Language Variation and Change, **29** (2017), 79–100.

doi:10.1017/S0954394517000023

Producing and perceiving the Canadian Vowel Shift: Evidence from a Montreal community

THOMAS KETTIG
University of Hawai'i at Mānoa

BODO WINTER
University of Birmingham

ABSTRACT

This paper investigates interspeaker variation in the mid and low short vowels of Jewish Montreal English, analyzing the Canadian Shift in both production and perception. In production, we find that young women are leading in the retraction of /æ/ and the lowering and retraction of /ɛ/. We furthermore find that across speakers, the retraction of /æ/ is correlated with the lowering and retraction of /ɛ/, providing quantitative evidence that the movements of these two vowels are linked. The trajectory implied by our production data differs from what was reported in Montreal approximately one generation earlier. In contrast to reliable age differences in production, a vowel categorization task shows widespread intergenerational agreement in perception, highlighting a mismatch: in this speech community, there is evidently more systematic variation in production than in perception. We suggest that this is because all individuals are exposed to both innovative and conservative variants and must perceptually accommodate accordingly.

The Canadian Shift (CS) is a systematic lowering and/or retraction of /ɪ/, /ɛ/, and /æ/ in Canadian English. Although the progress of the CS across Canada has been investigated extensively in production, perception of the CS has received comparatively little attention. In general, the relationship between production and perception in ongoing chain shifts is underresearched, leaving many questions unanswered regarding how changes in pronunciation correspond to changes in perception (Kendall & Fridland, 2012; Thomas, 2002). This paper

We would like to thank our wonderful, patient participants and the many people whose comments and suggestions contributed to the design and analysis of this study, including Christian Bentz, Charles Boberg, David Fleischer, Matt Hunt Gardner, Heather Goad, Becky Roeder, Morgan Sonderegger, Michael Wagner, audience members at the Sixth International Conference on Experimental Linguistics (Athens, Greece) and New Ways of Analyzing Variation (NWAV) 44 (Toronto, Ontario), and two anonymous reviewers. All errors remain our own.

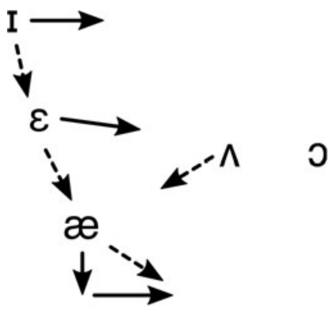

FIGURE 1. Dashed lines indicate the CS according to Clarke et al.'s (1995) impressionistic analysis; solid lines indicate Boberg's (2005) evidence from Montreal.

investigates the apparent-time trajectory of the /ɛ/ and /æ/ vowels in both production and perception.

Using data from a community in Montreal, we present new apparent-time evidence for the CS. In our production study, we elicited the nonhigh short (lax) /ɛ/, /æ/, /ʌ/, and /ɔ/ vowels from a group of 28 Anglophone Jewish Montrealers. We find age differences to be more pronounced for /ɛ/ than for /æ/. This suggests that the shift's earliest stage, the retraction of /æ/, is stabilizing. We furthermore show that if a speaker manifests a shift in /æ/, they also manifest a shift in /ɛ/, showing a link between the two vowels. In our perception study, the same participants judged synthetic vowel stimuli as belonging to categories represented by the words *bet, bat, but*, or *bought*. The results indicate that although people differ in the degree to which they produce innovative forms, they do not differ as much in vowel perception; even speakers who are leading the change (young women) exhibit only minor trends toward altered vowel perception. This is consistent with a view of incremental vowel change in which speakers may advance a shift's trajectory in production but retain flexibility in perception, presumably due to constant exposure to both conservative and innovative variants.

Both our production study and our perception study aim to add to the descriptive coverage of the CS. Clarke, Elms, and Youssef (1995) first reported that the /æ/, /ɛ/, and /ɪ/ vowels of English speakers from Ontario were involved in a chain shift, which they described as a lowering of the front lax vowels in apparent time (Figure 1). Many North American dialects such as Canadian and Californian English exhibit the merger of /a/ and /ɔ/ (LOT and THOUGHT, as well as PALM, in Wells's lexical set [1982]). Labov (1991) predicted that this merger would result in a relatively stable "third dialect," avoiding the ongoing chain shifts affecting the Northern Cities and the South. However, the low back merger instead created the conditions for /æ/ to lower and retract into the space vacated by /a/, with /ɛ/ and /ɪ/ subsequently moving. This shift has not only been noted in Canada, but also in other regions with the low back merger such as California (Eckert, 2008; Kennedy & Grama, 2012; Labov, Ash, & Boberg, 2006), Columbus, Ohio (Durian, 2012), Southern Illinois (Bigham, 2010), and Hawai'i (Drager, Kirtley, Grama, & Simpson, 2013).

Research following up on Clarke et al. (1995) has confirmed the existence of the CS across Canada. However, studies have reported a variety of different phonetic trajectories. To take just a few examples: Labov et al.'s (2006) *Atlas of North American English*, based on Canada-wide data from their Telsur project, described the CS as a chain shift involving first the retraction of merged /ɔ/,[1] followed by the retraction of /æ/, and finally the lowering and retraction of /ɛ/ into the space vacated by /æ/. Sadlier-Brown and Tamminga (2008) showed that among both Halifax and Vancouver speakers, /ɪ/ and /ɛ/ were retracting and lowering. On the other hand, they found /æ/ to be just retracting in Vancouver and both retracting and lowering in Halifax. Hoffman (2010) found young women to be leading the retraction of /æ/ in Toronto, while younger Torontonians as a whole were found to both retract and lower /ɛ/.

In Montreal, Boberg (2005) found /æ/ to have lowered before starting to retract in apparent time; he also found retraction of /ɛ/ and /ɪ/, but no statistically reliable degree of lowering. He argued that these findings are problematic for the definition of the CS as a pull shift, in which /æ/ might be expected to retract into the unfilled space created by the merger of the low back vowels, followed by /ɛ/ and /ɪ/ lowering into the space created by /æ/. Instead, the CS may manifest itself as a series of parallel retractions. This poses a dilemma for theories emphasizing vowel contrast maintenance (e.g., de Boer, 2001; Martinet, 1955) that would predict the low back merger and movement of /æ/ to primarily affect vowel height rather than the front-back dimension.

Following Boberg (2005), Roeder and Jarmasz (2010) provided a model of the CS as a systematically related series of retractions. Their study of Toronto English found strong retraction in /ɛ/ and /æ/, but they also reported retraction in /ɔ/, suggesting the following:

> rather than a chain shift, /æ/ and /ɛ/ are simultaneously redistributing within the reconfigured vowel space resulting from the low back vowel merger, and are engaged in a parallel shift that is motivated by the tension between forces of articulation, perception, and contrast. We propose that the vowels are ultimately moving towards equilibrium of a symmetrical vowel system. … The Canadian Shift comprises two stages. The first involves concurrent lowering and retraction and the second involves retraction only and includes retraction of /ɔ/. (397–398)

Roeder and Gardner (2013) elaborated on the phonological underpinnings of these movements by analyzing the feature specification of Canadian English vowels, with only contrastive features treated as active within the phonology. In their system, the /æ/ phoneme is specified as [–Peripheral, –High, +Low] and the merged /ɔ/ phoneme as [+Peripheral, –High, +Low]. Seen this way, the absence of a phonological feature specifying the "horizontal" (F2-related) dimension frees /æ/ to move along this dimension phonetically. Moreover, in their system, movement of /æ/ and /ɛ/ is thought to be affected by phonetic pressures, in particular with respect to perceptual dispersal accounts (de Boer, 2001; Liljencrants & Lindblom, 1972), which state that vocalic systems optimize the perceptual distance between vowels.

Varying descriptions of the CS may reflect differing stages of the shift, as well as differing methodological approaches, rather than any large regional distinctions

TABLE 1. *Gender and age composition of participant sample*

	Men	Women
Older group (1937–1961)	11	5
Younger group (1984–1995)	7	5

(Kettig, forthcoming). At present, many questions remain open with respect to the trajectory of the CS and its status as a series of parallel phonetic retractions or a chain shift triggered by an earlier phonological merger. We report on data from Montreal, precisely the city where Boberg (2005) found parallel retraction.

PRODUCTION STUDY

Montreal: Sample and historical context

Over the course of the 18th and 19th centuries, several Anglophone communities took root in Montreal as successive waves of immigrants from the British Isles and from Southern and Eastern Europe settled in the city and adopted (or retained) English as a home language. Quebec's "Quiet Revolution" in the 1960s made French the sole language of government, business, and schooling. This prompted an exodus of Anglophones from the province; since the establishment of these language laws, Montreal has lost over a third of its English-speaking population (Boberg, 2010). The 2011 Canadian census reported that in Greater Montreal, 9.9% of residents use English as their sole language at home, 9.5% are domestic English/French bilinguals, and nearly 5.2% use English and some other language at home. English speakers in Montreal are thus greatly outnumbered by those speaking French at home, who total over 56.6% of residents (Statistics Canada, 2011).

Boberg (2004) has described the three main ethnic varieties of Montreal English as British/Irish, Italian, and Jewish. These ethnolects differ along a number of linguistic dimensions, including the phonetic realization of vowels. This study was limited to the city's Jewish community (population ~80,000). In order to qualify for the present study, subjects needed to have been born and raised in Montreal with at least one Jewish parent and had to report English as a first and home language. None of the participants reported impaired hearing.

We carried out interviews in early 2013. Our sample comprised a total of 12 younger participants and 16 older participants, 10 women and 18 men (see Table 1).

Methodology of production study

Participants were recorded reading a list of 44 sentences containing words with the /ɛ/, /æ/, /ɔ/, and /ʌ/ vowels under primary stress, for example, "*He bought it at the **mall**, not at the supermarket"* (see Appendix A for full carrier sentences). We attempted to balance target words for the voicing, place, and manner of

TABLE 2. *Target words elicited in production study*

Following Consonant Type	ʌ	ɔ	ɛ	æ
Voiced stop	hug	knob	beg	lab
Voiced stop	mud	nod	bed	bad
Voiced fricative	buzz	cause	says	jazz
Alveolar nasal	run	gone	pen	pan
Bilabial nasal	gum	mom	gem	ham
Open syllable + lateral	gully	holly	belly	valley
Closed syllable + lateral	gull	mall	sell	Sal
Voiceless stop	stuck	sock	neck	stack
Voiceless stop	mutt	lot	bet	rat
Voiceless fricative	rough	soft	Stef	staff
Voiceless fricative	fuss	loss	Jess	pass

articulation of the consonant following the vowel (see Table 2); these linguistic factors have been shown to variably favor or inhibit aspects of the CS in production (De Decker & Mackenzie, 2000).[2]

Some interviews were recorded in sound-attenuated booths at McGill University using a head-mounted Logitech H390 microphone-headphone set for both the perception and production tasks. For the interviews that were conducted in participants' homes and offices, a professional high-definition USB recorder was used for the production task and a set of over-ear Sennheiser headphones was used for the perception task. In all cases, speech was recorded with Praat version 5.3.37 (Boersma & Weenink, 2013) at a sampling rate of 44,100 Hz. The subject was left alone in a silent room or inside the booth for both tasks.

All wave file recordings were processed with the Prosodylab-Aligner software (Gorman, Howell, & Wagner, 2011) to force-align phonemes. After alignment, each token was manually checked. In addition to analyzing the 44 tokens of stressed short vowels, we extracted 22 vowels from the surrounding carrier sentences in order to conduct normalization of the vowel space for all participants (see http://www.github.com/bodowinter/canadian_vowel_shift_analysis/). F1 and F2 data were extracted from the midpoint of each short vowel and the 33% and 66% points of each long/diphthongal vowel (Harrington & Cassidy, 1994), using a slightly modified Praat script by Lennes (2003).

F1 and F2 normalized values were calculated with the NORM online software suite (Thomas & Kendall, 2007), using the method described by Lobanov (1971) and following the best practice suggestions of Adank, Smits, and van Hout (2004). Normalized values were rescaled into Hertz in order to orient the vowels in relation to each other in a more familiar way. However, it should be noted that these rescaled values are not directly comparable to those of the synthetic vowels of the perception task, which represent actual formant values.

Results

All statistical analyses were conducted with R version 3.2.2 (R Core Team, 2015).

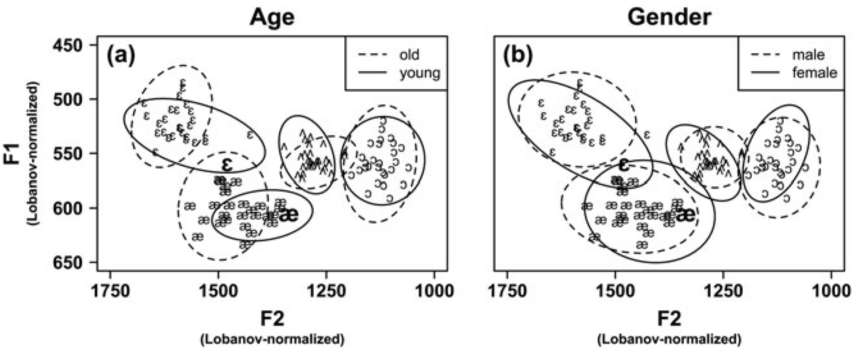

FIGURE 2. Lobanov-normalized F1/F2 speaker means with 95% confidence ellipses for (a) the two age groups and (b) men and women; the two large bold /æ/ and /ɛ/ speakers are the speakers selected as "shift leaders" in our analyses.

The R package "lme4" version 1.1.12 (Bates, Mächler, Bolker, & Walker, 2015) was used for the mixed-model analysis reported herein. The package "car" version 2.0.26 (Fox & Weisberg, 2011) was used for plotting vowel confidence ellipses. The package "dplyr" version 0.5.0 was used for data manipulation and processing (Wickham & Francois, 2015). The analysis and data can be retrieved through an openly accessible repository: http://www.github.com/bodowinter/canadian_vowel_shift_analysis/.

We first lay out the observed patterns in purely descriptive terms. Figure 2 displays the Lobanov-normalized and rescaled vowel space, with each vowel symbol representing the mean F1/F2 of one speaker. In Figure 2a, 95% confidence ellipses are drawn separately for the older speakers (dashed lines) and the younger speakers (bold lines). In Figure 2b, the confidence ellipses are drawn separately for men (dashed lines) and women (bold lines).

As can be seen, age differences are most pronounced for /æ/ and /ɛ/. The horizontal orientation of the ellipses indicates that the apparent-time change for /æ/ is primarily one in F2. For /ɛ/, the apparent-time change is expressed both in F1 and F2, although more strongly in F2; thus, /æ/ seems to only be retracting in apparent time, while /ɛ/ is mainly retracting with a tendency toward lowering.

Note the difference in the extent of the ellipses, an indicator of variation across speakers. The F1 dimension of /æ/ exhibits much less variability for the young group, suggesting that the change has stabilized along this dimension: younger speakers do not differ much with respect to the height of /æ/, but they do differ in the extent to which they participate in its retraction. Compared to /æ/ and /ɛ/, age differences in /ʌ/ and /ɔ/ are less pronounced. This is as we would expect given that most of the movement observed in the CS involves the short front vowels.

A look at Figure 2b reveals that there are by-gender differences that resemble the by-age differences. In particular, the solid ellipses (women) are more retracted for both /æ/ and /ɛ/. This finding is in line with general principles of sociolinguistic variation, which state that young females lead “change from below” (Eckert, 1989; Labov, 1990). Moreover, the stronger retraction for female speakers is consistent with the previous literature on the CS (Boberg, 2010; De Decker, 2010; Hoffman, 2010).

To assess vowel differences statistically, we first tested for an interaction between Age Group (two levels: young vs. old) and Vowel Type (four levels: ɛ, æ, ʌ, ɔ) using separate linear mixed-effects models, one with F1 as dependent measure, and another with F2 as dependent measure.[3] An interaction between the Age Group and Vowel Type factors would indicate that different vowels were produced differently by different age groups, in other words, some vowels were affected more by this sociolinguistic category than others. Indeed, for both F1 and F2, we found a statistically reliable Age Group * Vowel Type interaction (see Tables 3 and 4 in Appendix B for detailed statistical results). There were also Gender * Vowel Type interactions for both F1 and F2, indicating that men and women had different productions for some vowels, but not for others. There were no three-way interactions (Gender * Age * Vowel Type) for F1 or F2, nor were there any two-way interactions for Gender and Age. The absence of any Gender * Age interaction is theoretically meaningful with respect to the idea of “young females” as leaders of sound change: these results suggest that being female involves leading the change, as does being young, but these two factors are independent, that is, the influences of gender and age are additive rather than multiplicative.

We additionally performed individual tests for each vowel. Crucially, this analysis revealed that for /æ/, there was a difference between young and old speakers with respect to F2, but not with respect to F1. Thus, /æ/ was more retracted for younger speakers, but there were no age differences in vowel height. On the other hand, /ɛ/ exhibited statistically reliable differences between young and old speakers for both F1 and F2, indicating that it was both lowered and retracted among younger speakers. As expected, there were no age differences for /ɔ/ in F1 or in F2 (all p-values $> .05$); however, there was an effect of Age Group for the F2 of /ʌ/, with younger speakers having relatively more fronted /ʌ/.

Compared to the age differences, the gender differences in vowel production were found to be much smaller within the present dataset. There was a statistically reliable effect of Gender for the F2 of /æ/, but not for the F1 of /æ/. That is, women retracted /æ/, but there was no statistical support for them also lowering /æ/. For /ɛ/, the picture was reversed: there was a Gender effect for F1, but not for F2. We thus find that women were more likely to lower /ɛ/, but we do not find statistical evidence that they retracted /ɛ/ more than men did. Women also exhibited reliably more fronted realizations of /ʌ/.

Again, there was no indication of any Age * Gender two-way interactions for the analyses of individual vowels (all $p > .05$). This is another piece of evidence in support of the notion that age-related differences and gender-related differences are independent of each other.

The CS has been described as a pull shift, specifically a lowering and retraction of /æ/ followed by /ɛ/. However, so far, researchers have analyzed the connection between these two vowels only impressionistically, noting that where one is lowered/retracted, the other one is too. Here, we provide more stringent quantitative evidence for a connection between /æ/ movement and /ɛ/ movement. The link between these two vowels can be quantified by measuring the degree to which the position of an individual speaker's /æ/ corresponds to the speaker's position of /ɛ/. We selected two speakers with the most retracted/lowered mean realizations, which happened to be a male speaker for /æ/ and a female speaker for /ɛ/. We then calculated the Euclidian distance to the F1/F2 position of these shift leaders for each speaker and each vowel separately. The resulting measure characterized how advanced each speaker was with respect to the two leaders selected. Crucially, across speakers, Euclidian distance from the shift leader in /æ/ was found to be correlated with Euclidian distance in /ɛ/ [correlation test, $t(26) = 4.31$, $p = .0002$], as shown in Figure 3. The correlation between the two distances was quite high ($r = .65$). Figure 3 furthermore shows older speakers to be considerably further away from the /æ/ and /ɛ/ shift leaders (who are from the young group anyway), reflecting the age effect already reported. Finally, this visualization also suggests that the relationship between /æ/ and /ɛ/ was not qualitatively different for young and old speakers; in both groups the positions of the two vowels were associated with each other in a continuous fashion. A formal test of this idea in a linear regression (/ɛ/ distance as a function of /æ/ distance * Age Group) revealed no significant interaction [$F(3, 20) = .78$, $p = .52$], suggesting that the distance slopes were similar for both groups.

Figure 3 only shows the distance to the shift leaders (shown in Figure 2) and cannot be interpreted to show that the distances between the /æ/ and /ɛ/ means themselves differed by group. Examining the average Euclidian distance between /æ/ and /ɛ/ for each speaker revealed no main effects of Age [$F(1, 24) = 3.41$, $p = .078$] or Gender [$F(1, 24) = 3.07$, $p = .09$], and no interaction effect

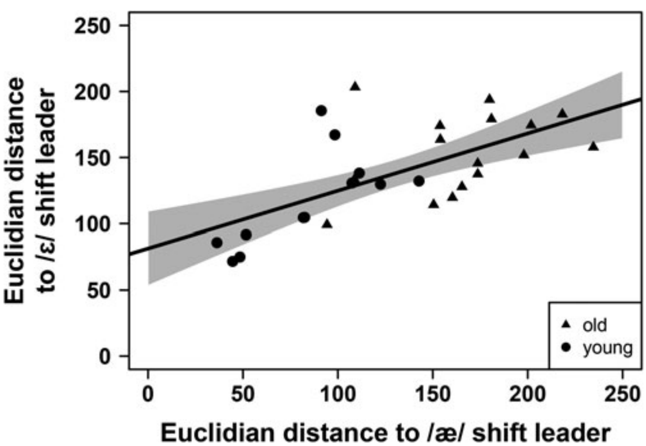

FIGURE 3. Correlation between /æ/ and /ɛ/ "shift-leadingness," with each data point representing a single speaker: speakers who were closer to the shift leader in /æ/ were also closer to the shift leader in /ɛ/.

[$F(1, 24) = 2.16$, $p = .15$]. Descriptive analysis of numerical trends revealed slightly higher /æ/ to /ɛ/ category distance for young speakers than for old speakers (174 Hz vs. 148 Hz) and also slightly higher distance for female speakers as opposed to male speakers (177 Hz vs. 149 Hz). These results are consistent with the idea that the diachronic progression of the CS involves dispersal, making /æ/ and /ɛ/ slightly more distinct; however, the absence of strong statistical support for increased category distance indicates that the pull chain movement of /ɛ/ following /æ/ is a stronger pattern than the concomitant dispersal.

Although these findings do not in and of themselves support a chain shift proposal (i.e., /æ/ and /ɛ/ could be moved along the trajectory because the same pressure to optimize perceptual dispersal acts on them simultaneously), they do at least show that the positions of the two vowels are associated within speakers, suggesting that *some* systemic pressure (chain shift, dispersal, etc.) is active, in other words, the movements of the two vowels are connected in apparent time.

Discussion

Figure 4 summarizes the movements observed by Boberg (2005) in Montreal and the present study, with a side-by-side comparison of the birth years of each study's subject groups. Boberg (2005) found lowering of /æ/ between his oldest and middle age groups followed by retraction in the youngest speakers, that is, two distinct movements, first downward and then back. In contrast, here we find /æ/ to be retracting without lowering. For /ɛ/, Boberg (2005) found retraction with marginal but nonsignificant lowering between his middle and youngest groups (represented by a dashed arrow in Figure 4), similar to our finding of more robust movement along the F2 dimension; however, we also find a smaller but statistically reliable amount of lowering of /ɛ/ along the F1 dimension. Thus, we find /ɛ/ to be both lowering and retracting, with retraction as the more dominant pattern.

Because the oldest speaker in our young age group was born after Boberg's (2005) youngest speaker, we interpret these differences as supporting his assertion that "/æ/ began to move lower among baby-boomers but reached the

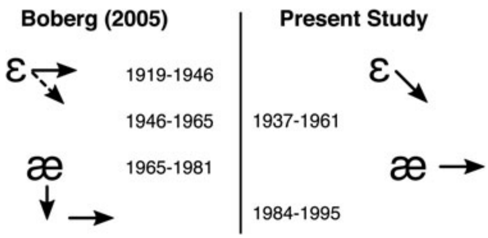

FIGURE 4. Comparison of results from Boberg (2005) and our production experiment. Birth years of the groups in each study are listed: Boberg's (2005) middle group is approximately equivalent to our older group, while our younger group represents a generation not covered by the previous study.

maximum extent of this shift (the bottom of the vowel space) by the mid-1960s" (144). In contrast, there was still room along the front-back dimension, allowing for further retraction. The more recently initiated movement of /ɛ/, on the other hand, exhibits the same pattern as was observed for Boberg's (2005) middle and young groups, with /ɛ/ showing significant retraction and (now statistically reliable) lowering. This temporal hypothesis corresponds with our finding of lower variability in /æ/ than in /ɛ/ for our young group, suggesting that the movement of /ɛ/ is still in full force even as /æ/ has fully stabilized in F1 and has slowed down its movement in F2. The finding that the younger group and women pronounce comparatively fronted /ʌ/ is not unprecedented among CS studies; we will return later to a discussion of the consequences of this possible increase in overlap between /ɛ/ and /ʌ/ in the vowel space.

PERCEPTION STUDY

Perception studies in the context of vowel shifts

In an early study of intercommunity variation in speech perception, Willis (1972) analyzed differences in vowel categorization between Fort Erie, Ontario, and Buffalo, New York. Western New York preserved the /a/–/ɔ/ distinction, while the low back merger had been well-reported across Canada by the time of Willis's study. Though neither Buffalo's Northern Cities Shift nor Ontario's CS had been noted in the sociolinguistic literature at the time of his investigation, Willis (1972:249) observed upstate New Yorkers' "peculiar pronunciation … variously described as fronting, lengthening, and diphthongization of /æ/." He focused on how each community's /ɛ/–/æ/ (*bet* vs. *bat*) and /æ/–/ɔ/ (*hat* vs. *hot*) spoken distinctions were reflected in their perceptual categorizations of vowels. His major finding was that the two groups of respondents indeed tended to divide their vowels differently depending on the pronunciations most often used within their communities.

De Decker (2010) tested vowel assignments of /æ/ and /ɔ/ among Ontario English speakers of various ages. He resynthesized the vowel in *sack* to create 19 different stimuli, with F1 held constant around 1000 Hz and F2 ranging from 2006 Hz to 1259 Hz. The result was a single continuum of sounds from *sack* to *sock*, which participants categorized as *sack, sock*, or "could be either." He found an effect of gender among the young and middle-aged listeners, with women more likely to accept more retracted stimuli (lower F2) as instances of /æ/.

In a handful of cases, perceptual investigations have been paired with analyses of speech production. Janson (1983, 1986) explored a vowel shift in Stockholm Swedish. He elicited participants' spoken /o:/ and /a:/ and administered a perceptual experiment involving a forced choice between the two vowels along a single 20-step continuum. He found that the differences in vowel production between young and old speakers were considerably larger than the differences in perceptual categorization, suggesting that a shift in perception may lag behind changes in production.

Kendall and Fridland (2010) investigated how individual variation in the production of /ɛ/ and /eɪ/ in the Southern Vowel Shift in Memphis, Tennessee, affects categorization of the two vowels along a single seven-step continuum. They found that listeners with greater degrees of /eɪ/ centralization in their own speech classified more central stimuli as /eɪ/ than nonshifted listeners did (see also Fridland & Kendall, 2012; Kendall & Fridland, 2012). This is one of the clearest demonstrations of an association between production and perception in ongoing vowel shifts; it supports the view, voiced by Roeder (2010:179), that "individuals have a more difficult time understanding pronunciations that they themselves do not use, even if those pronunciations reflect standard local norms."

In our perception experiment, we explore whether a similar coupling can be observed for the CS as well. Specifically, if a vowel shift in production goes along with a shift in perception, we would expect young female listeners to be most likely to accept highly retracted vowels as belonging to the /æ/ and /ɛ/ categories. If, however, as in Janson's (1983, 1986) analysis of Swedish vowels, perception lags behind production, we should see less pronounced differences in our perception study than in the production study. Specifically, following Janson (1983:31), it is plausible that younger speakers "still must classify the older generations' sounds correctly—something they learned when they were small children. Thus perception cannot shift too radically away from the parents' pattern."

Procedure

The listeners were the same 28 subjects that participated in the production study. The perception task was conducted immediately after the production task.

We created stimulus vowels covering a two-dimensional continuum of F1 and F2 values. Participants had the four short/lax monophthongs /ɛ/, /æ/, /ʌ/, and /ɔ/ as simultaneous response options (4-AFC task). These four categories were represented by the words *bat, bet, but*, and *bought* appearing on the screen as four large, labeled buttons of equal size. Each of the stimuli was presented once to all subjects in a single, randomized order.

We synthesized the vowel stimuli in Praat (Boersma & Weenink, 2013). Stimuli all had a falling fundamental frequency (F0) contour from 150 Hz to 100 Hz, making them sound like a human male voice. The average duration of stressed short vowels ranges from about 75 msec to 250 msec in North American English (Escudero & Polka, 2003; Wang & van Heuven, 2006). We therefore synthesized vowels with a duration of 250 msec; piloting the task revealed that this duration made stimuli still sound like short vowels while at the same time making them easy to perceive.

The stimuli ranged in steps of 50 Hz along an F1 continuum from 700 Hz to 950 Hz (6 steps) and along an F2 continuum from 1200 Hz to 1950 Hz (16 steps), yielding a total of 96 stimuli (6 F1 values × 16 F2 values). When comparing the F1 × F2 space spanned by these continua to measurements of Montreal English provided by Boberg (2005), the space spans the entirety of the distribution of /æ/ and /ʌ/, as well as most of /ɔ/ and the lower/backer section of /ɛ/.

Each trial started with a 250-msec masking tone of 150 Hz, followed by 250 msec of silence and then one 250-msec vowel stimulus. The experimental procedure was written in JavaScript and run on Firefox 3.0. The experimental session started with six practice stimuli.

Results

Overall, participants categorized most stimuli as /æ/ (~38%), followed by /ʌ/ (~22%), /ɛ/ (~21%), and /ɔ/ (~18%). These overall response proportions did not differ starkly between the two age groups for /æ/ (old: 38%, young: 39%) or for /ɔ/ (old: 25%, young: 19%). However, listeners from the younger group overall indicated hearing more /ɛ/ (24%) than listeners from the older group (19%) did; listeners from the older group overall categorized more tokens as /ʌ/ (25%) than listeners from the younger group (19%) did.

Figure 5 shows the most frequent categorization for each cell of the F1 × F2 space sampled by the synthetic vowel stimuli. First of all, it should be noted that there were many similarities among the four different groups. A noteworthy difference that appears to be somewhat systematic is that young female listeners were more likely to categorize vowels with low F2 (retracted) as /æ/. Moreover, for both men and women from the younger group, relatively lower F2 values were still accepted as /ɛ/. Finally, all listeners seemed to have difficulty localizing /ʌ/ at a consistent location along either the F1 or F2 dimension. Young female listeners in particular did not characterize many tokens as /ʌ/ at all.

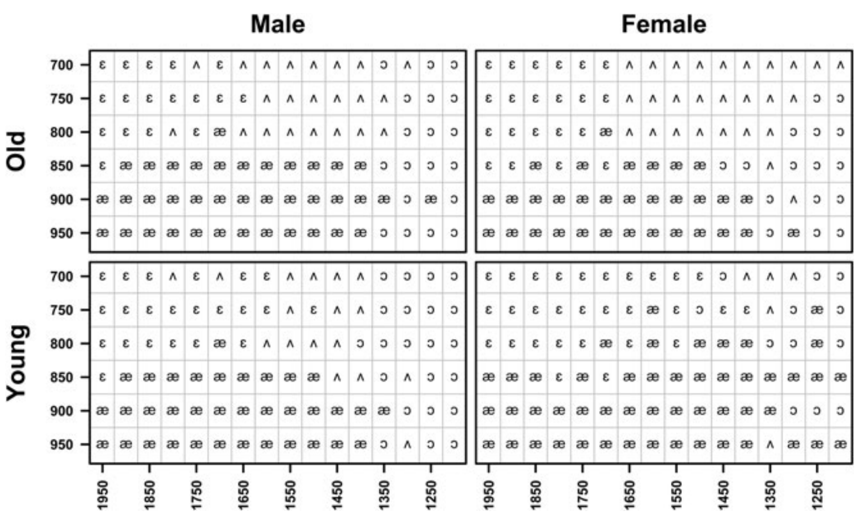

FIGURE 5. Most frequent categorization based on F1 × F2 cell of the synthetic vowel continuum, split up by male/female and young/old responses; *x*-axis represents F2; *y*-axis represents F1.

To assess categorization statistically, we partitioned the F1 continuum and the F2 continuum for each of the four vowels /ɛ/, /æ/, /ʌ/, and /ɔ/ and each listener separately, using the binary partitioning algorithm from the R package "party" version 1.0.25 (Hothorn, Hornik, & Zeileis, 2006). This algorithm tries to find the binary split point that results in the purest division between one category and another. For example, listener F1937 was estimated to have an F2 threshold for /ɛ/ at 1650 Hz: most vowels with F2 values above it were categorized as /ɛ/, and most vowels with F2 values below it were categorized as something else, so taking 1650 Hz as the threshold creates the clearest division between "/ɛ/" and "non-/ɛ/" responses. For /ɛ/, a meaningful split could be estimated for all listeners. For /æ/, 6 listeners did not have a clear threshold, for /ɔ/ this number was 4, and for /ʌ/ it was 13 (of 28 participants). The fact that it was impossible for the binary partitioning algorithm to determine /ʌ/ thresholds for a large proportion of our listeners suggests that this vowel in particular was difficult to localize in the perceptual space that we presented to them.

These listener thresholds were then analyzed with analyses of variance with the factors Age Group and Gender. Interestingly, there were no statistically reliable effects for either one of these factors for any of the vowels, regardless of whether F1 or F2 thresholds were analyzed (all $F < 3$). For /ʌ/ thresholds along the F2 dimension, there were age [$F(1,12) = 4.03$, $p = .067$] and gender differences [$F(1,12) = 3.6$, $p = .081$] that were almost statistically reliable. Younger listeners had F2 categorization thresholds on average 131 Hz lower than those of older listeners; female listeners had categorization thresholds on average 103 Hz lower than those of male listeners. This suggests that for young listeners and female listeners, the threshold between /ɛ/ and /ʌ/ is perceptually retracting, with the space categorized by /ʌ/ shrinking rather than expanding at the expense of /ɔ/. For young women, /ʌ/ categorization is disappearing altogether, while /æ/ is taking over the parts of the vowel space that characterize /ɔ/ for other listeners.

As noted in the results section of the production study, we quantified each speaker's "shift-leadingness" by measuring the Euclidian distance to two particularly retracted/lowered speakers (see Figure 2). To test whether this speech production-based measure predicted a speaker's categorization thresholds in perception, we performed correlations of /ɛ/ and /æ/ Euclidian distances with all categorization thresholds. There were no statistically reliable effects (all p-values $>$ .05); that is, we do not find evidence that a speaker's own production is related to their thresholds in perception.

Discussion

Our results show some tendencies for age and gender groups to pattern differently in perception. However, the picture is dominated by similarities between the groups, at least when compared to the large differences observed in the production study. It is possible that our failure to find reliable age, gender, or production effects in this perception experiment constitutes a type II error. For example, it could be that our design (or our analysis approach) did not have enough statistical power to investigate perceptual differences that were actually present in the population under study. However, we think that the present result can meaningfully be interpreted as a null result. First, we used the same number of speakers in both studies, and second, several ways of analyzing the perception data do not yield statistically reliable systematic differences between age groups or genders.[4] Thus, we are left to conclude that the differences in perception are at least smaller than the differences in production, consistent with Janson's (1983, 1986) statement that perception lags behind production in ongoing vowel shifts.

At the same time, younger women's acceptance of more retracted stimuli as /ɛ/ and /æ/ (Figure 4) demonstrates a tendency toward perception reflecting production patterns in this advanced group. Taken together with our finding that shiftedness in production does not correlate with shifted perceptual thresholds at the individual level, this raises the possibility that an individual's perceptual environment impacts their vowel categorization ability more than their own vowel production.[5] That is, perception is more strongly based on the surrounding sociolinguistic environment as a whole, which in most cases will include a fair degree of tokens from both young (shifted) and old (nonshifted) speakers. Although phonetic and psycholinguistic experiments have established links between perception and production (e.g., Jones & Munhall, 2000; Nielsen, 2011), the vowel categorization processes investigated in this experiment may be comparatively less tied to how a speaker pronounces their own vowels.

OVERALL DISCUSSION

Our production study demonstrates that /æ/ and /ɛ/ are shifting in apparent time in the vowel spaces of English-speaking Jewish Montrealers. An analysis of interspeaker variation shows ordered heterogeneity, with young women leading the change and older men retaining the most conservative pronunciations, the typical progression for a sound change advancing in a community below the level of consciousness (Eckert, 1989; Labov, 1990). It seems as though the operation of the CS in Montreal now involves the retraction of /æ/ without any accompanying lowering, whereas /ɛ/ is backing and slightly lowering in the vowel space.

As noted in the discussion of the production study, we interpret this to mean that the change has progressed in real time, as our participants represent an overall younger group than those in Boberg's (2005) study. Now, our apparent-time data suggests that there is no further lowering of /æ/ for younger speakers, but there is still variation along the front-back axis. Over time, /æ/ has "bottomed out," lowering as far in the vowel space as possible. This is consistent with Boberg's (2005) finding of /æ/ lowering between his oldest and middle groups and then retracting in his youngest group.

The /ɛ/ vowel, on the other hand, appears to still be backing with a small but now statistically reliable amount of lowering. The correlation between individuals' participation in /æ/ and /ɛ/ shifting (Figure 3) could be taken as evidence for /æ/ movement “pulling” /ɛ/, in line with a chain shift interpretation. However, other interpretations of the /ɛ/ ~ /æ/ correlation are possible. For example, the shift could be a form of phonetic analogy, in which movement in /æ/ along the F2 dimension is mirrored by movement in /ɛ/ because phonetic dimensions are biased toward similar expression across different phonemes (see Wedel, 2006; Winter & Wedel, 2016). Alternatively, the link could perhaps be explained as a consequence of a move toward perceptual dispersal (see Vaux & Samuels, 2015).

We favor the “pull chain” interpretation for several reasons. First, our findings suggest patterned variation along both F1 and F2 (not just a single dimension), and the apparent-time changes suggest that /æ/ has led the shift, followed by /ɛ/. In fact, /æ/ has stabilized, but /ɛ/ is still moving back and now shows stronger, not weaker, evidence of lowering. Parallel changes driven by phonetic analogy would be less likely to produce this sort of sequentiality. Second, an account that is merely based on phonetic dispersal is inconsistent with the observation that the distance between the /æ/ and /ɛ/ categories has not increased in a statistically reliable fashion, as well as the fact that /ɛ/ and /ʌ/ are moving *toward* each other, *decreasing* dispersion.[6] Third, phonologically neutral “phonetic drift” (Gardner, Roeder, & Childs, 2016) does not explain the systematicity of the /æ/ and /ɛ/ movements, with the two being linked and with /æ/ leading /ɛ/. Thus, we find that a pull chain interpretation accounts best for the movement of the front monophthongs.

Though we do not necessarily believe it reflects an apparent-time change of the same importance, we find fronting of /ʌ/ in our younger and female groups. The original formulation of the CS included the assertion that /ʌ/ is either centralizing or lowering (see Figure 1) (Clarke et al., 1995), and Eckert (2008) claimed that /ʌ/ is fronting in the structurally similar California Shift. Durian (2012) reported lowering/retraction of /æ/, /ɛ/, and /ɪ/ and separately noted ongoing /ʌ/ fronting for speakers in Columbus, Ohio. While, according to statistics presented by Boberg (2005:137), Labov et al.’s (2006) sample of 10 Ontario speakers may exhibit some degree of fronting, Boberg (2005) did not find statistically reliable /ʌ/ movement in his own Montreal sample. Other investigations since then have almost entirely disregarded /ʌ/ (Hoffman, 2010; Roeder & Jarmasz, 2009, 2010; Sadlier-Brown & Tamminga, 2008).

Traditional accounts of English phonology consider /ʌ/ as part of the same short vowel subsystem as /ɛ/ and /æ/; this makes the fronting of /ʌ/ puzzling, as it accelerates its collision course with /ɛ/ rather than moving in another direction to maintain perceptual dispersal. Theoretically, /ʌ/ could still maintain its status as an unmerged phoneme based on other phonetic cues, which would in turn suggest its membership in a subsystem structurally distinct from /ɛ/. Langstrof (2009), for instance, found evidence of duration being used as a primary cue differentiating /ɪ/ from /ɛ/ in archival recordings of New Zealand English, indicating one possible way for two vowels implicated in a shift to remain distinct at a stage when their F1/F2 ranges overlap.

As for categorization differences, our perception study finds some statistically weak indications of young females accepting more retracted vowels as /ɛ/ as opposed to /ʌ/, as well as a trend toward young females accepting more retracted stimuli as /æ/. The fact that the threshold difference for /ʌ/ is in the direction of the movement of retracting /ɛ/ rather than fronting /ʌ/ confirms our view that /ʌ/ movement is neither as perceptually salient nor as dramatically pronounced as the rest of the CS.

Overall, however, the picture in the perception study is dominated by similarities between the age and gender groups, mirroring Janson's (1983, 1986) findings of an ongoing sound change in Sweden. In particular, we suggest that while speakers participate in the shift in production, as listeners they must accommodate the fact that they are continuously exposed to both innovative and conservative variants in perception. We find no within-participant correlation between shiftedness in production and shiftedness in perception, but we do find trends in the direction of the shift at the group level. This indicates that an individual's perceptual performance may not simply be a reflection of their own production; instead, other factors such as their perceptual and sociolinguistic environment mediate this link.

A caveat with regard to the results of the perception experiment is that the vowel stimuli were played in isolation, with no surrounding consonantal context. Other successful perception studies have nested their vowel stimuli between consonants (De Decker, 2010; Rakerd & Plichta, 2010). Though several studies (Fox, 1989; Rakerd, 1984; Strange, Edman, & Jenkins, 1979; Strange, Verbrugge, Shankweiler, & Edman, 1976) indicated that a consonant-vowel-consonant (CVC) stimulus improves accuracy in vowel categorization tests, Macchi (1980:1641) "failed to provide evidence that vowels spoken in consonantal context are better identified than naturally produced isolated vowels." This finding is in line with the results of Strange, Jenkins, and Johnson (1983), who found error rates in phoneme mapping with isolated short vowel stimuli based on modified natural speech to be relatively low. Diehl, McCusker and Chapman (1980) found a slight advantage in selecting written CVC syllables if modified natural stimuli were reinserted between consonants, but they found no identification advantage using synthesized stimuli. Though we acknowledge that a CVC stimulus could have been a methodological improvement, we do not believe that the lack of consonantal context is a big concern for the present perception study.

Another potential shortcoming of the perception experiment is that the speaking voice was entirely decontextualized. The stimuli themselves were all produced by a single synthesizer script, so they were controlled for any non-F1/F2 phonetic features, such as the values of F0, F3, and breathiness, which have been shown to carry social and linguistic information marking gender and age (Johnson, Strand, & D'Imperio, 1999). The F0 contour of 150 Hz to 100 Hz made the synthetic voice seem relatively more male than female, but the stimuli were otherwise unmarked for any dialect or age group. Knowledge about a speaker's gender (Johnson et al., 1999; Strand, 1999), age (Drager, 2010), and dialect

(Hay, Nolan, & Drager, 2006; Niedzielski, 1999) has been found to influence speech perception behavior; perhaps hearing a “male” voice could have caused listeners to ascribe certain sociolinguistic features to the presented stimuli. For example, since our production study showed male speakers to be less advanced in the CS than female speakers are, listeners might have attributed conservative vowel positions to the speaking voice. Finally, the use of synthesized stimuli (though convincingly human-sounding) may have suggested a more formal variety of English, which could lead to a reduction of shift phenomena that are presumably associated more strongly with informal speech styles.

CONCLUSION

The evidence presented here contributes to the development of theories concerning the phonological and phonetic underpinnings of the CS. The results of our production task, when compared with Boberg’s (2005) apparent-time results for Anglophone Montrealers, indicate that changes seem to have occurred in the trajectories of /æ/ and /ɛ/: /æ/ has “bottomed out” in the F1 dimension and is now only retracting, while /ɛ/ exhibits both retraction and lowering.

Two results from our production study are worth highlighting. First, we find /æ/ and /ɛ/ positions to be linked within individuals. This, together with the fact that /æ/ appeared to lead the CS and is now stabilizing, suggests to us that the CS is most adequately characterized as a pull shift triggered by the merger of the low back vowels. Second, we found some degree of /ʌ/ fronting, although this was much less pronounced than any changes among the front vowels. This corroborates various reports of /ʌ/ fronting in other varieties of Canadian and American English.

The results of our paired-study methodology, which incorporated both production and perception, augment the literature on how ongoing vowel shifts are expressed in speaking and listening behavior, and how these two levels are related to each other. While the speakers in our sample varied in production as to their degree of participation in the shift, their vowel categorizations did not differ as much along either F1 or F2. Shift leaders were not observed to significantly shift in perception, despite younger speakers and women showing a weak trend toward accepting more retracted stimuli as /æ/ and /ɛ/. By collecting production and perception data from the same participants, we have demonstrated that while one might expect to find similar community-wide variability in vowel categorization as in vowel pronunciation, the need to accommodate to both innovative and conservative variants exerts a strong homogenizing effect in the perceptual domain.

NOTES

1. Following other studies of Canadian English, we use /ɔ/ to denote this merged LOT/THOUGHT/PALM class, though its phonetic reflexes may be more like [ɒ].

2. It should also be noted that unlike in other Canadian English varieties, speakers of Montreal English do not typically raise /æ/ before nasal or velar consonants (Boberg, 2010). It is therefore unnecessary to exclude any allophones of /æ/ from consideration in the CS.

3. We fitted random effects for speakers and items (Baayen, Davidson, & Bates, 2008). Following the guidelines of Barr, Levy, Scheepers, and Tily (2013), we fitted random slopes for the critical within-subjects or within-items effects in question. We did not fit random slopes for control variables that were not of primary interest in our study. In lme4 syntax, the model formula used was:

F1/F2 ~ Age Group + Gender + Vowel Type + LogFrequency + Age Group:Vowel Type + Consonant Voicing + MonoVsDisyllable + Consonant Manner of Articulation + Consonant Place of Articulation + (1 + Vowel|Speaker) + (1|Word) + (0 + Age|Word) + (0 + Gender|Word)

Models were fitted with maximum likelihood and *p*-values were generated using likelihood ratio tests. Visual inspection of Q-Q plots and plots of residuals against fitted values did not reveal any obvious deviations from normality and homoscedasticity. All continuous variables were centered and all categorical variables were deviation coded (Schielzeth, 2010).
4. In initial analyses, we calculated the response thresholds in different ways (using logistic regression fits, or using a simple 50% or 80% cut-off rule). Regardless of how categorization thresholds were calculated, we did not obtain any statistically significant age or gender differences. Moreover, an analysis of categorization behavior using logistic Generalized Additive Modeling with tensor product splines for F1 and F2 as predictors, that is, *te*(F1, F2, Age), also did not yield systematic age or gender differences.
5. We thank an anonymous reviewer for this insight.
6. We have to acknowledge, though, that dispersion ultimately is about the whole vowel system (cf. de Boer, 2001) and thus, to truly assess whether the system has dispersed or not, all vowels would have to be taken into account.

REFERENCES

Adank, Patti, Smits, Roel, & van Hout, Roeland. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America* 116(5):3099–3107.

Baayen, R. Harald, Davidson, Donald J., & Bates, Douglas M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4):390–412.

Barr, Dale J., Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3):255–278.

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Bigham, Douglas S. (2010). Correlation of the low-back vowel merger and TRAP-retraction. *University of Pennsylvania Working Papers in Linguistics* 15(2):4–31.

Boersma, Paul, & Weenink, David. (2013). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.37. Available at: http://www.praat.org/. Accessed February 8, 2016.

Boberg, Charles. (2004). Ethnic patterns in the phonetics of Montreal English. *Journal of Sociolinguistics* 8(4):538–568.

______. (2005). The Canadian Shift in Montreal. *Language Variation and Change* 17(2):133–154.

______. (2010). *The English Language in Canada*. Cambridge: Cambridge University Press.

Clarke, Sandra, Elms, Ford, & Youssef, Amani. (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change* 7(2):209–228.

de Boer, Bart. (2001). *The origins of vowel systems*. Oxford: Oxford University Press.

De Decker, Paul. (2010). Sounds shifty: Gender and age differences in perceptual categorization during a phonetic change in progress. *University of Pennsylvania Working Papers in Linguistics* 15(2):article 7. Available at: http://repository.upenn.edu/cgi/viewcontent.cgi?article=1129&context=pwpl. Accessed February 8, 2016.

De Decker, Paul, & Mackenzie, Sara. (2000). Slept through the ice: A further look at lax vowel lowering in Canadian English. *Toronto Working Papers in Linguistics* 18. Available at: http://www.ensani.ir/storage/Files/20101125165203-167.pdf. Accessed February 8, 2016.

Diehl, Randy L., McCusker, Susan Buchwald, & Chapman, Laura S. (1980). Perceiving vowels in isolation and in consonantal context. *Journal of the Acoustical Society of America* 69(1):239–248.

Drager, Katie. (2010). Sociophonetic variation in speech perception. *Language and Linguistics Compass* 4(7):473–480.

Drager, Katie, Kirtley, M. Joelle, Grama, James, & Simpson, Sean. (2013). Language variation and change in Hawai‘i English: KIT, DRESS, and TRAP. *University of Pennsylvania Working Papers in Linguistics* 19(2):41–50.

Durian, David. (2012). *A new perspective on vowel variation across the 19th and 20th centuries in Columbus, OH*. Ph.D. dissertation, The Ohio State University.

Eckert, Penelope. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1(3):245–267.

_______. (2008). Where do ethnolects stop? *International Journal of Bilingualism* 12(1–2):25–42.

Escudero, Paola, & Polka, Linda. (2003). A cross-language study of vowel categorization and vowel acoustics. In M.-J. Sole, D. Recansens, & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 861–864.

Fox, John, & Weisberg, Sanford. (2011). *An {R} companion to applied regression*. 2nd ed. Thousand Oaks: Sage.

Fox, Robert A. (1989). Dynamic information in the identification and discrimination of vowels. *Phonetica* 46(1–3):97–116.

Fridland, Valerie, & Kendall, Tyler. (2012). Exploring the relationship between production and perception in the mid front vowels of U.S. English. *Lingua* 122(7):779–793.

Gardner, Matthew Hunt, Roeder, Rebecca, & Childs, Rebecca. (2016). Social-moderation of a structural sound change? The Canadian Shift in four communities. Paper presented to the 2016 Meeting of the American Dialect Society, Washington, DC, January 7–10.

Gorman, Kyle, Howell, Jonathan, & Wagner, Michael. (2011). Prosodylab-Aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39(3):192–193.

Harrington, Jonathan, & Cassidy, Stephen. (1994). Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English. *Language and Speech* 37 (4):357–373.

Hay, Jennifer, Nolan, Aaron, & Drager, Katie. (2006). From fush to feesh: Exemplar priming in speech perception. *Linguistic Review* 23(3):351–379.

Hoffman, Michol. (2010). The role of social factors in the Canadian Vowel Shift: Evidence from Toronto. *American Speech* 85(2):121–140.

Hothorn, Torsten, Hornik, Kurt, & Zeileis, Achim. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3):651–674.

Janson, Tore. (1983). Sound change in perception and production. *Language* 59(1):18–34.

_______. (1986). Sound change in perception: An experiment. In J.J. Ohala & J.J. Jaeger (eds.), *Experimental phonology*. Orlando: Academic Press. 253–260.

Johnson, Keith, Strand, Elizabeth A., & D'Imperio, Mariapaola. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27(4):359–384.

Jones, Jeffery A., & Munhall, Kevin G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America* 109:1246–1251.

Kendall, Tyler, & Fridland, Valerie. (2010). Mapping production and perception in regional vowel shifts. *University of Pennsylvania Working Papers in Linguistics* 16(2):103–112.

_______. (2012). Variation in perception and production of mid front vowels in the U.S. Southern Vowel Shift. *Journal of Phonetics* 40(20):289–306.

Kennedy, Robert, & Grama, James. (2012). Chain shifting and centralization in California vowels: An acoustic analysis. *American Speech* 87(1):39–56.

Kettig, Thomas. (forthcoming). Now say ‘ah’: Internal factors of shifting and the English low vowel space. In E. Louviot & C. Delesse (eds.), *Studies in language variation and change 2: Shifting, switching and alternating patterns in the history of English*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Labov, William. (1990). The interaction of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2):205–254.

_______. (1991). The three dialects of English. In P. Eckert (ed.), *New ways of analyzing sound change*. New York: Academic Press. 1–44.

Labov, William, Ash, Sharon, & Boberg, Charles. (2006). *Atlas of North American English: Phonology and phonetics*. Berlin: Mouton de Gruyter.

Langstrof, Christian. (2009). On the role of duration in the New Zealand English front vowel shift. *Language Variation and Change* 21:437–453.

Lennes, Mietta. (2003). Collect_formant_data_from_files.praat [Praat script]. Available at: http://www.helsinki.fi/~lennes/praat-scripts/public/collect_formant_data_from_files.praat. Accessed February 8, 2016.

Liljencrants, Johan, & Lindblom, Björn. (1972). Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language* 48(4):839–862.

Lobanov, Boris M. (1971). Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America* 49(2B):606–608.

Macchi, Marian J. (1980). Identification of vowels spoken in isolation versus vowels spoken in consonantal context. *Journal of the Acoustical Society of America* 68(6):1636–1642.

Martinet, André. (1955). *Economie des changements phonétiques: Traité de phonologie diachronique*. Berne: Francke.

Niedzielski, Nancy. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1):62–85.

Nielsen, Kuniko. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39 (2):132–142.

R Core Team. (2015). R*: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/. Accessed February 6, 2016.

Rakerd, Brad. (1984). Vowels in consonantal context are perceived more linguistically than are isolated vowels: Evidence from an individual differences scaling study. *Perception and Psychophysics* 35 (2):123–136.

Rakerd, Brad, & Plichta, Bartłomiej. (2010). More on Michigan listeners' perceptions of /ɑ/-fronting. *American Speech* 85(4):431–449.

Roeder, Rebecca. (2010). Northern Cities Mexican American English: Vowel production and perception. *American Speech* 85(2):163–184.

Roeder, Rebecca, & Jarmasz, Lidia-Gabriela. (2009). The lax vowel subsystem in Canadian English revisited. *Toronto Working Papers in Linguistics* 31:1–12.

______. (2010). The Canadian Shift in Toronto. *Canadian Journal of Linguistics* 55(3):387–404.

Roeder, Rebecca, & Gardner, Matthew. (2013) The phonology of the Canadian Shift revisited: Thunder Bay & Cape Breton. *University of Pennsylvania Working Papers in Linguistics* 19(2):161–170.

Sadlier-Brown, Emily, & Tamminga, Meredith. (2008). The Canadian Shift: Coast to coast. In S. Jones (ed.), *Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association*. Available at: http://homes.chass.utoronto.ca/~cla-acl/actes2008/CLA2008_Sadlier-Brown_Tamminga.pdf. Accessed February 8, 2016.

Schielzeth, Holger. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* 1:103–113.

Statistics Canada. (2011). *Linguistic characteristics of Canadians: Language, 2011 Census of Population*. Government of Canada. Available at: http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf. Accessed February 6, 2016.

Strand, Elizabeth A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18(1):86–100.

Strange, Winifred, Edman, Thomas R., & Jenkins, James J. (1979). Acoustic and phonological factors in vowel identification. *Journal of Experimental Psychology* 5(4):643–656.

Strange, Winifred, Jenkins, James J., & Johnson, Thomas L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustic Society of America* 74(3):695–705.

Strange, Winifred, Verbrugge, Robert R., Shankweiler, Donald P., & Edman, Thomas R. (1976). Consonant environment specifies vowel identity. *Journal of the Linguistic Society of America* 60(1):213–224.

Thomas, Erik R. (2002) Sociophonetic applications of speech perception experiments. *American Speech* 77(2):115–147.

Thomas, Erik R., & Kendall, Tyler. (2007). *NORM: The vowel normalization and plotting suite*. Available at: http://ncslaap.lib.ncsu.edu/tools/norm/. Accessed February 6, 2016.

Vaux, Bert, & Samuels, Bridget. (2015) Explaining vowel systems: Dispersion theory vs natural selection. *Linguistic Review* 32(3):573–599.

Wang, Hongyan, & van Heuven, Vincent J. (2006). Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. In J.M. van de Weijer & B. Los (eds.), *Linguistics in the Netherlands*. Amsterdam: John Benjamins. 237–248.

Wedel, Andrew B. (2006). Exemplar models, evolution and language change. *Linguistic Review* 23:247–274.

Wells, John. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Wickham, Hadley, & Francois, Romain. (2015). *dplyr: A grammar of data manipulation*. R package version 0.4.2. See https://CRAN.R-project.org/package=dplyr. Accessed February 19, 2016.

Willis, Clodius. (1972). Perception of vowel phonemes in Fort Erie, Ontario, Canada, and Buffalo, New York: An application of synthetic vowel categorization tests to dialectology. *Journal of Speech, Language and Hearing Research* 15(2):246–255.

Winter, Bodo, & Wedel, Andrew B. (2016). The co-evolution of speech and the lexicon: The interaction of functional pressures, redundancy and category variation. *Topics in Cognitive Science* 8:503–513.

APPENDIX A

Participants in this study were recorded reading the following sentences. The 44 target words containing the vowels under investigation are presented in **bold**, as participants saw them. The 22 anchor vowels extracted to conduct normalization are also underlined here.

1. He bought it at the **mall**, not at the supermarket.
2. Don't make such a **fuss**, I can solve it.
3. Did you make a **bet**, or do you not gamble?
4. Are you on the **staff**, or a customer?
5. Give me a **hug**, it's been a bad day.
6. Did you sleep on a **bed**, or a cot?
7. Was the game a **loss**, or a win?
8. I want a **pen**, not a pencil.
9. I want them in a **stack**, not all messed up.
10. The door has a **knob**, not a handle.
11. Whatever he **says**, I don't believe.
12. He wants to **sell**, not buy.
13. Was your day **rough**, or easy?
14. He doesn't eat **ham**, because he keeps Kosher.
15. It's a **sock**, not a stocking.
16. Does your phone **buzz**, or ring?
17. Is the phenomenon a **cause**, or an effect?
18. That bird is a **gull**, not a crow.
19. Go past the **lab**, and then turn left.
20. If you don't have a **pass**, you can't enter.
21. He broke his **neck**, not his back.
22. When I **nod,** enter the room.
23. Clean up all that **mud**, your hands are filthy!
24. She grows **holly**, not juniper.
25. Has he **gone**, or is he still here?
26. Hold the cat by the **belly**, not the arms.
27. Are you **stuck**, or can you continue?
28. It's in the **valley**, not on the mountain.
29. It's my friend **Stef**, not Michael.
30. The pear was **soft**, not hard.
31. Was it her **mom**, or her dad?
32. The dog was a **mutt**, not a purebred.
33. Do you play **jazz**, or the blues?
34. His dinner was **bad**, so he sent it back.
35. The dog may **beg,** but I won't give him a treat.
36. He fell in a **gully,** but we got him out.
37. Hand me the **pan**, not the pot.
38. Is that from **Jess**, or from George?
39. Is that a **gem**, or a fake stone?
40. That's not a **rat**, it's a mouse!
41. The guy's name was **Sal**, not Mike.

42. Did you **run,** or walk here?
43. Are you chewing **gum**, or food?
44. He poured me a **lot**, not just a bit.

APPENDIX B: RESULTS FOR EXPERIMENT 1 (PRODUCTION)

TABLE 3. *Omnibus analysis (across all vowels) (1220 observations)*

	Fixed Effect	Test Statistic	*p*-value
F1	Age * Vowel Type	$\chi^2(3) = 11.2$	.01
	Gender * Vowel Type	$\chi^2(3) = 7.98$	.046
	Age * Gender	$\chi^2(1) = .11$	.74
	Age * Gender * Vowel Type	$\chi^2(3) = 4.06$	.26
F2	Age * Vowel Type	$\chi^2(3) = 22.1$	<.001
	Gender * Vowel Type	$\chi^2(3) = 9.16$	.027
	Age * Gender	$\chi^2(1) = 2.69$	.10
	Age * Gender * Vowel Type	$\chi^2(3) = 2.09$	.55

TABLE 4. *Analysis by individual vowel*

		Fixed Effect[a]	Test Statistic	*p*-value
æ (305 obs.)	F1	Age	$\chi^2(1) = 2.0$	.75
		Gender	$\chi^2(1) = .09$	.76
		Age * Gender	$\chi^2(1) = .71$	.40
	F2	Age	$\chi^2(1) = 25.89$	.001
		Gender	$\chi^2(1) = 6.0$	.014
		Age * Gender	$\chi^2(1) = .19$	.66
ɛ (309 obs.)	F1	Age	$\chi^2(1) = 6.99$	.008
		Gender	$\chi^2(1) = 3.97$	.046
		Age * Gender	$\chi^2(1) = 0.12$	.73
	F2	Age	$\chi^2(1) = 8.16$	.004
		Gender	$\chi^2(1) = .009$	.92
		Age * Gender	$\chi^2(1) = 0.79$	.37
ɔ (298 obs.)	F1	Age	$\chi^2(1) = 0.12$	.73
		Gender	$\chi^2(1) = 3.42$	.064
		Age * Gender	$\chi^2(1) = 2.78$	.095
	F2	Age	$\chi^2(1) = 0.55$	.46
		Gender	$\chi^2(1) = 0.34$	.56
		Age * Gender	$\chi^2(1) = 0.76$	.38
ʌ (308 obs.)	F1	Age	$\chi^2(1) = 2.73$	.098
		Gender	$\chi^2(1) = 2.29$	.13
		Age * Gender	$\chi^2(1) = 1.70$	.19
	F2	Age	$\chi^2(1) = 5.82$	.016
		Gender	$\chi^2(1) = 4.64$	.03
		Age * Gender	$\chi^2(1) = 0.42$	.52

[a]See Table 1 for gender and age composition of participant sample.