Open Library of Humanities

The effect of listener dialect experience on perceptual adaptation to and generalization of a novel vowel shift

Marie Bissell*, Department of Linguistics & TESOL, University of Texas at Arlington, Arlington, TX, USA, marie.bissell@uta.edu

Cynthia G. Clopper, Department of Linguistics, Ohio State University, Columbus, OH, USA, clopper.1@osu.edu

*Corresponding author.

When adapting to novel vowel shifts, listeners rely on the systematicity of novel variants across vowel categories (e.g., parallel shifts), even when only part of a novel system is presented. We examined how a listener's experience with components of a novel English front lax vowel backing shift impacts perceptual adaptation and generalization. First, listeners were exposed to no front lax vowels, shifted /ɪ/, shifted /ɪ æ/, or shifted /ɪ ɛ æ/. Then, listeners responded to items with shifted /ɪ ɛ æ/ in a lexical decision task. Listeners had varied dialect experience due to residential history. Westerners were experienced with /ɪ ɛ æ/ backing in the California Vowel Shift, Southerners were experienced with parallel movements of /ɪ ɛ/ but in another direction in the Southern Vowel Shift, and New Englanders were minimally experienced with front lax vowel shifts. Westerners and New Englanders endorsed more critical words in the /ɪ æ/ exposure condition than in the /ɪ/ exposure condition, consistent with a phonological feature theory of generalization, but Southerners endorsed fewer. Southerners' lack of familiarity with /æ/ shifting in parallel with /ɪ ɛ/ inhibited perceptual adaptation and generalization, suggesting that dialect experience affects listeners' perceptual processing of novel vowel shifts.

2

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

1. Introduction

1.1. Perceptual adaptation and generalization

Listeners are able to perceptually adapt very rapidly when exposed to novel vowel shifts (Maye et al., 2008; Weatherholtz, 2015). For example, Maye et al. (2008) reported that listeners adapted to two novel front vowel chain shifts after brief exposure. The listeners endorsed more shifted words in a lexical decision task after exposure to the chain shifts than after exposure to a control accent. This perceptual adaptation reflects the flexibility required of listeners to accommodate many-to-many mappings among acoustic characteristics and linguistic categories across talkers in the highly variable speech signal (Hillenbrand et al., 1995; Peterson & Barney, 1952).

In addition to perceptual adaptation, listeners also make phonological generalizations to new parts of a talker's accent that they have not yet been exposed to using the parts of the accent they have already heard (Finley & Badecker, 2009; McQueen et al., 2006; Skoruppa & Peperkamp, 2011; Weatherholtz, 2015). For example, Weatherholtz (2015) found that listeners exposed to two novel back vowel chain shifts both perceptually adapted to parts of the chain shifts included in an exposure phase and generalized the acoustic characteristics of the chain shift to a single vowel missing from the exposure in a subsequent lexical decision task. He proposed that listeners rely on knowledge about a novel accent's phonological characteristics, such as knowledge that nearly every member of a natural class with shared phonological features (e.g., [+back] vowels) is involved in a chain shift, to make informed inferences about unheard sounds in that natural class when they are produced in the same accent. Finley and Badecker (2009) similarly argued that listeners in their artificial language learning study used phonological knowledge about feature-based natural class membership to infer that a [+front] vowel left out of their exposure to a novel front vowel harmony pattern should also participate in the vowel harmony pattern. From this perspective, perceptual generalization is crucially dependent upon a listener combining their existing phonological knowledge with specific information from an unfamiliar talker's accent to make an informed generalization about an unheard sound in that accent (see also Skoruppa & Peperkamp, 2011, for similar arguments).

Perceptual generalization may also result from listeners perceptually broadening existing phonological categories (Babel et al., 2021; Baese-Berk et al., 2021; Maye et al., 2008; Zheng & Samuel, 2020). This mechanism involves listeners systematically loosening the usual boundaries of their phoneme categories in response to hearing a novel accent. This category broadening allows listeners to both perceptually adapt to parts of a novel accent included in the exposure phase and perceptually generalize to parts of a novel accent excluded from the exposure phase. Maye et al. (2008) reported some evidence that listeners who were exposed to a novel front vowel shift endorsed more shifted front vowels (i.e., perceptual adaptation), but also endorsed more shifted back vowels (i.e., generalization) in a lexical decision task as compared to a control group, even though canonical (i.e., unshifted) back vowels were included in the exposure to

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

3

the novel accent. Weatherholtz (2015) also included some evidence for perceptual category broadening in a similar study, in which listeners who were exposed to a novel back vowel lowering shift showed increased endorsement rates for both lowered back vowels (i.e., perceptual adaptation) and raised back vowels (i.e., generalization) in a lexical decision task as compared to a control group. Babel et al. (2021) proposed that listeners exposed to /s/ voicing, a rare change in English compared to the more common /z/ devoicing, used the noncanonical nature of the /s/ tokens they heard as the basis for relaxing their phonological category boundaries. That is, the noncanonical nature of the /s/ tokens was generalized to other sounds in the talker's accent. Taken together, these results suggest that exposure to accented speech, including just a single noncanonical variant, can prompt a listener to make broad modifications to (i.e., "loosen") the boundaries of their phoneme categories, including categories that both are and are not directly included in the exposure materials. For this perceptual broadening mechanism, tokens on the acoustic margins of phoneme categories become more acceptable. Unlike the phonological feature account, under this account there should be no relationship between how many sounds in a novel accent a listener has heard and the likelihood that they will generalize from the heard sound(s) to any unheard sound(s) in that novel accent.

These two potential mechanisms for perceptual adaptation and generalization, phonological features and category broadening, critically contrast in their predictions about how hearing tokens of more sound categories in a novel accent should affect a listener's likelihood to generalize to new sounds in that accent. In this study, we exposed listeners to different numbers of vowel categories produced by a talker with a novel front lax vowel backing shift to examine how the number of vowel categories in the exposure affected generalization to vowels left out of the exposure in a subsequent auditory lexical decision task.

In the current study, we predicted that the likelihood that a listener would make a phonological generalization from one front lax vowel to other front lax vowels (i.e., the natural class of front lax vowels) under the phonological feature-based account should be related to how much information from an unfamiliar talker's accent is available as the basis for that generalization. For example, a listener who has heard the acoustic characteristics of only one member of a natural class in a novel accent has accumulated less evidence for phonological generalization to that natural class than a listener who has heard the acoustic characteristics of all but one member of a natural class in a novel accent. A single sound has many phonological features simultaneously (e.g., [+ back], [+ round], [+ high], etc.) and therefore belongs to multiple natural classes, so a listener who has heard the acoustic characteristics of only one sound lacks sufficient information about which natural class should form the basis for their generalization. When a listener hears the acoustic characteristics of multiple sounds that share phonological features, those similarities form a stronger basis for generalization to other members of the natural class.

4

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

1.2. The role of dialect experience in perceptual adaptation and generalization

Dialect experience affects the speed and accuracy of lexical processing (Clopper et al., 2016; Clopper & Walker, 2017; Floccia et al., 2006; Impe et al., 2008). Dialect familiarity promotes speeded lexical processing, consistent with facilitation for familiar dialects relative to less familiar dialects (Clopper et al., 2016; Clopper & Walker, 2017). Therefore, we expect that there should be a positive relationship between a listeners' dialect experience with front lax vowel shifts and how much they perceptually adapt to shifted front lax vowels in the novel front lax vowel backing shift in the current study. Listeners who are familiar with front lax vowel shifts in their local dialect region should adapt more to our novel shift than listeners who are less familiar with front lax vowel shifts.

Dialect familiarity facilitates predictive lexical processing. Porretta et al. (2020) found that listeners with more experience with a non-native dialect more robustly predict upcoming words produced in that novel dialect in a visual-world eye-tracking task than listeners with less experience. By extension, in the current study, dialect experience may facilitate generalization from vowel variants in an exposure phase to vowel variants left out of the exposure phase to the extent that the novel shift is similar to existing shifts in a familiar dialect. Listeners who are familiar with front lax vowel shifts in their local dialect region should generalize more to other vowels in the shift than listeners who are less familiar with front lax vowel shifts.

Chain shifts (like the front lax vowel backing shift in our stimulus talker's dialect) frequently involve entire natural classes, but they may also include a subset of a natural class (Gordon, 2011; Labov, 1994). Chain shifts therefore provide an opportunity to assess the effect of dialect experience on generalization of a novel vowel shift from a phonological features perspective. In particular, listeners who are experienced with a dialect in which a chain shift affects an entire natural class may generalize from vowels in the exposure phase to vowels left out of the exposure phase either due to the vowels' shared phonological features or due to their experience with those vowels behaving similarly to one another in a familiar dialect (i.e., the natural class is involved in a single chain shift with which the listener is already familiar).

In contrast, listeners who are experienced with a dialect in which a chain shift affects only part of a natural class may exhibit a different pattern of feature-based generalization. In particular, these listeners make a robust perceptual generalization from vowels in the exposure phase to vowels left out of the exposure phase when the particular set of vowels behave similarly to each other in their dialect experience, but weaker perceptual generalization from vowels in the exposure phase to vowels left out of the exposure phase when the vowels do not behave similarly to each other in their dialect experience. That is, dialect experience with chain shifts may shape the robustness of the representations of natural classes, leading to

Bissell and Clopper: The effect of listener dialect experience on
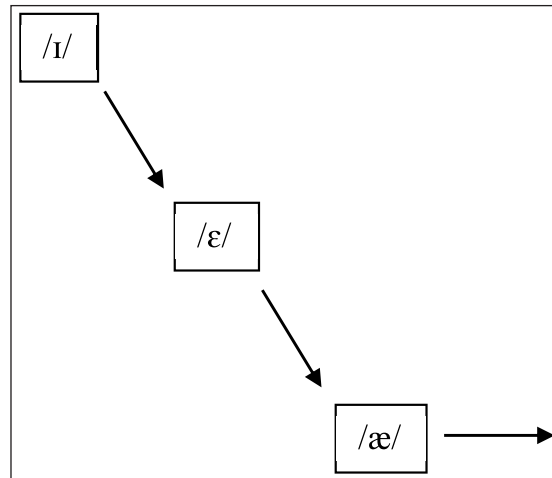perceptual adaptation to and generalization of a novel vowel shift

5

different feature-based generalization as a function of dialect experience. Critically, the category broadening mechanism for generalization should not depend on dialect experience, given that listeners exposed to a noncanonical pronunciation variant equally target other phonological categories for broadening (Babel at al., 2021). Under this account, listeners should robustly target all of a stimulus talker's phonological categories for broadening regardless of their experience with chain shifts.

In the current study, we examined the perceptual adaptation and generalization behaviors of listeners who had lifelong experience with three different American English dialects. Dialect experience with front lax vowel shifts, like the one employed in our talker's novel accent (backed /ɪ ɛ æ/, targeting the natural class of front lax vowels), was expected to vary across the three groups. Specifically, we expected that listeners from these three groups would have different amounts of general and specific familiarity with the front lax vowel shift present in our talker's novel accent. The three listener groups were American Westerners, Southerners, and New Englanders. **Figure 1** shows each of these regions, based on Labov et al. (2006).
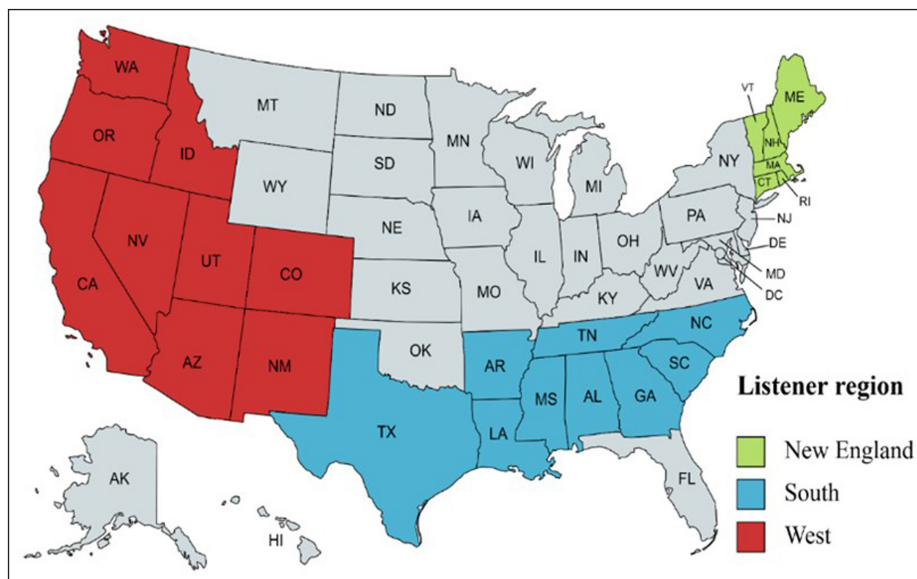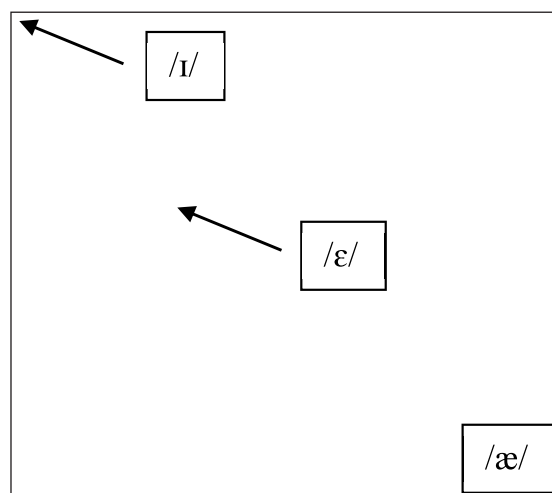

Figure 1: Western, Southern, and New England dialect regions in the United States.

We expected Western listeners to have lifelong exposure to the California Vowel Shift, which is characterized by backed and lowered /ɪ ɛ æ/ (Labov et al., 2006), as shown in **Figure 2**. The shift employed in the novel accent is similar, but not identical, to the California Vowel Shift, in that it involved front lax vowel backing but not lowering.

6

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

Figure 2: The California Vowel Shift.

We expected Southern listeners to have lifelong exposure to the Southern Vowel Shift, which is characterized by raised and fronted /ɪ ɛ/ (Labov et al., 1972), as shown in **Figure 3**. The experience of this listener group was different from the experience of the Western listener group in two important ways. First, the Southern Vowel Shift is characterized by only two of the three front lax vowels shifting in parallel (i.e., fronting and raising). The /æ/ vowel class diphthongizes, increasing its similarity to /eɪ/. Second, the Southern Vowel Shift is directionally different from the shift in the novel accent because it involves fronting and raising instead of backing. Therefore, Southern listeners were expected to have general familiarity with shifted /ɪ ɛ/, but not specific familiarity with /ɪ ɛ/ backing.

Figure 3: The Southern Vowel Shift.

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

7

We expected New England listeners to have lifelong exposure to the New England dialect, which is characterized by nonrhoticity and the low back merger; front lax vowels have not been described as involved in a shift in New England (Labov et al., 2006; Nesbitt & Stanford, 2021). This listener group was assumed to have the least experience with a dialect similar to the novel accent in the current study.

2. Methods

2.1. Participants

Three hundred and eight adult participants were recruited online using the Prolific Academic platform. Thirteen participants who reported a history of speech, hearing, or language disorders and three participants who did not fill out any of the demographic survey were excluded from the analysis. Forty two geographically mobile participants, defined as participants who self-reported residential history of two or more dialect regions as described by Labov et al. (2006), were excluded from the analysis, Additionally, ten participants with overall real word lexical decision accuracy below 75% were excluded from the analysis. Real word lexical decision accuracy was defined using participants' proportions of "word" responses to maximal real control words (i.e., real words not implicated in the novel vowel shift), such that participants who provided 'word' responses to less than 75% of these real words were excluded. The 240 remaining participants (56 men, 162 women, 12 nonbinary, and 10 who did not provide their gender) were between 18 and 64 years old ($M = 29.25$ years, $SD = 10.21$ years). All of the 240 participants reported being monolingual with English as their only native language.

Of the 240 participants, 80 listeners were lifelong residents of each of the three target dialect regions. Lifelong residence was determined by asking participants to self-report all of the places they have lived, and only participants who reported living in a single dialect region were included in the analysis. Westerners were defined as lifelong residents of Arizona, California, Colorado, Idaho, Nevada, New Mexico, Oregon, Utah, and/or Washington. Southerners were defined as lifelong residents of Alabama, Arkansas, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and/or Texas. New Englanders were defined as lifelong residents of Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and/or Vermont.

Although participants were likely to have some familiarity with other dialects, we sought to minimize the influence of this familiarity by selecting only monolingual participants with lifelong residence in the target regions (maximizing familiarity with their native dialect), following Clopper and Walker (2017). Given that listeners from a dialect region may have a variety of experiences with other dialects and that this variety is consistent across different dialect regions, any effects of familiarity with other dialects should be evenly distributed across the three listener groups. Additionally, listeners from border regions within each target dialect region were included (e.g., one participant was from El Paso, Texas near the New Mexico border

8

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

and was included in the Southern group). While these border regions may exhibit transitional dialect features (Cramer, 2010), any effects of these border regions should be evenly distributed across the three listener groups, given that border regions are present in all of the dialect regions we studied.

2.2. Stimulus materials

The stimulus materials in the perceptual adaptation task comprised 467 English monosyllabic words with initial obstruents, nasals, or laterals, and final obstruents (e.g., *fix*, *guess*, *trap*). There were three stimulus types in the experiment: 179 manipulated/test words, 216 control words, and 72 maximal nonwords. The manipulated/test words had the same characteristics, but we refer to them separately because manipulated words were featured in the first phase of the experiment (exposure) and test words were featured in the second phase of the experiment (test). Manipulated/test words consisted of words with /ɪ/ without a competing /ʊ/ minimal pair, words with /ɛ/ without a competing /ʌ/ minimal pair, and words with /ae/ without a competing /ɔ/ or /ɑ/ minimal pair. This minimal pair restriction was designed to prevent listeners from misidentifying the lexical item in the lexical decision task (e.g., misidentifying *bat* with a backed /æ/ as *bot* and providing a "word" response). For example, we chose /blɪts/ as a manipulated/test word because it does not have an /ʊ/ minimal pair competitor */blʊts/. All control words either contained vowels uninvolved in these pairs (/o u/) or back vowels (/ʊ ʌ ɔ ɑ/) without a competing front lax minimal pair. For example, /sɔft/ was selected as a control word because it does not have an /æ/ minimal pair competitor */sæft/. Maximal nonwords are phonotactically legal words with English sounds and syllable structures. Maximal nonwords contrast with test words because test words are real words with a singularly shifted vowel (i.e., no feature change), while maximal nonwords contain one or more feature changes from a real word (Connine et al., 1997). Maximal nonwords contained the full set of vowels from the test and control word lists. Maximal nonwords containing test vowels (/ɪ ɛ æ/) had no competing back lax real-word minimal pair (/ʊ ʌ ɔ ɑ/). For example, the maximal nonword /glɛd/ was selected because its /ʌ/ minimal pair competitor */glʌd/ is also not a real English word; in contrast, the maximal nonword /kɛb/ was not selected because its /ʌ/ minimal pair competitor /kʌb/ is a real English word. The distribution of vowels within each stimulus type, along with example stimulus words, is shown in **Table 1**. The vowel categories were unevenly distributed within each stimulus type due to the lexical restrictions described above. For example, English has more /ʌ/ monosyllables that lack an /ɛ/ minimal pair competitor than /ʊ/ monosyllables that lack an /ɪ/ minimal pair competitor. Despite this uneven distribution, each listener was exposed to all six control vowel categories during the exposure phase.

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

9

| Stimulus type	Vowel distribution	Example stimuli
Manipulated/test	83 /ɪ/ 48 /ɛ/ 48 /æ/	*fix* *guess* *trap*
Control	64 /o/ 41 /u/ 8 /ʊ/ 60 /ʌ/ 29 /ɑ/ 14 /ɔ/	*both* *choose* *good* *much* *stop* *toss*
Maximal nonword	8 /ɪ/ 8 /ɛ/ 8 /æ/ 19 /o/ 13 /u/ 3 /ʊ/ 4 /ʌ/ 5 /ɑ/ 4 /ɔ/	/gɪsp/ /flɛt/ /dæsk/ /bok/ /dʒub/ /kʊf/ /plʌt/ /sɑθ/ /plɔg/

Table 1: Vowel distribution and example stimuli for each stimulus type.

A 20-year-old female native American English speaker with specialized training in reading the International Phonetic Alphabet (IPA) was recorded producing the stimuli in citation form. The recording was made with a Shure stand microphone positioned approximately 5 inches from the stimulus talker and connected to a desktop computer in a sound-attenuated booth. The sampling rate for the recording was 44,100 Hz with 16-bit quantization. The stimulus talker was a lifelong resident of the U.S. Midland dialect region, which encompasses much of central and southern Kansas, Missouri, Illinois, Indiana, and Ohio. The Midland dialect is characterized by the lack of a front lax vowel shift, in contrast to regions like the South and West, and by back vowel fronting (Labov et al., 2006). The uneven distribution of control vowel categories shown in **Table 1**, in which /o u ʌ/ are disproportionately represented relative to /ʊ ɑ ɔ/, may have exposed listeners disproportionately to vowel categories that are noncanonical in the Midland dialect (i.e., /o u/ fronting).
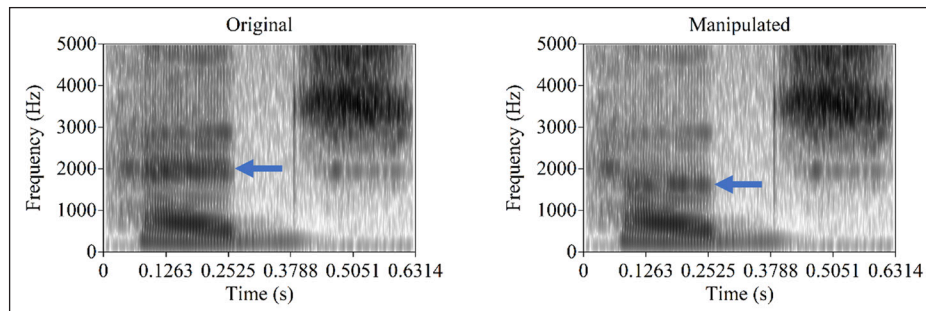
Manipulated/test words (N = 179) and maximal nonwords (N = 24) containing front lax vowels were segmented in preparation for acoustically manipulating the vowels. Segmentations

were completed following the conventions described by Peterson and Lehiste (1960), with boundaries placed at zero crossings. All front lax vowels were manipulated using Praat (Boersma & Weenink, 2022) to have F2s lowered by 300 Hz. To make these changes, the vowel from each word was extracted and the Vocal Toolkit plugin for Praat (Corretge, 2022) was used to lower F2 by 300 Hz throughout the vowel. This 300 Hz modification is consistent with the empirical magnitude of front lax vowel backing in American English as reported in a style-shifting task by Villarreal (2018). The vowel was then reinserted into its consonantal frame. These manipulated/ test words were reasonably naturalistic, as reinsertion was done at zero crossings to avoid acoustic artifacts. Participants did not report that the stimuli sounded unnatural (e.g., robotic). While front lax vowel backing shifts like the California Vowel Shift empirically feature both backing and lowering (Eckert, 2008), we employed an F2 manipulation only to preserve the novelty of the front lax vowel shift in the stimulus talker's dialect for listeners from all three dialect regions.

Figure 4 shows F1 (Hz) and F2 (Hz) means for the stimulus talker's original (left) and modified (right) vowels. In this figure, the front lax vowels are 300 Hz backer on the right panel, but the rest of the vowels are the same as on the left panel. The vowel symbols in the plots represent means over manipulated/test words, control words, and maximal nonwords containing each vowel. Figure 5 shows spectrograms for the word *hedge* before and after manipulation, with the key difference being the F2 of the vowel between 0.1263 s and 0.2525 s.
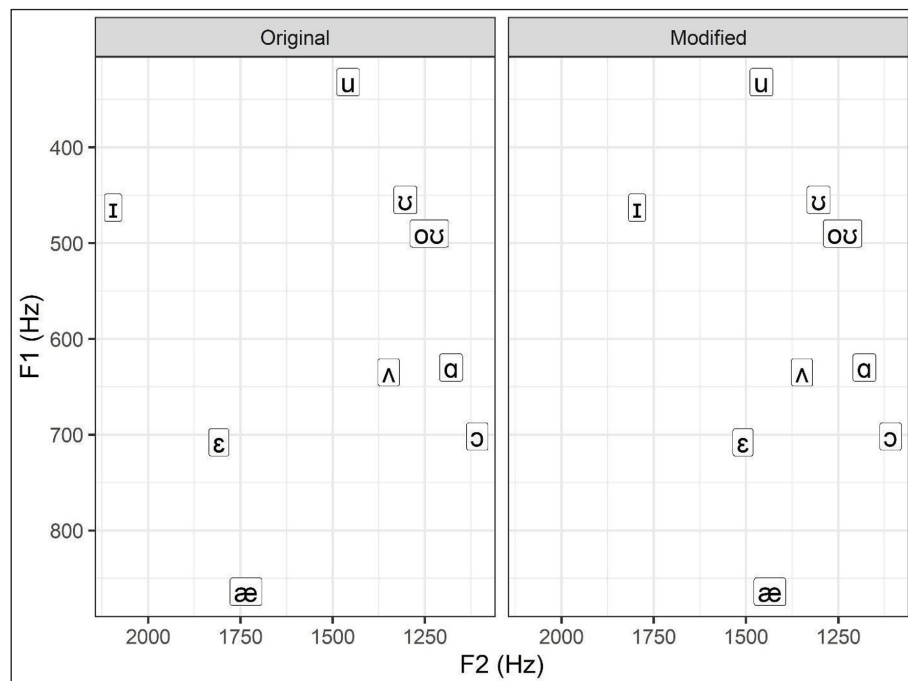
Figure 4: Mean F1 (Hz) and F2 (Hz) values for the stimulus talker's vowels, with original vowels shown in the left panel and manipulated vowels shown in the right panel.

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

11

Figure 5: Spectrograms for the word *hedge,* with the original spectrogram shown on the left panel and the manipulated spectrogram shown on the right panel. Arrows indicate F2.

2.3. Procedure

The experiment was conducted fully online. Listeners completed a headphone check before participating, in which they were asked to type the word they heard over the headphones. Additionally, all participants self-reported that they were wearing headphones before beginning the experiment. The experiment was built using lab.js (Henninger et al., 2020), an open-source JavaScript experiment builder, and hosted on a secure server at the Department of Linguistics at Ohio State University.

The main experiment consisted of an exposure phase immediately followed by a speeded auditory lexical decision phase. The purpose of the exposure phase was to familiarize listeners with the novel accent to prompt perceptual adaptation and generalization in the lexical decision phase. In the exposure phase, listeners heard the stimulus talker produce 144 words in citation form in a randomized order. Listeners were shown the orthography of each word on their computer screen as the word played over their headphones. No response was required of participants during the exposure phase and the trials advanced automatically with a 500 ms intertrial interval. Listeners were not provided with any explicit information about the stimulus talker's accent or identity. There was an attention check every 30 words, for which listeners were asked to type the word they had just heard. The attention checks also functioned as brief breaks from the exposure task, so their consistent spacing served to minimize participant fatigue. All participants whose data were analyzed passed the headphone check, self-reported wearing headphones during the experiment, and passed all attention checks throughout the exposure phase.

There were four between-subject exposure conditions, with each condition containing a different number of the stimulus talker's manipulated vowel categories (none, /ɪ/, /ɪ æ/, and /ɪ ɛ æ/). The 80 listeners from each dialect region were distributed evenly among the four exposure conditions, with 20 listeners per region per condition. For the no exposure condition, which we treat as the baseline condition, the exposure phase consisted of 144 control words. In the other

conditions, the exposure consisted of 72 manipulated words and 72 control words. The exposure phase never contained any maximal nonwords. The number of words included per test vowel varied depending on the exposure condition (/ɪ/ condition: 72 /ɪ/, /ɪ æ/ condition: 36 /ɪ/ and 36 /æ/, and /ɪ ɛ æ/ condition: 24 /ɪ/, 24 /ɛ/, and 24 /æ/). This design therefore entailed consistent overall exposure to the novel accent, even though exposure to individual vowels varied across conditions.

The lexical decision phase contained 180 total stimuli, consisting of 72 control words, 72 maximal nonwords, and 36 test words (12 with each front lax vowel). All test words in the lexical decision task were new (i.e., not included in the exposure phase). Listeners were instructed to decide whether what they heard was a real English word as quickly as possible without sacrificing accuracy. The lexical decision task was auditory only, and no orthography was provided on the screen during the task. Listeners pressed the "f" key on their keyboard to respond "word" and the "j" key on their keyboard to respond "nonword."[1] Trials were separated by a 500 ms intertrial interval, and listeners took self-timed breaks after every 60 trials. The experiment took 10–15 minutes to complete.

The exposure-test design employed in our study provided a between-subjects baseline via the no exposure condition. This between-subjects baseline is an alternative to a within-subjects baseline, in which the design would be test-exposure-test to compare pretest with posttest to assess perceptual adaptation and generalization. Weatherholtz (2015) argued in favor of a between-subjects baseline because it avoids stimulus repetition in pretest and posttest lists, avoids nonword priming effects from pretest to posttest, and reduces participant fatigue by minimizing the number of lexical decision trials. The last benefit was especially important for this study, which was run entirely online. Therefore, we follow Weatherholtz (2015) in treating the no exposure condition as a between-subjects baseline.

2.4. Statistical analysis

Prior to analyzing the data, we excluded lexical decision responses that were faster than 500 ms because they were likely too fast to be in response to the auditory stimulus. We also excluded lexical decision responses that were slower than 5000 ms because they likely indicated lack of participant attention to the task of responding as quickly as possible without sacrificing accuracy. These exclusion conventions follow Clopper and Walker (2017).

To explore how both short-term exposure in the experiment itself and lifetime dialect exposure due to residential history affected perceptual adaptation and generalization behavior for each vowel, we fit a logistic mixed-effects regression model with the *lme4* package (Bates

---

[1] We recognize that this design made it such that participants were more likely than not using their non-dominant (left) hand to respond "word." However, as we did not analyze response times in this study, we did not treat participant handedness as relevant.

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

13

et al., 2015) to predict "word" responses to test words from exposure condition (none, /ɪ/, /ɪ æ/, /ɪ ɛ æ/), listener region (New England, South, West), and vowel (/ɪ/, /ɛ/, /æ/) in a fully factorial design. A scaled covariate of neighborhood density based on the Hoosier Mental Lexicon (Nusbaum et al., 1984) was also included, given that lower neighborhood density has been associated with faster and more accurate lexicality judgments in auditory lexical decision tasks (Vitevitch & Luce, 1999). We also included a scaled covariate of lexical frequency based on the Hoosier Mental Lexicon (Nusbaum et al., 1984), as more frequent words have been shown to elicit more "word" responses than less frequent words in lexical decision tasks (Balota & Chumbley, 1984; Scarborough et al., 1977). Finally, this model also included a covariate corresponding to each listener's mean maximal nonword endorsement rate in the lexical decision task (range = 1.39%–43.06%) to account for potential response biases (Babel et al., 2019; Clarke-Davidson et al., 2008; Weatherholtz, 2015). This covariate represented how often listeners provided "word" responses to maximal nonword stimuli. Listeners who endorse more maximal nonwords may be generally more biased towards providing "word" responses than listeners who endorse fewer maximal nonwords and this covariate captures this variation in bias across listeners. Because trial order was fully randomized for each participant separately, any spillover effects of accuracy on the preceding trial should be randomly distributed across participants and items, so we did not include a covariate to capture spillover effects to avoid overly complexifying the model. We fit the maximal random effect structure that achieved convergence for this model (Barr et al., 2013), which was by-subject and by-item random intercepts.

The significance of fixed-effect main effects and interactions was determined using log-likelihood comparisons. This procedure employs pairwise comparisons of models that differ in a single term (i.e., main effect or interaction), similar to backwards stepwise approaches, with the step order determined manually based on log-likelihood comparisons. In particular, the log-likelihood comparisons were conducted by comparing the maximal model to a model in which the next most complex interaction has been removed. For example, the three-way interaction was tested by comparing the maximal model to a model in which the three-way interaction had been removed (leaving all two-way interactions, covariates, and random effects). For all remaining two-way interactions and fixed effects, each term was compared to the model without the three-way interaction. For example, the exposure condition × listener region calculation was computed by comparing a model with all possible two-way interactions, covariates, and random effects to a model that was identical except that the exposure condition × listener region interaction was removed. This process continued until all interactions and fixed effects had computed log-likelihood comparison values, unless fixed effects were involved in a higher-order interaction, in which case those fixed effects were not tested individually. This manual procedure produced identical results to the buildmer() function in the *buildmer* package (Voeten, 2020). Where applicable, post hoc tests of estimated marginal means using the *emmeans* package (Lenth, 2020) were used to investigate contrasts involved in significant interactions.

The data and analysis scripts are available at the OSF repository for the project (https://osf.io/faw4s/).

3. Results

On average, listeners endorsed or provided "word" responses to 14% of maximal nonwords (SD = 2.6%), 76% of critical words (SD = 4.5%), and 92% of control words (SD = 1.2%). The 14% endorsement rate for maximal nonwords is consistent with nonword endorsement rates that have been reported elsewhere for similar perceptual adaptation experiments with lexical decision tasks, including 12% by Babel et al. (2019) and 15% by Weatherholtz (2015). Overall, these results suggest that the critical words were less recognizable as real words than the unmanipulated control words, but more word-like than the maximal nonwords. Our analysis focused on the endorsement rates for the critical words.

To illustrate how short-term exposure in the experiment itself affected perceptual adaptation and generalization behavior for each vowel, **Figure 6** shows the endorsement rates (i.e., proportions of "word" responses in the lexical decision task) for the critical words in each vowel category in each exposure condition. Endorsement of critical /æ/ words was consistent across all four exposure conditions, whereas endorsement of the critical words with the other two vowels varied by exposure condition. Generally, endorsement of critical /ɛ/ words was lowest, perhaps because the F2 manipulation in the novel accent caused /ɛ/ to be very close to /ʌ/, as shown in **Figure 4**.
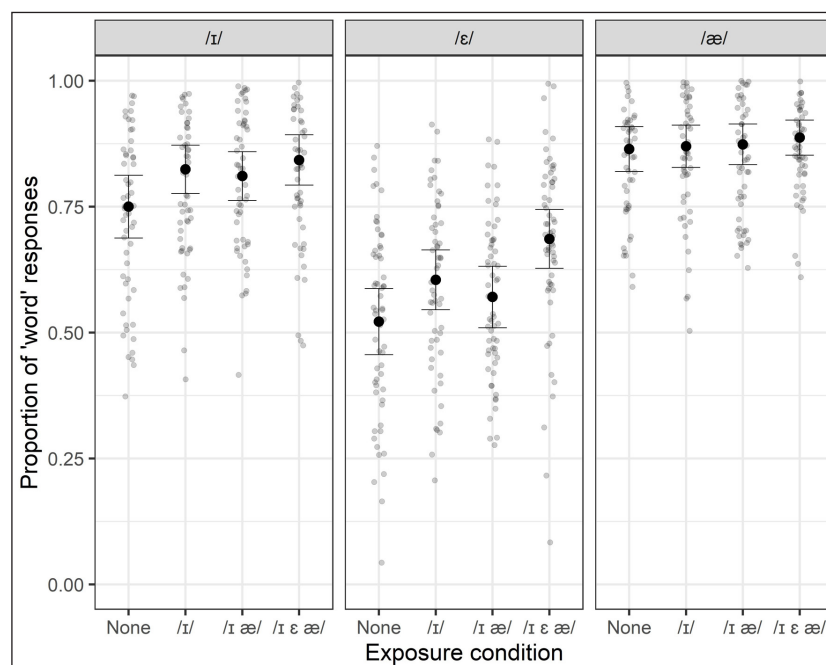
Figure 6: Endorsement rates for critical words across exposure conditions by vowel, with overall subject means shown in black and individual subject means shown in gray. Error bars show standard errors of by-subject means.

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

15

To illustrate how lifetime dialect exposure due to residential history affected perceptual adaptation and generalization behavior for each vowel, **Figure 7** shows endorsement rates for critical words by exposure condition for each listener region. Western listeners showed a steady increase in endorsement rates as the number of vowels in the exposure condition increased, while Southern listeners showed a decrease in endorsement rate in the /ɪ æ/ exposure condition relative to the /ɪ/ exposure condition. New England listeners showed no effect of exposure condition on endorsement rates.
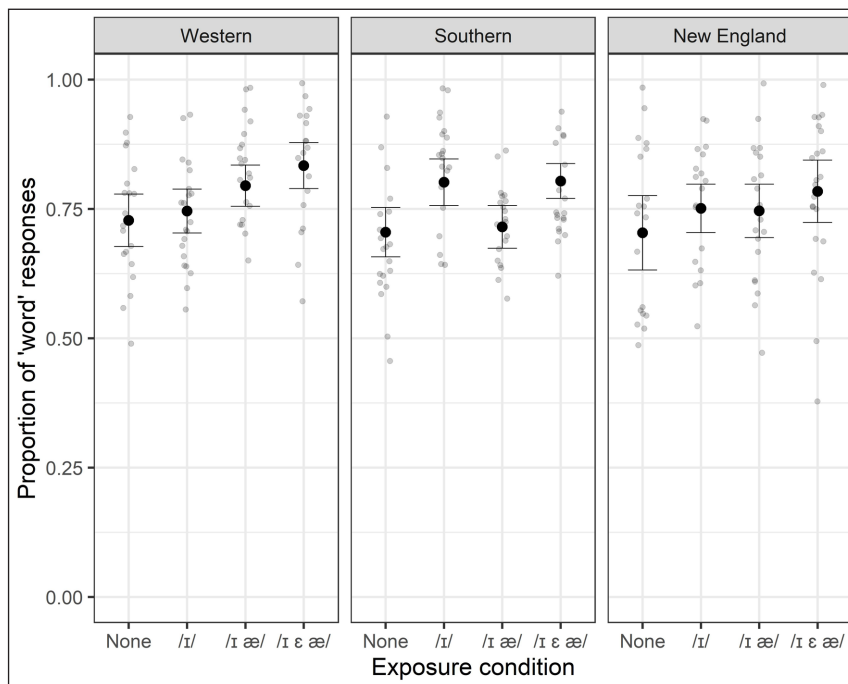
Figure 7: Endorsement rates for critical words across exposure conditions by listener region, with overall subject means shown in black and individual subject means shown in gray. Error bars show standard errors of by-subject means.

The results of the log-likelihood comparisons for the logistic model predicting endorsement rate from exposure condition, listener region, and vowel category for the critical words are shown in **Table 2**. The analysis revealed significant exposure condition × vowel category and exposure condition × listener region interactions, as suggested by **Figures 6** and **7**. The maximal nonword endorsement rate covariate was also significant, such that listeners with higher maximal nonword endorsement rates were more likely to endorse critical words, consistent with an overall "word" bias. Additionally, there was a significant effect of word frequency, such that more frequent words were more likely to elicit "word" responses. The exposure condition × listener region × vowel category and listener region × vowel category interactions were not

16

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

significant. The neighborhood density covariate was also not significant. No other main effects were evaluated for significance in this model, since they were all involved in significant higher-order interactions.

Coefficient	X^2 value	df	*p* value
Exposure condition × listener region × vowel category	12.43	12	.412
Exposure condition × listener region	13.69	6	.033
Listener region × vowel category	2.03	6	.730
Exposure condition × vowel category	16.29	6	.012
Neighborhood density	0.81	1	.369
Maximal nonword endorsement rate	5.29	1	.021
Word frequency	12.92	1	< .001

Table 2: Log-likelihood comparisons for the logistic model predicting endorsement rates from exposure condition, listener region, and vowel category for test words only.

We performed post hoc tests of estimated marginal means with Tukey adjustments for multiple comparisons for each of the significant interactions in the model, focusing on differences across exposure conditions within vowel categories and listener regions to answer our research questions about short-term and lifetime exposure, respectively. The analysis of the exposure condition × vowel category interaction revealed several significant pairwise differences across exposure conditions within individual vowel categories in critical words. Critical words with /ɪ/ were endorsed more in the /ɪ/ exposure condition ($z = 3.26$, $p = .006$), /ɪ æ/ exposure condition ($z = 2.61$, $p = .045$), and the /ɪ ɛ æ/ exposure condition ($z = 4.21$, $p < .001$) than in the no exposure condition (i.e., when /ɪ/ was not included in the exposure). This result is consistent with perceptual adaptation, as expected (Maye et al., 2008; Weatherholtz, 2015). No other pairwise comparisons between exposure conditions for critical words with /ɪ/ were significant.

Critical words with /ɛ/ were endorsed more in the /ɪ ɛ æ/ exposure condition (i.e., when /ɛ/ was included in the exposure phase) than in the /ɪ æ/ exposure condition ($z = 4.33$, $p < .001$), the /ɪ/ exposure condition ($z = 3.24$, $p = .007$), and the no exposure condition ($z = 5.78$, $p < .001$). This result is also consistent with perceptual adaptation, as expected (Maye et al., 2008; Weatherholtz, 2015). Additionally, listeners endorsed critical words with /ɛ/ more in the /ɪ/ exposure condition than in the no exposure condition ($z = 2.59$, $p = .048$), consistent with generalization from /ɪ/ to /ɛ/. No other pairwise comparisons between exposure conditions for critical words with /ɛ/ were significant. None of the pairwise comparisons between exposure

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

17

conditions for critical words with /æ/ were significant. This lack of significant differences across exposure conditions for /æ/ reflects the near-ceiling endorsement rates for /æ/ across all exposure conditions (M = 87.4%).

The post hoc analysis for the exposure condition × listener region interaction likewise revealed several significant pairwise differences across exposure conditions within listener groups. Western listeners endorsed more critical words in the /ɪ ɛ æ/ exposure condition than in the /ɪ/ exposure condition (z = 3.29, p = .006) and no exposure condition (z = 3.55, p = .002), consistent with perceptual adaptation. No other pairwise comparisons between exposure conditions were significant for Western listeners.

Southern listeners also endorsed more critical words in the /ɪ ɛ æ/ exposure condition than in the no exposure condition (z = 2.72, p = .034). Additionally, Southern listeners endorsed more critical words in the /ɪ/ exposure condition than in the /ɪ æ/ exposure condition (z = 3.08, p = .011) and no exposure condition (z = 3.25, p = .006), suggesting an inhibitory effect for Southerners in the /ɪ æ/ exposure condition relative to the /ɪ/ exposure condition. No other pairwise comparisons between exposure conditions were significant for Southern listeners. None of the pairwise comparisons between exposure conditions for critical words were significant for New England listeners.

4. Discussion

4.1. Effects of short-term exposure on perceptual adaptation and generalization

Our main finding about short-term exposure is that exposure to only /ɪ/ facilitated endorsement of critical words with /ɛ/ across listener groups. The robustness of this finding suggests that these two vowels may have a unique perceptual relationship across dialects of American English. Dialectologists have noted that /ɪ ɛ/ frequently shift together in chain shifts in regional varieties of North American English, including in the California Vowel Shift (Eckert, 2008), the Southern Vowel Shift (Labov et al., 1972), the Northern Cities Vowel Shift (Gordon & Strelluf, 2016), and the Canadian Vowel Shift (Boberg, 2019). Listener knowledge about the characteristics of /ɪ ɛ/ in these dialects could have motivated the perceptual generalization pattern we observed. Alternatively, the relationship between these vowels in English (due to their shared phonological features, phonetic similarity, etc.) could motivate both the results of the current study and the chain shift patterns that are empirically observed in English, consistent with Labov's (1994) proposal that phonological features underlie these vowels' parallel shifts in many American English dialects. These two possibilities are difficult to differentiate empirically and require more attention in future research.

The short-term exposure design that we used in this study prioritized consistent overall exposure to the novel accent by making the total number of words in the exposure phase

equivalent across exposure conditions. This design resulted in uneven amounts of exposure to words with each vowel category across exposure conditions, such that the number of tokens per front lax vowel decreased as the number of unique front lax vowels in the exposure condition increased (e.g., 36 /ɪ/ tokens in the /ɪ æ/ exposure condition versus 24 /ɪ/ tokens in the /ɪ ɛ æ/ exposure condition). It is possible that this relationship could have affected the results we report here. For example, it could be the case that increasing the number of vowel categories in the exposure condition actually decreases that amount of perceptual adaptation and generalization observed due to diminishing tokens of each vowel category as the number of unique vowel categories increases. However, we did not observe this pattern in our results, as evidenced by the vowel-specific results shown in **Figure 6** (e.g., endorsement of critical words with /æ/ is consistent across exposure conditions even with variable token counts). Therefore, we acknowledge that this design choice is a possible factor in our results, but we did not find specific evidence of its effect.

4.2. Effects of dialect experience on perceptual adaptation and generalization

Our analyses revealed variation in perceptual adaptation and generalization across the three listener groups. New England listeners showed no effect of exposure condition on endorsement rates for test words, which is consistent with a perceptual broadening mechanism for perceptual generalization. That is, endorsement rates of test words with specific vowels were comparable regardless of whether listeners were perceptually adapting to a vowel included in the exposure phase or perceptually generalizing to a new vowel in the test phase. This endorsement behavior, in which New England listeners provided comparable rates of "word" responses to test words across all exposure conditions, was present even in the no exposure condition, when there were no noncanonical test vowels included in the exposure phase. This result suggests that New England listeners employ category broadening that does not depend on specific exposure to any particular noncanonical front lax vowel(s) produced by an unfamiliar talker. That is, listeners who were the most unfamiliar with the characteristics of the novel accent showed evidence of a category broadening mechanism because they lack the basis for generalization based on phonological features; rather, their ultimate generalization is that all vowel categories in the novel accent are noncanonical.

However, this sort of category broadening may have been prompted by exposure to unshifted Midland vowels in the control words in the exposure phase (see **Figure 4**), consistent with the idea that specific exposure (i.e., hearing a noncanonical or unfamiliar vowel variant) prompts category broadening. For example, New England listeners may have been unfamiliar with /u/ fronting in the stimulus talker's Midland accent, leading to category broadening for all vowel categories in the lexical decision task. Another possibility is that New England listeners would employ category broadening in response to all unfamiliar talkers, not just this one. Testing this

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

19

possibility would require exposing New England listeners to additional stimulus talkers with different accents, which is one potential direction for future research.

In contrast to the New England listeners, Western listeners endorsed more test words in the /ɪ ɛ æ/ exposure condition than in the /ɪ/ and no exposure conditions. This result is consistent with perceptual adaptation findings in the previous literature, such that perceptual adaptation to test vowels included in the exposure phase is more robust than perceptual generalization to test vowels excluded from the exposure phase (Maye et al., 2008; Weatherholtz, 2015). However, this result is not fully consistent with either a perceptual broadening or a phonological feature account of perceptual generalization. There is an effect of exposure condition, but test word endorsement rates in the /ɪ/ and no exposure conditions only contrast with test word endorsement rates in the /ɪ ɛ æ/ exposure condition (i.e., the condition requiring only perceptual adaptation without perceptual generalization). We did not find evidence that increasing the number of test vowels in the exposure condition increases endorsement rates in the lexical decision task, as evidenced by comparable endorsement rates of test words in the no, /ɪ/, and /ɪ æ/ exposure conditions. Thus, these listeners showed very weak evidence of perceptual generalization and their results therefore do not support either mechanism for perceptual generalization.

Finally, Southern listeners endorsed more test words in the /ɪ ɛ æ/ exposure condition than in the no exposure condition, but also endorsed more test words in the /ɪ/ exposure condition than in the /ɪ æ/ and no exposure conditions. Generally, listeners would be expected to endorse more test words containing a particular vowel when that vowel is included in the exposure phase than when it is not because then the task involves perceptual adaptation to vowels in the exposure phase rather than perceptual generalization to new vowels in the test phase. However, Southern listeners endorsed more test words in the /ɪ/ exposure condition than in the /ɪ æ/ exposure condition, which is not consistent with perceptual adaptation that has been observed in previous studies (Maye et al., 2008; Weatherholtz, 2015). In the absence of a significant three-way interaction, we did not explore endorsement rates separately for each vowel category in each exposure condition for the Southern listeners. However, given the near-ceiling endorsement rates for critical words with /æ/ across exposure conditions (see **Figure 6**), the difference in overall endorsement rates for the Southern listeners in the /ɪ/ and /ɪ æ/ exposure conditions likely reflects primarily differences in endorsement rates for critical words with /ɪ ɛ/.

The results for the Southerners are most consistent with a phonological feature mechanism for perceptual generalization, but crucially one that is directly influenced by dialect experience. The phonological feature mechanism is reliant upon generalization within natural classes, which are conventionally defined by shared phonological features (e.g., [+front] and [–tense]). Previous perception studies (e.g., Chládková et al., 2017; Weatherholtz, 2015) have shown that language users have implicit knowledge of natural classes. However, we propose that Southern listeners treat the three front lax vowels as a natural class less robustly due to lifelong exposure

to the Southern Vowel Shift, which is characterized by fronting and raising of /ɪ ɛ/ but not /æ/ (Labov et al., 1972). From this perspective, we argue that Southern listeners showed reduced endorsement of test words in the /ɪ æ/ exposure condition as compared to the /ɪ/ exposure condition because /ɪ æ/ patterning together is inconsistent with their dialect experience with the Southern Vowel Shift, even though /ɪ æ/ are both in the natural class of front lax vowels. In the /ɪ/ exposure condition, Southern listeners encountered positive evidence that the novel accent was distributionally similar to the Southern Vowel Shift (i.e., had noncanonical /ɪ/). That is, Southern listeners have general familiarity with noncanonical /ɪ/, although not specific familiarity with backed /ɪ/. This general familiarity permits Southern listeners to make the appropriate generalization. However, in the /ɪ æ/ exposure condition, Southern listeners encountered both positive and negative evidence that the novel accent was distributionally similar to the Southern Vowel Shift (i.e., had a noncanonical /ɪ/, but also an /æ/ vowel class that patterned with /ɪ/). That is, the novel accent had both generally familiar and unfamiliar patterns for Southern listeners. The Southern listeners' response to this mixed evidence was reduced perceptual adaptation and generalization across all test vowels, suggesting that their weaker natural class of front lax vowels inhibited both perceptual adaptation and generalization.

The Southern listeners' results thus demonstrate that familiarity with aspects of a novel accent may not always facilitate lexical processing. Instead, general familiarity (i.e., familiarity with /ɪ ɛ/ shifting in parallel but in a different direction than in the novel accent) inhibits lexical processing of the novel dialect when unfamiliar components are introduced (i.e., /æ/ backing). That is, Southern listeners' familiarity with a front lax vowel shift that does not include /æ/ inhibited their processing of a front lax vowel shift that does include /æ/ because inclusion of /æ/ in the novel front lax vowel shift mismatched Southern listeners' expectations about which front vowels will be noncanonical in the novel front lax vowel shift and, crucially, noncanonical in the same way (i.e., backed). Moreover, Southern listeners performed similarly in the /ɪ/ exposure condition and the /ɪ ɛ æ/ exposure condition because natural classes are relevant to perceptual generalization but not to perceptual adaptation. Noncanonical /ɪ/ forms the basis for perceptual generalization in the former case, whereas in the latter case no generalization is required. These results add nuance to existing findings showing that dialect familiarity promotes speeded lexical processing (Clopper & Walker, 2017; Floccia et al., 2006; Impe et al., 2008), suggesting that even generally familiar phonological structures can inhibit processing if they mismatch a listener's specific dialect exposure.

Taken together, these results suggest that dialect experience affects perceptual adaptation and generalization of a novel front lax vowel shift. New England listeners show robust perceptual adaptation and generalization regardless of exposure condition, consistent with a perceptual broadening mechanism. Western listeners show weak evidence of perceptual generalization, supporting neither a perceptual broadening mechanism nor a phonological feature mechanism. Southern listeners show evidence of a phonological feature mechanism that is affected by their

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

21

dialect experience with front lax vowel patterns, leading to inhibition of both perceptual adaptation and generalization in the exposure condition that is inconsistent with their dialect experience.

This mixed pattern is consistent with previous work, which provides evidence for both phonological feature and category broadening mechanisms of perceptual adaptation and generalization for a range of variants, including vowel raising/lowering (Maye et al., 2008; Weatherholtz, 2015) and fricative voicing/devoicing (Babel et al., 2021). Babel et al. argued that these mechanisms are both available, but crucially arise in response to stimuli with specific characteristics. Our results provide additional evidence that a single mechanism does not underlie perceptual adaptation and generalization behaviors for speech. They further suggest that listeners' experiences with phonological patterns in their native dialects and the phonological patterns present in a novel accent interface to elicit a particular perceptual adaptation and generalization mechanism. It is therefore possible that one factor driving the variation observed in perceptual generalization behaviors in previous work is individual dialect exposure, where listeners with exposure to a dialect with similar phonological structures to a novel accent adopt a phonological feature mechanism and listeners with exposure to a dialect with dissimilar phonological structures to a novel accent adopt a category broadening mechanism. Crucially, the interface between a listener's dialect exposure and the characteristics of a novel accent determines which mechanism underlies perceptual generalization in a particular case. For example, we predict that the listeners in our study would show different perceptual generalization mechanisms when encountering a different kind of novel accent, making the mechanism a function of the relationship between dialect exposure and the novel accent rather than simply a property of the listeners or the variants themselves.

4.3. Conclusion

The results of the current study reveal that both short-term exposure and lifetime dialect experience affect perceptual adaptation to and generalization of a novel vowel shift. The perceptual generalization relationship observed between /ɪ ɛ/ (i.e., that exposure to only backed /ɪ/ facilitates endorsement of backed /ɛ/) suggests that these two vowels have a special relationship, which may be due to their shared phonological features or their empirical behaviors in American English vowel shifts. Further research is required to disentangle these two explanations from one another. In terms of lifetime experience, listeners with the least amount of dialect experience with similar vowel shifts exhibited a category broadening mechanism of perceptual generalization, while listeners with dialect experience with very similar vowel shifts showed little evidence of perceptual generalization. Listeners with dialect experience with dissimilar front lax vowel shifts exhibited a phonological features mechanism of perceptual generalization. Future investigations of the role of participant demographic factors like age may yield further insights into how lifetime exposure to variation affects perceptual adaptation and generalization.

22

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

Ethics and consent statement

Acknowledgements

Competing interests

Author contributions

References

Babel, M., Johnson, K. A., & Sen, C. (2021). Asymmetries in perceptual adjustments to noncanonical pronunciations. *Laboratory Phonology, 12*(1). https://doi.org/10.16995/labphon.6442

Babel, M., McAuliffe, M., Norton, C., Senior, B., & Vaughn, C. (2019). The Goldilocks zone of perceptual learning. *Phonetica, 76*(2–3), 179–200. https://doi.org/10.1159/000494929

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *Journal of the Acoustical Society of America, 133*, EL174–EL180. https://doi.org/10.1121/1.4789864

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance, 10*, 340–357. https://doi.org/10.1037//0096-1523.10.3.340

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

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

Boberg, C. (2019). A closer look at the short front vowel shift in Canada. *Journal of English Linguistics, 47*(2), 91–119. https://doi.org/10.1177/0075424219831353

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer [Computer program]. Version 6.2.11. http://www.praat.org/

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

23

Chládková, K., Podlipský, 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. https://doi.org/10.1037/xhp0000333

Clarke-Davidson, C. M., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias? *Perception & Psychophysics, 70*(4), 604–618. https://doi.org/10.3758/pp.70.4.604

Clopper, C. G., Tamati, T. N., & Pierrehumbert, J. B. (2016). Variation in the strength of lexical encoding across dialects. *Journal of Phonetics, 58*, 87–103. https://doi.org/10.1016/j.wocn.2016.06.002

Clopper, C. G., & Walker, A. (2017). Effects of lexical competition and dialect exposure on phonological priming. *Language and Speech, 60*(1), 85–109. https://doi.org/10.1177/0023830916643737

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language, 37*(4), 463–480. https://doi.org/10.1006/jmla.1997.2535

Corretge, R. (2022). Praat Vocal Toolkit [Computer program]. http://www.praatvocaltoolkit.com/

Cramer, J. (2010). *The effect of borders on the linguistic production and perception of regional identity in Louisville, Kentucky* [Doctoral dissertation]. University of Illinois at Urbana-Champaign.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism, 12*(1–2), 25–42. https://doi.org/10.1177/13670069080120010301

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

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276–1293. https://doi.org/10.1037/0096-1523.32.5.1276

Gordon, M. J. (2011). Methodological and theoretical issues in the study of chain shifting. *Language and Linguistics Compass, 5*(11), 784–794. https://doi.org/10.1111/j.1749-818x.2011.00310.x

Gordon, M. J., & Strelluf, C. (2016). Working the early shift: Older Inland Northern speech and the beginnings of the Northern Cities Shift. *Journal of Linguistic Geography, 4*(1), 31–46. https://doi.org/10.1017/jlg.2016.7

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2020). lab.js: A free, open, online study builder: https://doi.org/10.5281/zenodo.597045

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97*(5), 3099–3111. https://doi.org/10.1121/1.411872

Impe, L., Geeraerts, D., & Speelman, D. (2008). Mutual intelligibility of standard and regional Dutch language varieties. *International Journal of Humanities and Arts Computing, 2*(1–2), 101–117. https://doi.org/10.3366/e1753854809000330

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

24

Bissell and Clopper: The effect of listener dialect experience on
perceptual adaptation to and generalization of a novel vowel shift

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology and sound change.* Walter de Gruyter.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A quantitative study of sound change in progress.* Philadelphia, PA: U.S. Regional Survey.

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* R package version 1.5.2-1. https://CRAN.R-project.org/package=emmeans

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. https://doi.org/10.1080/03640210802035357

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113–1126. https://doi.org/10.1207/s15516709cog0000_79

Nesbitt, M., & Stanford, J. N. (2021). Structure, chronology, and local social meaning of a supra-local vowel shift: Emergence of the low-back-merger-shift in New England. *Language Variation and Change*, 1–27. https://doi.org/10.1017/s0954394521000168

Nusbaum, H. C. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, 10*, 357–376.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*(2), 175–184. https://doi.org/10.1121/1.1906875

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693–703. https://doi.org/10.1121/1.1908183

Porretta, V., Buchanan, L., & Järvikivi, J. (2020). When processing costs impact predictive processing: The case of foreign-accented speech and accent experience. *Attention, Perception, & Psychophysics, 82*(4), 1558–1565. https://doi.org/10.3758/s13414-019-01946-7

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Learning & Memory, 3*, 1–17. https://doi.org/10.1037//0096-1523.3.1.1

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

Villarreal, D. (2018). The construction of social meaning: A matched-guise investigation of the California Vowel Shift. *Journal of English Linguistics, 46*(1), 52–78. https://doi.org/10.1177/0075424217753520

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374–408. https://doi.org/10.1006/jmla.1998.2618

Voeten, C. C. (2020). *Buildmer: Stepwise elimination and term reordering for mixed-effects regression.* R package version 1.7.1. https://CRAN.R-project.org/package=buildmer

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

Zheng, Y., & Samuel, A. G. (2020). The relationship between phonemic category boundary changes and perceptual adjustments to natural accents. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(7), 1270–1292. https://doi.org/10.1037/xlm0000788