THE EFFECTS OF THE FRENCH VOWEL INVENTORY ON VOWEL PRODUCTION IN SPANISH SPEAKERS

LEONOR TELES
OLIVIA HUEY
Department of Linguistics, University of Rochester

Abstract

This study examines acoustic and auditory aspects of how speakers of a small-vowel inventory language, in this case Madrid Spanish, produce vowels from Parisian French, which has a larger vowel space containing non-peripheral vowels. Although both languages belong to the same language family, the Madrid Spanish vowel inventory contains 5 exclusively peripheral vowels, while the Parisian French vowel inventory contains 15 peripheral and non-peripheral vowels. The Spanish speakers were asked to pronounce 5 French vowels /y/, /ɛ̃/, /ø/, /œ/ and /ɛ/. Vowel contrasts in an F1-F2 vowel space, kmeans-clustering, and the Bark Difference Metric were implemented to analyze the acoustic characteristics, while the auditory model UR Ear was used to analyze how the characteristics were encoded in the auditory pathway. We predict that the Spanish speakers would 'cluster' the French vowels, meaning they would produce the non-peripheral French vowels close to a known peripheral Spanish vowel. The acoustic results show that Spanish speakers implement different strategies to pronounce unfamiliar non-peripheral vowels: they either 'cluster' some vowels or replace roundedness with backness, thus producing a more back version of their back vowel /u/. The auditory model results echo the vowel contrasts regarding F1 and F2 in the acoustic vowel space analysis. In addition, the auditory model results show that the dips and peaks were more aligned between Spanish speakers' pronunciation of French vowels and the authentic French vowels in formants higher than F2. This suggests that the differences between the two types of speaker's pronunciations may be mainly manifested in aspects reflected by F1 and F2.

1 Introduction

In this project, we investigate how Spanish speakers from Madrid attempt to pronounce vowels of Parisian French. Specifically, we investigate whether these speakers turn the vowels into known vowels, which we refer to as 'clustering', or produce a vowel outside of their vowel space by over-articulating one aspect of the vowel in an attempt to compensate differences they hear in a different aspect, which we refer to as 'overshooting'. While learning a new language can be an exciting challenge, it also involves different forms of speaking. The first impact of any language comes from the spoken word, as the basis of all languages is sound. Pronunciation accuracy can become a major barrier to communication as a non-native speaker. It is then helpful to understand the relation and trends of phenomena that occur in our language systems as we attempt to acquire new languages. The processes of how non-native speakers acquire a new language and battle with pronunciation pertain to phonetics and can depend on our perception and sociocultural backgrounds.

A significant portion of phonological studies have investigated how language-specific factors affect and warp the perceptual distance between sounds. For instance, English speakers can easily differentiate sounds that cross a phonemic boundary (e.g., /l/ and /r/) but speakers of a very different language such as Japanese fail to do so with the same ease (Kuhl 2000). Production studies have also shown the effect of inventory size on the acoustic distance between phonemes. Bradlow (1995) and Jongman et al. (1989) propose that speakers of languages with larger inventories such as English or German have an expanded acoustic space as compared to speakers of languages with smaller inventories such as Modern Greek or Madrid Spanish. This expansion is achieved by producing the same target phonemes acoustically at a greater distance from one another. However, there has been limited research in this aspect for languages from the same language family. We have chosen French and Spanish, both members of the Romance language family. We examine the ways in which Madrid Spanish speakers attempt to pronounce novel French vowels, where and how they place the French vowels in a vowel space, and how this is reflected in the auditory system using an auditory model.

Romance languages share many of the same consonant phonemes, and there is little to no variation in their production. However, a distinct difference between Romance languages is in their vowel inventories (Calebrese 2002). Parisian French possesses 15 vowels, whereas Madrid Spanish has 5 peripheral vowels, as seen in Figure 1 (Collins and Mees 2013). Along with distinct vowel phonemes, the production of these vowels in different languages also varies. As a result, speakers of one Romance language, while having no difficulty understanding written passages from other closer Romance languages, often struggle to produce native-sounding French vowels (Towell and Hawkins 1994). This creates what most would call an accent. While accents may not interfere with the communication in many languages, it can become problematic in languages such as French, where slight differences in phoneme production could lead to a completely different vowel that alters the meaning of the entire word.

Vowel spaces in language systems evolve by building up peripheral vowels first, and, as more vowels become part of a language inventory, non-peripheral vowels are added. Vowel contrasts are represented in a space defined by F1-F2, i.e., the first and second formants. Dispersion Theory (DT), first proposed by Liljencrants and Lindblom (1972), suggests that if the number of vowels in an inventory is known, then their phonetic qualities can also be identified. This is due to the existence of strong regularities in the way vowel systems develop and are designed cross linguistically. Vowels are dispersed within an area defined by the first two vowel formants, F1 and

F2, regardless of the number of vowels present in the system. In addition, the way internal vowels increase as language inventories are built is very distinct from the way peripheral vowels are added. This was later reiterated from the idea that the requirement of perceptual contrast predicts vowel inventories (Lindblom 1986).

Thus, we analyze how speakers from a language without non-peripheral vowels handle non-peripheral vowels and how this is encoded in the auditory system. Our hope is that analyzing this phenomena could help expand our knowledge of second-language acquisition and its implications. Mainly, we interpret the French vowels as spoken by Spanish speakers and assess the differences between French vowels produced by a Spanish speaker as compared to those produced by a French speaker. We specifically look at how non-native French speakers handle vowels not in their native language, such as the French non-peripheral vowels /y/, /ø/, /œ/ and peripheral vowels not included in the Spanish vowel space /ɛ̃/ and /ɛ/. We assume that the vowel produced by the Spanish speakers is what they have perceived and thus that they believe it sounds the same as the original French vowel. This assumption was made in order to remove complex cognitive analysis, which is not part of the goal of this study.

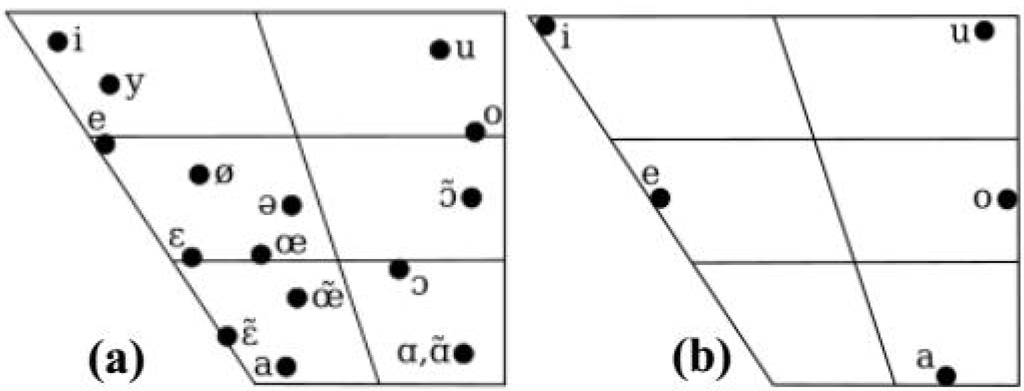

Figure 1. IPA Vowel Space of (a) Parisian French and (b) Madrid Spanish (Collins and Mees 2013)

To expand the analysis further, we look at how the auditory system encodes and processes the information in the acoustic signal of some of these vowels. If a vowel seems to cluster with another or overshoot, would the auditory system reflect this as well? To answer this, the UR Ear computational model (Carney and McDonough 2018) is implemented to illustrate two levels of auditory pathways' response to the vowels in question based on existing physiological data. These two levels are reflected by two models within the UR Ear system. The first model encodes the responses of auditory-nerve (AN) fibers, incorporating realistic nonlinear properties associated with the basilar membrane, inner hair cells (IHCs), and the IHC-AN synapse (Zilany et al 2014). The auditory nerve has a greater dynamics range with lower spontaneous rate. The second model predicts behavior in the Inferior Colliculus (IC), accounting for the fact that the IC is very sensitive to fluctuations in a stimulus (Nelson and Carney 2004). IC responses are based on how much the AN is fluctuating, and thus dips in band-enhanced average rates show the formants of the vowel.

Our main goal is to investigate how Spanish speakers pronounce French vowels, where in the vowel space these are placed, and the auditory system's response. To do so, we asked six Madrid Spanish speakers to pronounce five words in French containing vowels not present in the Spanish vowel inventory. If a vowel has been placed in an area statistically close to another, we call this clustering. If a vowel has been placed far from where it should have been placed in the French vowel space, we call this overshooting. We hope to answer the following questions: (1) Do Spanish speakers cluster the French vowels into the existing vowels in the Spanish vowel space? (2) Which vowels cluster, and which vowels are significantly moved into a different quadrant of the vowel space? (3) Is the same behavior observed when the third formant is accounted for? (4) What are the responses of the auditory system to these new vowels?

In the following sections, we attempt to answer these questions. Our hypothesis is that since vowel systems are built with peripheral vowels first, Madrid Spanish speakers will cluster the French vowels into one vowel that exists in the Spanish vowel space. Moreover, we hypothesize that speakers will hyper-articulate front-rounded vowels by producing a more extreme version of the back vowel to handle an unknown non-peripheral vowel. Our null hypothesis is that Spanish speakers would be able to distinguish between all French vowels and place them accordingly in the vowel space.

2 Methods

Six Spanish speakers (3 males and 3 females) from Madrid who have never learned French were recruited on a voluntary basis through networking. The region was specified to minimize regional pronunciation differences between speakers. A wordlist with French words shown in Table I containing the vowels of interest /y/, /ɛ̃/, /ø/, /œ/ and /ɛ/ were drawn from the UCLA Phonetics Archive and concatenated as a single audio file. Each word was repeated twice with a pause in between for the subjects to repeat after. A wordlist with Spanish vowels was created based on a wordlist found in Cao (2014), shown in Table 2. The audio file containing French word recordings along with an instruction document were sent to subjects that volunteered to participate in the experiment. The instruction document is found in Appendix C. The speakers were asked to record themselves according to the instructions provided with Voice Recorder Pro by © BejBej Apps, repeating after the French words in one audio file and pronouncing the Spanish wordlist in another.

The two audio files for each speaker were then trimmed in Praat (Boersma and Weenink 2020) to each vowel's steady state, and formant values of each vowel of each speaker were extracted with Praat's formant listing function. The recording files were saved as .wav files and uploaded to our computers using built-in features of Voice Recorder Pro. The Spanish speakers' pronunciation of the French words were then trimmed by hand in Praat to preserve only steady states of the vowels. Afterwards, formant values of each vowel for each speaker were extracted with Praat's formant listing function. Only the first recording of each vowel repetition was used. It was assumed that the word pronounced by the Spanish speakers with the corresponding vowel represented exactly what they had perceived the French word and vowel had sounded.

Table 1. French Wordlist Selected for Analysis with Corresponding Vowels and IPA

Vowel	French word	IPA
y	du	dy
ɛ̃	bain	bɛ̃
ø	deux	dø
œ	de	dœ
e	mais	me
ɛ	taie	tɛ

Table 2. Spanish Wordlist Selected for Analysis with Corresponding Vowels and IPA

Vowel	French word	IPA
a	casa	kasa
e	pero	peɾo
i	piso	piso
o	toro	toɾo
u	tubo	tuβo

Vowel spaces with the French and Spanish vowels combined were plotted for each speaker using MATLAB and are shown in Figure 2. Black lines were drawn between Spanish vowels to construct each speaker's Spanish vowel space, as the peripheral Spanish vowels are the five contrastive vowels in the Spanish vowel space. This figure was assembled so that vowels that were overshot could be identified. Then, the Elbow Method for finding an optimal number of vowel clusters was applied before determining which vowels had clustered together (Syakur et al 2018). The clusters are regions with vowels that have been placed in an area statistically close to one another. The Elbow method considers the total within-cluster sum of squares as a function of the number of clusters, and is a direct method for finding the optimal number of clusters. The location of a bend (knee) in the plot of total within-cluster sum of squares versus number of clusters indicates the appropriate number of clusters where the bend lies. Once the Elbow Method had found the total number of clusters in each speaker's vowel space, kmeans (Syakur et al 2018) was applied to ultimately determine which vowels had clustered together. This technique finds points that are clustered such that the total intra-cluster variation, or total within-cluster sum of squares, is minimized. This technique was used in each speaker's resultant vowel space with both French and Spanish vowels they had recorded. MATLAB was used for the plots shown in Figure 3.

Then, the Bark Difference Metric was used as provided by the NORM vowel normalization and plotting suite (Thomas and Kendall 2007), in order to analyze the effects of F3 that previous methods do not account for. The Bark Difference Metric is a vowel-intrinsic method modified from the formula developed by Syrdal and Gopal (1986). NORM converts the formant values given to bark formula:

(1) $$Z_i = \frac{26.81}{(1+\frac{1960}{F_i})} - 0.53$$

In (1), F_i is the value for a given formant i and Z_i is the converted bark for the given formant i. Following the conversion, NORM computes the differences Z3-Z1, Z3-Z2, and Z2-Z1. Z3-Z2 is used to plot the normalized front-back dimension conventionally represented by F2 values in a standard vowel space plot, and Z3-Z1 is used to plot the normalized height dimension conventionally represented by F1 in a standard vowel space plot.

The Bark Difference Method was chosen due to its advantage of not requiring measurement of all vowels for all speakers to be included in a study. Other methods provided by NORM are all vowel-extrinsic and work optimally when the entire vowel system is measured, which we did not do in this study. However, the overall shape of the vowel space constructed using the Bark Difference method often appears distorted. Therefore, observations were made in conjunction with formant value listings extracted from Praat for quantitative analysis, instead of solely using the Bark Difference method. Appendix E includes the tab-delimited text file used.

Finally, for the auditory analysis, two models were used: the Zilany et al. (2014) model for the AN, and the SFIE model for the IC, first introduced by Nelson & Carney (2004). The AN model computes neural responses in the AN for different Characteristic Frequencies (CFs). The IC model accounts for differences in the temporal envelope, measuring these changes, to predict how IC neurons respond to amplitude-modulated stimuli. These neurons possess an average firing rate tuned to modulation frequency. The SFIE model thus considers the physiological mechanisms of these neurons as combinations of excitatory and inhibitory inputs with matched CFs, serving as a detection strategy for separating the tone from noise without cross-frequency integration (Gai and Carney 2006). For the purposes of this study, only the band-enhanced cells of the IC were considered. These show the highest average response at the modulation frequency and fluctuate depending on how AN fibers respond. Thus, the AN average response has a flatter envelope when the CF of the AN fiber is closer to a formant and, as the IC is sensitive to these changes, it will have a lower rate. At moderate-to-high sound pressure levels (SPL), there is an average rate saturation in the AN fibers, hence vowels were assessed at 65 db SPL. In summary, the SFIE model from the computational model code was used with 40 CFs and at 65 db SPL. The trends we looked for included assessing how behavior in the AN was reflected into the IC; how the dips in the band-enhanced cells reflected formants that were congruent with the vowel space plots; and, although F1 and F2 did not align between Spanish and French speakers for the same vowel, if the IC responses matched at F3 and beyond.

4 Results Section

4.1 F1-F2 Vowel Space

Figure 2 shows the vowel space plots for each of the 6 speakers with both Spanish and French vowels pronounced. Female speakers are shown at the top and male speakers at the bottom. French vowels are outlined in blue, Spanish vowels are in green, and vowels outside the Spanish vowel space are highlighted in red. Black lines connecting the Spanish vowels were drawn to create the speakers' Spanish vowel spaces.

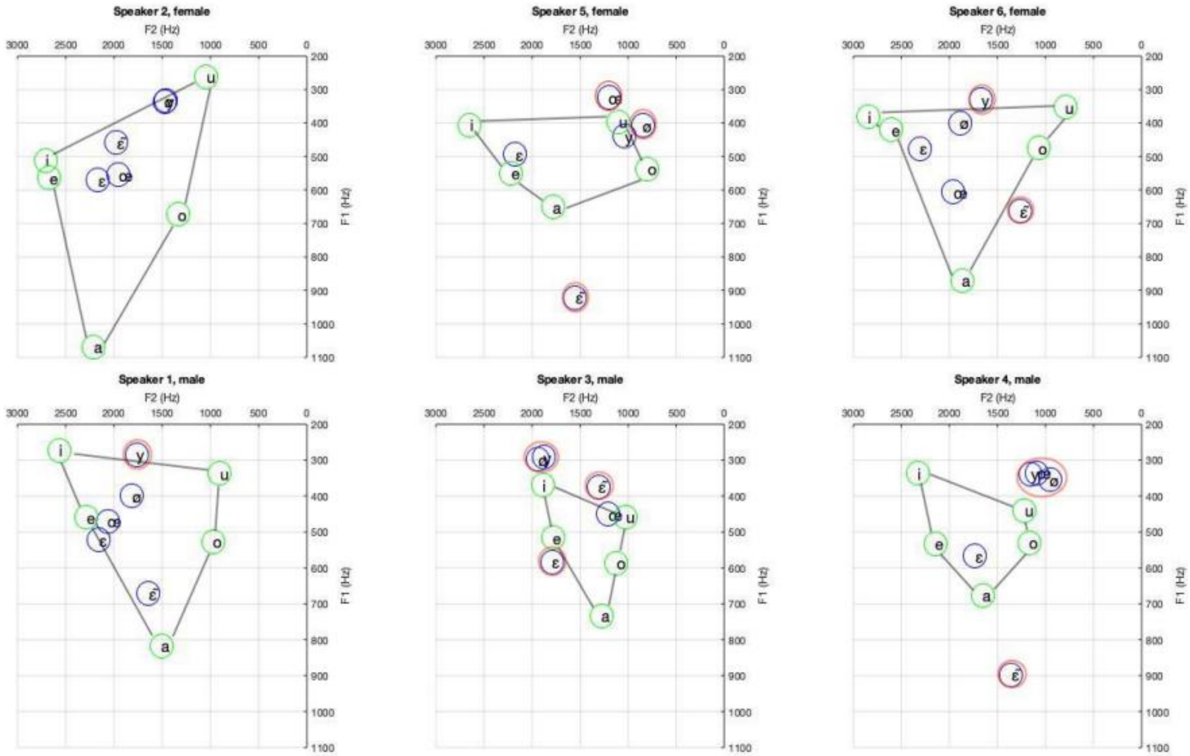

Figure 2. Vowel space plots of Spanish and French vowels pronounced by all speakers, with Spanish vowels highlighted in green, French vowels outlined in blue, and vowels outside the vowel space in red. Female speakers at the top and male speakers at the bottom.

From Figure 2 above, it can be seen that all speakers except speaker 1 placed French vowels outside of their Spanish vowel space. Moreover, all speakers except speaker 3 placed the French vowels /y/ and /ø/ at a lower F2. Speakers 4 and 5 placed /ɛ̃/ very far from the Spanish vowel space, at a lower F2, and speakers 1 and 4 placed this vowel closer to a. Finally, speakers 1, 3, 5 and 6 placed the French vowel /ɛ/ very close to the Spanish vowel /e/.

Figure 3 below shows plots of the (a) French vowels pronounced by all Spanish speakers, as well as (b) their respective Spanish vowels.

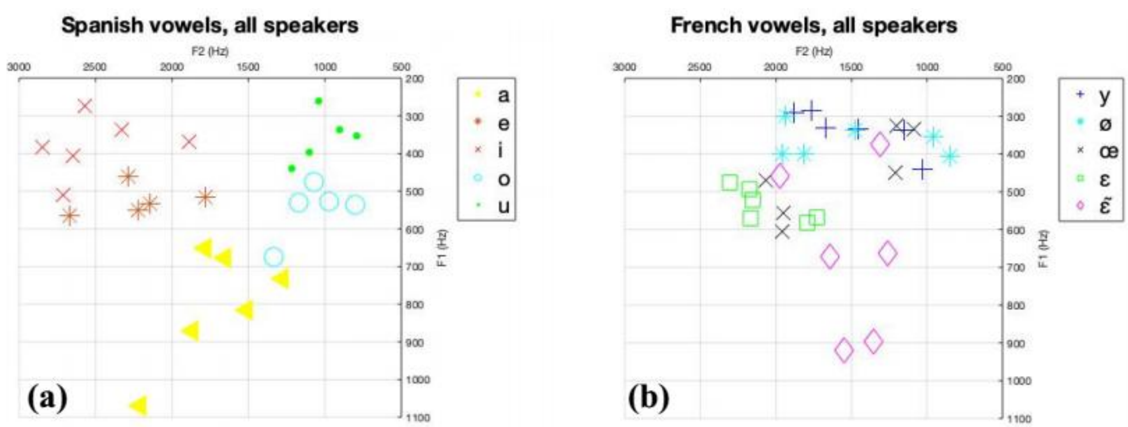

Figure 3: Vowel plots of (a) French vowels pronounced by all speakers, and (b) Spanish vowels pronounced by all speakers

Figure 3 shows the variability across all speakers, although Spanish vowels are more consistently placed. The French vowel ɛ also appears to be placed in a similar area for all speakers.

To further confirm the clusters, the Elbow Method and kmeans were applied for each speaker. Summary of results from kmeans clustering for each speaker are shown in Table 3 below. Raw data from each cluster is found in Appendix B. The result of the Elbow method yielded 4 clusters for all speakers, except speaker 5.

Table 3. Summary of Phenomena Observed in each Speaker

Speaker	Cluster e and ɛ	Front-rounded vowel placed as a back vowel	Mid-front-rounded vowel placed in a much lower F2	Place ɛ̃ in a lower F2; cluster ɛ̃ and a
1	yes	yes	no	yes; yes
2	no	yes	yes	yes; no
3	yes	no	no	yes; no
4	no	yes	yes	yes; yes
5	yes	yes	yes	yes; no
6	yes	yes	no	yes; no

4.2 Bark Difference Method

Figure 4 below shows the placement of each vowel on the vowel space using the Bark Difference Method as provided on the NORM vowel normalization and plotting suite (Thomas and Kendall 2007). Vowels were labeled as follows for display purposes: y = y, ø = o/, œ = oe, ɛ = E, ɛ̃ = E~ .

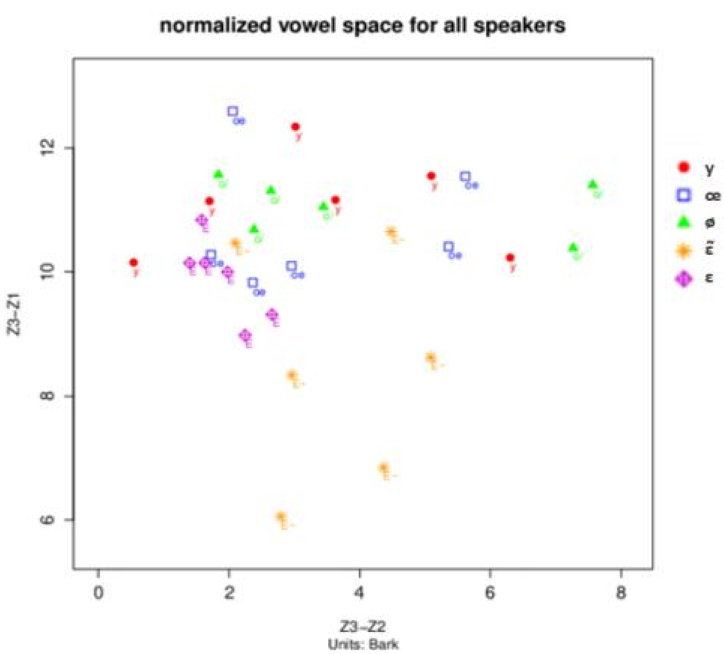

Figure 4: Bark Difference method results displaying the placement of each vowel on the vowel space accounting for F1, F2 and F3.

From the Bark Difference Method results shown above, one can see even more variance in /ø/ across all speakers, as well as that /ɛ/ was placed around the same area for all speakers.

4.3 UR Ear

The SFIE model was run for selected speakers and vowels, and results are shown in Figure 5 below. Speaker 5 had a /y/ and /ø/ very close to the /u/, and this was analyzed in the auditory model. The same was done for speaker 1 that had an /e/ very close to the /ɛ/, and the same for /œ/. Figure 5 below shows the results from the model.

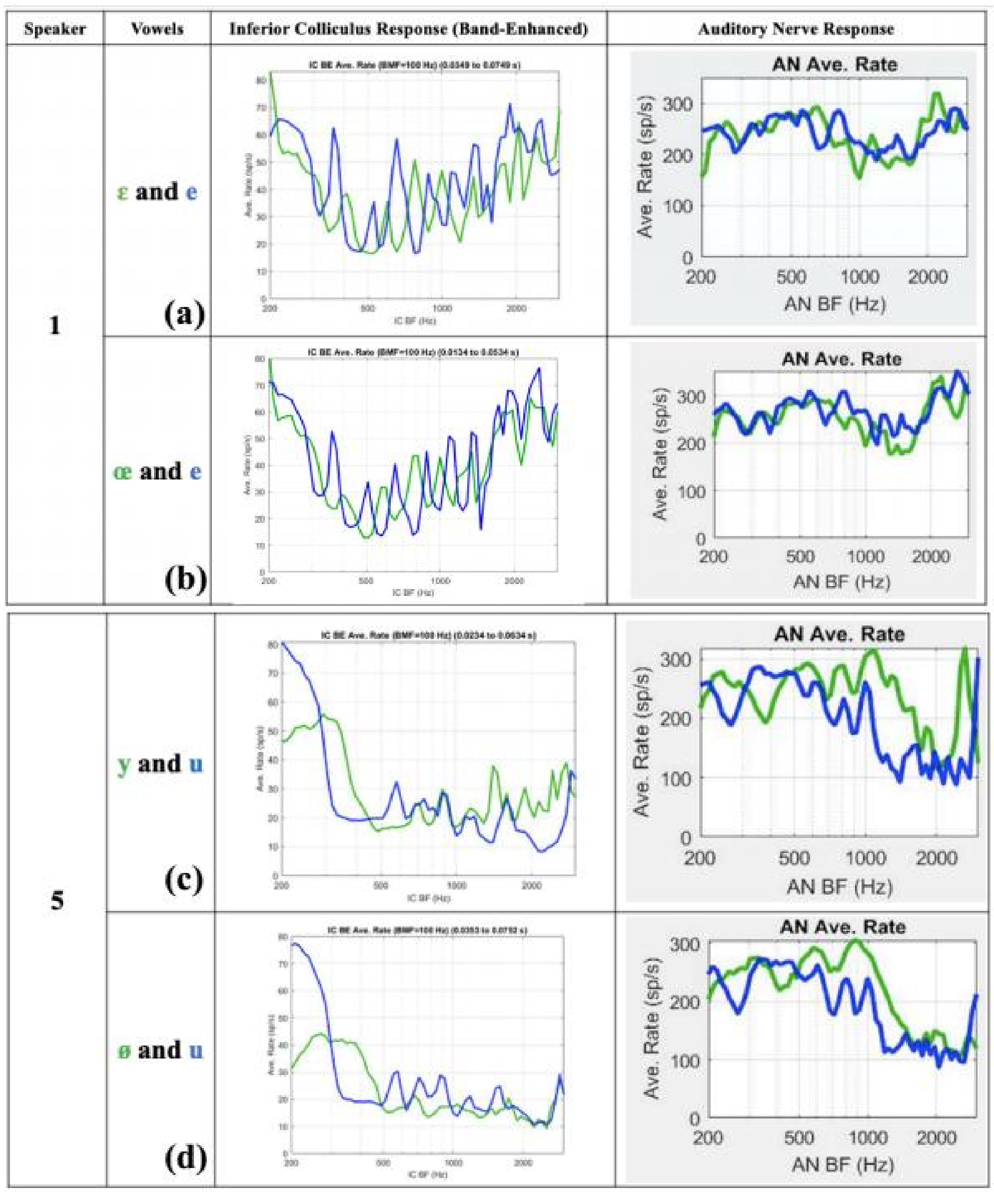

Figure 5: SFIE model results, showing the IC and AN responses for selected Spanish speakers' vowels. (a) is speaker 1's /ɛ/ and /e/ response; (b) is speaker 1's /œ/ and /e/ response; (c) is speaker 5's /y/ and /u/ response; and (d) is speaker 5's /ø/ and /u/ response

Figure 5 shows how fluctuations in the AN result in large dips in the IC response. On the other hand, the model has encoded the vowels, since dips in the IC correspond to formants that are plotted in Figure 2 above. However, as seen for all of speaker 5's responses, the AN fluctuates more for the Spanish vowels than for the French vowels, seen in the large dip at the start. Moreover, while F1 and F2 are somewhat similar in the vowel spaces for Speaker 1's /ɛ/, /e/ and /œ/, some of their dips in the IC response align at higher formants. For (c), while the vowel /u/ shows deeper dips, /y/ shows more fluctuation. This dip reflects the fluctuation in the AN response at around 200 Hz. The dips, although slightly different for Speaker 1, are somewhat constant between the vowels, further suggesting their clustering. Figure 6 below shows the vowel /y/ that did not match the French /y/ in the Spanish speakers, compared to the French speaker /y/. Speaker 3, however, did manage to accurately pronounce this vowel, and the auditory results are also shown below.

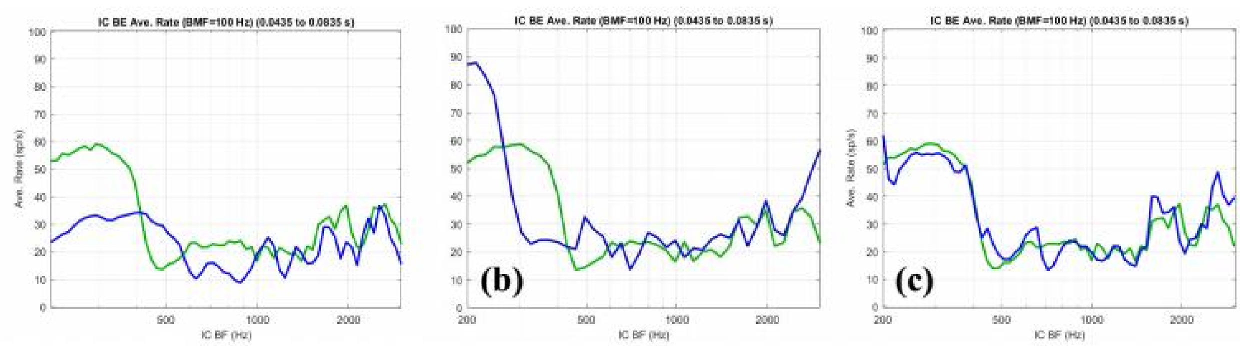

Figure 6: SFIE model results, showing the IC responses for selected Spanish speakers' /y/ and French speaker's /y/ vowels for (a) speaker 2, (b) speaker 3, and (c) speaker 6

These results show that, although F1 and F2 for speakers 2 and 6, seen in the first 2 dips in the IC response, do not exactly match those of French speaker, the behavior matches in later formants, as seen in the more consistent and aligned dips. Speaker 3, shown in (b) aligns even more as this speaker placed the /y/ far from the Spanish /u/, what had been observed in most other speakers, and accurately pronounced the vowel.

5 Discussion

5.1 Vowel Space

Results from Figures 2 and 3 suggest that for all of the speakers except 1 and 2, the Spanish vowel space is much smaller in range compared to the French vowel space. In addition, some French vowels pronounced by the Spanish speakers were placed in either 'overshoot' areas or merged close to the Spanish vowel that is closest in the Spanish vowel space, as summarized in Table 3. The high rounded French vowel /y/ was pronounced as a back vowel for most speakers. The nasalized vowel /ɛ̃/ caused pulses in formant structure of oral vowels and as a mid front vowel was interpreted as a back, low F2 vowel. /ɛ/ clustered with Spanish vowel /e/. /ø/ was placed as a mid back vowel as opposed to mid front. Behavior for the vowel /œ/ varied across all speakers, as some correctly placed it as a mid front vowel, while speakers 4 and 5 placed this French vowel in the back with a low F2.

The results agree with our hypothesis: Spanish speakers either merge the French vowel, perceiving them as another already existing Spanish vowel (/ɛ/ merging to /e/), or pronounce front rounded French vowels with a low F2 to account for their inability to pronounce and perceive high front rounded vowels (/y/), mid front rounded vowels (/ø/), and mid front nasalized vowels /ɛ̃/. Overshoot was thus observed when speakers placed these vowels further from their vowel spaces, as indicated in vowels outlined in red in Figure 2.

A greater vowel dispersion and a least consistent inter-vowel distance observed for French vowels is consistent with the Adaptive Dispersion Theory mentioned in the introduction. This theory also helps explain why Spanish has a smaller vowel space dispersion and less space between its peripheral vowels. Vowel space dispersion indicates the overall expansion or compactness of the tokens from each speaker. This dispersion correlates to vowel space area and intelligibility. According to this theory, Spanish exhibits the smallest range between peripheral vowels and the most equidistant vowels as it only possesses a vowel inventory with 5 vowels.

Regarding the clustering of /ɛ/ into /e/, these results may be linked to how the Spanish language evolved (Pidal 1958). The palatal semivowel /j/, often called *yod*, has been involved in a very large number of sound changes throughout the history of the Spanish language. Metaphony occurred in a regressive manner, as /j/ always came after the vowel that underwent raising. The vowel /ɛ/ is one of the vowels that was capable of being affected, and was raised to /e/.

5.2 Accounting for F3 using NORM

As seen in Figure 5, when F3 is accounted for, the results showed the Madrid Spanish speakers generally still perceived /ɛ̃/ as a more central vowel than the actual French vowel but /ɛ/ as a front mid vowel. This could mean that the main discrepancy between the Spanish speakers' and the French speaker's pronunciations is the height of their vowels, encoded in F1. However, with F3 accounted for, the subjects' perceptions of /y/, /ø/, and /œ/ were all more front and therefore closer to being accurate than when only F1 and F2 were analyzed. This general trend, in conjunction with the formant values listed in Appendix D, indicate that the main source of discrepancy was the backness encoded by F2, but that the Spanish speakers compensated this by adjusting the roundness encoded by F3, resulting in vowels that are more accurate than when only F1 and F2 were analyzed. This is confirmed by the raw data tables present in Appendix D and reflected in the plots, which suggest that the Spanish speakers attempted to correct this difference in backness by what we call a universal overshooting in the other formants: raising their values significantly above the accurate or 'needed' values.

5.3 UR Ear

The model used encoded the vowels, as seen in the fluctuations in the AN model being reflected in the IC model responses, as well as the fact that the formants from Praat plotted in Figure 2 match with dips in the IC. While the results in Figures 2 and 3 show some discrepancies between the French vowels pronounced by the Spanish speakers and the corresponding vowels pronounced by the French speaker, many of the audio files sounded more similar than the created vowel spaces indicated. We speculate that this was due to the vowel spaces only accounting for F1 and F2, thus disregarding important information encoded in F3 and formants beyond that. This conjecture was confirmed with the IC model results in Figure 6. While the first two formants differ between the Spanish and French speakers, the behavior in higher formants aligns. Although Figure 2 did not show the French speaker and Spanish speaker /y/'s clustering together on the vowel space plots, behavior of the fibers shows more similarity in the location and timing of overall trends of the IC band-enhanced average responses after F2. Despite the fact that F1 and F2 had different timing and location, the dips that indicate F3, F4 and F5 almost completely align.

A possible explanation is that rounding in the French vowel system is abundant while rounding in the Spanish vowel system is relatively rare. Since F1 indicates the height, F2 indicates the backness, and F3 indicates the roundness of the vowel, the results suggest that Spanish speakers were imitating other characteristics such as roundness of the vowel to compensate for the difference in height and backness. This further suggests that speakers switched F2 and F3 as an attempt to pronounce the vowels.

5.4 Limitations

To narrow the analysis criterion, we did not focus on stressed versus unstressed vowels. We also assumed that the French words Spanish speakers pronounced and recorded were exactly what they had perceived and heard. Although no cognitive analysis has been made in this study, Hacquard, Walter and Marantz (2007) have shown that speakers of languages with larger inventories (French speakers in this case) perceive the same sounds as less similar than speakers with smaller inventories. This suggests that once a cognitive aspect is included, speakers that have been exposed to a larger vowel inventory can perceive differences between vowels and place them more accurately in the vowel space.

Additional limitations to this study include the small sample size, high physiological variability between speakers, differences in duration of the vowels in each speaker which affected the quality and availability of steady-state portions, and differences in pronunciation habits even in the subjects' native tongue despite selecting subjects within the same region. Finally, the research-tasked speech evoked was out of context and may not be fully representative of daily speech under conversational settings.

5.5 Future Work and Significance

For subsequent studies, a larger sample size and other Romance languages could be considered. For example, Italian has an inventory of 5 vowels, Romanian 7, and Portuguese 9. It would be interesting to explore if there is a significant threshold value for, or relationship between, inventory size along with vowel location and the occurrence of clustering and overshooting. For instance, what number of vowels or location characteristics in the native language vowel space do non-native French speakers require to become equipped to pronounce unknown French vowels? Does the number of vowels or location characteristics play a larger role in inducing the phenomenon discussed?

Nonetheless, this study has provided insight into how nonnative speakers of a small-vowel-inventory Romance language manage pronunciation of another language with a larger vowel inventory in the same language family, and their auditory system responses. These findings, along with further studies, could provide suggestions for educational methods in second-language acquisition. Specifically, the studies could abet guides handling barriers of pronunciation, which could drastically change the accuracy and efficiency in languages with large vowel inventories and acute nuances such as French.

6 Conclusion

This study was done with the selected French vowels /y/, /ɛ̃/, /ø/, /œ/ and /ɛ/. Results suggest that Spanish speakers exposed to French for the first time may cluster the unknown vowels to one they know, and that they tend to overshoot known aspects of vowels in the process. We speculate that the unfamiliarity and struggle related to vowel aspects such as backness could be due to the small range and purely peripheral Spanish vowel space as opposed to the large range of the Parisian French vowel space. The auditory models confirmed the discrepancy between height and backness of the Spanish speaker pronounced French vowels. In addition, the auditory models showed that it is plausible Spanish speakers were compensating the height and backness differences by mimicking other vowel characteristics such as roundness encoded in F3 and formants beyond.

The findings in this study could provide insights into the learning process of a speaker trying to learn a language with a larger and less peripheral vowel inventory than their own. Limitations such as small sample size deem this study as preliminary at best, and further studies are required for a more cogent analysis.

Acknowledgments

This project was advised by Dr. Joyce McDonough and Dr. Laurel Carney at the University of Rochester for the course Speech on the Brain (BME/LIN 216).

References

Boersma, P. and D. Weenink. 2020. Praat: doing Phonetics by Computer. Version 6.1.09. Accessed from http://www.fon.hum.uva.nl/praat/. [Computer Program].

Bradlow, A. 1995. A comparative acoustic study of English and Spanish vowels. *JASA* 97.3: 1916-1924.

Cao, R. 2014. *Stress, dispersion, and variability of Catalan, French, and Spanish vowels.* BA thesis, University of Michigan.

Calebrese, A. 2002. On the Evolution of the Short High Vowels of Latin into Romance. *Romance Linguistics* 63-94.

Carney, L.H. and J.M. McDonough. 2018. Nonlinear auditory models yield new insights into representations of vowels. *Attention, Perception, & Psychophysics*, 81.4:1034-1046.

Collins, B. and I. M. Mees. 2013. *Practical Phonetics and Phonology: A Resource Book for Students*, 3rd ed. Routledge.

Gai, Y. and L. H. Carney. 2006. Temporal Measures and Neural Strategies for Detection of Tones in Noise Based on Responses in Anteroventral Cochlear Nucleus. *Journal of Neurophysiology* 96.5:2451-2464.

Hacquard, V., M. Walter, and A. Marantz. 2007. The effects of inventory on vowel perception in French and Spanish: An MEG study. *Brain and Language* 100.3:295-300.

Jongman, A., M. Fourakis, and J. Sereno. 1989. The acoustic vowel space of Modern Greek and German. *Language and Speech* 32:221-248.

Kuhl, P.K. 2000. Language, mind, and brain: experience alters perception. M.S. Gazzaniga (ed.), *The New Cognitive Neurosciences*, 2nd ed., 99-115. MIT Press, Cambridge, MA.

Liljencrants, J. and B. Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48:839–862.

Lindblom, B. 1986. Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (eds.), *Experimental Phonology*, 13–44. Orlando: Academic Press

MATLAB. 2019. Version 9.7 (R2019b). Natick, MA: The MathWorks Inc. [Computer Program].

Nelson, P. and L. H. Carney. 2004. A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *JASA* 116.4:2173-2186.

Pidal, R.M. 1958. *Manual de Gramática Histórica del Español*, 10th ed. Madrid: Espasa Calpe.

Syakur, M.A. et al. 2018. Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering* 336:12017.

Thomas, E.R. and T. Kendall. 2007. NORM: Vowel Normalization Suite 1.1 Methods. lingtools.uoregon.edu. [Online].

Towell, R. and R. Hawkins. 1994. *Approaches to Second Languages Acquisition*. Bristol: Multilingual Matters Ltd.

UCLA Phonetics Lab Archive. 2020. Los Angeles, CA. [Online]

Zilany, M., I. Bruce, L. H. Carney. 2014. Updated Parameters and expanded simulation options for a model of the auditory periphery. *JASA* 135.1:283-386.

Appendices

Appendix A. Extra Figures

Figure 7 below show plots of the raw data for the French vowels pronounced by the Spanish speaker as well as their respective Spanish vowels. MATLAB was used for the plots.

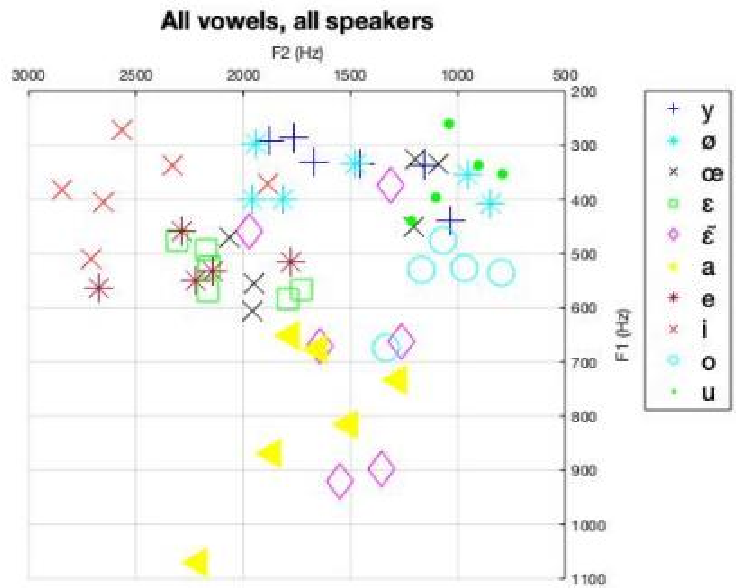

Figure 7. All vowels (French and Spanish) from all speakers plotted together in the F1-F2 vowel space.

Appendix B. Kmeans Cluster Results

Table 4 below shows the cluster results for each speaker using kmeans. Clusters are in no particular order.

Table 4. Summary of Phenomena Observed in each Speaker

Speaker	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	o, u	yes	no	yes; yes
2	y, ø, o, u	yes	yes	yes; no
3	y, ø, i	no	no	yes; no
4	ɛ, a	yes	yes	yes; yes
5	yes	yes	yes	yes; no
6	yes	yes	no	yes; no

Appendix C. Instruction Sheet for Spanish Speakers

Thank you for participating in our group project. The data we will be collecting will be used for our project only. Our goal is to analyze the pronunciation of French words in Spanish speakers and investigate how that reflects the vowel space the speaker is born in. Below are step-by-step instructions for recording yourself:

1. Find a quiet place without background noise or noisy vents or an echoe-y room.
2. Please avoid having have earphones/headphones on
3. Download the phone app "Voice Record Pro" (it is free on the Store App)
4. Watch this video for how to record yourself and sending the file to your laptop
5. Listen to the French words in the attached audio file, one at a time
6. Make sure to speak into the microphone (under the phone)
7. Make the recording of yourself trying to imitate these French words as naturally you can.
8. Listen to one word at a time and imitate them once (one trial only)
9. You can keep the app recording as you go from one word to the other
10. Send the final recorded audio to lteles@u.rochester.edu

Thank you. If any questions arise, please do not hesitate to reach out.

Appendix D. Formant Values of French Vowels as Extracted from Praat Formant Listing Function

Speaker	French	1	2	3	4	5	6
vowel	y	y	y	y	y	y	y
F1	295	297	334	331	292	473	359
F2	2134	1806	1457	1984	1154	1014	2018
F3	3601	2846	2514	2149	2501	2664	2605

Speaker	French	1	2	3	4	5	6
vowel	œ	œ	œ	œ	œ	œ	œ
F1	606	471	555	420	367	337	607
F2	1973	2064	1953	1126	1195	2299	1982
F3	3111	2675	2788	2544	2800	3151	3105

Speaker	French	1	2	3	4	5	6
vowel	ø	ø	ø	ø	ø	ø	ø
F1	393	397	336	296	398	428	393
F2	1884	1795	1478	1916	888	822	1886
F3	2807	2566	2479	2623	2868	2562	2806

Speaker	French	1	2	3	4	5	6
vowel	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃
F1	634	671	459	404	873	922	629
F2	1236	1641	1975	1313	1311	1550	1176
F3	2536	2558	2707	2578	2530	2356	2543

Speaker	French	1	2	3	4	5	6
vowel	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ
F1	475	515	570	567	583	483	476
F2	2307	2174	2164	1779	1791	2160	2308
F3	2937	2781	2920	2489	2669	2666	2936

Appendix E. Tab-delimited Text File of Bark Difference Metric Results

```
Speaker	Vowel	Context	Z3-Z1	Z3-Z2	Z2-Z1	Z3-Z1	Z3-Z2	Z2-Z1
						gl	gl	gl
1	y	NA	12.346	3.018	9.328	NA	NA	NA
2	y	NA	11.162	3.632	7.53	NA	NA	NA
3	y	NA	10.154	0.535	9.619	NA	NA	NA
4	y	NA	11.551	5.094	6.457	NA	NA	NA
5	y	NA	10.231	6.304	3.927	NA	NA	NA
6	y	NA	11.145	1.699	9.446	NA	NA	NA
1	oe	NA	10.279	1.722	8.557	NA	NA	NA
2	oe	NA	9.827	2.362	7.465	NA	NA	NA
3	oe	NA	10.41	5.358	5.052	NA	NA	NA
4	oe	NA	11.545	5.616	5.929	NA	NA	NA
5	oe	NA	12.595	2.056	10.539	NA	NA	NA
6	oe	NA	10.099	2.957	7.142	NA	NA	NA
1	o/	NA	10.683	2.383	8.3	NA	NA	NA
2	o/	NA	11.049	3.446	7.603	NA	NA	NA
3	o/	NA	11.569	1.836	9.733	NA	NA	NA
4	o/	NA	11.406	7.57	3.836	NA	NA	NA
5	o/	NA	10.387	7.27	3.117	NA	NA	NA
6	o/	NA	11.309	2.64	8.669	NA	NA	NA
1	E~	NA	8.341	2.962	5.379	NA	NA	NA
2	E~	NA	10.465	2.096	8.369	NA	NA	NA
3	E~	NA	10.652	4.478	6.174	NA	NA	NA
4	E~	NA	6.847	4.362	2.485	NA	NA	NA
5	E~	NA	6.057	2.794	3.263	NA	NA	NA
6	E~	NA	8.628	5.086	3.542	NA	NA	NA
1	E	NA	10.147	1.627	8.52	NA	NA	NA
2	E	NA	10.001	1.973	8.028	NA	NA	NA
3	E	NA	8.982	2.244	6.738	NA	NA	NA
4	E	NA	9.312	2.657	6.655	NA	NA	NA
5	E	NA	10.145	1.394	8.751	NA	NA	NA
6	E	NA	10.839	1.579	9.26	NA	NA	NA
```