Perceptual learning and convergence are extended phonologically across vowels

Chelsea Sanker

chelsea.sanker@yale.edu

Yale University, New Haven, 06511, Connecticut, USA

Abstract: Exposure to sounds with altered characteristics alters speakers' subsequent production and perception of those sounds. Moreover, shifts can be extended to sounds that were not presented during exposure. However, previous work has not established which shifts are extended to which other sounds. Establishing patterns of how shifts extend could help clarify the underlying changes that are reflected in these shifts. Vowels provide a useful test because they have several contrasts along several dimensions, and some differ in their phonological similarity and phonetic similarity. By exposing listeners to a single shifted vowel quality, it is possible to narrow down possible explanations. Two perception tasks and one production task tested how exposure to shifted F1 or F2 in a single vowel quality influences other vowels that match either in the characteristic being manipulated or do not. The results suggest that shifts in production and perception extend to vowels that share phonological features in the domain of manipulation, even when the sounds do not share acoustic targets. In Experiment 1, raised or lowered F1 in shadowing stimuli with /ɛ/ results in correspondingly shifted production of F1 in both /ɛ/ and /ʌ/. In Experiment 2, shifted F1 in /ɪ/ training items produces corresponding perceptual shifts in the boundary between high and mid vowels. In Experiment 3, shifted F2 in /u/ training items produces corresponding perceptual shifts in the boundary between front and back vowels. Vowels that do not share these phonological features are not impacted.

1.0 Introduction

The phonetic details in speakers' representations of sounds can be shifted based on exposure to those sounds with altered characteristics; these shifts are reflected both in perception (e.g. Kraljic & Samuel, 2005; Maye, Aslin, & Tanenhaus, 2008; Mitterer, 2006; Norris, McQueen, & Cutler,

2003) and production (e.g. Babel, 2010; Delvaux & Soquet, 2007; Nielsen, 2011; Ross, Lilley, Clopper, Pardo, & Levi, 2021). Work on consonants has demonstrated that shifts in one sound can be extended to other sounds with shared characteristics, but there is less work on extension within vowel systems. Vowels provide a useful test in two ways. First, the vowel space includes contrasts along several dimensions, so it is possible to test if shifts are extended to segments with parallel characteristics, cause a chain of shifts in vowels threatened by the first shift, or cause the whole vowel system to shift based on normalization to the speaker's apparent vocal tract. Second, some vowel contrasts differ in their phonological similarity and phonetic similarity, which makes it possible to distinguish between the two as the basis for extension.

Previous work has not established how vowel shifts are extended, though it is clear that exposure to a shifted realization of one sound can impact more than just that one sound. Convergence studies on vowels often use natural speech and include many different vowels during exposure (e.g. Babel, 2010; Ross et al., 2021), so effects of one vowel on later productions of another vowel cannot be distinguished from effects of each vowel on later productions of the same vowel. Perceptual learning studies on vowels usually include several vowels in exposure (Ladefoged & Broadbent, 1957; Maye et al., 2008; Mitterer, 2006; Weatherholtz, 2015) and test a set of novel vowels differing in just one dimension from the training items, usually backness (Maye et al., 2008; Weatherholtz, 2015), or test a single novel vowel (Chládková, Podlipskỳ, & Chionidou, 2017; Sjerps, McQueen, & Mitterer, 2013). With the relatively large number of exposure vowels and small number of novel test vowels, it is difficult to establish what drives the results.

This study uses a shadowing task and two perception tasks to test how exposure to shifted F1 or F2 in a single vowel quality influences production and perception of other vowels that match either in the characteristic being manipulated or in other characteristics, and whether the alignment that matters is having shared phonological features or having shared acoustic or articulatory characteristics. The results suggest that shifts in production and perception extend to vowels that share phonological features in the domain of manipulation, even when the sounds do not share acoustic targets. These results also provide an additional line of evidence for the phonological features structuring vowel systems.

1.1 Extension based on shared characteristics

Work in perceptual learning illustrates extension across related contrasts, suggesting that these shifts can be at the level of the feature rather than the segment. Differences in extension across different contrasts can be informative in narrowing down possible explanations for the mechanism underlying these shifts. In perceptual learning, participants' phonological categories are shifted when they hear words in which an otherwise ambiguous sound is disambiguated by the lexical item it appears in (e.g. Kraljic & Samuel, 2005; Norris et al., 2003). Not all shifts are extended in the same way. A shifted VOT contrast trained in one place of articulation was extended to other places (Kraljic & Samuel, 2006), and a place contrast in fricatives was extended across voicing categories (Schuhmann, 2014). On the other hand, a place contrast trained in voiced stops did not extend to nasals (Reinisch & Mitterer, 2016; Reinisch, Wozny, Mitterer, & Holt, 2014). Extension might depend on both series of sounds using the same set of cues; nasals include place cues not present in oral stops. Such a dependency could suggest that extension is based on shared acoustics rather than or in addition to shared phonological features. Most patterns that are consistent with features could also be predicted by acoustic similarity, because they are closely correlated, particularly for consonants. Some evidence for extension being shaped by phonological similarity rather than acoustic similarity comes from phonological processes in artificial languages. Backness harmony was extended to novel vowels of new height in an artificial language (Finley & Badecker, 2009), but listeners did not learn harmony rules that differed by vowel height (Skoruppa & Peperkamp, 2011). Vowels have the potential to provide a useful test of the changes that underlie shifts in perception and production, because some vowels differ in their phonological similarity and phonetic similarity due to how tense/lax contrasts interact with other dimensions of contrast. However, this has not been directly tested in previous work.

Convergence can also extend across sounds with the same feature, which helps demonstrate the nature of the shifts that are reflected in perception; the changes in perceptual boundaries between sounds accompany a shift in the target pronunciation of those sounds. For example, Nielsen (2011) demonstrates that exposure to lengthened VOT at one place of articulation results in lengthening at other places of articulation, paralleling the extension of VOT shifts in perceptual learning..

Vowels also allow for examination of how different dimensions of similarity influence extension of shifts, which in turn can shed light on the phonological representation of vowel features. Several perceptual learning studies establish that shifts are extended beyond the vowels present in training, though it is sometimes unclear how dependent the results are on the particular set of training vowels presented during exposure. The exposure phase usually presents several vowel qualities, including the vowel or vowels used in the testing phase. In a seminal paper, Ladefoged and Broadbent (1957) demonstrate that a vowel ambiguous in height between /ɪ/ and /ɛ/ was more likely to be perceived as /ɛ/ when following a carrier phrase with a lower F1, with similar shifts in other contrasts. This is interpreted as reflecting an overall shift in the realization of vowel height. Maye et al. (2008) conducted a similar experiment using a longer exposure passage, with an artificial vowel space in which the front vowels were categorically lowered. Listeners learned this artificial "accent", as reflected in responses in a subsequent lexical decision task. There was some evidence that the front vowel lowering was also extended to back vowels in one of their experiments, but not in the other experiment. Weatherholtz (2015) also found that back vowel lowering was extended to front vowels. However, exposure to this novel "accent" with lowered back vowels also increased how often listeners identified words in a way consistent with a raised back vowel system, suggesting that listeners were broadening their vowel categories rather than shifting them. These experiments suggest extension based on the realization of shared phonological features, but it is unclear whether the overall shift in the vowel space is dependent on the presence of several exposure vowels manipulated in corresponding ways.

Some work has used restricted sets of exposure vowels, which make it possible to narrow down which aspects of the exposure phase are necessarily in eliciting shifts that are extended to additional sounds. The results from such studies are mixed. Mitterer (2006) demonstrates that identifications of a Dutch word with ambiguous F2 on a continuum between /e/, /ø/, and /o/ are influenced by the F2 manipulations of a carrier sentence. When the carrier sentence included both front and back vowels, manipulations with lower F2 in the vowels resulted in more identifications of ambiguous test vowels as front and as non-round. When the exposure phase only included front vowels, the manipulation condition affected the boundary in the roundness contrast between the front vowels /e-ø/, but not the boundary in the backness contrast between /ø-o/. This effect would be consistent with the manipulation not influencing the boundary per se, but influencing the range of expected values for /e/, so only decisions between the impacted vowel and neighboring vowels will be impacted. This provides evidence for the extension of backness and roundness across vowel heights, at least when the exposure phase represents most of the vowel space. Sjerps et al. (2013) presented Dutch listeners with vowels ambiguous in height between /ɪ-ɛ/ in the context of a nonce word also containing /a/ and /u/, with either raised F1 or lowered F1. /ɪ/ responses were significantly more likely in the high F1 condition and /ɛ/ responses were more likely in the low F1 condition, suggesting that the shift in vowel height for the /ɪ-ɛ/ boundary did not depend on the shifted context containing vowels with shared acoustic characteristics; it may be sufficient that /u/ and /ɪ/ match in height, though they differ in tenseness. Chládková et al. (2017) provide evidence for vowel height contrasts extended across backness in Greek, with a single vowel in the exposure phase. They compared exposure conditions with raised /e/ versus lowered /i/, and found that subsequent differences in perception between the two exposure conditions were reflected not only in /i-e/ decisions but also more weakly in /u-o/ decisions. However, they did not test whether additional contrasts were impacted.

1.2 Normalization

An alternative phonetic account for extension of formant shifts across vowels is that they reflect normalization for the apparent vocal tract of the voice (Sjerps et al., 2013). Some evidence for normalization comes from manipulations that reduce naturalness and might increase the difficulty of isolating particular formants or identifying their phonological environments. For example, Watkins and Makin (1994) demonstrate that shifted perceptual boundaries can be elicited even when the exposure phase with the manipulated formant uses backwards speech or if speech is altered with a filter to reshape the long-term average spectra based on a high-F1 or low-F1 carrier phrase. They interpret these results as indicating that shifts reflect normalization based on the overall distribution of frequencies, rather than adaptation specifically to F1. In this way, extension could be explained as the result of exposure to acoustic values that suggest the length of the speaker's vocal tract, rather than a shift in the expected realization of particular phonemes or particular features.

The question of whether normalization drives extension of shifts in convergence and perceptual learning is separate from the question of whether normalization plays a role in perception more generally. There is substantial evidence for some sort of normalization underlying listeners' ability to associate very acoustically distinct vowels with the same category (Joos, 1948; Ladefoged & Broadbent, 1957) and accurately identify vowels produced by unfamiliar speakers (Mullennix, Pisoni, & Martin, 1989; Verbrugge, Strange, Shankweiler, & Edman, 1976). Further evidence for normalization comes from the interaction between perceived vowel quality and believed speaker size; lower formants are expected from larger speakers (Barreda, 2020). This association is also reflected in influences of F0 and formants higher than F2 on perceived vowel quality (Barreda, 2020; Fujisaki & Kawashima, 1968). However, the fact that listeners perceptually normalize for differences between speakers does not necessarily indicate that normalization is responsible for the shifts observed in convergence and perceptual learning, particularly when considering that participants accommodate to differences in characteristics such as VOT, which are not caused by vocal tract differences.

If extension across vowels is due to normalization, exposure to any shifted vowel is likely to change all other vowels in the system, regardless of whether or not they share features with the testing vowels. This analysis also predicts that a shift in one formant will shift other formants, because they are all similarly influenced by vocal tract length.

1.3 Contrast preservation

Another possible prediction is that a shift in one sound will alter neighboring sounds in order to preserve contrasts. Speakers might maintain categories by shifting both the exposed category and also the category being threatened by the shift; historical chain shifts are explained as the result of such processes (Labov, 1994, Ch. 9; Gordon, 2002) and there are also synchronic processes that seem to be chain shifts (Moreton & Smolensky, 2002). Thus, it is possible that changing one segment could shift other segments that contrast along the dimension being manipulated; for example, lowering a high vowel would also result in lowering of the corresponding mid vowel. There is some evidence in natural speech that perceptual contrastiveness is a factor shaping dispersal in vowel systems (Becker-Kristal, 2010; Lindblom, 1986). However, this influence is not absolute; other work has found no effect of phonological inventory on how variable vowel production is (Bradlow, 1995; Flege, 1989). Of course, historical mergers also exist (Labov, 1994, Ch. 20). Shifting one segment might be possible without any change to other segments that are being encroached on as a result of the shift. For example, Mitterer (2006) found that shifting the F2 of front unrounded vowels changes the perceptual boundary of the roundness contrast between /e/ and /ø/, but does not influence the boundary between /ø/ and /o/.

There is no clear evidence that shifting one sound produces corresponding shifts in the sounds that it is shifted towards or away from. Some work finds that exposure to stimuli that would reduce the distance between categories simply fails to generate convergence, at least in VOT (Nielsen, 2011) and vowel duration (Podlipskỳ & Simácková, 2015). Given that both of these studies involve duration and specifically failure to shorten a long category, it is worth considering whether some aspect of shortening poses a particular challenge. In favor of this explanation, speakers exhibit more compensatory lengthening based on shortened auditory feedback than compensatory shortening based on lengthened auditory feedback (Karlin, Naber, & Parrell, 2021; Oschkinat & Hoole, 2020). However, speakers do converge in speech rate, and the speaker's baseline speech rate is not a significant predictor of convergence; both slow and fast speakers converge to their interlocutor's speech rate (Cohen Priva, Edelist, & Gleason, 2017). Failure to converge to manipulations that reduce the distance between categories might be the result of ambiguity of the training items; if the stimuli were sometimes miscategorized, the effect on the target category would be weakened.

2.0 Experiment 1

Experiment 1 tests whether exposing listeners to shifted F1 in a single vowel quality influences their subsequent productions of F1 in other vowels, using a shadowing paradigm. As previous work has shown, listeners converge to the acoustic characteristics of speech that they hear from an interlocutor or pre-recorded stimuli. Nielsen (2011) demonstrates that convergence in VOT is extended across stops with different places of articulation, so phonetic convergence may be at the level of features, rather than the level of individual sounds. However, it is unclear how convergence might extend across vowels. Although convergence in formants has been found in previous work (e.g. Babel, 2010; Delvaux & Soquet, 2007; Ross et al., 2021), such work includes multiple vowels during exposure, which makes it impossible to definitively identify how a shift in one vowel might influence production of other vowels. A shadowing paradigm is used here in order to strictly control for the input that listeners receive, exposing them to only a single manipulated vowel quality.

2.1 Stimuli

Stimuli were made from recordings of one female American English speaker reading monosyllabic English words, elicited individually in randomized order with PsychoPy (Peirce, 2007) and recorded in a quiet room with a stand-mounted Blue Yeti microphone in the Audacity software program and digitized at a 44.1 kHz sampling rate with 16-bit quantization.

In the shadowing task, participants heard 15 acoustically manipulated words with the mid front unrounded lax vowel /ɛ/. Manipulations were done in Praat using the Change formants function in the Vocal Toolkit, and were manually checked to ensure that the output was as intended. The only change was in F1, in order to focus on a single feature and avoid potential confounds. Stimuli in both conditions were made from the same recordings; thus, the stimuli in the two conditions differed only in F1.

There were two manipulation conditions. In the lowered F1 condition, F1 of stimulus vowels was lowered on average by 48 Hz, and in the raised F1 condition, F1 was raised on average by 98 Hz.

The size of each manipulation was based on pilot testing of perceived vowel quality, to maintain the perceived vowel quality /ɛ/.[1]

2.2 Experiment design

24 monolingual native speakers of American English participated in the study (mean age 22.0 years); all participants were female. Having only female participants and a female model talker was done in order to avoid potentially confounding the question of normalization, which could produce different predictions for cross-gender pairs than same-gender pairs.

There were three stages of the experiment, all of which were recorded in a quiet room with a stand-mounted Blue Yeti microphone in the Audacity software program and digitized at a 44.1 kHz sampling rate with 16-bit quantization: (1) pretask reading of the full set of words, (2) shadowing of the manipulated words with /ɛ/, (3) post-task reading of the full set of words.

First the participants read a set of monosyllabic English words twice in randomized order. The second set of productions were used for each speaker's baseline, in order to reduce disfluencies that are likely to appear in initial readings; these are the pretask productions used to evaluate each speaker's starting point prior to any shift elicited by the shadowing task. The word list included 60 target words and 60 filler items. The 60 target words included the 15 shadowing items with /ɛ/, 15 test-only items with /ɛ/, and 10 test-only items each of /æ/, /ɪ/, and /ʌ/. (A list of the items is given in the Appendix.) The test-only items were not present in the shadowing task. The test-only items with /ɛ/ were included in order to examine potential effects of novelty, that is, whether the effects of the manipulation condition are different for words that were present among the manipulated stimuli in the shadowing task or not, so that this potential effect could be distinguished from effects of the vowel quality. These vowel qualities were selected as they share different features with the training vowel; /æ/ and /ɪ/ share frontness with /ɛ/, while /ʌ/ shares height. Different competing hypotheses make different predictions about how these vowels will be

[1]A vowel identification task was used to select the stimulus for each word in each condition (raised F1, lowered F1). Four versions of each stimulus were tested, with F1 shifts of approximately 20 Hz, 50 Hz, 75 Hz, and 105 Hz, raised and lowered respectively for each condition. The item with that largest shift that was still identified as having /ɛ/ by at least 14 of the 16 pilot participants was selected to be used in Experiment 1.

impacted by the manipulation of /ɛ/.

In the shadowing task, participants repeated after 15 acoustically manipulated target words (as described above), given in randomized order; each was presented three times. The exposure items in this task only had the vowel /ɛ/. There were two conditions (12 participants in each condition): half of participants heard these words with a raised F1 in the /ɛ/ and half heard a lowered F1.

After the shadowing task, participants read all words again a single time, in randomized order. These are the posttask productions; they were compared to the pretask productions to measure how each speaker's production of F1 was influenced by the shadowing task.

Participants' F1 was measured for each word before and after the shadowing task (pretask and posttask). Reported statistics come from a regression model with change in F1 as the dependent variable, calculated with the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). *p-values* were calculated by the lmerTest package (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2015).

2.3 Hypotheses and predictions

Hypothesis 1: Previous work demonstrates that speakers will converge to vowel formants (e.g. Babel, 2009; Pardo, Urmanche, Wilman, & Wiener, 2017). If listeners are exposed to tokens of a vowel that are shifted in a particular direction, they are more likely to subsequently produce those vowels with the same shift. Thus, exposing listeners to /ɛ/ with a raised F1 is likely to increase their F1 in subsequent productions of /ɛ/. This is likely to be apparent both for words that were present as stimuli in the shadowing task and also for novel words that were not present in shadowing (cf. Nielsen, 2011).

Moreover, shifting the boundaries of one vowel is likely to influence more than just that vowel. Several possible patterns in the vowel space might be predicted.

Hypothesis 2a: Shifts within the vowel space might be contrast preserving. Some work in convergence has suggested that shifts are resisted when the result would encroach on a category boundary

(Nielsen, 2011; Podlipskỳ & Simácková, 2015). If the vowel space changes as a system to preserve contrasts when the realization of one category is altered, shifting /ɛ/ downwards encroaches on /æ/, which will in turn shift downwards, and shifting /ɛ/ upwards encroaches on /ɪ/, which will in turn shift upwards.

Hypothesis 2b: Shifts within the vowel space might be based on featural or acoustic parallels (Chládková et al., 2017; Maye et al., 2008; Mitterer, 2006). Exposing listeners to a raised F1 in /ɛ/ is likely to also increase the F1 in their productions of /ʌ/, which aligns phonologically as a mid vowel and aligns phonetically in having a similar F1.

Hypothesis 2c: Shifts within the vowel space are based on convergence to the speaker's assumed vowel space, as proposed in normalization accounts (Sjerps et al., 2013; Watkins & Makin, 1994). This would predict that exposure to raised F1 in /ɛ/ will also result in raised F1 for all other vowels.

2.4 Results

Figure 1 provides a visualization of the results for each vowel in Experiment 1. Note that the figure is based on the raw data, pooled across participants and across words, so the means and confidence intervals are not identical to what is seen in the regression model.

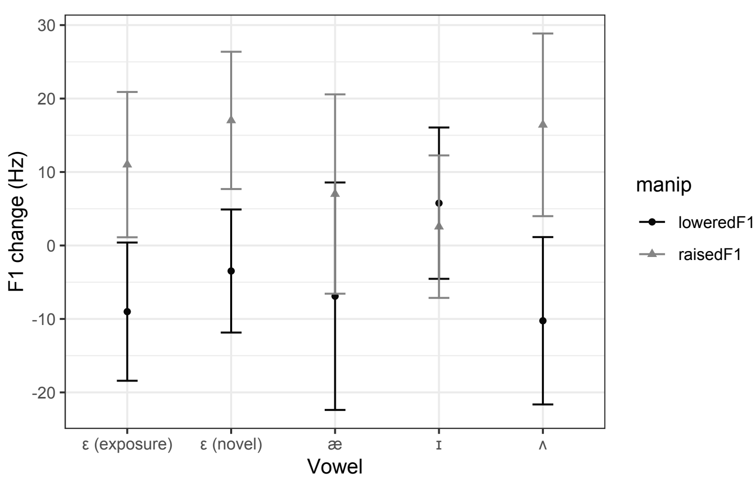

Figure 1: F1 change (posttask-pretask).

Table 1 presents the summary of a mixed effects linear regression model[2] for change in F1. Change in F1 is used rather than raw F1 to reduce noise produced by the between-subjects design; participants' baseline F1 values varied, so it is desirable to factor out that variation when interpreting the post-task productions.

The fixed effects were vowel quality (ɛ, æ, ɪ, ʌ) and the interaction between vowel and manipulation condition in the training phase (raised F1, lowered F1). Note that manipulation was not included as a main effect; the interaction between vowel and manipulation is testing whether there was an effect of manipulation for each vowel contrast.

There was a random intercept for participant and for word, and a random slope for manipulation by word. There was no random slope for participant, because each participant was in a single manipulation condition.

Table 1: Model for F1 change. *Intercept: Manipulation = lowered F1, Vowel = /ɛ/*

	β	*SE*	*t-value*	*p-value*
(Intercept)	-6.24	5.3	-1.18	0.247
Vowel /æ/	-0.664	6.77	-0.098	0.922
Vowel /ɪ/	12.0	6.77	1.77	0.0769
Vowel /ʌ/	-4.01	6.77	-0.592	0.554
Manip RaisedF1 : Vowel /ɛ/	20.3	7.63	2.66	0.0117
Manip RaisedF1 : Vowel /æ/	13.9	10.4	1.34	0.185
Manip RaisedF1 : Vowel /ɪ/	-3.2	10.4	-0.307	0.759
Manip RaisedF1 : Vowel /ʌ/	26.7	10.4	2.57	0.012

There were no significant differences in F1 change by vowel quality.

There was a significant interaction between manipulation condition and vowel /ɛ/; that is, the manipulation condition was a significant predictor of F1 change in /ɛ/ words. Participants raised their F1 in the raised F1 condition and lowered it in the lowered F1 condition. There was also a significant effect of the manipulation for /ʌ/, the other mid vowel used in the experiment. There was no significant effect of manipulation for /æ/ or /ɪ/.

Half of the words were present as stimuli in the shadowing task and half were not. Novelty, whether or not the item had been heard during the shadowing task, was examined as a possible factor in F1

[2] $glmer(F1{\sim}vowel * manip - manip + (1|ParticipantID) + (manip|word), data = FormantConvergence)$

change; however, a model including novelty and the interaction between novelty and manipulation condition did not differ significantly from a model without novelty at all ($\chi^2 = 1.24$, $df = 2$, $p = 0.538$), so it was omitted.

3.0 Experiment 2

Does exposing listeners to one vowel with shifted acoustic characteristics influence category boundaries for other vowels? Experiment 2 uses a perceptual learning task to investigate how exposure to F1 manipulations in the high front unrounded lax vowel /ɪ/ influence subsequent perception of the vowel space, testing the boundaries of several vowel contrasts that involve F1.

3.1 Stimuli

Stimuli were made from recordings of one female American English speakers reading monosyllabic English words, elicited individually in randomized order with PsychoPy (Peirce, 2007) and recorded in a quiet room with a stand-mounted Blue Yeti microphone in the Audacity software program and digitized at a 44.1 kHz sampling rate with 16-bit quantization.

The training stimuli were 36 monosyllabic English words with /ɪ/. There were two conditions for the training stimuli, differing only in the F1 manipulation. Manipulations for these training stimuli and for the testing stimuli described below were done in Praat using the Change formants function in the Vocal Toolkit, and were manually checked to ensure that the output was as intended. The only change was in F1, in order to focus on a single feature and avoid potential confounds. Stimuli in both conditions were made from the same recordings; thus, the stimuli in the two conditions differed only in F1.

In the raised F1 condition, listeners heard these words with a raised F1 (mean 640 Hz), and in the lowered F1 condition, listeners heard them with a lowered F1 (mean 425 Hz). The manipulations were based on 2/3 of the distance within the unmanipulated recordings between F1 of /ɪ/ and /ɛ/ for the raised F1 and between F1 of /ɪ/ and /ɪ/ for the lowered F1, calculated in Bark. The speaker's naturally produced F1 range for /ɪ/ was 445-580 Hz. The goal of these manipulations was to create items that were slightly past the expected F1 for /ɪ/ without being entirely in the F1 range of a neighboring vowel.

The testing stimuli were continua of ambiguous vowels based on 28 pairs of monosyllabic English words, differing only in the vowel. There were seven vowel contrasts, all differing in F1, either with height contrasts or tenseness contrasts: /i-ɪ/, /i-ei/, /ɪ-ei/, /ɪ-ɛ/, /ei-ɛ/, /ɛ-æ/, and /ʊ-ɔ/.[3] Each contrast was reflected in four pairs of words. For each contrast, all words had four F1 manipulations: F1 matching the speaker's mean F1 for each of the two vowel qualities, and two intermediate values, equally spaced in the Bark scale. Both vowels for each contrast were used to create stimuli, producing two continua (eight items): for example, the *pit, pet* items included four manipulations made from naturally produced *pit* and four manipulations made from naturally produced *pet*, with the same F1 values in both continua. No other characteristics were altered, so the two continua for each pair often differed substantially in F2.

A list of the items is given in the Appendix.

3.2 Experiment design

Participants were 128 monolingual native speakers of American English (85 male, 42 female, 1 nonbinary; mean age 37.3). 16 participants were excluded and replaced based on having accuracy below 75% for identifications of training items; the training items differed in onsets or codas, which had not been manipulated and thus were generally unambiguous.

The study was run online, with participants recruited and paid through the Amazon Mechanical Turk system and the experiment presented through Qualtrics.

Participants were instructed that they would hear English words and identify each one as matching one of two associated response options. The items were presented as a list; listeners clicked on each item to hear the stimulus. Responses were given by clicking on one of the written words. Within

[3] All of these /ɔ/ words have /ɔ/ for speakers without the /ɔ-ɑ/ merger. The speaker who produced the recordings speaks a dialect in which both /ɔ/ and /ɑ/ are merged to /ɔ/.

a block, the order of items was randomized. The order of the two response options was balanced across participants.

There were two blocks. First, listeners completed a training block in which they made consonant decisions, deciding between two response options with the same vowel quality; this was the exposure phase to shift the expected realization of the vowel /ɪ/. Across participants, there were two different F1 manipulation conditions for the training phase. Second, listeners completed a testing block in which they made vowel decisions; this tested how their perceptual categories for the vowels had shifted.

In the training block, listeners heard a set of 36 items (18 pairs), all monosyllabic English words with /ɪ/, and were presented with response options that differed either in the onset (e.g. *ship, chip*) or the coda (e.g. *dig, did*). Listeners heard both items of each pair. Half of participants heard these words with a raised F1 (mean 640 Hz), and the other half heard them with a lowered F1 (mean 425 Hz). These items served to establish the shifted /ɪ/ category by providing response options which both had this vowel.

In the testing block, listeners heard a set of 28 pairs of monosyllabic words along continua of ambiguous vowels, and were presented with response options that differed in the vowel. The endpoints of the continua were seven vowel contrasts, all differing in height or tenseness, which is reflected in F1 differences: /i-ɪ/, /i-ei/, /ɪ-ei/, /ɪ-ɛ/, /ei-ɛ/, /ɛ-æ/, and /ʊ-ɔ/. For each contrast, there were four F1 manipulations, ranging from the speaker's mean F1 of one vowel quality to the speaker's mean F1 for the other vowel quality, with the intermediate values equally spaced in the Bark scale. Each participant heard four stimuli for each pair: the first and third F1 step from one base recording and the second and fourth F1 step from the other base recording. Both vowel endpoints for each contrast were used to create stimuli, so there were 224 items; to limit the length of the experiment, each participant heard only half of them. Exposure was balanced across participants, so each stimulus was identified the same number of times. Identifications of items in these vowel continua served to test the effect of the manipulations that listeners were exposed to in the training block.

Statistical results are from a logistic mixed effects model, calculated with the lme4 package in R (Bates et al., 2015). *p-values* were calculated by the lmerTest package (Kuznetsova et al., 2015).

3.3 Hypotheses and predictions

Hypothesis 1: Consistent with effects in production, previous work has demonstrated that training in shifted F1 alters listeners' perceptual boundaries for those vowels (Ladefoged & Broadbent, 1957; Maye et al., 2008). If listeners are exposed to tokens of a vowel that are shifted in a particular direction, they are more likely to subsequently perceive tokens in that formerly ambiguous space as belonging to that category. Thus, exposing listeners to /ɪ/ with a raised F1 is likely to increase their identifications of ambiguous items as /ɪ/ rather than /ɛ/, and exposing listeners to /ɪ/ with a lowered F1 is likely to increase their identifications of ambiguous items as /ɛ/ rather than /ɪ/.

As observed with convergence in Experiment 1, shifting the realization of one vowel is likely to influence more than just that vowel, and thus the perceptual boundaries are likely to shift for multiple vowels. There are several competing hypotheses for how such shifts behave and thus what response patterns will be observed.

Hypothesis 2a: Shifts within the vowel space might be contrast preserving. Although some work has suggested that shifts are resisted when the result would encroach on a category boundary (Nielsen, 2011; Podlipskỳ & Simácková, 2015), Mitterer (2006) found that training listeners in a shifted realization of one series in a three-way contrast only impacted the contrast between that series and the neighboring one. If the vowel space changes as a system to preserve contrasts when the realization of one category is altered, shifting /ɪ/ downwards encroaches on /ɛ/, which will in turn shift downwards, and thus exposing listeners to a raised F1 is likely to not only increase their identifications of ambiguous items as /ɪ/ rather than /ɛ/, but also increase their identifications of ambiguous items as /ɛ/ rather than /æ/. Contrasts not threatened by the original shift should not be impacted.

Hypothesis 2b: Shifts within the vowel space might be based on featural parallels, as is suggested by the production data in Experiment 1. There is some evidence that perceptual training in a height difference is extended from front vowels to back vowels (Chládková et al., 2017; Maye et al., 2008). Under this analysis, exposure shifts the representation for a phonological feature rather than changing the representation for a specific vowel. Thus, exposing listeners to a raised F1 in /ɪ/ is likely to not only increase their identifications of ambiguous items as /ɪ/ rather than /ɛ/, but also increase their identifications of ambiguous items as /ʊ/ rather than /ɔ/ and /i/ rather than /ei/. Other contrasts should not be impacted, except as a side effect of shifts in F1 of the high vowels.

Hypothesis 2c: Shifts within the vowel space might be based on shared acoustic targets in addition to or instead of shared phonological features. The results from Maye et al. (2008) and Chládková et al. (2017) do not distinguish between this account and a phonological account. Under this analysis, exposure shifts F1 for all vowels that have an approximately shared acoustic target, rather than just changing the target for a specific vowel. Thus, exposing listeners to a raised F1 is likely to not only increase their identifications of ambiguous items as /ɪ/ rather than /ɛ/, but also increase their identifications of ambiguous items ambiguous items as /ʊ/ rather than /ɔ/, possibly with a similar shift in /ei/ rather than /ɛ/.

Hypothesis 2d: Shifts within the vowel space might reflect normalization, as is proposed by Watkins and Makin (1994) and Sjerps et al. (2013). If this is the case, then exposing listeners to a raised F1 in /ɪ/ will cause the whole vowel space to expand; for all contrasts, the raised F1 condition will increase responses of the vowel with the lower F1.

3.4 Results

Figure 2 provides a visualization of the results for each vowel contrast in Experiment 2. Note that the figure is based on the raw data, pooled across participants and across words, and does not include factors other than contrast and manipulation, so the means and confidence intervals are not identical to what is seen in the regression model.

Table 2 presents the summary of a mixed effects logistic regression model[4] for responses of the

[4] $glmer(responsetype \sim F1step + origbasetype + contrast * manip - manip + (1|ParticipantID) +$

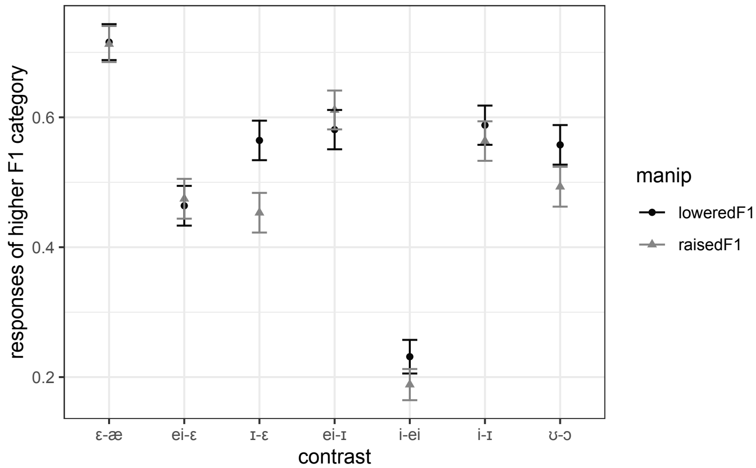

Figure 2: Proportion of responses of the vowel with the higher F1, by F1 manipulation and vowel contrast. In most cases, this is the phonologically lower vowel. When the only contrast is tenseness, it is the lax vowel. For /ɪ-ei/, this is /ɪ/, even though it is phonologically higher.

higher F1 vowel quality (e.g. /ɛ/ rather than /ɪ/). The responses were categorized in this way in order to allow comparison across vowels.

The fixed effects were the vowel contrast (/i-ɪ/, /i-ei/, /ɪ-ei/, /ɪ-ɛ/, /ei-ɛ/, /ɛ-æ/, /ʊ-ɔ/); F1 manipulation step (1-4, with increasing F1); quality of the vowel in the base recording (higher F1 category, lower F1 category); and an interaction between vowel contrast and the manipulation that listeners were exposed to during training (raised F1, lowered F1). Note that manipulation was not included as a main effect; the interaction between vowel contrast and manipulation is testing whether there was an effect of manipulation for each vowel contrast.

There was a random intercept for participant and for word pair, and a random slope for manipulation by word pair. Recall that stimuli were all manipulations along a continuum, where each endpoint vowel quality made a real English word (e.g. *set, sat*); "word pair" indicates which specific continuum the items came from. There was no random slope for participant, because each participant was in a single manipulation condition.

The reference level for the base vowel (the vowel in the recording that the manipulated stimulus was made from) was always the vowel with the higher F1. For most contrasts, this is also the phonologically lower vowel, though some vowel pairs phonologically match in height, and for /ɪ-

$(manip|wordpair), data = VowelSystemsF1[which(VowelSystemsF1\$Task == \text{“}testing\text{”}),], family = binomial, control = glmerControl(optimizer = \text{“}bobyqa\text{”}))$

ei/, /ɪ/ has the higher F1 but is phonologically higher.

Table 2: Model for responses for the proportion of responses of the higher F1 vowel category. *Intercept: Contrast = /ɛ-æ/, Manip = lowered F1, BaseVowel = higher F1 category*

	β	*SE*	*z-value*	*p-value*
(Intercept)	1.31	0.268	4.89	< 0.001
F1 Step	0.695	0.0217	32.0	< 0.001
BaseVowel Lower F1 Category	-3.04	0.0525	-57.9	< 0.001
Contrast /ei-ɛ/	-1.77	0.369	-4.81	< 0.001
Contrast /ɪ-ɛ/	-1.09	0.368	-3.0	0.00302
Contrast /ɪ-ei/	-0.979	0.369	-2.66	0.00788
Contrast /i-ei/	-3.45	0.372	-9.28	< 0.001
Contrast /i-ɪ/	-0.936	0.369	-2.54	0.0111
Contrast /ʊ-ɔ/	-1.14	0.368	-3.1	0.00195
Contrast /ɛ-æ/ : Manip RaisedF1	-0.0141	0.132	-0.106	0.915
Contrast /ei-ɛ/ : Manip RaisedF1	0.073	0.124	0.587	0.557
Contrast /ɪ-ɛ/ : Manip RaisedF1	**-0.751**	**0.125**	**-6.01**	**< 0.001**
Contrast /ɪ-ei/ : Manip RaisedF1	0.209	0.125	1.67	0.0958
Contrast /i-ei/ : Manip RaisedF1	**-0.381**	**0.139**	**-2.74**	**0.00622**
Contrast /i-ɪ/ : Manip RaisedF1	-0.165	0.125	-1.32	0.185
Contrast /ʊ-ɔ/ : Manip RaisedF1	**-0.433**	**0.124**	**-3.48**	**< 0.001**

The F1 step along the continuum was also a significant predictor of responses. Stimuli with a higher F1 were more likely as being the category with a higher F1, for example /æ/ responses for /ɛ-æ/ decisions were more frequent when the F1 for that item was higher.

The identity of the vowel in the base recording used for the manipulations was also a significant predictor. Stimuli were more likely to be identified as having the vowel of the base recording. For example, continua of vowels that were originally /æ/ were more likely to be identified as having /æ/ than continua of vowels that were originally /ɛ/.

The contrast was a significant predictor of responses. Some contrasts elicited more responses of the higher F1 vowel category, while others elicited more responses of the lower F1 category.

The interaction between manipulation condition and contrast type was the factor of interest. There were significant interactions between contrast and manipulation condition for /ɪ-ɛ/, /i-ei/, and /ʊ-ɔ/. Exposing listeners to raised F1 in /ɪ/ in the training phase of the experiments results in fewer /ɛ/ identifications for decisions along the /ɪ-ɛ/ continuum, as compared to the higher number of /ɛ/ responses in condition of exposure to lowered F1 /ɪ/ items. Notably, there was also a significant effect for the other two high∼low contrasts present in the experiment. Even though participants had not been exposed to /ʊ/ or /ɔ/ in the training phase, the effects of the manipulation paralleled the effects for /ɪ-ɛ/ decisions: Listeners in the raised F1 condition were less likely select /ɔ/ as the response in /ʊ-ɔ/ decisions. A similar effect is also seen for /i-ei/, even though these vowels differ in tenseness from the training vowel /ɪ/ and thus do not phonetically align with the F1 values in /ɪ-ɛ/. There was no significant interaction between contrast /ɛ-æ/ and manipulation condition. There was similarly no significant interaction for /ei-ɛ/, /ɪ-ei/, or /i-ɪ/.

4.0 Experiment 3

Experiment 2 demonstrated that exposure to shifted F1 in one vowel quality produces parallel perceptual shifts in the category boundaries for other vowels with the same features. Experiment 3 tests whether a similar effect can be found based on exposure to shifted F2.

4.1 Stimuli

Stimuli were made from recordings of one female American English speakers reading monosyllabic English words, elicited individually in randomized order with PsychoPy (Peirce, 2007) and recorded in a quiet room with a stand-mounted Blue Yeti microphone in the Audacity software program and digitized at a 44.1 kHz sampling rate with 16-bit quantization.

The training stimuli were 36 monosyllabic English words with the high back rounded tense vowel /u/. There were two conditions for the training stimuli, differing only in the F2 manipulation. Manipulations for these training stimuli and for the testing stimuli described below were done in Praat using the Change formants function in the Vocal Toolkit, and were manually checked to ensure that the output was as intended. The only change was in F2, in order to focus on a single feature and avoid potential confounds. Stimuli in both conditions were made from the same recordings; thus, the stimuli in the two conditions differed only in F2.

In the raised F2 condition, listeners heard all these words with a raised F2 (mean 2050 Hz), and in the lowered F2 condition, listeners heard them with a lowered F2 (mean 1310 Hz). For many speakers of American English, /u/ can be substantially fronted. For this speaker, F2 in /u/ across words ranged from 1160 Hz to 1877 Hz. The goal of these manipulations was to create items that were slightly past the expected F2 for /u/ without being entirely in the F2 range of a neighboring vowel.

The testing stimuli were continua of ambiguous vowels based on 24 pairs of monosyllabic English words, differing only in the vowel. There were six vowel contrasts. There were five contrasts differing in frontness, with most also differing in roundness: /u-i/, /ʊ-ɪ/, /ou-ei/, /ɔ-ɛ/, /ʌ-ɛ/. /u-ou/ was also included, as an example of a contrast which should not be impacted by shifting F2 boundaries.

Each contrast was reflected in four pairs of words. For each contrast, there were four F2 manipulations: F2 matching the speaker's mean F2 for each of the two vowel qualities, and two intermediate values, equally spaced in the Bark scale. Both vowels for each contrast were used to create stimuli, producing two continua (eight items): For example, the *beast, boost* items included four manipulations made from naturally produced *beast* and four manipulations made from naturally produced *boost*, with the same F2 values. No other characteristics were altered.

A list of the items is given in the Appendix.

4.2 Experiment design

Participants were 128 monolingual native speakers of American English (86 male, 42 female; mean age 37.1); participation was restricted to individuals who had not participated in Experiment 2. 31 participants were excluded and replaced based on having accuracy below 75% for identifications of training items; the training items differed in onsets or codas, which had not been manipulated and thus were generally unambiguous.

The study was run online, with participants recruited and paid through the Amazon Mechanical Turk system and the experiment presented through Qualtrics.

The procedure was the same as in Experiment 2: Participants heard individual English words and identified each one as matching one of two written options. As before, there were two blocks. First, listeners completed a training block in which they made consonant decisions, deciding between two response options with the same vowel quality; this was the exposure phase to shift the expected realization of the vowel /u/. Across participants, there were two different F2 manipulation conditions for the training phase. Second, listeners completed a testing block in which they made vowel decisions; this tested how their perceptual categories for the vowels had shifted.

In the training block, listeners heard a set of 36 items (18 pairs), all monosyllabic English words with /u/, and were presented with response options that differed either in the onset (e.g. *goose, juice*) or the coda (e.g. *scoot, scoop*). Listeners heard both items of each pair. Half of participants heard these words with a raised F2 (mean 2050 Hz Hz), and the other half heard them with a lowered F2 (mean 1310 Hz). These items served to establish the shifted /u/ category by providing response options which both had this vowel.

In the testing block, listeners heard a set of 24 pairs of monosyllabic words along continua of ambiguous vowels, and were presented with response options that differed in the vowel. The endpoints of the continua were six vowel contrasts, most differing in frontness and thus F2: /u-i/, /ʊ-ɪ/, /ou-ei/, /ɔ-ɛ/, /ʌ-ɛ/, and /u-ou/, each represented by four pairs of words. For each contrast, there were four F2 manipulations, ranging from the speaker's mean F2 of one vowel quality to the speaker's mean F2 for the other vowel quality, with the intermediate values equally spaced in the Bark scale. Each participant heard four stimuli for each pair: the first and third F2 step from one base recording and the second and fourth F2 step from the other base recording. Both vowel endpoints for each contrast were used to create stimuli, so there were 224 items; to limit the length of the experiment, each participant heard only half of them. Exposure was balanced across participants, so each stimulus was identified the same number of times. Identifications of items in these vowel continua served to test the effect of the manipulations that listeners were exposed to in the training block.

Statistical results are from a logistic mixed effects model, calculated with the lme4 package in R

(Bates et al., 2015). *p-values* were calculated by the lmerTest package (Kuznetsova et al., 2015).

4.3 Hypotheses and predictions

Hypothesis 1: As in Experiment 2, it is expected that exposing listeners to manipulated F2 will shift the category boundary of the exposure vowel: Exposing listeners to /u/ with a raised F2 is likely to increase their identifications of ambiguous items as /u/ rather than /i/, and exposing listeners to /u/ with a lowered F2 is likely to increase their identifications of ambiguous items as /i/ rather than /u/.

As demonstrated in Experiment 2, shifting the boundaries of one vowel is likely to influence more than just that vowel. In addition to showing that similar effects are found with F2, the goal of this experiment is to further examine whether parallel shifts are likely to be driven by shared phonological features or shared phonetic targets.

Hypothesis 2a: Shifts within the vowel space might be contrast preserving. This would only predict additional shifts if there are also shifts based on featural parallels, because English high tense vowels just have a two-way contrast. While there are few three-way F2 contrasts in English, there is one such contrast included in this experiment, among the mid vowels. If shifting the boundary of one contrast puts pressure on another contrast boundary along the same dimension, then exposure to altered F2 in back rounded vowels in the current experiment could influence decisions about any F2 contrast, not just F2 contrasts that involve back rounded vowels: Raised F2 will increase identifications of ambiguous items as /ɔ/ rather than /ɛ/ and also /ʌ/ rather than /ɛ/. If there are no such chain effects, then both contrasts might exhibit no effects, because only the /ɔ-ʌ/ boundary is affected. However, the lack of contrastive roundness among other back vowels in English might alter how these contrasts behave.

Hypothesis 2b: Shifts within the vowel space might be based on featural parallels: Rather than changing the representation for a specific vowel, training shifts the representation for a specific feature. Thus, exposing listeners to a raised F2 is likely to not only increase their identifications of ambiguous items as /u/ rather than /i/, but also increase their identifications of ambiguous items as /ʊ/ rather than /ɪ/, /ou/ rather than /ei/, and /ɔ/ or /ʌ/ rather than /ɛ/. Other contrasts should not be impacted, except as a side effect of shifts in F2 of back round vowels.

Hypothesis 2c: Shifts within the vowel space might be based on shared acoustic targets, in addition to or instead of shared phonological features, shifting F2 for all vowels which have an approximately shared acoustic target. In this case, the manipulation might have more of an effect on back vowels that are also fronted; many speakers have fronted /u/ and /o/ but preserve the backness of other back vowels (Labov, Ash, & Boberg, 2008, Ch. 10). An acoustic relationship might also drive an effect for /ʌ/, which has an F2 close to the F2 of partially-fronted /u/. If the grouping is based on similarity of tongue position, high or high-mid vowels may be more likely to exhibit an effect of /u/ manipulations than low vowels.

Hypothesis 2d: Shifts within the vowel space might reflect normalization, as is proposed by Watkins and Makin (1994) and Sjerps et al. (2013). If this is the case, then exposing listeners to a raised F2 in /u/ will cause the whole vowel space to expand; for all contrasts, the raised F2 condition will increase responses of the vowel with the lower F2. In addition, expansion of the vowel space will increase expected F1, so the /u-ou/ contrast will produce a larger number of /u/ responses.

4.4 Results

Figure 3 provides a visualization of the results for each vowel contrast in Experiment 3. Note that the figure is based on the raw data, pooled across participants and across words, and does not include factors other than contrast and manipulation, so the means and confidence intervals are not identical to what is seen in the regression model.

Table 3 presents the summary of a mixed effects logistic regression model[5] for responses of the higher F2 vowel quality (e.g. /ɪ/ rather than /ʊ/). The responses were categorized in this way in order to allow comparison across vowels.

The fixed effects were the vowel contrast (/u-i/, /ʊ-ɪ/, /ou-ei/, /ɔ-ɛ/, /ʌ-ɛ/, and/u-ou/); F2

[5] $glmer(responsetype \sim F2step + origbasetype + contrast * manip - manip + (1|ParticipantID) + (manip|wordpair), data = VowelSystemsF2[which(VowelSystemsF2\$Task == \text{“}testing\text{”}),], family = binomial, control = glmerControl(optimizer = \text{“}bobyqa\text{”}))$

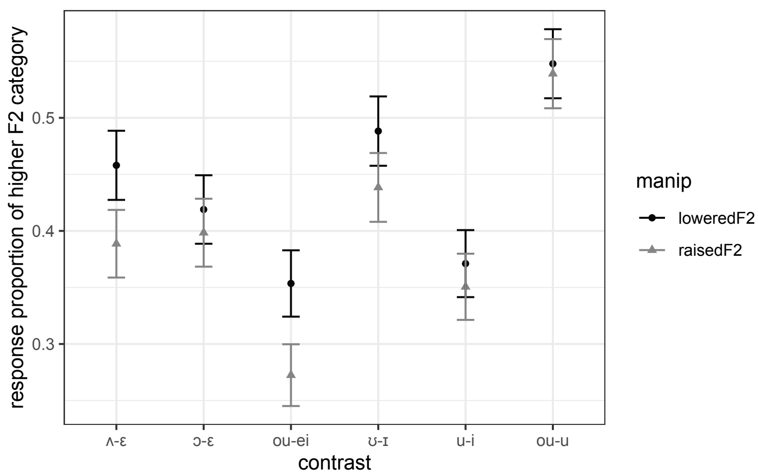

Figure 3: Proportion of responses of the vowel with the higher F2, by F2 manipulation and vowel contrast. In most cases, this is a back vowel, as compared to a corresponding front vowel.

manipulation step (1-4, with increasing F2); quality of the vowel in the base recording (higher F2 category, lower F2 category); and an interaction between vowel contrast and the manipulation that listeners were exposed to during training (raised F2, lowered F2). Note that manipulation was not included as a main effect; the interaction between vowel contrast and manipulation is testing whether there was an effect of manipulation for each vowel contrast.

There was a random intercept for participant and for word pair, and a random slope for manipulation by word pair. Recall that stimuli were all manipulations along a continuum, where each endpoint vowel quality made a real English word (e.g. *foot, fit*); "word pair" indicates which specific continuum the items came from. There was no random slope for participant, because each participant was in a single manipulation condition.

The reference level for the base vowel (the vowel in the recording that the manipulated stimulus was made from) was always the vowel with the lower F2. In most cases, this difference also aligned with differences in backness.

The F2 step along the continuum was also a significant predictor of responses. Stimuli with a higher F2 were more likely as being the category with a higher F2, for example /i/ responses for /u-i/ decisions were more frequent when the F2 for that item was higher.

The identity of the vowel in the base recording used for the manipulations was also a significant predictor. Stimuli were more likely to be identified as having the vowel of the base recording. For

Table 3: Model for responses for the proportion of responses of the higher F2 vowel category. *Intercept: Contrast = /u-ou/, Manip = lowered F2, BaseVowel = higher F2 category*

	β	*SE*	*z-value*	*p-value*
(Intercept)	-1.49	0.269	-5.53	< 0.001
F2 Step	1.13	0.0245	45.9	< 0.001
BaseVowel Lower F2 Category	-2.07	0.0513	-40.4	< 0.001
Contrast /u-i/	-1.12	0.369	-3.02	0.00253
Contrast /ʌ-ɛ/	-0.563	0.369	-1.53	0.127
Contrast /ɔ-ɛ/	-0.797	0.369	-2.16	0.0307
Contrast /ou-ei/	-1.22	0.369	-3.3	< 0.001
Contrast /ʊ-ɪ/	-0.368	0.369	-0.999	0.318
Contrast /u-ou/ : Manip RaisedF2	-0.0539	0.149	-0.363	0.717
Contrast /u-i/ : Manip RaisedF2	-0.146	0.152	-0.963	0.336
Contrast /ʌ-ɛ/ : Manip RaisedF2	**-0.47**	**0.152**	**-3.09**	**0.00198**
Contrast /ɔ-ɛ/ : Manip RaisedF2	-0.13	0.149	-0.872	0.383
Contrast /ou-ei/ : Manip RaisedF2	**-0.567**	**0.154**	**-3.69**	**0.000228**
Contrast /ʊ-ɪ/ : Manip RaisedF2	**-0.321**	**0.15**	**-2.14**	**0.0326**

example, continua of vowels that were originally /u/ were more likely to be identified as having /u/ than continua of vowels that were originally /i/.

The contrast was a significant predictor of responses. Some contrasts elicited more responses of the higher F2 vowel category, while others elicited more responses of the lower F2 category.

The interaction between manipulation condition and contrast type was the factor of interest, which tests whether there was an effect of manipulation for each vowel contrast. There was no significant interaction between manipulation condition and contrast /u/-/ou/, consistent with F2 not being a major difference between these two vowels. More surprisingly, there was no significant interaction between contrast /u-i/ and manipulation condition. However, as will be seen subsequently, other vowels did exhibit significant effects of manipulation condition, which makes it unlikely that participants' /u/ category was truly unaffected. The lack of significant effect might be due to substantial cross-speaker variation in where this boundary lies, perhaps in combination with the wide spacing between steps on this continuum. There was also no significant interaction for /ɔ-ɛ/ decisions. There might be an influence of speakers' vowel inventories; vowels which have an F2 intermediate between that of /ɔ/ and /ɛ/ often sound more like /ʌ/ than either of the two response options /ɔ/ and /ɛ/.

The results might also be influenced by what vowels exist in each participant's dialect. If the analysis is restricted to participants from regions that lack the /ɔ-ɑ/ merger (primarily excluding western states), the effect of the manipulation on identification of items from the /ɔ-ɛ/ is greater, though it still does not reach significance (β = -0.345, SE = 0.20, z = -1.72, p = 0.0849). This model is presented in the Appendix. It is perhaps also noteworthy that the effect of the manipulation on the /u-i/ continuum becomes significant in this model (β = -0.493, SE = 0.206, z = -2.39, p = 0.0169); this may indicate that /u/ fronting in the western United States influenced how listeners responded to the /u-i/ F2 continuum, or that regional variation in the realization of /u/ creates enough variation across participants that the effect of the manipulation was not apparent in the full dataset.

On the other hand, the model containing all participants found significant interactions between contrast and manipulation condition for /ʌ-ɛ/, /ou-ei/, and /ʊ-ɪ/. Even though participants had not been exposed to /ʌ/ or /ɛ/ in the training phase, listeners in the raised F2 condition were less likely select /ɛ/ as the response in /ʌ-ɛ/ decisions. A parallel effect is also seen for /ou-ei/ and /ʊ-ɪ/.

5.0 Discussion

Consistent with previous work (Ladefoged & Broadbent, 1957; Maye et al., 2008; Mitterer, 2006), these experiments demonstrate that exposure to tokens of a vowel in which a characteristic is shifted will shift both production and perception of that vowel. Experiment 1 demonstrates a shift of /ɛ/ productions based on exposure to /ɛ/ with shifted F1: Exposure to /ɛ/ with a raised F1 increases F1 in subsequent productions of /ɛ/, while exposure to /ɛ/ with a lowered F1 decreases F1. Experiment 2 similarly shows shift of the category boundary for /ɪ-ɛ/ contrasts based on exposure to /ɪ/ with shifted F1: Participants exposed to /ɪ/ with a raised F1 made fewer /ɛ/ identifications for decisions along the /ɪ-ɛ/ continuum relative to identifications made by participants exposed to /ɪ/ with a lowered F1. Experiment 3 does not as clearly show whether exposure to shifted F2 alters category boundaries for the vowel heard in the exposure phase, but results for vowels that were not part of exposure demonstrate that the manipulation condition did influence listeners' categories, so the lack of significant effects for the /i-u/ continuum are likely a result of limited ambiguity of the particular testing items.

There was no effect of whether the word had been presented during the exposure phase or not. This result is consistent with previous work demonstrating extension across words with the same sound, both in convergence (Nielsen, 2011) and perceptual training (Ladefoged & Broadbent, 1957; Maye et al., 2008; Mitterer, 2006). In Experiment 1, F1 change was the same for novel words with /ɛ/ as for the /ɛ/ words that were present among the shadowing stimuli.

5.1 Shared Features

The results also indicate that shifts are extended beyond the vowel that is present during exposure. Moreover, the experiments suggest that shifts are extended based on shared phonological features in the domain of manipulation. While shared phonological features often align with shared acoustic characteristics, some of them do not, and when phonological and phonetic alignment disagrees within these experiments, the results are better captured by shared phonological features. These results provide an additional type of evidence for phonologically active classes. Even though the sets of vowels patterning together in each change do not undergo or trigger a shared categorical process, as is often an identifying criterion of a phonologically active class (cf. Mielke, 2008), they still behave together for these non-categorical processes that shift the phonetic realization.

In Experiment 1, convergence to the F1 manipulation in /ɛ/ was extended to the other mid vowel included in the test items, /ʌ/, suggesting that the shift in target F1 is at the featural level, rather than specific to a particular vowel quality. The effect was even a similar size for both vowel qualities. The manipulations did not produce corresponding shifts in other vowel heights. While it is desirable to replicate these results with a larger sample size, it is noteworthy that the manipulation condition had a significant effect for /ɛ/ and /ʌ/ despite the small sample size.

Phonological representations can include detailed phonetic implementation rules, based on differences in the language-specific realizations of the same sound (e.g. Cho & Ladefoged, 1999; Cohn, 1993; Keating, 1985; Lieberman, 1970). Work in convergence (e.g. Nielsen, 2011) and perceptual learning (e.g. Kraljic & Samuel, 2006) demonstrates that these phonetic details can be shifted. I propose that these patterns, along with the results of the present study, can be explained by extending Feature Theory with detailed acoustic targets for each feature. Having an acoustic target shared by mid vowels might suggest that mid vowels are best characterized with a [mid] feature, rather than just an absence of [high] or [low] features. This feature [mid] has a bundle of target phonetic details, crucially including a specification for F1. The manipulation of F1 in stimuli presented to participants results in a shift in that F1 specification, which is reflected both in production and in perceptual expectations. This interacts with other factors that also influence F1, both characteristics of the vowel itself like tenseness, and characteristics of the neighboring sounds, like voicing. These might be specified as modifiers that increase or decrease the baseline F1 specified by vowel height.

Experiment 2 demonstrates the perceptual effects of exposure to stimuli with shifted F1 targets, consistent with the trend observed by Maye et al. (2008), Weatherholtz (2015), and Chládková et al. (2017) for a height difference to extend from front vowels to back vowels. Experiment 2 demonstrates that exposure to instances of /ɪ/ with a manipulated F1 produces corresponding changes in other high versus mid contrasts, even when the phonetic realization of that contrast is dissimilar. Exposure to /ɪ/ with a raised F1 results in significantly fewer /ɔ/ identifications for decisions along the /ʊ-ɔ/ continuum, as well as fewer /ei/ identifications for decisions along the /i-ei/ continuum. The shifted boundaries probably reflect the shifted prototype of high vowels; the frequency with which vowels are identified as mid changes not because the prototype of the mid-vowel has changed, but because their acceptability as high vowels has changed. These results are consistent with a shift in the F1 target of [high] vowels.

In Experiment 3, exposure to /u/ with raised F2 results in fewer /ɛ/ identifications for decisions along the /ʌ-ɛ/ continuum, fewer /ei/ identifications for decisions along the /ou-ei/ continuum, and fewer /ɪ/ identifications for decisions along the /ʊ-ɪ/ continuum. These results are consistent with what Mitterer (2006) found for F2 shifts with an exposure phase that included multiple front vowels and a test phase using a single vowel height. The results in this experiment can be explained by a shift in the F2 target of the [back] feature.

The results of these studies might suggest an important point of consideration in studies using artificial dialects that include several shifted vowels; some artificial dialects might be more natural or easier to learn than others, based on speakers' default expectations about how one vowel should be realized given the realization of a different vowel. For example, lowering /ɪ/ towards /ɛ/ in Experiment 2 went along with lowering of /i/ towards /ei/. If an experiment has exposure stimuli where /ɪ/ becomes /ɛ/ and /i/ becomes /ɪ/, the exposure to shifted /ɪ/ stimuli will set expectations for /i/ that might be in conflict with the actual /i/ stimuli. Weatherholtz (2015) uses a shift mixing tense and lax vowels: /u/ > /ʊ/, /ʊ > /o/, /o/ > /ɔ,ɑ/. Maye et al. (2008) use a shift with primarily lax vowels but also including /i/: /i/ > /ɪ/, /ɪ/ > /ɛ/, /ɛ/ > /æ/. The naturalness of the shifted system might influence whether listeners broaden their categories (cf. Weatherholtz, 2015) or shift them while maintaining the breadth of the category.

5.2 Issues with Alternative Accounts

The results do not provide any evidence that shifting one vowel additionally shifts vowels that are being encroached on. Previous work in convergence has suggested that shifts which would threaten a category boundary are prevented or weakened (Nielsen, 2011; Podlipskỳ & Simácková, 2015). However, no convergence work has found evidence for chain shifts. In Experiment 1, F1 convergence was not extended to the high and low front vowels that the shifted /ɛ/ shifted towards; rather, the shift was extended to /ʌ/, keeping it aligned with the height of /ɛ/. In Experiment 2, shifting the expected F1 values of /ɪ/ influenced the /ɪ-ɛ/ contrast, but the change in that category boundary did not produce a corresponding change in the /ɛ-æ/ contrast. Nonetheless, contrast pressures will not necessarily be the same in other contexts; previous work has demonstrated that there are situations in which contrast preservation does influence phonological behavior.

The results are also not consistent with normalization as an explanation for extension (cf. Sjerps et al., 2013; Watkins & Makin, 1994); there is no evidence that exposure to shifted formants in one vowel changes expectations about the whole vowel space. The results do not seem to be due to the limited exposure to a single vowel. Previous work demonstrates that exposure to a single vowel improves accuracy of identification of other vowels produced by the same speaker (Morton, Sommers, & Lulich, 2015), and accuracy of vowel identifications is not dependent on the particular set of vowels previously heard from that speaker (Verbrugge et al., 1976). Although exposure to a single speaker facilitates normalization and thus accuracy of identification decisions, listeners are substantially above chance accuracy when identifying vowels from several speakers mixed in the same block (Mullennix et al., 1989; Verbrugge et al., 1976). When one formant is manipulated in a way that could suggest a different vocal tract, other formants still reflect the speaker's actual vocal tract. Given that other evidence establishing the vocal tract, listeners may interpret the shifted formant as reflecting a particular phonological target rather than providing evidence for vocal tract length.

5.3 Possible Dialect Effects

This study was not designed to examine how convergence and perceptual learning are influenced by dialectal differences. However, the results suggest that these shifts in perception and production might be affected by how closely the phonological system of the speaker maps onto the phonological system of the listener; this is an important topic for future work. Previous work in convergence has established that speakers can converge to acoustic characteristics that differ between their dialect and the dialect of a model talker for real dialects (e.g. Babel, 2010; Delvaux & Soquet, 2007) and artificial dialects (e.g. Maye et al., 2008; Weatherholtz, 2015). Because the vowel space is continuous, it is not clear if convergence in these cases indicates that the vowels are phonologically the same in both dialects despite their differences in phonetic realization, or if convergence occurs despite a difference in phonological features.

There is less work on convergence to dialectal variants that differ in the number of contrastive sounds; however, Babel, McAuliffe, and Haber (2013) demonstrate that it is possible for speakers to decrease the distance between merging categories based on exposure to a non-merged dialect, which might suggest that lack of phonological alignment is not an absolute constraint on phonetic convergence. Although possible effects of the merger are not the goal of their study, Ross et al. (2021) demonstrate cross-dialectal convergence to /ɔ/ between Mid-Atlantic speakers and "General American" speakers, who were mostly from regions without the /ɔ-ɑ/ merger, which provides evidence that the /ɔ-ɑ/ merger does not prevent convergence to specific realizations of these vowels when the target vowel is itself present in the exposure items. Much of the literature on perceptual learning does not report participants' dialects or include dialect as a predictor, so it remains unclear how perceptual shifts might be influenced by dialectal differences.

In Experiment 3, the results differed when omitting participants from regions with the /ɔ-ɑ/ merger. If all vowels with the [back] feature are fronted in the raised F2 condition, both /ʌ/ and /ɔ/ should be fronted, so both the /ʌ-ɛ/ and /ɔ-ɛ/ decisions should have increased /ɛ/ identifications. However, in the full model, the /ɔ-ɛ/ continuum did not exhibit an effect of manipulation. This might be the result of many American English speakers lacking /ɔ/ in their phonological inventories, as it has lowered to merge with /ɑ/ (Labov et al., 2008). Based on the states that the participants reported being from, about 65% of them were from regions with this merger. If speakers only have /ɑ/, then this vowel would differ from /ɛ/ in height, which might explain why this contrast does not behave like the other F2 contrasts; shifts are only extended to vowels that match phonologically. Consistent with this explanation, the effect of the manipulation on the /ɔ-ɛ/ approaches significance if the analysis is restricted to participants from regions that lack the /ɔ-ɑ/ merger, despite the much smaller sample size. However, other explanations are also possible. Acoustic similarity may be crucial in combination with phonological matching: /ʌ/ is back but not round, which results in an F2 close to the F2 of partially-fronted round /u/, while /ɔ/ has a lower F2 than the typical American English /u/.

6.0 Conclusions

Exposing listeners to shifted formants in a single vowel quality shifts their representations not just for the exposure vowel but also for other vowels which phonologically align along the dimension of manipulation, indicating that the representations are being altered at the featural level rather than the segmental level. The set of vowels impacted by each manipulation suggests that the shift is based on shared phonological representations, rather than shared acoustic characteristics. These shifts are reflected in both F1 and in F2, and in both altered category perception and altered production. This featural pattern provides support for phonological features having detailed acoustic targets, and suggests some aspects of the structure of vowel features.

References

Babel, M. (2009). *Phonetic and social selectivity in speech accommodation* (Unpublished doctoral dissertation). University of California, Berkeley.

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*, 437–456. doi: 10.1017/S0047404510000400

Babel, M., McAuliffe, M., & Haber, G. (2013). Can mergers-in-progress be unmerged in speech accommodation? *Frontiers in Psychology*, *4*, Article 653. doi: 10.3389/fpsyg.2013.00653

Barreda, S. (2020). Vowel normalization as perceptual constancy. *Language*, *96*(2), 224–254. doi: 10.1353/lan.2020.0018

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Becker-Kristal, R. (2010). *Acoustic typology of vowel inventories and Dispersion Theory: Insights from a large cross-linguistic corpus* (Unpublished doctoral dissertation). University of California, Los Angeles.

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, *97*(3), 1916–1924. doi: 10.1121/1.412064

Chládková, K., Podlipskỳ, V. J., & Chionidou, A. (2017). Perceptual adaptation of vowels generalizes across the phonology and does not require local context. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(2), 414–427. doi: 10.1037/xhp0000333

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*, 207–229. doi: 10.1006/jpho.1999.0094

Cohen Priva, U., Edelist, L., & Gleason, E. (2017). Converging to the baseline: Corpus evidence for convergence in speech rate to interlocutor's baseline. *Journal of the Acoustical Society of America*, *141*(5), 2989–2996. doi: 10.1121/1.4982199

Cohn, A. C. (1993). Nasalisation in English: Phonology or phonetics. *Phonology*, *10*, 43–81.

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, *64*, 145–173. doi: 10.1159/000107914

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437. doi: 10.1016/j.jml.2009.05.002

Flege, J. E. (1989). Differences in inventory size affect the location but not the precision of tongue positioning in vowel production. *Language and Speech*, *32*(2), 123–147. doi: 10.1177/002383098903200203

Fujisaki, H., & Kawashima, T. (1968). The roles of pitch and higher formants in the perception of vowels. *IEEE Transactions on Audio and Electroacoustics*, *16*(1), 73–77. doi: 10.1109/TAU.1968.1161952

Gordon, M. J. (2002). Investigating chain shifts and mergers. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 244–266). Oxford: Blackwell.

Joos, M. (1948). Acoustic phonetics. *Language*, *24*(2), 1–136.

Karlin, R., Naber, C., & Parrell, B. (2021). Auditory feedback is used for adaptation and compensation in speech timing. *Journal of Speech, Language, and Hearing*, *ePub Ahead of Issue*.

Keating, P. (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 115–131). Academic Press.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178. doi: 10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*(2), 262–268. doi: 10.3758/BF03193841

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2015). lmertest: Tests in linear mixed effects models [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=lmerTest (R package version 2.0-29)

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.

Labov, W., Ash, S., & Boberg, C. (2008). *The atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*(1), 98–104. doi: 10.1121/1.1908694

Lieberman, P. (1970). Towards a unified phonetic theory. *Linguistic Inquiry*, *1*(3), 307–322.

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

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562. doi: 10.1080/03640210802035357

Mielke, J. (2008). *The emergence of distinctive features*. Oxford: Oxford University Press.

Mitterer, H. (2006). Is vowel normalization independent of lexical processing? *Phonetica*, *63*(4), 209–229. doi: 10.1159/000097306

Moreton, E., & Smolensky, P. (2002). Typological consequences of local constraint conjunction. In *Proceedings of the 21st west coast conference in formal linguistics* (pp. 306–319).

Morton, J. R., Sommers, M. S., & Lulich, S. M. (2015). The effect of exposure to a single vowel on talker normalization for vowels. *Journal of the Acoustical Society of America*, *137*(3), 1443–1451. doi: 10.1121/1.4913456

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*(1), 365–378. doi: 10.1121/1.397688

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142. doi: 10.1016/j.wocn.2010.12.007

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238. doi: 10.1016/S0010-0285(03)00006-9

Oschkinat, M., & Hoole, P. (2020). Compensation to real-time temporal auditory feedback perturbation depends on syllable position. *Journal of the Acoustical Society of America*, *148*(3), 1478–1495. doi: 10.1121/10.0001765

Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, *79*(2), 637–659. doi: 10.3758/s13414-016-1226-0

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Podlipskỳ, V. J., & Simácková, S. (2015). Phonetic imitation is not conditioned by preservation of phonological contrast but by perceptual salience. In *Proceedings of the 18th international congress of phonetic sciences (ICPhS), Glasgow.*

Reinisch, E., & Mitterer, H. (2016). Exposure modality, input variability and the categories of perceptual recalibration. *Journal of Phonetics*, *55*, 96–108. doi: 10.1016/j.wocn.2015.12.004

Reinisch, E., Wozny, D. R., Mitterer, H., & Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics*, *45*, 91–105. doi: 10.1016/j.wocn.2014.04.002

Ross, J. P., Lilley, K. D., Clopper, C. G., Pardo, J. S., & Levi, S. V. (2021). Effects of dialect-specific features and familiarity on cross-dialect phonetic convergence. *Journal of Phonetics*, *86*, Article 101041. doi: 10.1016/j.wocn.2021.101041

Schuhmann, K. S. (2014). *Perceptual learning in second language learners* (Unpublished doctoral dissertation). Stony Brook University.

Sjerps, M. J., McQueen, J. M., & Mitterer, H. (2013). Evidence for precategorical extrinsic vowel normalization. *Attention, Perception, & Psychophysics*, *75*(3), 576–587. doi: 10.3758/s13414-012-0408-7

Skoruppa, K., & Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, *35*(2), 348–366. doi: 10.1111/j.1551-6709.2010.01152.x

Verbrugge, R. R., Strange, W., Shankweiler, D. P., & Edman, T. R. (1976). What information enables a listener to map a talker's vowel space? *Journal of the Acoustical Society of America*, *60*(1), 198–212. doi: 10.1121/1.381065

Watkins, A. J., & Makin, S. J. (1994). Perceptual compensation for speaker differences and for spectral-envelope distortion. *Journal of the Acoustical Society of America*, *96*(3), 1263–1282. doi: 10.1121/1.410275

Weatherholtz, K. (2015). *Perceptual learning of systemic cross-category vowel variation* (Unpublished doctoral dissertation). Ohio State University.

7.0 Appendices

Table 4: Words in Experiment 1

ɛ (shadowing)	ɛ (novel)	ʌ	ɪ	æ
best	chess	bud	bit	back
bet	deck	bus	did	bat
dead	ebb	bust	fit	cat
debt	edge	but	hiss	fad
fed	fret	cub	kid	gas
guess	head	hut	list	lass
less	jet	luck	miss	pass
mess	led	mud	pick	sad
met	let	nut	pit	sat
net	mesh	what	rid	stab
pet	neck			
red	pest			
set	step			
test	vet			
wet	web			

Table 5: Training items in Experiment 2

bit - kit	chip - ship	dip - zip	fig - jig	kick - sick	kiss - this
sit - spit	stick - stitch	thick - tick	bid - big	did - dig	dish - ditch
fib - fizz	fish - fit	hiss - hitch	pick - pig	pit - pitch	skip - skit

Table 6: Test items in Experiment 2

ɛ-æ	ei-ɛ	ɪ-ɛ	ɪ-ei	i-ei	i-ɪ	ʊ-ɔ
guess - gas	age - edge	hid - head	fit - fate	beak - bake	deep - dip	foot - fought
peck - pack	chase - chess	itch - etch	give - gave	keep - cape	heat - hit	hook - hawk
said - sad	date - debt	pit - pet	kiss - case	feed - fade	peak - pick	soot - sought
set - sat	shade - shed	tick - tech	tip - tape	piece - pace	seep - sip	took - talk

Table 7: Training items in Experiment 3

booze - ooze	coo - goo	do - zoo	food - sued	goose - juice	moot - newt
shoe - who	shoot - suit	soup - stoop	boot - booth	chewed - choose	duke - dupe
hoop - hoot	mood - move	scoop - scoot	stewed - stews	toot - tooth	tube - two

Table 8: Test items in Experiment 3

ʌ-ɛ	ɔ-ɛ	ou-ei	ʊ-ɪ	u-i	u-o
bud - bed	bought - bet	goat - gate	foot - fit	boost - beast	boot - boat
chuck - check	saws - says	pose - pays	hood - hid	food - feed	booth - both
duck - deck	sought - set	soak - sake	could - kid	suit - seat	choose - chose
pup - pep	talk - tech	showed - shade	took - tick	tooth - teeth	hoop - hope

Table 9: Model for responses for the proportion of responses of the higher F2 vowel category, based only on participants from regions that lack the /ɔ-ɑ/ merger (57 participants). *Intercept: Contrast = /u-ou/, Manip = lowered F2, BaseVowel = higher F2 category*

	β	*SE*	*z-value*	*p-value*
(Intercept)	-1.52	0.29	-5.26	< 0.001
F2 Step	1.12	0.0368	30.4	< 0.001
BaseVowel Lower F2 Category	-2.15	0.0776	-27.7	< 0.001
Contrast /u-i/	-0.878	0.386	-2.28	0.0227
Contrast /ʌ-ɛ/	-0.264	0.385	-0.686	0.493
Contrast /ɔ-ɛ/	-0.424	0.384	-1.1	0.27
Contrast /ou-ei/	-1.02	0.386	-2.63	0.00845
Contrast /ʊ-ɪ/	-0.255	0.385	-0.662	0.508
Contrast /u-ou/ : Manip RaisedF2	0.0501	0.2	0.251	0.802
Contrast /u-i/ : Manip RaisedF2	**-0.493**	**0.206**	**-2.39**	**0.0169**
Contrast /ʌ-ɛ/ : Manip RaisedF2	**-0.65**	**0.205**	**-3.17**	**0.00151**
Contrast /ɔ-ɛ/ : Manip RaisedF2	-0.345	0.2	-1.72	0.0849
Contrast /ou-ei/ : Manip RaisedF2	**-0.743**	**0.209**	**-3.56**	**< 0.001**
Contrast /ʊ-ɪ/ : Manip RaisedF2	**-0.434**	**0.202**	**-2.15**	**0.032**