Articulatory strategies in acquisition of the French /y/-/u/ contrast

Madeleine Oakley

November 13, 2019

Abstract

This study uses Ultrasound Tongue Imaging and acoustic data to show that learners do not reuse L1 articulatory gestures in L2 vowel productions, which has implications regarding the mental representation of sound. Previous studies examining the relationship between L1 and L2 sound inventories have found that learners reuse L1 phone categories in L2 productions if the phones are perceived as similar (Flege, 1987; Kamiyama and Vaissière, 2009). Studies to date have solely focused on acoustic data, however, and have not examined the nature of the categories reused in second language productions. Theories of second language production differ as to whether it is the articulatory gesture or the acoustic-phonetic category that is the basic unit of sound perception and production (Flege, 2005; Best and Tyler, 2007). The present study finds that L1 English learners of French do not reuse L1 articulatory gestures of English /u/ to produce French /u/. Furthermore, learners are acoustically target-like in their productions of the 'new' French phone /y/, but are not articulatorily target-like in /y/ productions. These findings suggest that vowels may have an acoustic target in L2 production, rather than articulatory target, and have implications for theories of L2 category formation.

1 Introduction

Second language production data in the present study shows that learners are reusing L1 acoustic, but not articulatory, categories in L2 productions, which suggests that vowels may have an acoustic rather than articulatory target. L2 data of this type provides an avenue to explore the mental representation of sounds by investigating how L1 and L2 sound systems interact. The interaction of two sound systems, and the transfer (or lack thereof) between them, can clarify the representation of units.

It is well documented that a learner's L1 will impact their perception and production of a second language (Flege, 1987; Darcy et al., 2012; Ingram and Park, 1997; Major, 2008). For L2 perception, a learner's ability to discriminate two L2 phones depends on how these phones are mapped to L1 phones (Best and Tyler, 2007). L2 sound production also depends on how L2 sounds are perceived in relation to L1 sounds (Flege, 2005; Best and Tyler, 2007; Kamiyama and Vaissière, 2009).

The Speech Learning Model (SLM) proposes that when an L2 phone is perceived as similar to an L1 phone, learners equate the two categories, which leads to learners reusing L1 categories in L2 productions (Flege, 1987, 2005). SLM postulates that the acoustic-phonetic category is the relevant representation of sound. Therefore, SLM hypothesizes

that L1 acoustic categories will be reused in L2 productions if the L1 and L2 phones are perceived as similar. For L2 phones that are not perceived as similar to L1 phones, SLM proposes that learners will create new categories for these phones, leading to target-like productions. Again, because SLM proposes that the relevant unit for speech sounds is the acoustic-phonetic category, new phones are predicted to be *acoustically* target-like.

The Perceptual Assimilation Model-L2 (PAM-L2) is similar to SLM, but importantly differs as to what is the basic phonetic unit of sound. PAM-L2 suggests that it is the articulatory gesture, rather than the acoustic-phonetic category, that is the basic unit of speech perception (Best and Tyler, 2007). L2 perceptual learning, therefore, depends on how L2 gestures are perceived in relation to L1 gestures. Similar to SLM, PAM-L2 predicts that learners will have difficulty perceiving the difference between an L2 and an L1 phone that are mapped to the same category, and it is hypothesized that no perceptual learning will occur when the two phones are mapped to the same category (Best and Tyler, 2007). Taking the assumption that the articulatory gesture is the basic object of speech perception, it can thus be predicted that in production, learners will use L1 articulatory gestures to produce L2 phones if the L1 and L2 phones are perceived as similar.

Production data of L1 English-L2 French learners provides an avenue to test the hypotheses of PAM-L2 and SLM because these learners have both new and similar phones to learn. French /y/ is different from L1 English vowels. French /u/ is similar to English /u/, but is typically produced with a lower F2 value (Flege and Hillenbrand, 1984). This study addresses whether learners reuse L1 acoustic or articulatory categories in L2 productions, and whether new phoneme categories are more target-like in both acoustic and articulatory productions.

Before discussing the results of the production task of the present study, background on L2 phoneme category formation, and hypotheses of these models are discussed. This is followed by an overview of several production studies looking at the acquisition of French round vowels. In sections 3 and 4, results of the acoustic and articulatory study are presented. I conclude with a discussion of the implications of these findings for L2 phoneme category formation.

1.1 L2 Category Formation

Models of L2 category formation emphasize how the L1 phonological system greatly influences perception and production of L2 phones for adult second-language learners. Additionally, it has been assumed that perception of non-native contrasts precedes production (Maye, 2000; Darcy et al., 2012). One model of category formation that has received much attention in the discussion of L2 category formation is the SLM. SLM posits that the L1 and L2 phonetic systems exist in the same phonological space, meaning L1 phoneme categories will influence L2 phone perception and production (Flege, 2005). A non-native phone will be perceived as either 'similar' or 'different' to an L1 phone, and the perception and production of the L2 phone depend on how 'similar' this phone is to any L1 category (Flege, 1987). One hypothesis that stems from SLM is that the greater the dissimilarity between an L1 and an L2 phone, the more likely a new category will be formed for the L2 phone (Flege, 2005). If an L2 phone is perceived as different from any L1 phone, then a listener is likely to build up a new category for this phone. A second hypothesis is that an L2 phone will assimilate to an L1 category if it is perceptually similar to an L1 phone (Flege, 2005). Finally, SLM predicts that the degree

of perceptual dissimilarity between two phones strongly influences how "accurately" an L2 phone will be produced (Flege, 2005; Major, 2008). If an L2 phone is perceptually similar to an L1 phone, then an L2 learner will produce the L1 phone. Alternatively, if an L2 phone is perceptually different from an L1 category, the learner will not use L1 categories to produce this phone. Because SLM assumes the acoustic-phonetic cues to subsume L2 phonological learning, these categories are acoustic-phonetic categories.

The Perceptual Assimilation Model (PAM) has also been central to discussions about non-native phone perception. PAM is importantly a model of naive listener perception. In this model, a listener maps non-native phones to the phoneme categories in their L1 (Best, 1991). A listener's ability to discriminate non-native sound contrasts depends on how these sounds are mapped to phoneme categories in their L1. If two sounds are mapped to two separate phoneme categories in their L1 (Two-Category assimilation), a listener will be excellent at perceiving the contrast between the non-native phones. If two sounds are mapped to the same category in their L1 (Single-Category assimilation), a listener will display poor discrimination of these two sounds. If two non-native phones are mapped to the same L1 category, but one phone is a better exemplar of this category (Category-Goodness assimilation), listeners will vary in how well they can discriminate these phones. Non-native phones may also be "uncategorized" (Tyler et al., 2014). In such a situation, a non-native sound is so different from any L1 category that it will not be assimilated to any L1 phone, and listeners will have excellent discrimination abilities of an uncategorized phone compared to a categorized phone.

PAM is designed to make explicit predictions about how naive listeners perceive non-native phones, but does not make explicit predictions about *L2* perception. PAM-L2 extends the predictions proposed by PAM to L2 perceptual learning, while adhering to some of the central tenets of SLM (Best and Tyler, 2007). Best and Tyler (2007) discuss central postulates of SLM, and where the assumptions of PAM differ. Important to the present discussion, SLM states that language specific aspects of sounds are stored in memory, and are called phonetic categories. PAM and PAM-L2 do not agree with SLM about what aspects of speech sound are stored as representations. SLM assumes that sound categories are formed through perception of acoustic-phonetic cues, whereas PAM and PAM-L2 take a direct-realist approach to speech perception, in that listeners perceive articulatory gestures of speech sounds. In this view, listeners directly perceive articulatory gestures, and learning happens when learners perceive "higher-order invariant information in the speech signal" (Best and Tyler, 2007). Thus, PAM-L2 assumes two levels of representation are involved in L2 perceptual learning; learners may detect differences and similarities between L1 and L2 sounds on the phonetic level or the phonological level. The phonetic level is defined as perception of gestural information that are sub-lexical, such as dialectal differences in gestures, or allophonic differences. The phonological level is defined as perception of gestural information that is relevant to lexical contrasts, akin to the "higher-order invariant information in the speech signal", discussed by Best and Tyler (2007). Perceptual similarity between an L1 and an L2 on the phonetic level implies that a learner does not perceive differences between the two sounds, meaning they do not perceive a difference in the articulatory gestures of the two sounds. Perceptual similarity between an L1 and an L2 phone on the phonological level does not necessarily imply that the two phones are perceived as equivalent. One example used by Best and Tyler (2007) to show the distinction between similarity on the phonetic and phonological level is the French and English rhotics. French /r/ is produced as a uvular fricative, and the English /r/ is produced as an alveolar approximate, and although these two sounds are

often not confusable to learners on the phonetic level, learners tend to equate these two phonological categories (Best and Tyler, 2007).

Keeping in mind these differences from SLM, PAM-L2 also makes explicit predictions about L2 perceptual learning. First, if only one L2 phone is perceived as phonetically similar to an L1 category, learners are not predicted to have further perceptual learning. This L2 phone will be mapped to the corresponding L1 phone category. The next case is that learners may perceive an L2 phone as phonologically similar, but phonetically deviant, from an L1 phone. This is the case with /r/ for English learners of French. Because French /r/ is phonetically different from English /r/, learners will build a new phonetic category for French /r/ within the same phonological category (Best and Tyler, 2007). The next possibility is that two L2 phones will be assimilated to the same L1 phonological category, but one phone will be considered 'deviant', similar to the Category-Goodness assimilation of PAM. PAM-L2 predicts that a new phonological and phonetic category will be formed for the deviant phone, whereas no new category will be formed for the 'better fit' phone. Finally, two L2 phones may be considered equally good or equally bad instances of an L1 phone, in which case they would be assimilated to the same L1 phonetic and phonological category, similar to Single Category assimilation. Whether or not learners can learn to perceive the difference between L2 phones that are assimilated to the same category depends on whether the phones are good or poor examples of an L1 phone.

PAM-L2 is a model for L2 perceptual category formation, and is not explicitly a model of L2 production. However, because the articulatory gesture is considered the basic phonetic unit, it can be assumed that production will follow from the L2 perceptual categories that have been formed. If a learner perceives an L2 phone as phonetically and phonologically similar to an L1 phone, then it can be hypothesized that the learner will recruit L1 articulatory gestures to produce the L2 phone. If a learner perceives two L2 phones as phonetically and phonologically similar to an L1 phone, but one L2 phone is more 'deviant' (a Category-Goodness assimilation pattern), then L1 gestures may only be recruited for the more similar L2 phone.

The present study uses articulatory data of L2 learners' productions of high round vowels to compare the assumptions of PAM-L2 and SLM, and finds that learners do not reuse L1 articulatory gestures in productions of 'similar' L2 phones. L1 and L2 French learners are recorded producing the high vowels /i, u/ in English, and /i, y, u/ in French. Even when learners were non-target like in their acoustic productions of French /u/, they did not appear to be using their L1 articulatory category to produce the L2 phone. For the phonetically and phonologically 'new' phone /y/, English-French learners remain non-target like in fronting of their tongue, showing that learning an entirely new category remains difficult for learners.

1.2 Acquisition of French high vowels

Acquiring the contrast between high round vowels is a well-documented difficulty for L1 English learners of French (Flege, 1987; Levy and Strange, 2008; Kamiyama and Vaissière, 2009; Darcy et al., 2012; Pillot-Loiseau et al., 2015). Perceptually, beginner learners have a tendency to merge the high front round and high back round vowel categories (Levy and Strange, 2008). Because English does not have front round vowels, learners tend to perceive French /y/ as similar to /u/. This is sensitive to context effects, and learners may be more likely to hear the /y/-/u/ contrast following bilabial consonants (Levy and

Strange, 2008).

Production of the /y/-/u/ contrast has also been shown to be difficult for learners, as evidences through acoustic data. Importantly, this contrast is difficult for learners to produce because French /u/ tends to have non-target like acoustic values for L1 English speakers (Flege, 1987). French /u/ has a low F2, particularly compared to English /u/ (Flege, 1987). Production studies have found L2 phones that are similar phones to L1 phones to be more difficult for adult learners to acquire than completely new phones (Flege, 1987). Acoustic data shows that beginner, intermediate, and advanced learners all tend to produce target-like /y/, but learners of all proficiency levels have difficulty producing French /u/. Flege (1987) argues that these results support a theory of equivalence classification. Equivalence classification suggests that learners will classify non-native phones as L1 phones. This will lead to production errors when the L2 phones are not acoustically identical to L1 phones. Learners produce French /u/ as if it were their L1 category. This causes production errors, and French /u/ to be produced with a higher F2 value. Alternatively, French /y/ is more target-like for English-French learners. The front round vowel cannot be classified as an L1 phone because there is no category similar to /y/ in English. Therefore, learners will make an entirely new category for /y/. These findings support the claim that new phones are easier to acquire than similar phones.

Similar results have been found for L1 Japanese L2 French learners. In a similar production study, Kamiyama and Vaissière (2009) compared Japanese-French learner productions of /u/, /y/, and /ø/. This study found that Japanese-French learners had difficulty producing target-like /u/, similar to English-French learners. Japanese learners were acoustically producing /u/ as they would in their L1, which led to production errors. For the front round vowels, Japanese-French learners produced target-like /y/ and /ø/ (Kamiyama and Vaissière, 2009). /y/ and /ø/ do not have comparable categories in Japanese, making these phonemically and phonetically new phones. These results indicate that phonetically similar but phonemically new vowels are difficult to produce (Kamiyama and Vaissière, 2009).

Taken together, the results discussed in this section have supported equivalence classification, as proposed in SLM, highlighting the importance of L1 categories in shaping the production of L2 phones. However, little work has been done to understand how L1 articulatory strategies affect L2 productions. Equivalence classification predicts that learners will produce at L2 phone as an L1 phone if the two phones are similar, suggesting learners may use L1 articulatory strategies to produce L2 phones. The present study is designed to address this question.

The results of the studies discussed here focus in particular on the production of L2 phones that are perceived as similar to L1 phones. However, if an L2 phone does not have an L1 counterpart, or is perceived as a deviant exemplar of an L1 phone, this vowel will have its own category created in a speaker's phonological space. This claim is supported by acoustic data (Flege, 1987; Kamiyama and Vaissière, 2009). From the acoustic data, it appears that when English-French learners have an acoustically merged /y/-/u/ category, /u/ is produced with an F2 that approaches /y/. How, then, is this contrast realized if a 'new' phone has been formed for /y/? Additionally, the results from previous production studies do not discuss the articulatory strategies used to produce novel phones. Presumably, L1 English L2 French learners are producing French /u/ with the same articulatory strategies as used to produce English /u/. However, for new phones such as /y/, and learner must use novel combinations of articulatory gestures. A

learner must maintain the tongue position of /i/ while producing a rounding lip normally associated with /u/. As previous results have shown that learners can produce target-like /y/, they must be using a combination of novel articulatory gestures to reach this acoustic target.

This study addresses how L2 category formation occurs articulatorily, and will seek to answer the following research questions:

1. How do English-French learners' productions of high vowels differ from L1 French speakers? It has been discussed that English-French learners have difficulty producing French /u/ because they tend to have a higher F2 value. This acoustic tendency may be caused by multiple articulatory mechanisms, namely tongue fronting and lip rounding. This study seeks to understand how, and if, learners deviate from L1 French speakers articulation of these vowels. Additionally, SLM hypothesizes that /y/ will be acoustically target-like for learners, since it is a new category, but it remains to be seen whether learners produce /y/ articulatorily target-like.

2. Do English-French learners have a merged /y/-/u/ category? Previous production studies have addressed this question via acoustic data, but to-date, there is no data showing how articulatory gestures are involved in maintaining this contrast. As discussed above, the French /y/ and /u/ may both be assimilated to the same L1 category, but /y/ may be a more deviant example, following CG assimilation pattern. If learners have a merged /y/-/u/ category, are these phones produced with different articulatory gestures?

3. Do learners use L1 articulatory gestures to produce L2 phones? PAM-L2 suggests that learners perceive L2 phones as articulatory gestures, and these sounds will be mapped to L1 categories accordingly. If L2 French /u/ is mapped to L1 English /u/, than it is hypothesized that learners will recruit L1 gestures to produce French /u/.

To address these questions, this study uses Ultrasound Tongue Imaging and video recordings of lip rounding of 6 L1 English L2 French learners and one native French speaker.

2 Methodology

2.1 Participants

7 participants, 6 L1 English L2 French learners and 1 L1 French speaker, completed a production task. 6 L2 French speakers (EN02-EN07) are currently enrolled in Intermediate 2 or Advanced Intermediate French courses at a university in the United States. All the L2 French speakers are self-reported native speakers of English. The L1 French speaker (FR01) is from Northern France, and is a proficient speaker of English. All participants are female and between the ages 19 and 28.

2.2 Procedure

Each participant completed a production task in French and a production tasks in English. Participants were instructed to read a word list, designed to elicit tokens of /i/, /u/, and

/y/. Filler words were included as distractors. Participants read each word in isolation. The French word list included 70 words, and can be found in Appendix A. The English word list consisted of 50 words, and can be found in Appendix B. All words were read by the participants once. The present study analyses 5 tokens of each target vowel, /i/ and /u/ from the English word list, and /i/, /y/, and /u/ from the French word list. Each vowel was balanced for phonological context to control for coarticulatory effects. Participants also completed a reading passage task after reading the word list, but the present study only analyzes tokens from the word list task, and thus the reading passage will not be discussed further.

Participants completed the tasks in a 45 minute recording session. Participants wore a stabilized headset, which was adjusted to fit participants' heads. An Ultrasound probe was attached to the headset beneath the speakers' jaw to record a midsagittal view of the tongue movements, at a frame rate of 84 frames/second. A camera was attached to the side of the headset and was adjusted to capture lip protrusion. Finally, participants wore an AKG C544L headset microphone and acoustic data was recorded using a Marantz PMD661 recorder. Ultrasound data was recorded using Articulate Instruments software (AAA, 2012), and tongue contours were extracted using EdgeTrak (Li and Stone, 2005).

Acoustic data was analyzed using PRAAT (Boersma and Weenink, 2017). F1 and F2 were calculated at the mid-point of the vowels. To measure lip protrusion, still frames were extracted from the videos at the maximum point of constriction. The maximum point of constriction is defined here as the video frame after movement towards a position is reached, and before movement away from the position begins. Following Havenhill et al. (2017), lip protrusion was measured as the horizontal distance between the corners of the mouth, and the vertical distance from the upper and lower lips. Distances were measured using Inkscape's measurement tool that converts pixels to millimeters. Finally, still frames of tongue contours were taken at the point of maximum constriction, again defined as the point in time after a maximum point is reached before movement away from the constriction begins. Maximum point of constriction was chosen as the measurement for vowels in the present study because this study is comparing vowel target gestures. The maximum point of constriction represents the 'target' of the vowel gestures in this sense. For each vowel token, 100 points were extracted along the contour in Edgetrak. Contours were hand corrected before extraction. Tongue contours of each speaker's vowels were compared using SSANOVA in R.

Articulatory results are compared within each speaker. Because each speaker has a different size vocal tract, and the ultrasound probe and video camera are inevitably placed at slightly different angles for each speaker, comparing results across participants is not feasible. Rather, each speaker's productions of the vowel contrasts are compared. This way, it is possible to compare a speaker's tongue position for /i/, /y/, and /u/, and to compare how a speaker produces L1 /u/ to L2 /u/. SSANOVA is used to compare amount of overlap between these vowel tongue positions.

3 Acoustic Results

The contrast between /y/ and /u/ lies primarily on the F2 dimension. Previous studies have found that F2 is a more reliable measurement of the /y/-/u/ contrast than F1 or F3 (Flege and Hillenbrand, 1984; Flege, 1987). The present study will therefore use F2 measurements to compare L2 French speakers' productions of /y/ and /u/ to L1 French

speaker's productions.

3.1 French Vowel Acoustic Results

Acoustic results were first compared across speakers to address the first research question of how English-French learners differ from native French speakers in production of the high round vowels. To compare L2 French productions of /y/ and /u/ to L1 productions of /y/ and /u/, F1, F2, and F3 formant values were used to normalize across speakers, using the Lobanov method. Normalization of vowel formants was done using the Vowels package in R (Kendall et al., 2013). A t-test was run to compare each learners' normalized F2 values of /y/ and /u/ to the F2 values of the L1 speaker's (FR01) /y/ and /u/ tokens. The p-values of each t-test are presented in Table 1.

	Mean (SD)	t-value	p-value
FR01	.4928 (.2934)		
EN02	.1658 (.81)	.8413	.4349
EN03	-.6006 (1.168)	2.03	.1043
EN04	.5446 (.3707)	-.245	.8129
EN05	-.5132(.9478)	2.267	.0754
EN06	-.4024 (.5868)	3.0512	**.0239***
EN07	.4738 (.2629)	.10784	.917

Table 1: Normalized /y/ F2 Values: Comparing learners to FR01 (n=5)

	Mean (SD)	t-value	p-value
FR01	-1.553 (.2962)		
EN02	-.795 (.6452)	-2.3892	**.056***
EN03	-.3108 (.7145)	-3.5926	**.014***
EN04	-.9348 (.7957)	-1.6296	.263
EN05	-.6772(.5358)	-3.2	**.017***
EN06	-.662 (.2212)	-5.367	**.0008***
EN07	-1.073 (.431)	-2.0551	.078

Table 2: Normalized /u/ F2 Values: Comparing learners to FR01 (n=5)

Cross-speaker acoustic results support the findings of Flege (1987), that learners are more likely to have target-like /y/ productions than /u/. 5 of the 6 learners do not produce a significant difference in the F2 value of /y/ tokens compared to the L1 speaker, suggesting these learners are producing target-like /y/. EN06 is the only French learner who produces a significantly different F2 value for /y/ than the L1 speaker, meaning their /y/ productions are not target-like. The mean F2 value for EN06's /y/ is lower than FR01. We will return to the discussion of EN06 below.

Learners do seem to have difficulty producing target-like French /u/, a phonetically similar phone to English /u/. Three of the learners (EN03, EN05, EN06) produce /u/ with a different F2 value from the native speaker (p=.014, p=.017, p=.0008). One of the speakers (EN02) approaches a significant different from the native speaker in the F2 value of /u/ (p=.056). As will be discussed below, EN02 patterns with the other three learners who do not produce target-like /u/ in that she does not produce /y/ and /u/

with different F2 values. Because EN02 appears to have an acoustically merged /y/-/u/ category, and has marginally significant difference in F2 of /u/ compared to the native speaker, EN02 is considered to pattern with EN03, EN05, and EN06 in having non-target like French /u/ productions.

Overall, cross-speaker acoustic results show that learners are more likely to produce non-target like /u/ than /y/. However, 2 speakers (EN04 and EN07) are target-like in productions of /y/ and /u/, meaning they do not produce a significantly different F2 value from the L1 speaker's /y/ and /u/.

Turning now to test whether learners have acoustically merged /y/ and /u/ categories, a t-test was run comparing the F2 values for each learners' /y/ and /u/ productions. Table 3 shows the results of the t-test. Because no measurements were made across speakers, t-test were run using raw hertz values for F2.

	/y/ Mean (SD)	/u/ Mean (SD)	t-value	p-value
FR01	1902.114 (739.09)	1086.068 (380.01)	-2.442	**.0436***
EN02	2208.273 (460.106)	1396.34 (252.52)	2.074	.0735
EN03	1457.53 (761.65)	1646.36(465.81)	.479	.6514
EN04	2346.95 (253.411)	1335.44 (544.07)	-3.768	**.0103***
EN05	1667.52 (540.68)	1573.89 (305.44)	-.3371	.7469
EN06	1483.883 (329.62)	1335.607 (124.2)	-.9412	.3889
EN07	2093.43 (136.86)	1287.92 (224.411)	6.614	**.0003***

Table 3: Comparison of /y/-/u/ F2 values (Hz) within each speaker (n=5)

The L1 French speaker does produce a significant difference in the F2 value of /y/ and /u/, suggesting these categories are distinct. EN04 and EN07 also produces a significant difference between French /y/ and /u/, suggesting they produce an acoustic difference between the two high round vowels. These are the same two learners who did not reach a significantly different F2 value French /u/ compared to FR01. The other four learners (EN02, EN03, EN05, and EN06) do not produce a significant difference between /y/ and /u/ F2 values. The the same four learners who produced /u/ significantly differently from FR01 did not produce a significant difference between /y/ and /u/.

3.2 French vs. English Acoustic Results

We now turn to the cross-language acoustic comparisons to address the question of whether learners use L1 acoustic categories to produce L2 phones. English /u/ is typically produced with a higher F2 value than French /u/ (Flege, 1987). In order to test the hypothesis that English /u/ is causing learners' French /u/'s to have a higher F2 value, a t-test was run comparing the F2 values, in hertz, of each learners' French /u/ category to English /u/ category. The results are presented in Table 4.

	English /u/ Mean (SD)	French /u/ Mean (SD)	t-value	p-value
EN02	1396.34 (252.52)	1462.78 (366.37)	-.333	.7482
EN03	1926.88 (209.1066)	1646.36(465.81)	-1.2285	.2688
EN04	1068.85 (133.74)	1335.44 (544.07)	1.064	.3413
EN05	1738.4 (314.212)	1573.89 (305.44)	-.8394	.4256
EN06	1775.59 (359.5)	1335.607 (124.2)	-2.586	**.0495***
EN07	1797.86 (586.01)	1287.92 (224.411)	-1.817	.1272

Table 4: Comparison of learner's English /u/ and French /u/ F2 values (Hz) within each speaker (n=5)

Most of the speakers do not have a significant acoustic difference in F2 value between French /u/ and English /u/. Interestingly, EN06 does produce a significant difference between French /u/ and English /u/. This speaker has a high average F2 value for English /u/ tokens, and slightly lower F2 value for French /u/ tokens.

3.3 Interim Discussion

Based on how learners pattern with regard to acoustic results, the learners will be categorized in two groups; Group A consists of EN04 and EN07. These two learners do not have an acoustically merged /y/-/u/ category, and produce French /u/ with a comparable F2 value to FR01. Group B consists of EN02, EN03, EN05, and EN06. These speakers do not have a significant acoustic difference between /y/ and /u/, and also do not have target-like F2 values for /u/. EN02, who barely produced /u/ significantly differently from FR01 does not have a significant difference in the /y/-/u/ F2 contrast. EN02 thus patterns with the other learners in Group B. It should be noted that these groups were not created based on any external proficiency measures, and are only included to more easily compare acoustic and articulatory results.

As shown in the cross-speaker acoustic results, EN06 produced /y/ with a significantly lower F2 value than FR01, and was the only learner to do so. The results in Table 3 show that the low F2 value for /y/ is causing much of the overlap between the /y/-/u/ category. While the other learners in Group B seem to have difficulty producing target-like /u/ because they have a higher F2 value for French /u/, EN06 is also producing /y/ with a lower F2 value. EN06 also produces French /u/ with a higher average F2 value than FR01. This means that EN06 appears to have a merged /y/-/u/ category that is has a lower F2 value than the native-French speaker /y/, but higher than the native-French speaker /u/.

The acoustic results suggest that learners in Group B have a merged /y/-/u/ contrast in French. This is because learners in Group B are producing /u/ with a higher F2 value than the L1 French speaker. Most learners in Group B have similar F2 values for English /u/ and French /u/ tokens. Alternatively, learners in Group A have acquired the contrast between /y/ and /u/, and produce French /u/ with an appropriately low F2. There are two major articulatory strategies that cause F2 to lower: backing of the tongue, and protruding the lips. We now turn to results of articulatory data to better understand the articulatory strategies causing Group B to produce non-target like /u/, and appear to have a merged /y/-/u/ category.

4 Articulatory Results

Initial results suggest that tongue position is responsible for the difference in L1 and L2 productions of high round vowels. L2 French speakers do not differ from the L1 French speaker in degree of lip rounding for the three vowels /i/, /y/, and /u/.

4.1 Lip Rounding Results

Learners are rounding both /y/ and /u/ to the same degree as the native French speaker. Figure 1 shows the distance of FR01's vertical and horizontal opening of the lips, in millimeters, for each of the French high vowels. Figures 2-7 show the lip measurements, in millimeters for the learners' French and English vowels. A lower value for openness corresponds to more lip rounding, following Havenhill et al. (2017).

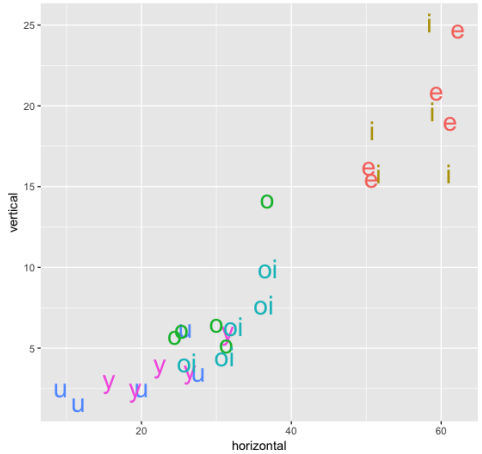

Figure 1: Vertical and horizontal distance of lip opening for FR01's production of French Vowels (mm)

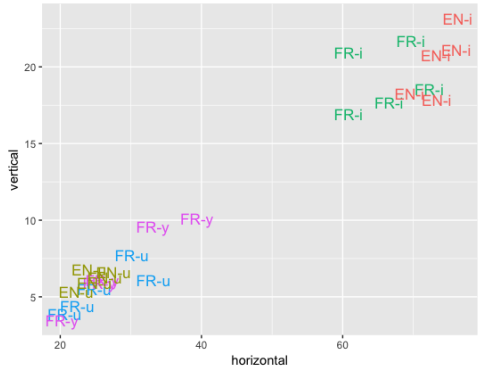

Figure 2: Vertical and horizontal distance of lip opening for EN01's productions of English and French Vowels (mm)

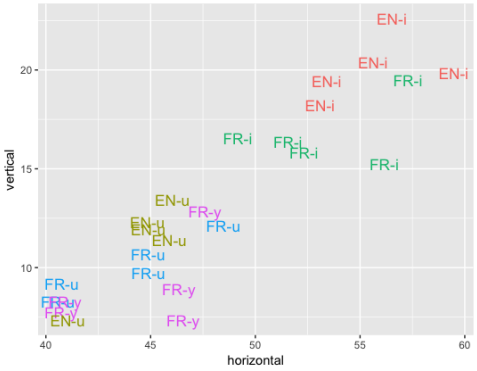

Figure 3: Vertical and horizontal distance of lip opening for EN03's productions of English and French Vowels (mm)

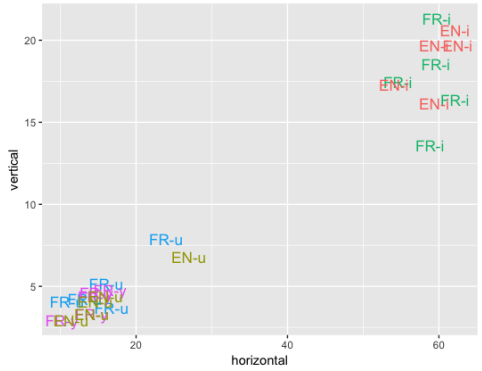

Figure 4: Vertical and horizontal distance of lip opening for EN04's productions of English and French Vowels (mm)

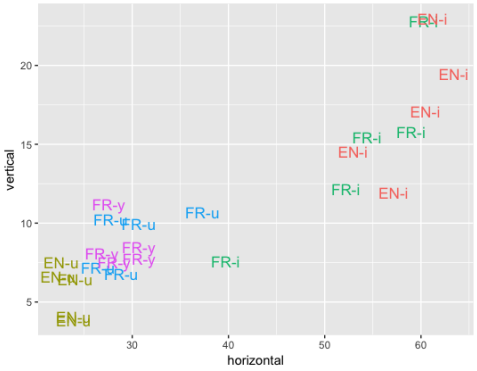

Figure 5: Vertical and horizontal distance of lip opening for EN05's productions of English and French Vowels (mm)

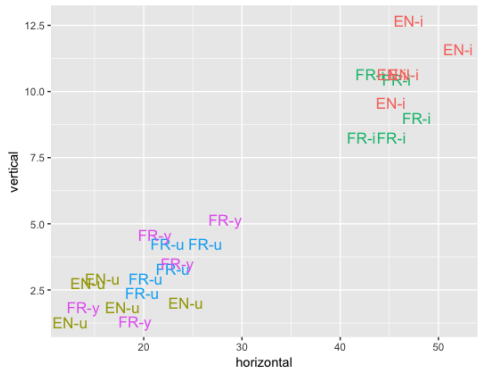

Figure 6: Vertical and horizontal distance of EN06's productions of English and French Vowels (mm)

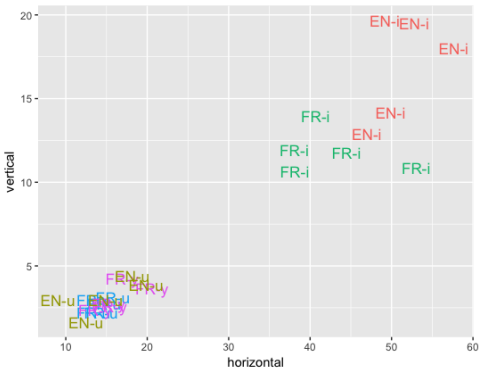

Figure 7: Vertical and horizontal distance of EN07's productions of English and French Vowels (mm)

Two one-way ANOVA tests were performed for each speaker to see if the degree of opening (and thus lip rounding) is different for each vowel within each speaker. One ANOVA looking at the interaction between vowel category and horizontal distance of lip opening was run for each speaker, and one ANOVA looking at the interaction between vowel category and vertical distance of lip opening was run for each speaker. English and French high vowels were included as separate categories in each ANOVA, in order

to compare the degree of lip rounding across each vowel category in each language. A post-hoc Tukey HSD test was run for each speaker to see which vowels had significantly different degrees of lip rounding. Again, it is necessary to compare only within each speaker, as each speaker has a different size vocal tract. Despite only comparing within speakers, it is still possible to see how speakers are rounding for the different vowel categories. The results of the ANOVA's are presented in Table 5.

Vertical opening

	F	p-value
FR01	60.08	5.61e-07*
EN02	70.21	1.74e-11*
EN03	31.83	2.05e-08*
EN04	80.42	4.92e-12*
EN05	9.776	.00015*
EN06	61.36	6.04e-11*
EN07	77.85	6.66e-12*

Horizontal opening

	F	p-value
FR01	48.97	1.69e-06*
EN02	116.3	1.5e-13*
EN03	17.93	2.14e-06*
EN04	168	4.36e-15*
EN05	60.61	6.76e-11*
EN06	74.52	1e-11*
EN07	95.02	1.02e-12*

Table 5: One-way ANOVA results comparing lip opening and vowel category (n=5)

The results of each speaker's ANOVA show a main effect of vowel category on the distance of lip opening on both the horizontal dimension and the vertical dimension for all speakers. The Tukey HSD tests reveal that, in general, there is no difference in amount of rounding of the round vowels /y/ and /u/, but both round vowels are more round that /i/.

The first research question asks whether learners and the native French speaker produce the French high vowels similarly. All learners show similar lip rounding patterns to the native speaker, in that the French round vowels /y/ and /u/ are more round than French /i/. To look at the specific opening measurements, the distance of the horizontal opening of the lips is smaller for /u/ and /y/ than /i/ for all speakers (learners and the native speaker). On the vertical dimension, we find similar results. The native speaker and the learners have a smaller vertical distance in lip opening for /y/ and /u/ than /i/. Only one speaker (EN05) does not round French /u/ or /y/ more than French /i/ on the vertical opening. As can be seen in Figure 5, French /i/ tends to have a slightly smaller opening on the vertical dimension, suggesting it is slightly more rounded. Despite this finding, EN05 still shows as significant difference between French round and unround vowels on the horizontal dimension. Additionally, EN05 does round French /u/ and /y/ significantly more than English /i/ on the horizontal and vertical dimension, suggesting /y/ and /u/ are indeed rounded. All other learners pattern similarly to the native speaker in rounding the round French vowels more than the unround French vowel.

Next, we turn to a comparison of whether learners have a merged category for French /y/ and /u/. Tukey HSD tests show that all learners round French /y/ and /u/ to a similar degree. Learners do not produce a difference in degree of lip opening on the horizontal or vertical dimension for /u/ or /y/. This is similar to the native speaker, who produces a similar degree of lip rounding for /y/ and /u/.

Turning to the third research question comparing whether learners use L1 English categories for L2 French productions, learners do appear to use the same lip posture for L1 and L2 round vowels. Tukey HSD tests show that none of the learners have a significant difference in degree of vertical or horizontal lip openings for the round vowels.

Learners are rounding French /u/, French /y/, and English /u/ to a similar degree.

4.2 Tongue Position Results

Smoothing Spline curves were generated for the tongue contours of each vowel for each speaker. The smoothing splines represent the averages for 5 vowel tokens each. SSANOVAs are used to determine if there are differences between sets of data. This data is typically a continuous factor, such as time or distance. SSANOVAs are typically used to compare articulatory data, as tongue positions for different categories can be compared. This is particularly useful in observing vowel contrasts, or lack there of, for speakers. Smoothing splines are generated for the best fit of each group, and Bayesian 95% confidence intervals are generated along the smoothing splines (Davidson, 2006). The difference between the groups is significant at any point where the confidence intervals do not overlap. SSANOVAs are thus graphically represented, and it can be determined which splines have a different shape based on where the curves do not overlap. For the present study, each speaker's productions of /i/, /y/, and /u/ were compared, and subsequently productions of French /u/ and English /u/ were compared. Distance is measured in pixels from origin coordinates extracted from EdgeTrak. Bayesian 95% confidence intervals are represented by dotted lines around the curves. Any point where the confidence intervals do not overlap suggests the tongue positions are significantly different at that point. The tongue tip is positioned at the left in Figure 8-20, and the tongue root is to the right. For more on smoothing spline curves, see Davidson (2006).

4.2.1 French Vowel Tongue Positions

Figures 8-14 show the SSANOVA results for each speaker's French vowels. Because the visibility of the entire tongue contour varies from token to token, variation at the tip of the tongue and the root of the tongue is an expected artifact of Ultrasound Tongue Imaging.

We first compare how the native-speaker produces the contrast between /i/-/y/-/u/ to the learners. As can be seen in Figures 8-10, the native speaker and Group A learners produce a significant degree of overlap in tongue position for the front vowels /i/ and /y/, while the tongue position for /u/ does not overlap with any of the front vowels. For the native speaker in Figure 8, the maximum point of constriction for /u/ is further back that /y/. At the tip of the tongue and at the root, /y/ and /i/ are not significantly different, suggesting the two front vowels have a similar tongue position.

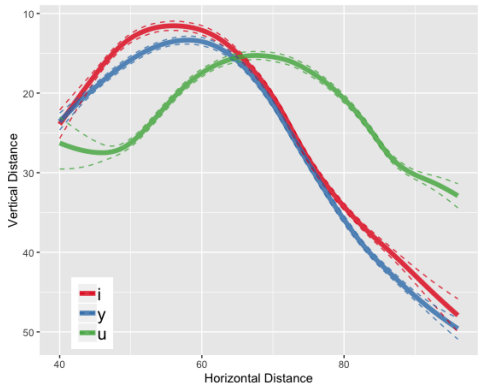

Figure 8: SSANOVA results comparing the tongue position of French high vowels produced by native speaker FR01

The SSANOVAs for the learners in Group A, who did produce an acoustic contrast between French /y/ and /u/ are presented in Figures 9-10, and show similar results to the native speaker. The round vowels /y/ and /u/ differ significantly along the entire curve of the tongue for both speakers in Group A. However, EN07 does produce a difference in tongue position for /i/ and /y/ along the tongue contour. Despite this finding, the maximum point of constriction remains similarly fronted for EN07's /i/ and /y/ and these two vowels do not overlap with the back vowel /u/.

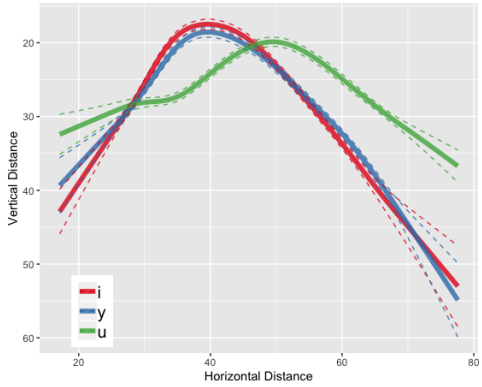

Figure 9: SSANOVA results comparing the tongue position of French high vowels produced by learner EN04

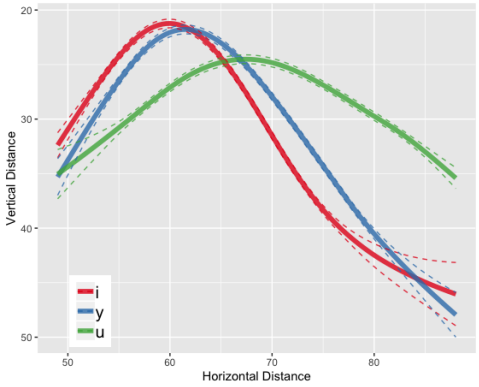

Figure 10: SSANOVA results comparing the tongue position of French high vowels produced by learner EN07

In general, Group A learners display similar patterns to the native French speaker in terms of tongue position for French vowel contrasts.

The results for the Group B learners (the non-target like acoustic group) show a lot of individual differences in tongue positions, but generally indicate that these learners have a fronted tongue position for /u/, and a more backed tongue position for /y/ compared to the native speaker's vowel contrasts. The SSANOVAs for the learners in Group B are presented in Figures 11-20. Recall from Figure 8 that the native French speaker had a significant degree of overlap in tongue position for the vowels /i/ and /y/, and no overlap in tongue position of the front vowels and the back vowel /u/. Only one learner in Group B (EN03, shown in Figure 12), shows overlap of /i/ and /y/. Learners EN05 (Figure

13) and EN06 (Figure 14) produce /y/ further back than /i/. Furthermore, two out of four learners in Group B (EN03 and EN05) show overlapped tongue positions at some point along the contour between the round vowels /y/ and /u/, which importantly differs from the native speaker's productions. On an individual level, no learners in Group B follow native speaker patterns in having overlap in tongue position for /i/ and /y/, and additionally no overlap along any point of the curve for /u/.

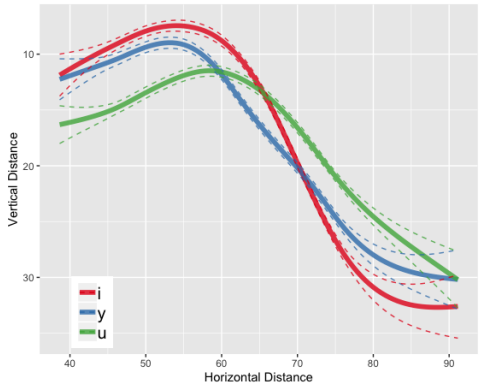

Figure 11: SSANOVA results comparing the tongue position of French high vowels produced by learner EN02

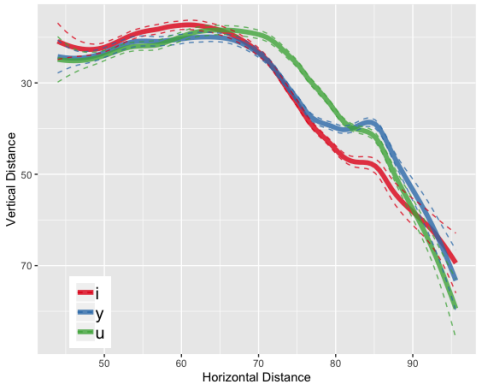

Figure 12: SSANOVA results comparing the tongue position of French high vowels produced by learner EN03

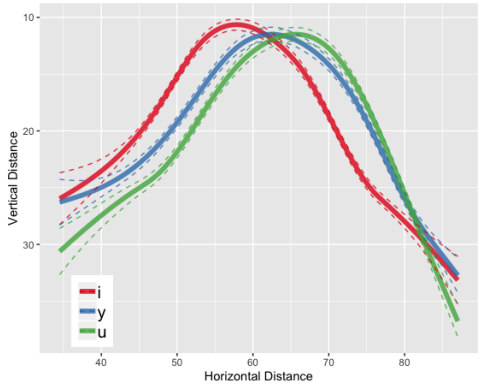

Figure 13: SSANOVA results comparing the tongue position of French high vowels produced by learner EN05

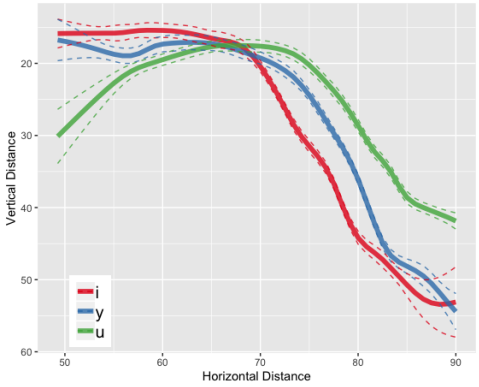

Figure 14: SSANOVA results comparing the tongue position of French high vowels produced by learner EN06

Turning to discussion of whether learners have a merged tongue position for the round vowels /y/ and /u/, results show that learners in Group A and in Group B produce the contrast between /y/ and /u/ in tongue position. Both speakers in Group A (Figure 8 and Figure 9) show no overlap between /y/ and /u/ along any point of the tongue contour. Group B learners, who importantly all have acoustically merged /y/ and /u/, generally produce at least some difference in tongue position for /y/ and /u/. EN03 (Figure 12) shows a slightly backed position for /u/ compared to /y/ at the root of the tongue, although this contrast is relatively small. EN05 (Figure 13) produces a difference in tongue position for /y/ and /u/ at the front of the tongue, but does show overlap along the root of the tongue contour. The other two speakers in Group B (Figure 11 and Figure 14) show no overlap between the round vowels /y/ and /u/, and back /u/ more than /y/. Overall, learners tend to produce /y/ and /u/ with different tongue position, regardless of whether they produce an acoustic difference between these vowels.

4.2.2 Cross-Language Tongue Position Comparison

The cross-language SSANOVA results show that learners are not using the same tongue position to produce French /u/ and English /u/. Figures 15-16 show the comparison of Group A learner tongue position for their French /u/ category compared to their English /u/ category. In Figure 15, we can see that EN04 fronts her French /u/ significantly more than English /u/ throughout the entire length of the tongue. French /u/ is more front than her English /u/. Interestingly, these results are not apparent looking at the acoustic results. Looking at Figure 16, EN07 has a slight difference in tongue position between French /u/ and English /u/. Around the tip of the tongue and near the root of the tongue, EN07 has a more backed French /u/ than English /u/. Near the maximum point of constriction, the two vowels are overlapped.

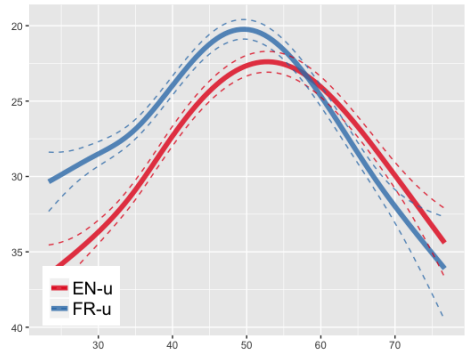

Figure 15: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN04

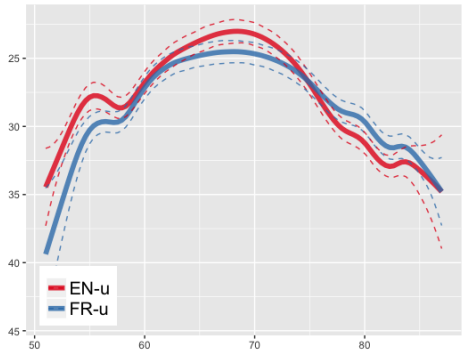

Figure 16: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN07

Turning to Group B results, recall from section 3.1 that 3 of the 4 learners in Group B did not produce a significant difference in their English and French /u/ categories. Figures 17-20 show the comparisons of Group B's French /u/ and English /u/ tongue positions. Interestingly, most of the learners in this group do show significantly different tongue positions for the two languages /u/ categories.

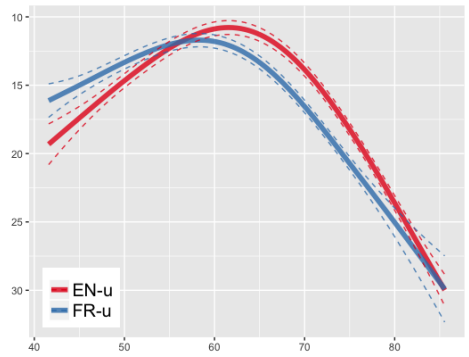

Figure 17: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN02

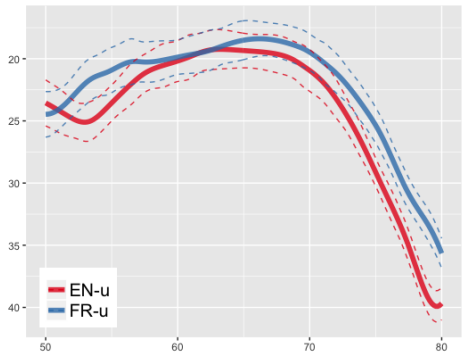

Figure 18: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN03

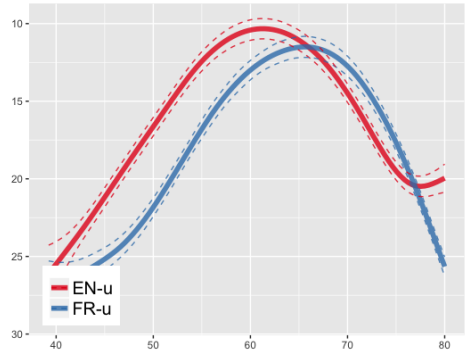

Figure 19: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN05

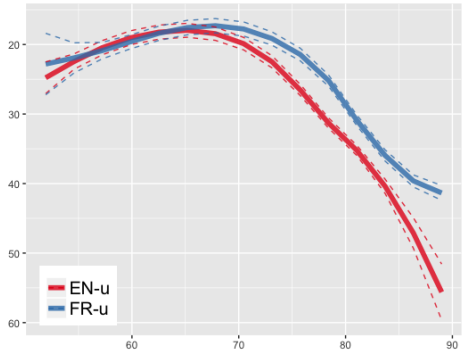

Figure 20: SSANOVA results comparing the tongue positions of French /u/ and English /u/ produced by learner EN06

EN02 has similar results to EN04 in Group A. EN02 has a more fronted tongue position for French /u/ than English /u/. Because French /u/ has a lower F2 value than English /u/, production of a target-like French /u/ category would require the opposite result.

EN03 does not produce a significant difference in French /u/ and English /u/ tongue positions. These two vowels are not significantly different anywhere along the curve.

Learners EN05 and EN06 both produce French /u/ with their tongue further back than English /u/. EN05 has a more backed tongue position along the entire length of the tongue for French /u/ compared to English /u/. EN06 has a further back position of the tongue root for French /u/ than English /u/.

Group A and Group B SSANOVA results both show that learners tend to use different tongue positions to produce L1 and L2 /u/.

5 Discussion

The acoustic and articulatory data of the present study provide insight into the relationship between L1 and L2 phonetic categories. First, the L1 English-L2 French learners tended to produce English /u/ with the same F2 value as French /u/, showing that learn-

ers were using the same acoustic category in L1 and L2 productions. The articulatory results, on the other hand, showed that learners do not have the same tongue positions for French and English /u/. This result suggests that learners may be reusing L1 acoustic categories, rather than articulatory gestures, in production of L2 vowel categories. Secondly, the results show that learners are acoustically target-like in production of the phonetically and phonologically new phone /y/, but were not consistently target-like in their tongue position of /y/. Again, this result suggests that learners have an acoustic, rather than articulatory, target in production of L2 vowels. Three research questions were outlined in Section 1 to specifically test hypotheses posed by theories of L2 category formation, which I return to here.

The first research question asks if learners produce the /y/-/u/ contrast differently than a native French speaker. Acoustically, learners are target-like in production of the new L2 phone /y/, but are non-target like producing the similar L2 phone /u/. This supports the hypothesis of SLM that new L2 phones are easier to acquire than L2 phones that are similar to L1 phones. Articulatorilly, both phones are non-target like, which is not expected if L2 phones have an articulatory target. The L1 French speaker has a significant amount of tongue position overlap between the two front vowels /i/ and /y/, and did not have any overlap between /u/ and /y/. Recall that every learner, except EN06, produced target-like acoustic productions of /y/. Even though most of the learners produce a target-like /y/ acoustically, learners have difficulty fronting the tongue as far forward as /i/. Only one speaker, EN04, produces a target-like degree of overlap in tongue position between /i/ and /y/. EN04 has a significant degree of overlap between these vowels along the entire length of the curve. One other speaker, EN03, has a significant amount of overlap between /i/ and /y/ along the front of the tongue, but /y/ is remains further back towards the root of the tongue. The other learners produce /y/ further back than /i/. If learners were using articulatory gestures as the targets in L2 productions, it is anticipated that learners would be target-like in the articulatory tongue position of new phones. The fact that learners are target-like in acoustic productions of the new phone /y/, but are non-target like in tongue position, suggests that vowels have an acoustic, rather than articulatory, target in L2 productions.

The second research question asked whether L1 English L2 French speakers have a merged /y/-/u/ category. Acoustic results show that 4 out of 6 French learners did not produce a significant difference in the F2 value of /y/ and /u/. These learners were referred to as Group B in this study. The merged /y/-/u/ category for Group B learners was caused by the high F2 value of /u/. Group A learners did produce an acoustic contrast between French /y/ and /u/. These acoustic results support Flege (1987), who found that when L1 English speakers learning French appear to have a merged /y/-/u/ category, French /u/ is not produced with an appropriately low F2 value.

Turning to articulatory differences for learners /y/-/u/ productions, every learner produced a difference in tongue position for /y/ and /u/ at least at some point along the tongue contour. Similar to native speaker results, learners did not produce a difference in the degree of lip rounding for /y/ and /u/. The native speaker results suggest that the articulatory contrast between /y/ and /u/ is in tongue position. The evidence that all 6 learners in the present study produced /y/ and /u/ with different tongue positions supports the claim that learners produce an articulatory difference between the French high round vowels, and therefore do have a contrast for /y/ and /u/. As discussed in Section 1, learners may initially have a CG perceptual assimilation pattern for these two high round vowels, meaning French /u/ is assimilated as a 'good' example of English /u/,

and /y/ is a 'deviant' example of English /u/. The articulatory results of the present study show that learners do produce these vowels with different tongue positions, but the degree to which learners discriminate these vowels varies, as is evidenced by Group A and Group B. Acoustically, Group A learners distinguish French /y/ and /u/, and articulatorily, Group A learners have a significantly different tongue positions along the entire curve of the tongue. This shows Group A learners do not have a merged /y/-/u/ category.

Group B learners did not produce an appropriately back tongue position for French /u/ in many cases, but still produced /y/ and /u/ with different tongue positions. EN02 and EN06 produce /y/ and /u/ with different tongue positions along the entire length of the contour, despite not producing an acoustic difference between /y/ and /u/. The other two learners in Group B, EN03 and EN05, did have significant overlap between /y/ and /u/ at various points along the contours, but did not have complete overlap between these two tongue contours. Group B learners did not produce an acoustic difference between /y/ and /u/, but importantly showed a tendency to produce an articulatory difference between these vowels. Including the articulatory data shows that these learners do, in fact, have a contrast between these two vowel categories. Overall, tongue positions of /y/ and /u/ are distinct for Group B learners. The intermediate learners in Group B likely have the contrast between these vowels, but have not yet achieved the appropriate difference in F2 for /y/ and /u/. This finding would be obscured if only acoustic data were analyzed.

The final research question of this study concerns whether learners use L1 categories to produce similar L2 phones, and whether these categories are acoustic or articulatory in nature. SLM posits that learners may have difficulty producing L2 phones that are phonetically 'similar' to L1 categories, because they are producing the phones as if they are L1 phones. Equivalence classification suggests that learners will classify L2 phones as if they were L1 phones. If an L2 phone is categorized as an L1 phone, it will be produced as if it were an L1 phone. Similarly, PAM-L2 predicts that L2 phones will be mapped to L1 categories, thus suggesting the learners in the present study may map French /u/ to their L1 English /u/ category. Importantly, PAM-L2 posits that the articulatory gesture is the basic unit of perception. Because of this assumption, it can be hypothesized that learners will use L1 gestures to produce L2 phones. The articulatory results of the present study do not fully support this prediction. Only two of the six learners (EN03 and EN07) have a significant degree of overlap in tongue position of French /u/ and English /u/, and only one of these learners (EN03) has a significant degree of overlap along the entire contour. The other five learners produce French /u/ and English /u/ with different tongue positions, showing that learners are not reusing L1 articulatory gestures in L2 productions. However, 5 out of 6 learners did not produce an acoustic difference between French and English /u/. This result on its own may suggest support for Equivalence Classification, that learners are using L1 categories to produce L2 phones. Learners appear to be using the same acoustic-phonetic category to produce French /u/ as English /u/, but the present results suggest that learners are not using L1 tongue gestures to produce L2 phones.

The findings show that learners tend to use L1 acoustic categories to produce L2 phones that are phonetically similar, and learners are acoustically target-like in their productions of the new French phone /y/. Taken together, these results indicate that learners use acoustic, rather than articulatory, targets in L2 vowel productions. This finding has interesting implications for theories of L2 category formation that suggest

the articulatory gesture is the basic phonetic unit, namely PAM-L2. Learners do not appear to be using L1 gestures to produce L2 phones, and furthermore, the fronted tongue position of /i/ is used by intermediate learners to produce /y/ (as evidenced by Group B learners). However, learners do appear to use L1 acoustic categories to produce L2 phones, and do produce target-like acoustic /y/. This finding suggests that acoustic-phonetic categories are used in L2 learning, and thus theories of L2 category formation need to account for learning of acoustic categories, particularly in discussing vowel categories.

Finally, this study shows that the novel use of articulatory data is important to having the complete picture of L2 production categories. Learners in Group B acoustically merged French /u/ and /y/. However, the articulatory data show that learners did not use the same tongue position for these vowels. Only considering the acoustic data may lead to incorrect assumptions that these learners do not have the contrast between these vowels, when in fact they do not produce them with equivalent articulatory positions. Group B learners are not quite acoustically target-like in creating the contrast between /u/ and /y/, and may be in the intermediate stages of acquiring this contrast, having not yet appropriately back their tongue to for French /u/ to produce the F2 difference between /y/ and /u/. Importantly, the development of vowel contrasts in L2 learners should be studied using both articulatory and acoustic measures to reveal all possible developmental stages of acquisition.

One of the speakers in the present study had slightly different results from learners in Group A and Group B, and will briefly be discussed here. EN06 is the only learner who produces an acoustic difference between French and English /u/. According to tongue position results, EN06 is backing French /u/ significantly more than English /u/. EN06 is not using the acoustic or articulatory L1 category of English /u/ to produce French /u/, suggesting she has not mapped, or Equivalence Classified, French /u/ as L1 English /u/. This speaker is also the only learner to produce a non-target like /y/ acoustically. Unlike the other speakers in Group B, EN06 has a merged French /y/-/u/ category because she produces /y/ with a relatively low F2 value, encroaching on the acoustic space of /u/. EN06 is non-target like in both acoustic and articulatory productions of the new phone /y/. Although SLM suggests that new phones are easier to acquire than similar L2 phones to L1 phones, learners still may have difficulty producing new phones at certain stages of acquisition. This data provides additional insight about the range of individual differences involved in L2 pronunciation.

6 Limitations

There are several limitations in this study that are often associated with processing ultrasound data in general. The first limitation is that data from seven speakers were analyzed. This is common with ultrasound tongue imaging research (Lee-Kim et al., 2013; Lee-Kim, 2014; Mielke et al., 2011; De Decker and Nycz, 2012; Allen et al., 2013), but results and implications would be more generalizable with a larger sample. However, SSANOVA, which is the most appropriate measure to compare tongue positions of vowel contrasts, does not allow for direct comparison across speakers. Thus, a smaller number of speakers does provide detailed data about how individual speakers produce vowel contrasts. Although learners can only be compared descriptively, ultrasound data is necessary in order to understand how learners are using articulatory gestures in speech,

and to test the assumptions of PAM-L2 and other models of category formation. Studies should continue to use ultrasound tongue imaging to provide a broader picture of the articulatory strategies used in L2 production.

Additionally, the number of tokens analyzed for each speaker was small. Participants become fatigued when wearing the headset for more than 45 minutes at a time, making it difficult to collect more tokens of each vowel. However, effort was made to have consistent productions of each vowel by controlling for coarticulatory effects across all vowel categories. The standard deviation for the acoustic measurements is relatively small for each vowel. Additionally, the confidence intervals for the SSANOVA contours is relatively small for each vowel, with the exception of the very tip and root of the tongue. More tokens may have helped with some of the variation at the edges of the contours, but the maximum point of constriction was relatively consistent for speakers.

7 Conclusion

Using ultrasound tongue imaging, video recordings of lip protrusion, and acoustic data, this study finds that learners do not reuse L1 articulatory gestures in L2 vowel productions, but do use L1 acoustic categories in L2 productions. Additionally, the articulatory results do not show support for the hypothesis that learners are more accurate in producing phonetically and phonologically new L2 phones (i.e. French /y/ for English learners) than phonetically new but phonologically similar L2 phones (i.e. French /u/ for English learners). Learners in the present study remained non-target like in their articulatory productions of French /y/, even if their acoustic productions were target-like. The Speech Learning Model and Perceptual Assimilation Model-L2 both argue that L1 and L2 phonetic systems exist in the same phonological space, and thus will interact. However, SLM posits that phonetic categories are acoustic, while PAM-L2 argues that articulatory gestures are the basic unit of speech perception (Flege, 2005; Best and Tyler, 2007). The findings of the present study indicate that vowel representations have an acoustic target for learners in intermediate stages of acquisition. Theories of L2 category formation that exclusively refer to articulatory gestures, without acoustics, as the phonetic unit in L2 learning do not adequately describe the nature of L2 vowel productions.

References

AAA. 2012. Articulate instruments ltd.: Articulate assistant advanced user guide: Version 2.14.

Allen, Blake, Douglas Pulleyblank, and O Ajiboye. 2013. Articulatory mapping of yoruba vowels: an ultrasound study. *Phonology* 30:183–210.

Best, Catherine T. 1991. The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words* 167:233–277.

Best, Catherine T, and Michael D Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege* 1334:1–47.

Boersma, Paul, and David Weenink. 2017. Praat: doing phonetics by computer (version 5.1.13). URL `http://www.praat.org`.

Darcy, Isabelle, Laurent Dekydtspotter, Rex A Sprouse, Justin Glover, Christiane Kaden, Michael McGuire, and John HG Scott. 2012. Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in l1 english–l2 french acquisition. *Second Language Research* 28:5–40.

Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance a. *The Journal of the Acoustical Society of America* 120:407–415.

De Decker, Paul M, and Jennifer R Nycz. 2012. Are tense [æ] s really tense? the mapping between articulation and acoustics. *Lingua* 122:810–821.

Flege, James E. 2005. Origins and development of the speech learning model. *Retrieved December* 13:2005.

Flege, James Emil. 1987. The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics* 15:47–65.

Flege, James Emil, and James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76:708–721.

Havenhill, J., E. Zsiga, O. Tlale Boyer, and S. Petersen. 2017. The labio-coronal fricative of setswana: Its features and articulation. 2017 Annual Meeting on Phonology.

Ingram, John CL, and See-Gyoon Park. 1997. Cross-language vowel perception and production by japanese and korean learners of english. *Journal of phonetics* 25:343–370.

Kamiyama, Takeki, and Jacqueline Vaissière. 2009. Perception and production of french close and close-mid rounded vowels by japanese-speaking learners. *Acquisition et interaction en langue étrangère* 9–41.

Kendall, Tyler, Erik R Thomas, and Maintainer Tyler Kendall. 2013. Package 'vowels'. *R-package version* 1.

Lee-Kim, Sang-Im. 2014. Revisiting mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association* 44:261–282.

Lee-Kim, Sang-Im, Lisa Davidson, and Sangjin Hwang. 2013. Morphological effects on the darkness of english intervocalic /l/. *Laboratory Phonology* 4:475–511.

Levy, Erika S, and Winifred Strange. 2008. Perception of french vowels by american english adults with and without french language experience. *Journal of phonetics* 36:141–157.

Li, Kambhamettu-C., M., and M. Stone. 2005. Automatic contour tracking in ultrasound images.

Major, Roy C. 2008. Transfer in second language phonology: A review. phonology and second language acquisition, ed. by jette g. hansen edwards and mary l. zampini, 6394.

Maye, Jessica C. 2000. The acquisition of speech sound categories on the basis of distributional information .

Mielke, Jeff, Kenneth S Olson, Adam Baker, and Diana Archangeli. 2011. Articulation of the kagayanen interdental approximant: An ultrasound study. *Journal of Phonetics* 39:403–412.

Pillot-Loiseau, Claire, Takeki Kamiyama, and Tanja Kocjančič Antolík. 2015. French/y/-/u/contrast in japanese learners with/without ultrasound feedback: vowels, non-words and words. In *International Congress of Phonetic Sciences (ICPhS) 2015*, volume 1, 1–5. International Phonetic Association: London.

Tyler, Michael D, Catherine T Best, Louis M Goldstein, and Mark Antoniou. 2014. Investigating the role of articulatory organs and perceptual assimilation of native and non-native fricative place contrasts. *Developmental psychobiology* 56:210–227.

Appendices

A French word list

/i/		/y/		/u/		/a/	
titre	‘title	tutrice	‘tutor’	toute	‘all’	tatouage	‘tattoo’
qui	‘who’	cumuler	‘to add’	couper	‘to cut’	camarade	‘friend’
bible	‘bible	but	‘goal’	bout	‘end’	bas	‘low’
niveau	‘new’	numerique	‘digital’	nouveau	‘new’	navette	‘shuttle’
mise	‘appearance’	musée	‘museum’	moustache	‘mustache’	masquer	‘to hide’
vie	‘life’	vu	‘saw’	vous	‘you (pl)’	vache	‘cow’
fiche	‘form’	futur	‘future’	fou	‘crazy’	faux	‘false’
militaire	‘military’	multiplier	‘multiply’	mouler	‘to form’	mal	‘bad’
guise	‘as you like’	culture	‘culture’	goûter	‘to taste’	gâteau	‘cake’
midi	‘noon’	sud	‘south’	soudain	‘suddenly’	madame	‘madame’

Filler Words

neuf	‘nine’	sait	‘knew’	ses	‘their’
boeuf	‘beef’	bête	‘silly’	bébé	‘baby’
oeuf	‘egg’	même	‘same’	mes	‘my’
coeur	‘heart’	guepe	‘wasp’	quai	‘platform’
deuxième	‘second’	dessus	‘top’	déjeuner	‘lunch’
ceux	‘this’	nostalgie	‘nostalgia’	nos	‘our’
peu	‘a little’	bonne	‘good’	beau	‘beautiful’
eux	‘them’	mauve	‘mauve’	mauvais	‘bad’
que	‘that’	école	‘school’	coller	‘to stick’
deux	‘two’	d’ature	‘another’	tomber	‘to fall’

B English word list

/i/	/u/	/a/
teeth	tooth	tot
keep	coo	cop
bee	boo	bop
knee	noon	novice
measly	moose	moss
fee	voodoo	volleyball
feed	food	father
meal	pool	mall
geek	google	goggles
mead	mood	modern

Filler words	
never	November
beverage	bow
member	moment
kept	cope
demonstrate	dome
name	gnaw
babe	pawn
main	operate
okay	collar
tame	dawn