them because the sounds are consistently used interchangeably already, so at least from a perceptual perspective, they should be the same, which would likely drive the physical distances to be reduced as well. While I expected the phonetic differences of quasi-phonemes to be somewhere in the middle, I did not know whether to expect the sounds to be more similar to true phonemes or allophones from a perceptual point of view. However, since I had already analyzed the waveforms in a previous study, I already knew them to be physically distinct.

Another point of interest that drives some of the tests in this study is that the Vietnamese language does not contain a [t͡ʃ] phoneme, while English and Turkish both do. This creates the potential for Vietnamese speakers to distinguish between [c] and [k] by making [c] sound more similar to [t͡ʃ], which is not an option in the other two languages. In fact, it is actually quite common to slightly affricate [c], which is a way to help distinguish the sound from [k] [3]. For this reason, many of the tests we perform will also include data points for the [t͡ʃ] phoneme even though it is not being directly analyzed, so that we have a reference as to where [c] falls relative to both [k] and [t͡ʃ] since Turkish and English both limit how similar it can be to [t͡ʃ], while Vietnamese has no such limit.

III. Methodology

This experiment was conducted using three native monolingual English speakers, three native Turkish speakers who are English bilingual, and three native Vietnamese speakers who are English bilingual. Each speaker participated in an interview that lasted approximately 30 minutes to provide all the data required to analyze the difference between the [c] and [k] phonemes in their language. We started by asking a few simple questions to get a sense of the speakers' language proficiencies. The questions are as follows:

1. How long have you been speaking English/Turkish/Vietnamese, and is it your first language?
2. Where did you grow up?
3. Would you consider yourself fluent in English/Turkish/Vietnamese?
4. (For the Turkish/Vietnamese bilinguals) Do you feel more comfortable speaking English or Turkish/Vietnamese, or are you equally comfortable with both languages?

These questions helped to verify that all of our bilinguals were in fact native speakers, and provided the opportunity for follow-up questions so we could take note of any important information that may affect the data. Generally, the answers to the questions were as expected, but it was good to validate this before proceeding.

After answering these questions, the subjects were given explanations of the three tests they were about to do: producing recordings of words containing the [c], [k], and [t͡ʃ] phonemes, producing palatograms of the same words, and performing a phoneme confusion test. The detailed procedures for each test are outlined in the sections below.

A. Recordings

In order to take recordings of each of the phonemes of interest, each speaker was presented with three simple words on a sheet of paper in their native language, one containing the phoneme [c], one containing the phoneme [k], and one containing the phoneme [t͡ʃ] (in this order). The Turkish and Vietnamese speakers were then asked to translate each of the words in their list to English, to make sure that they knew the word and could pronounce it naturally. Afterwards, they used a voice recording program on the test laptop to record themselves saying each word.

Since the nature of this study dictates that all of these sounds are not necessarily present in the same form in each of the languages, the chosen test words in each language were slightly different in a few ways. Firstly, the Vietnamese language does not have a true [t͡ʃ] phoneme (typically the closest thing is a slightly affricated [c]), so in order to elicit this sound, we simply had the Vietnamese speakers pronounce that sound in an English word. Additionally, since [c] and [k] are quasi-phonemes in Turkish, the words to produce these sounds must contain the orthographic ‘k’ before a front vowel and back vowel respectively. In English, the [c] phoneme tends not to exist at all, but for the sake of consistency, we tried to produce as similar a sound as possible by placing the typical [k] sound before a high vowel to see if there was any measurable difference, since for some speakers, words like “key” become slightly palatalized. Lastly, since both the Turkish and English test words use /k/ before a front vowel in an attempt to produce a realized [c], we also considered placing the contrastive Vietnamese [c] and [k] before front and back vowels to see if that produced any phonetic difference within the language itself. While we were able to find Vietnamese words that included [c] before front and back vowels, we were unable to find words in which [k] preceded a front vowel, which could be indicative of the fact that [k] is a more natural sound to produce before a back vowel. Because of this, we did add a fourth word to the Vietnamese test set to account for [c] before front and back vowels, but since we couldn’t do the same for [k], we did not plan to do extensive analysis on the differences between these sounds within the Vietnamese language.

The test words for each language were selected based on these restrictions, and are presented in the following table.

Language	Phoneme	Word	English Translation	IPA
English	[c]	key	key	[ci]
	[k]	cop	cop	[kap]
	[t͡ʃ]	cheap	cheap	[t͡ʃip]
Turkish	[c]	iki	two	[ici]
	[k]	kapı	gate/door	[kapɨ]
	[t͡ʃ]	çim	grass	[t͡ʃim]
Vietnamese	[c] (before front vowel)	chia	divide	[cia]
	[c] (before back vowel)	cha	dad	[ca]
	[k]	cao	high	[kao]
	[t͡ʃ]	English(cheap)	cheap	[t͡ʃip]

B. Palatograms

Another benefit of the words chosen above is that, when pronounced, the only time the tongue makes contact with the roof of the mouth is during the phoneme of interest because all other sounds are either vowels or bilabial stops. This was intentional because it allows us to measure exactly where this contact occurs using palatograms. This section of the test was optional, since it is a little more involved and posed potential covid risks for subjects who may have felt uncomfortable. However, all but one Turkish speaker chose to participate in this portion of the study.

Before the interview began, I prepared a mixture of charcoal powder (poured from the inside of a charcoal capsule purchased at a local pharmacy) and canola oil to create a viscous black fluid. Once the subject arrived and we reached this part of the interview, they would apply a layer of this mixture to their tongue using a cotton swab so their entire tongue was black. Then, they pronounced one of the words from the list in the previous section. Finally, they rested a mirror on their bottom lip so the roof of their mouth was visible, and I took a picture of it so that the image could be analyzed later. The subjects were instructed not to let their tongue make contact with the roof of their mouth between applying the mixture and taking the photo (other than when the actual word was being said), so that no extraneous marks appeared on the roof of the mouth. This process was repeated for all words on the list of test words, and the subject would rinse out their mouth with water in between each palatogram.

C. Phoneme Confusion Test

While the previous two tests were meant to analyze the physical distance between sounds, a phoneme confusion test is a way to measure the perceptual distance between sounds. Since we are primarily looking to measure the distance between the voiceless stops [c] and [k], and would also like to obtain similar information for [t͡ʃ], we decided to include all consonants from the series of voiceless stops and affricates going from alveolar to velar ([c, k, t͡ʃ, ʃ, t]), in order to account for all potential sources of confusion. Each consonant was then paired with a vowel from the set: [a, o, u, ɨ, ɔ, ʌ, ɛ, æ, ø, y, e, i]. This vowel list deliberately contains an even number of front and back vowels since this environment has the potential to influence how certain sounds are perceived, especially when these vowels are the conditioning environment for Turkish quasi-phonemes. The consonants were paired with the vowels by placing them in word initial position. Essentially, the samples were all sounds like [kɛ] or [ʃø], but nothing like [cuk] or [ty]. All 60 initial consonant and vowel pairs were then recorded on the test laptop and saved to a wav file with background noise added to induce errors when the test was taken.

To prepare the actual tests for the speakers to take, we generated the python scripts Confusion-English.py, Confusion-Turkish.py, and Confusion-Vietnamese.py. The majority of the code was identical for each of the three tests. They read through all wav files in a directory (in this case we pass in the directory with the 60 consonant-vowel pair recordings we just generated), play the sound and present a series of options to the subject to select a word in their language containing the same initial consonant as the sound they heard. The program then tallies the results and outputs a file containing information on the played sound, the selected sound, and correctness. The difference between the three files was the list of words that displayed on the screen for the participants to select from. I wanted these to be in the subject's native language because I wanted to introduce as little bias as possible into the experiment.

However, there were some issues with choosing words to display as selection options. Particularly in Vietnamese, we were unable to find words that began with [t͡ʃ] or [ʃ]. For these sounds, we simply included English words that began with these sounds and added a small note to the option indicating the word was in English. After running the first Vietnamese test, I also learned that Vietnamese speakers will often not aspirate their [t], and since some of the recordings contained aspiration, the recording sounded like it did not match any of the available options. Since the [t] sound was not an important data point for the purposes of this project, I decided to rectify this by adding a disclaimer before running the test that both aspirated and unaspirated [t] sounds should have the unaspirated Vietnamese [t] selected as that was the only possible [t] choice. Lastly, English doesn't actually contain the sound [c], so in order to allow English speakers to select this option, we included the word "key" and explained beforehand to only select this option if they believed the sound they heard was articulated farther forward than the default [k] sound. We then provided an exaggerated example of this sound being pronounced as [c], so they could hear the difference at least once before taking the test. In addition to these disclaimers before Vietnamese and English tests, all participants were instructed to only select the ??? option if they did not hear the target sound at all. Despite all of these challenges, we were finally able to decide on the following words to present as options for the speakers of each language.

Language	Phoneme	Word	English Translation
English	[c]	key	key
	[k]	car	car
	[t͡ʃ]	chat	chat
	[ʃ]	ship	ship

	[t]	tap	tap
Turkish	[c]	kir	dirt
	[k]	kol	arm
	[t͡ʃ]	çok	a lot
	[ʃ]	şeker	sugar
	[t]	tuz	salt
Vietnamese	[c]	cha	dad
	[k]	cơm	rice
	[t͡ʃ]	English(chat)	chat
	[ʃ]	English(ship)	ship
	[t]	tên	name

Once all the instructions were given, the subjects used the test laptop to take the confusion test, and the results file was saved to the laptop.The procedure to make this data usable for analysis is discussed in the following section.

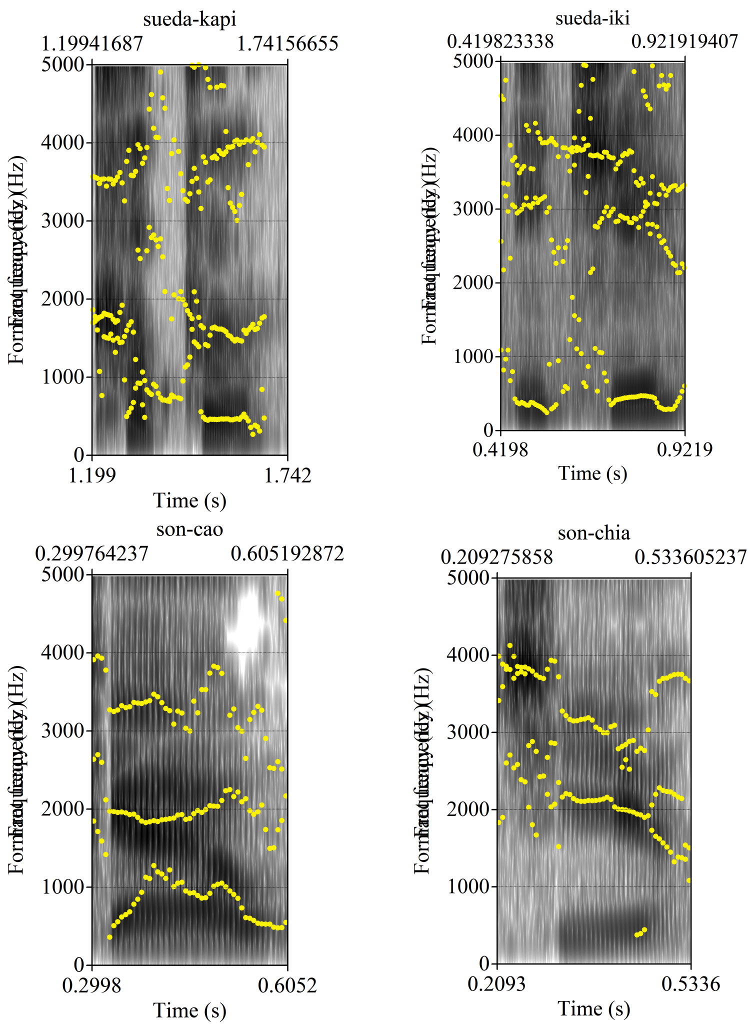
IV. Analysis and Results

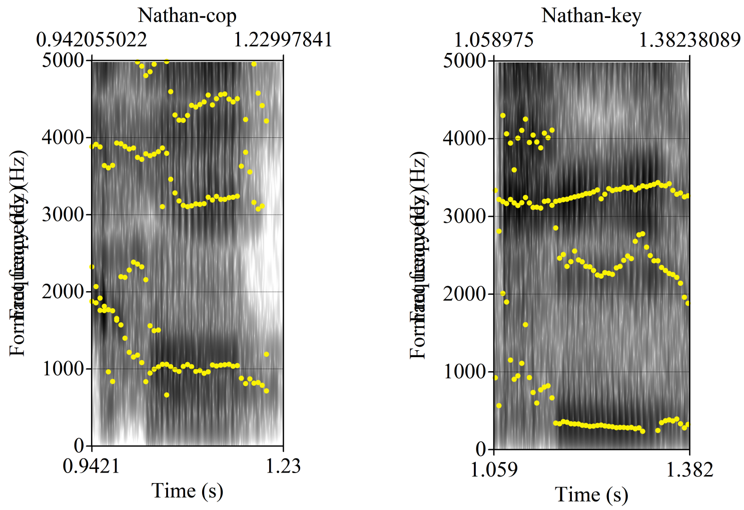
A. Spectral Analysis and Center of Gravity Calculations

To prepare the recording results for spectral analysis, each recording sample was converted into .wav format and loaded into the program Praat. To generate the spectrograms,we simply use Praat's automatic function to do this, and zoom to the portion of the recording where the word was being pronounced to draw the spectrogram and formant structure and save to a png file. To generate the power spectrum for each sound, we select a moment in time in the spectrogram where the consonant of interest was being pronounced, and let Praat generate the spectrum at that time to save to a png file. At this same point in time, we also record the center of gravity.

Here we qualitatively analyze the level of palatalization of the [c] and [k] sounds from each language by looking at the spectrogram and power spectra. The main marker of a velar sound in a spectrogram is the "velar pinch" which is when the second and third formants of a vowel adjacent to a velar consonant pinch together at the boundary of the consonant (here we will look for pinches going into the vowel since most of our test words contain initial consonants) [4]. A palatal sound should not show this pinching. Another way to distinguish between velar and palatal sounds in the power spectrum is to look for a more diffuse spectrum with more energy in higher frequencies for [c] because it is pronounced further forward. This idea can also be modeled by the center of gravity of the spectrum at these points. So before doing the tests, I would hypothesize that the palatal sounds in both Turkish and Vietnamese would show less pinching in formants 2 and 3, and have a higher center of gravity.

After looking at the final spectrogram images, the formant structure was unfortunately somewhat inconsistent and difficult to analyze. While some of the better quality spectrograms showed evidence of pinching for velar sounds and no pinching for palatal sounds for certain speakers of both Turkish and Vietnamese as shown below, this was not generalizable to all spectrograms (see the Spectrograms folder in the project github for all images), meaning that we cannot make any real claims about the physical difference of these sounds based on the spectrogram images alone.





Palatograms that do show evidence of velar pinching (left) as opposed to no velar pinching (right) in Turkish, and Vietnamese (as read from top to bottom). The bottom row shows English sounds in "cop" and "key", both of which contain the velar pinch. Note that the formant lines do not always perfectly line up with F2 and F3.

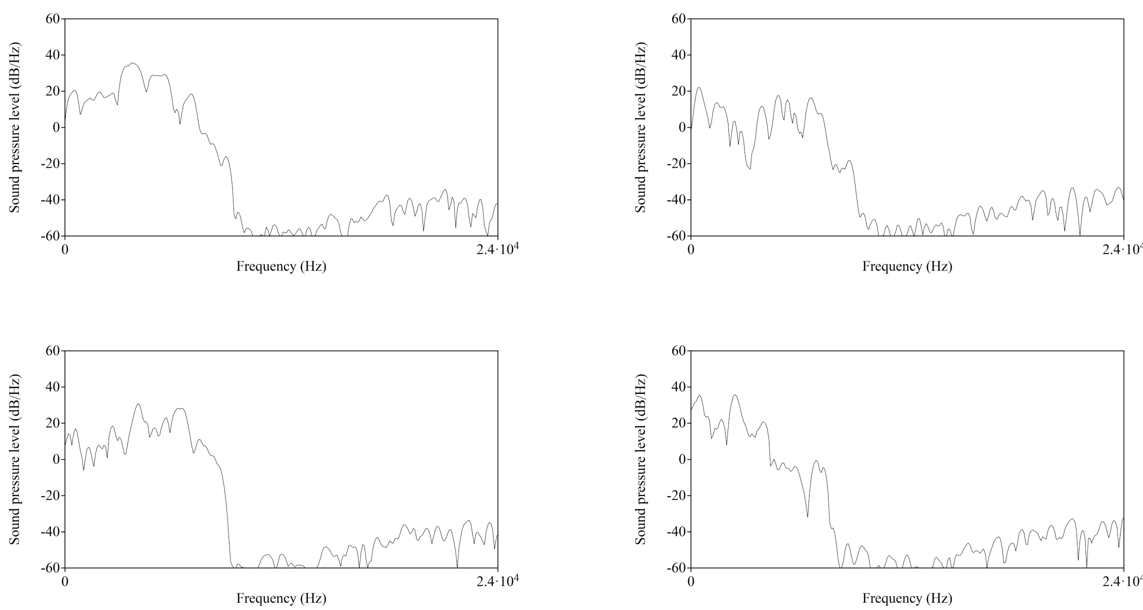
Using the data from the power spectra, we then calculated the average centers of gravity of [c], [k], and [t͡ʃ] in English, Turkish, and Vietnamese. Note that one of the English speakers' recordings was very noisy, which caused the center of gravity to be unreasonably high, so this value was left out of the average calculation for the English [c] sound. The table below contains the results of all average centers of gravity.

	[c]	[k]	[t͡ʃ]
English	3865.39	3140.697	4363.21
Turkish	4164.843	2293.77	4671.44
Vietnamese	4759.642	1854.78	4664.727

As expected, the [c] phoneme had a higher center of gravity for all three languages, but more significantly for Turkish and Vietnamese, likely because they are actually sounds in the language, whereas the difference in English is probably just a result of the surrounding vowel sound. What is also interesting is that the difference between the Vietnamese [c] and [k] is much larger than the difference between the same two sounds in Turkish. While both seem to be clearly different sounds, it is more important for

Vietnamese speakers to disambiguate between the sounds because they need to be contrastive, so the greater distance here helps with that purpose.

We also ran t-tests comparing each phoneme between each of the languages. Due to the small sample size (made even smaller by having to remove one of the English data points), all of the results proved to be not statistically different. However, the difference between the Turkish and Vietnamese [c] was almost statistically significant with a p value of 0.075, and would have most likely shown a significant difference had we tested with more people. This is actually the most important difference in sound we care about for this study though because it is the slightly affricated pronunciation of this [c] in Vietnamese that probably contributes to the greater distance (which is further confirmed by the higher COG for [c] in Vietnamese too). Representative spectrograms for a Turkish [c] (top left) and Vietnamese [c] (bottom left) are shown below compared with their corresponding [k] (top/bottom right). For the rest of the power spectra, see the Power Spectra folder in the project github.



Representative power spectra for Turkish [c] (top left), Turkish[k] (top right), Vietnamese[c] (bottom left), and Vietnamese[k] (bottom right)

B. Palatogram Analysis

To analyze the palatogram data, we first needed to calculate normalized values for the distance from the constriction to a reference point on the palate (referred to as “h”), as well as the width of the constriction. To generate these values, we loaded each palatogram image into Microsoft Paint, and drew a horizontal line at a consistent reference level (h) of the fifth tooth from the center of the mouth. We then drew a vertical line extending from h to the bottom lip. Still using Paint, we then took the pixel values at the top and bottom of this vertical line, as well as the top and bottom of the constriction. Assuming pixel number increases in the down direction, the normalized constriction location was calculated as (top of constriction - top of vertical line) / (bottom of vertical line - top of vertical line). The normalized constriction width was calculated as (bottom of constriction - top of constriction) / (bottom of vertical line - top of vertical line). Essentially, these measurements take whatever length we are interested in, and divide that by the length of the vertical line extending from the lower lip to h to account for differences in things like camera angle and mouth size.

Note that palatograms are inherently an imprecise measurement, and while they can give a good idea of general patterns of where sounds are articulated, it is nearly impossible to get significant results on small measurements, especially with a sample size of three speakers from each language. Additionally, since one of the Turkish speakers opted to not participate in this portion of the project, and three data points had to be thrown out due to the image clearly showing an incorrect point of contact or no contact at all (1 English [k], 1 English [c], and 1 Turkish [k], all from different speakers), we decided to analyze the general patterns we saw in the data rather than trying to make any specific claims about how the palatograms proved any new facts about the physical difference between sounds. This analysis is therefore meant to support the findings in the previous section where the physical difference between the [c] and [k] sounds are different between Vietnamese and Turkish.

Keeping this in mind, the overall summary of the palatogram data is presented below. It includes the average contact location (measured as distance from h to the top of the contact) and the average width of contact, for [c], [k], and [t͡ʃ] in English, Turkish, and Vietnamese.

	[c]	[k]	[t͡ʃ]
English location:	0.195728	0.369402	0.435964
Turkish location:	0.187108	0.119048	0.466602
Vietnamese location:	0.431737	0.308802	0.521058
English width:	0.239974	0.102392	0.241149
Turkish width:	0.287282	0.119048	0.264961
Vietnamese width:	0.234431	0.221645	0.192773

Looking at this data, some of the themes we see are that in most cases, [c] is pronounced further forward than [k], except for the English speakers. This supports the idea that Turkish and Vietnamese speakers differentiate these sounds, while they are total allophones in English. Additionally, in all three languages, [c] had a larger contact area than [k], which is expected at least for Vietnamese because of the tendency to affricate that sound, but the fact that the difference was even greater in Turkish was surprising, though potentially an error. Lastly, looking at the location of [c] in particular, for English and Turkish speakers, the sound is closer to [k], while in Vietnamese, it is closer to [t͡ʃ]. This is likely because in Turkish and English, the sound [c] is paired with as its corresponding allophone/quasi-phoneme is [k], while in Vietnamese [c] and [k] are contrastive with the [t͡ʃ] sound not existing at all. This makes it easier to conceptualize and then pronounce [c] and [k] closer together in English and Turkish, and [c] and [t͡ʃ] closer together in Vietnamese.

An example of a successful series of palatograms is presented below for one of the Turkish speakers, showing the clear differences in contact location between [t͡ʃ], [c], and [k]. While all palatograms were not this clear, these images can give an idea of what the expected results were for Turkish with each sound being pronounced progressively further back. For the rest of the palatogram images, see the Palatograms or Palatograms with Lines folders in the project github.

Palatogram series for a Turkish speaker pronouncing çim, iki, and kapı.
From left to right we see the contact locations for the phonemes [t͡ʃ], [c], and [k].

C. Phoneme Confusion and Multidimensional Scaling

The raw results files from the phoneme confusion were not very useful when it came to analyzing data, so the first step after we generated the data was to put them into the form of confusion matrices. For each speaker, we manually went through the results file and tallied up each time the subject confused a particular sound for each of the others. This data was then combined to create three general confusion matrices for all English, Turkish, and Vietnamese speakers. Some other post-processing was done to convert these values into percentages, and to remove the ??? option from the matrix and normalize the percentages to the slightly smaller sample size when those were not taken into account. All confusion matrices (tallies, percents, percents with the ??? option removed) are available for all speakers on the project github, but the final combined results for each of the languages are displayed here. Note that the row indicates the sound stimulus, and the column indicates the percentage of time the subject selected the above sound as what they heard.

English

	"c"	"k"	"t͡ʃ"	"ʃ"	"t"
[c]	0.055556	0.333333	0.277778	0	0.333333
[k]	0.058081	0.883838	0.030303	0	0.027778
[t͡ʃ]	0	0	1	0	0
[ʃ]	0	0	0.388889	0.611111	0
[t]	0.027778	0.083333	0.027778	0	0.861111

Turkish

	"c"	"k"	"t͡ʃ"	"ʃ"	"t"
[c]	0.260101	0.169192	0.143939	0.027778	0.39899
[k]	0.388889	0.35	0.083333	0.055556	0.122222
[t͡ʃ]	0.027778	0	0.888889	0.027778	0.055556
[ʃ]	0	0.027778	0.058081	0.886364	0.027778
[t]	0.055556	0.027778	0.083333	0.083333	0.75

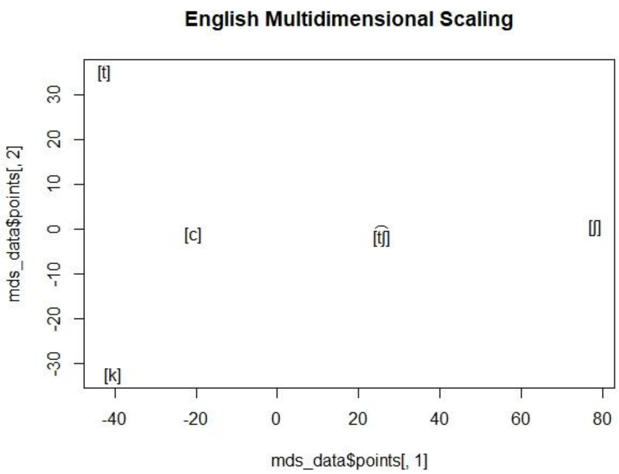
Vietnamese

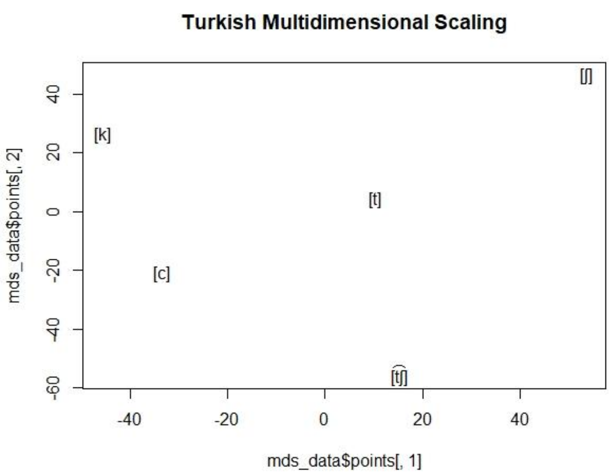
	"c"	"k"	"t͡ʃ"	"ʃ"	"t"
[c]	0.216162	0.154545	0.238384	0.055556	0.335354
[k]	0.095238	0.661376	0	0	0.243386
[t͡ʃ]	0.611111	0	0.388889	0	0
[ʃ]	0.088384	0	0.143939	0.737374	0.030303
[t]	0.030303	0.027778	0.033333	0.058081	0.850505

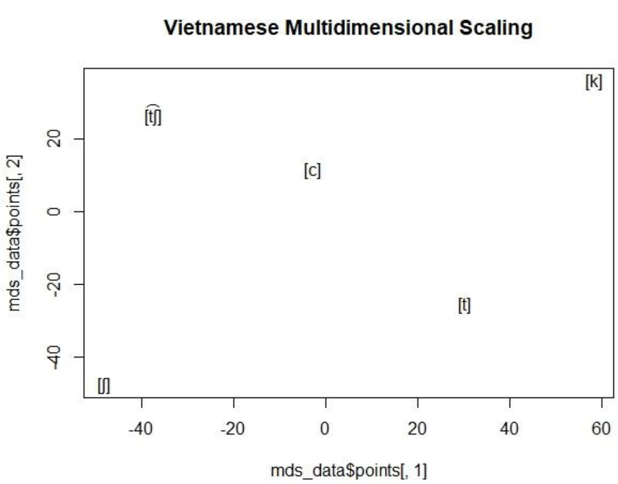
From the confusion matrices, we can see that, for example, English speakers commonly confuse [c] as “k”, while this is not as much the case for Turkish and Vietnamese speakers because the sounds are more distinct in those languages. Additionally, we can see that in Vietnamese, the [t͡ʃ] and [c] sounds were frequently confused which is to be expected based on the [c] affrication and lack of a true [t͡ʃ].

While the confusion matrices help us to point out a few interesting facts about which sounds speakers of each language most easily confuse, they do not provide a clear representation of true perceptual distance. The best way to do this is to perform multidimensional scaling on the confusion data to reduce the dimensionality and display each sound on a 2D grid where their distances on that grid directly correspond to their perceptual distances. There is a simple R script available to generate these plots, but it does not take the confusion matrix directly as input. Instead, we first need to generate distance matrices, which includes the same row and column labels as the confusion matrix, but stores a number from 0 to 100 indicating how far apart the sounds are from each other (making it symmetric along the diagonal). These can be automatically generated from the confusion matrices using a python script, and the distance matrices for all speakers (and compiled ones for each of the three languages) are available for viewing in the project github.

After generating the distance matrices and running the R script to produce the multidimensional scaling plots for each speaker, we obtained the following results corresponding to the perceptual distances between each of the consonants of interest in English, Turkish, and Vietnamese.


English Multidimensional Scaling
[t]
[c]
[tʃ]
[ʃ]
[k]
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
60
80
mds_data$points[, 2]
mds_data$points[, 1]


Turkish Multidimensional Scaling
[ʃ]
[k]
[t]
[c]
[tʃ]
40
20
0
-20
-40
-60
-40
-20
0
20
40
mds_data$points[, 2]
mds_data$points[, 1]


Vietnamese Multidimensional Scaling
[k]
[tʃ]
[c]
[t]
[ʃ]
20
0
-20
-40
-40
-20
0
20
40
60
mds_data$points[, 2]
mds_data$points[, 1]

From these plots, we can see that [c] and [k] were perceptually most spread out for Vietnamese speakers, followed by Turkish, then English. This is exactly the same order we found for the physical distances as well. Another interesting result we can see from the MDS plots are that [c] and [t͡ʃ] are most perceptually similar in Vietnamese, which is also consistent with both prior research and the physical results we found previously from spectrum and ap.palatogram analysis.

V. Conclusion

After analyzing the results of recorded spectra, palatograms, and a phoneme confusion test for English, Turkish, and Vietnamese speakers, we found that there is a slight difference between languages in the physical and perceptual distances of the [c] and [k] phonemes that corresponds to whether they are true phonemes, quasi-phonemes, or allophones. The phonemes were furthest apart in Vietnamese where they are true phonemes, and closest together in English where they are allophones, which aligns with the initial prediction that true phonemes would have the greatest phonetic distance because this is the only case where distinguishing between the sounds is vital to meaning. We also found that in many cases the [c] and [t͡ʃ] phonemes in Vietnamese were most similar, because [t͡ʃ] does not exist in Vietnamese, so there is no reason to distinguish between the sounds to convey meaning. This could also have been a factor in allowing the [c] and [k] phonemes to be more spread in Vietnamese.

Of the tests performed, the analysis of the centers of gravity in the power spectra proved to be the most useful and significant when determining differences in physical distance between languages. We could clearly see a difference between the distances in Turkish, Vietnamese, and English when it came to the COG values, and the [c] sound was shown to have an almost statistically significant difference in pronunciation between Turkish and Vietnamese (and would likely have been significant if we had tested more speakers). The analysis of images, including spectrograms, power spectra, and palatograms were not very impactful in reaching the final conclusion for this paper since they tended to be imprecise and difficult to analyze, but they still provided some interesting data, and could have potentially further supported our results if we had more subjects. The phoneme confusion test was the only test that measured perceptual distance, and the final multidimensional scaling plot helped to confirm that the perceptual distances were related to the nature of phonological contrast in the same way physical distances were. While the confusion matrices also provided this same data, reducing the dimensionality like this made the results much more clear and easy to understand.

One of the biggest challenges in this project was finding native speakers of Turkish and Vietnamese. Eventually, we were able to find three of each, but this took some time, and was still a small overall sample size that made analyzing the data difficult, because it was hard to tell if results were because of a feature of the language or an individual speaker. Another challenge (that would have actually been made worse by more subjects) was the sheer amount of data that needed to be parsed by hand. There were 30 different words that all needed to have palatograms, spectrograms, power spectra, and center of gravity generated and analyzed. Additionally, to go through all the steps to take the phoneme confusion results for every speaker and convert them into a multidimensional scaling plot took a lot of precise formatting of the data that also needed to be done by hand in order to make sure that the scripts produced accurate results. Lastly, the quality of the palatograms tended to correspond with creating a good viscosity of the charcoal and canola mixture, which was not very consistent between subjects.

To address these challenges in future studies, I would have interviewed many more speakers from each language to get more statistically significant results in all of our tests, and to have the freedom to more easily throw out outlier data. I would also generate computer programs to automatically parse and make calculations based on the raw data, so that more effort could be spent on the analysis portion of the project. Lastly, I would just pay more attention to the viscosity of the palatogram mixture, and give myself more practice making it, so that the palatograms would have a more consistent better quality. It would also be interesting to conduct the same study for a different pair of phonemes on a different set of languages, to see if the result is the same.

Overall, this was a very fun study to conduct and led me to answer a question that genuinely interested me. While I wish that I was able to do more rigorous statistical analysis of the data to get a definitively significant result, the final finding was still clear and fascinating in the sense that it showed how phonetic distance between sounds and the form of phonological contrast those sounds have in a language are actually somewhat related.

VI. References

1. https://www.laits.utexas.edu/phonology/turkish/turk_dorsal_pal2.html
2. https://www.alotrip.com/about-vietnam-language/linguistic-characteristics-vietnamese-language
3. http://www.lel.ed.ac.uk/~jkirby/docs/kirby2011vietnamese.pdf
4. http://www.linguisticsnetwork.com/phonetics-the-basics-about-acoustic-features-of-consonants-in-standard-english/