CANADIAN RAISING: CONTRAST NEUTRALIZATION IN MEAFORD, ON AND VANCOUVER, BC*

Geoff Fullerton
University of British Columbia

1. Introduction

1.1 Background

One of the more commonly cited characteristics of Canadian English is Canadian Raising. As is well documented, the diphthongs /aɪ/ and /aʊ/ have different allophones, dependent on the voicing of the following segment: specifically, when followed by a tautosyllabic voiceless consonant, the diphthongal nucleus is mid-central [ʌ], while it is low elsewhere (Joos 1942, Chambers 1973). Though Chambers (1973) named it "Canadian Raising", he noted that the name is one of convenience only, and that Canadian Raising is by no means unique to Canada, nor to Canadian English (Chambers 1973; see also Trudgill 1986). In fact, Canadian Raising does not even characterize all of Canada: some parts of the Maritime provinces and Newfoundland in particular don't exhibit it at all (Trudgill 1986). Raising has also been examined and documented by a wide variety of studies outside of Canada, for example in the United States (Vance 1987, Allen 1989, Dailey-O'Cain 1997, Niedzielski 1999, Fruehwald 2007, Hualde et al. 2017), Northern England (Milroy 1996), the English Fens (Britain 1997), and the Caribbean (Trudgill 1986), among others. What is striking, however, about many of these studies (e.g. Vance 1987, Milroy 1996, Dailey-O'Cain 1997, Hall 2005) is the inconsistency with which Canadian Raising actually occurs (or doesn't) in the relevant environment. The present paper explores and attempts to quantify this inconsistency within two specific populations: one in Meaford, ON, and the other in Vancouver, BC.

1.2 The "Rule"

In addition to preceding a voiceless consonant, the diphthong must carry greater stress than the following syllable (Chambers 1973) for raising to occur. The voiceless consonant may be ambisyllabic; that is, associated with both the stressed preceding and unstressed following syllables (Kahn 1980). In a (trochaic) foot-based account, a following voiceless consonant compels raising only if it occurs in the same foot as the diphthong (Jensen 2000, Bermúdez-Otero 2003), so a following stressed syllable cannot trigger raising. Further, as Fruehwald (2007)

* The author gratefully acknowledges the assistance and support of Kathleen Currie Hall, Murray Schellenberg, Bryan Gick, Amanda Cardoso, the Interdisciplinary Speech Research Lab (UBC), and all participants.

Actes du congrès annuel de l'Association canadienne de linguisitique 2019.
Proceedings of the 2019 annual conference of the Canadian Linguistic Association.

and others point out, raising does not usually occur at morpheme boundaries, nor does it occur across word boundaries.

Other papers have looked at a variety of other factors and potential conditions of raising as well, leading to a general consensus that it may occur in more widely distributed environments than the traditional account would predict. In a study of speakers from southern Manitoba, Onosson (2014) argues that raising is conditioned by the sonority of the following segment rather than its voicing. Milroy (1996) showed pre-[+voice] and pre-[+nasal] raising in a speaker from Saginaw, Michigan, and suggested that it was a weak following syllable conditioning raising among those speakers, rather than the voicing of the following consonant. Vance (1987) found among a small sample of speakers from the northeastern US that raising was occurring in some words before voiced consonants, most notably before /d/ and /ɹ/. Dailey-O'Cain (1997) also found raising before /ɹ/, as well as before nasal clusters in a study of speakers from Michigan – though perhaps it should be noted that Chambers (2006) argues that this may reflect a process other than raising. Finally, Hall (2005) provided evidence that raising may also be affected by the *previous* segment, as well as other factors such as lexical frequency and neighbourhood effects. Each of these studies documents raising in different geographical regions and dialects of English, and each finds slightly different patterns in the raising data. What they do have in common, however, is that the classic rule for raising does not seem to fully account for all instances of it they find.

1.3 Opacity

While Canadian Raising is considered a classic phonological rule (albeit one whose conditions have been shown to be less than regular), it also famously gives rise to opacity as a result of its interaction with the neutralization of unstressed intervocalic /t/ and /d/ to [ɾ] (Hualde et al. 2017, inter alia), as in *writer* [ɹʌɪɾəˑ] and *rider* [ɹaɪɾəˑ]. As a result, Joos (1942) proposed phonemic status for the raised and unraised variants, and others since have likewise suggested that the distribution of [aɪ] ~ [ʌɪ] is in fact contrastive, including Vance (1987) and Mielke, Armstrong, and Hume (2003), who claim that complementary distribution here is a "historical relic…the artifacts of their previously allophonic distribution still linger in the lexicon" (2003: 132). However, this account is disputed by Harris (1960), Chomsky (1964), and Idsardi (2006), among others. In Chomsky (1964)'s analysis, the processes of raising and flapping are independently motivated, and the rule which raises /aɪ/ to [ʌɪ] applies *before* the rule that changes /t/ to [ɾ], producing derivations like /ɹaɪtəˑ/ → ɹʌɪtəˑ → [ɹʌɪɾəˑ]. Certainly it is a contested and problematic issue; as Harris points out "Nowhere else in English do we have phonemes with just such a distribution, nor is it elegant to have two phonemes which are complementary through so much of their distribution" (Harris 1960: 70-71). If the phonemic/allophonic status of these variants is in fact unclear, perhaps it is not wholly surprising that so many of the studies on Canadian Raising have found the process to be so inconsistently attested.

1.5 The current study

While examining the question of opacity in Canadian Raising, Hall (2005) conducted a study in Meaford, Ontario, during which she found some unusual patterns of raising among the speakers whose productions she analyzed there. Specifically, the three male speakers over the age of 65 who were analyzed produced nearly a third of their /aɪ/ variants contrary to what would be

expected, given the traditional account of raising (see Hall 2005 for detailed results and discussion).

The current study aims first to replicate these findings, by analysis of more of Hall (2005)'s data, and to determine any generalizable patterns amongst speakers of Meaford. The second aim of this paper is to use an experimental design similar to that of Hall (2005) to examine patterns of raising in Vancouver-born speakers and to compare them with those found in Meaford. Implications for the current understanding of raising will be discussed.

It should be noted that although Canadian Raising involves both /aɪ/ and /aʊ/ diphthongs, this paper will confine its focus to the /aɪ/ diphthong. Not only did Hall (2005)'s study similarly restrict its investigation to productions of /aɪ/, but it has been argued that /aɪ/- and /aʊ/-raising are in fact two distinct processes (see e.g. Vance 1987, Chambers 1989, Dailey-O'Cain 1997, Boberg 2008).

2. Production Experiment

2.1 Participants

Hall (2005) conducted her experiment with 20 participants, three of whose recorded data she analyzed for that paper (see Hall 2005 for details). To replicate and add to her findings, read word list data from six speakers of the original 20 speakers in that experiment (ages 53-76; male, n = 3; female, n = 3) were analyzed alongside that of six new Vancouver speakers.

These new participants were adults (ages 55-75; male, n = 4; female, n = 2), born and raised in Vancouver, British Columbia. Participants were recruited such that they were of the same age range at the time of data collection as those in Hall (2005)'s participant pool were when they were recorded. Prior to the experiment, each completed a questionnaire inquiring as to birthplace, places of residence, other languages spoken, as well parents' birthplaces and languages spoken. All participants had lived in the Greater Vancouver area almost all of their adult lives (two participants reported having lived for a year or two as adults elsewhere in Canada). All participants were native speakers of Canadian English, and none reported speaking any language other than English at more than a high-school level. None reported any speech or hearing disorders. The task took about 30 minutes and participants were compensated $15 CAD.

2.2 Materials

Stimuli for the word list task in Hall (2005)'s production experiment were chosen from the CELEX corpus (Baayen et al. 1995). Hall's word list consisted of 299 /aɪ/-containing stimuli and 255 fillers. Target words were chosen to represent all combinations of factors attributed to Canadian raising: i.e. voicing of the following segment, stress, and syllabification. Within these contexts, and with the expectation of word frequency effects, stimuli were also selected as representative of high-, mid- and low-frequencies as given in the CELEX corpus (Baayen et al. 1995; see Hall 2005 for complete methodology), for 24 contexts in all.

Stimuli for the Vancouver production experiment were chosen to present a task similar to that in Hall (2005), but with more stringent controls of the relevant contexts. 96 target words and 100 fillers, selected from the IPhOD corpus (Vaden et al. 2009), constituted the word list (see Appendix for conditions and stimuli). The software *Phonological CorpusTools* (Hall et al. 2016) was used to search the IPhOD corpus for words containing /aɪ/ in the same contexts used by Hall

(2005). Of those words, three frequency bins (high-, mid-, and low-) were created based on the raw frequency data given in the IPhOD corpus. The 25% most frequent words were treated as "high frequency", the 25% least frequent words were treated as "low frequency", and words in the quartile centered around the median (37.5%-62.5%) were treated as "mid frequency". Each of the three frequency bins was then divided into the same eight possible contexts, for 24 categories (Figure 1, below).

FREQUENCY	High/Mid/Low							
TARGET SYLLABLE STRESS	Stressed				Unstressed			
SYLLABLE TYPE	Open		Closed		Open		Closed	
FOLLOWING SEGMENT VOICING	+	-	+	-	+	-	+	-

Figure 1: Contexts used for stimuli selection

While this methodology was essentially the same as that of Hall (2005), a few more steps were taken to further control environments. Each of the 24 categories contained four words. Where possible, words were chosen such that all four words in each condition were similar in terms of the number of syllables, and the syllable in which the target vowel occurred. This was done in an attempt to maximize the comparablility of environments. Several constraints were then imposed upon the contexts themselves.

First, for simplicity of stimulus selection, the syllable containing the target vowel was rated as "stressed" only if it carried primary stress. However, later predictions of raising or non-raising in words considered secondary stress as well. Following Hall (2005), stressed syllables that could be considered closed by an ambisyllabic segment (see Kahn 1980) were classified as "open". While this defines "open" fairly liberally, and "closed" fairly conservatively, it was necessary for the purposes of finding enough stressed open-syllable stimuli. Almost all closed syllables containing the target vowel were word-final. The few words in which the closed target syllable was not word-final (e.g. "sidestep, "limestone", "blindfold", "typeset") were chosen such that the coda was phonotactically precluded from being the onset of the following syllable, thus unambiguously closing the syllable.

Second, words with word-final open-syllable /aɪ/ were not used, since raising is blocked across word boundaries (Bermúdez-Otero 2003). Because these stimuli would be produced in isolation, no following segment or word boundary would exist to license or prevent raising, so it was decided to restrict the investigation to the effects of within-word conditions.

Third, words with two instantiations of /aɪ/ (e.g. "hypothesize", "lifelike") were also excluded from the Vancouver stimuli for the sake of minimizing confounding effects of one vowel influencing the other (note that Hall (2005)'s stimuli included several such words).

Fourth, as a frame of reference and for the purpose of comparison to previous work, some word pairs typical of studies of Canadian raising were included (*writer* / *rider*, *sighting* / *siding*, *sited* / *sided*, *write* / *ride*) as well as other near-minimal pairs or groups with similar

neighbourhood environments like *undecided* / *recited* / *chided*, *lysine* / *dicey*, *dissected* / *digress* / *divert* / *dicey*.

Finally, several words (or their morphological relatives) which elicited unexpected productions for Hall (2005) were specifically included for the purposes of replication and comparison. Some of these include *gigantic*, *Siberia*, *divert*, *digress*, *psychology*, and *ninth.*

Stimuli and fillers were randomized in Powerpoint, and further arranged by hand such that no more than two stimuli occurred successively. Where two stimuli did occur consecutively, every effort was made to have the words be as different as possible in terms of location of the target vowel within the word, stress, and syllable type, in order to minimize the salience of the target vowel across tokens. Four separate randomizations were created, each comprising three differently randomized blocks of the same stimuli.

Of these four versions, one was presented to each participant. Each block contained all 196 words (96 stimuli/100 fillers); thus, each word was produced three times by each participant. In a methodological departure from Hall (2005), whose participants produced each stimulus only once, collection of multiple tokens from each participant was done in anticipation of possible intra-speaker variability.

2.3 Procedure

The experiment took place at the University of British Columbia's Interdisciplinary Speech Research Lab. Participants were seated in front of a computer screen in a sound-attenuated booth and instructed to read aloud the words as they appeared on the screen in a normal, relaxed tone of voice. A short practice round of three filler words preceded the first block to familiarize participants with the procedure. Each word appeared, one at a time, on the screen for 2.4 seconds, a duration which, after piloting, was determined to provide participants enough time to read and produce each word comfortably while moving through the experiment without undue delay. In this way, the task differed from that used by Hall (2005), whose participants read the wordlist from paper at their own pace. As a result, speech rates in that study vary between and within speakers from word to word. It was decided for the Vancouver study that a consistent delivery of stimuli was preferable to control for speech rate, anticipating durational effects on vowel height.

Participants were recorded using a personal computer equipped with LogicPro X (Apple 1993) using an Audio-Technics AT2020 studio condenser microphone through a Focusrite Scarlett 2i2 preamp, at a sampling rate of 44.1 kHz. Data from Hall (2005)'s participants, recorded at the time onto digital audio tape (DAT), were digitized using a TASCAM DA-40 DAT machine into Audacity 2.3.0 (2018).

2.4 Data Analysis

Data from one of the female Vancouver participants was excluded due to experimenter error in recording; as a result, data from only five speakers are presented (male, n = 4; female, n = 1). Although each speaker recorded three blocks of productions, only the first block of each speaker's data was used for the present paper, due to time and resource constraints (apart from one speaker, all three of whose blocks were analyzed). It was decided to use the first block of productions to most closely replicate the data collected by Hall (2005), in which participants recorded each word only once.

Analysis of all recorded data was performed in Praat (Boersma & Weenink 2009). Following Hall (2005) and others (e.g. Sadlier-Brown 2012), a single measurement of the first and second formants (F1 and F2) of each target /aɪ/ vowel was taken at the point of maximum nuclear F1 (representing the lowest point of the diphthongal nucleus). This measurement was made by hand. Where no clear such point was obvious, the measurements were taken at the midpoint of the vocalic steady state.

Words produced haltingly, or mispronounced altogether due to unfamiliarity or production error, were excluded. In all, 2369 tokens of /aɪ/ from the two experiments were measured.

3. Results

3.1 Predicted distribution of vowels

Given the traditional rules governing Canadian Raising as laid out above, any stressed /aɪ/ preceding a tautosyllabic, voiceless segment would be expected to surface as the raised (hereafter referred to as "high") variant [ʌɪ], while [aɪ] (hereafter "low") should be the "elsewhere" surface variant. Further, if the traditional account of Canadian Raising is correct, we would expect a clear distinction between the two categories in the vowel space, with predicted high and low vowels occupying their own, discrete clusters, as shown in Figure 2.

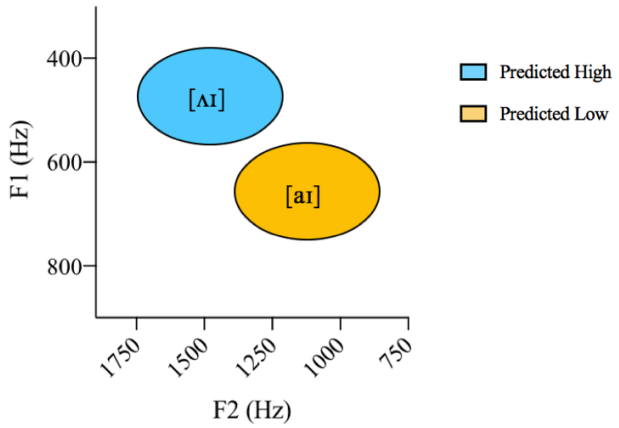

Figure 2: Hypothetical distribution of predicted high and low vowels

3.2 Actual distribution of vowels

The results of the present study, however, do not conform to either of these expectations. The graphs below present the actual distribution of the participants' productions. Each plot represents the vowel space of an individual speaker; the first six show the productions of the Meaford speakers; the last five, the Vancouver speakers.

Two observations are immediately evident. The first is that the high and low variants are not, in fact, distributed in distinct and separate clusters according to their respective predictions, but in a single cluster, comprising both vowels predicted to be high and those predicted to be low.

In some cases (Meaford participant #5, hereafter M5, for example), this singular cluster of tokens is close and dense, with both high and low vowels occupying a limited area. In others (e.g. Vancouver participant #301, hereafter V301), the distribution is remarkably diffuse. The second observation is that some vowels predicted to be low are surfacing high, and vice versa.

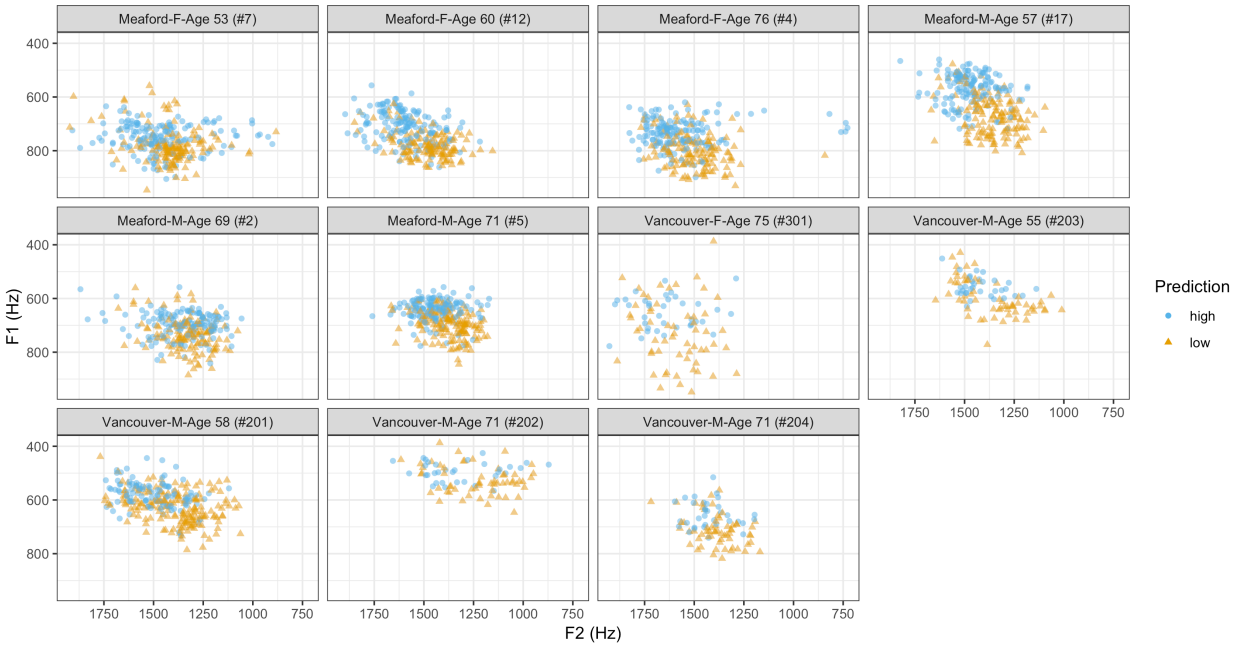

Figure 3: Individual vowel plots for six Meaford and five Vancouver speakers

Of course, this raises the question: what exactly may be considered "high" and "low" for each speaker? In other words, what exactly constitutes raising on a speaker-by-speaker – or even token-by-token – basis? Further, the degree of overlap between those vowels predicted to surface high and those predicted to surface low makes such a demarcation guesswork at best. One could, in theory, take F1 and F2 measurements from non-target central vowels as a sort of landmark against which to compare target vowels, and that might seem reasonable given the variation in the locations of each individual participant's data cluster. But it would be difficult to determine the "coverage" of this central region for each speaker, and therefore the limits beyond which a vowel may be considered raised (or not) might be hard to identify.

Thus, in order to normalize these data somewhat, and to impose less ambiguously "high" and "low" distinctions on them, the present analysis follows Hall (2005) in taking upper and lower quartiles from the data, and disregarding the middle 50%. These quartiles were determined by performing a Principal Components Analysis (PCA), and taking the upper and lower 25% from the first component. A PCA allows measurement of both F1 and F2 variation with a single variable (see Gould 1996 for background; also see Hall 2005 for details with respect to these data). As Hall (2005) points out, the decision to use quartiles in particular is an arbitrary one – one could be even more conservative and look only at, say, the upper and lower 10%, or less so by looking at the upper and lower thirds. Following Hall (2005), then, I submit that using quartiles here admits a reasonable amount of data into each category, while dispensing with the middle 50% of productions and avoiding having to make the kind of judgments described above. Removing the middle 50% yields the plots in Figure 4 below.

Subsequent analysis is then restricted to the data in these two quartiles. In other words, the ambiguous middle 50% of productions is discarded and the remaining 50% comprises the whole of what will be referred to as the "considered" data.

From these plots, a clearer picture emerges. It is, for example, apparent that for almost all participants, some vowels predicted to surface with the high variant (blue circles) appear in the lowest quartile of productions. Vancouver participant #203 (V203) is the only exception to this – his lowest quartile does in fact consist exclusively of low-predicted vowels (yellow triangles). In the opposite direction, all speakers produced vowels predicted to be low in the upper 25% of their data. These results suggest that raising is being "over-applied" in both groups of speakers. While low-predicted vowels constitute most of the lower quartile, high-predicted vowels constitute a comparatively smaller proportion of the upper. In other words, a notable proportion of vowels surfacing in the upper quartiles for both Vancouver and Meaford speakers are predicted to be low, while the converse does not occur to the same degree.

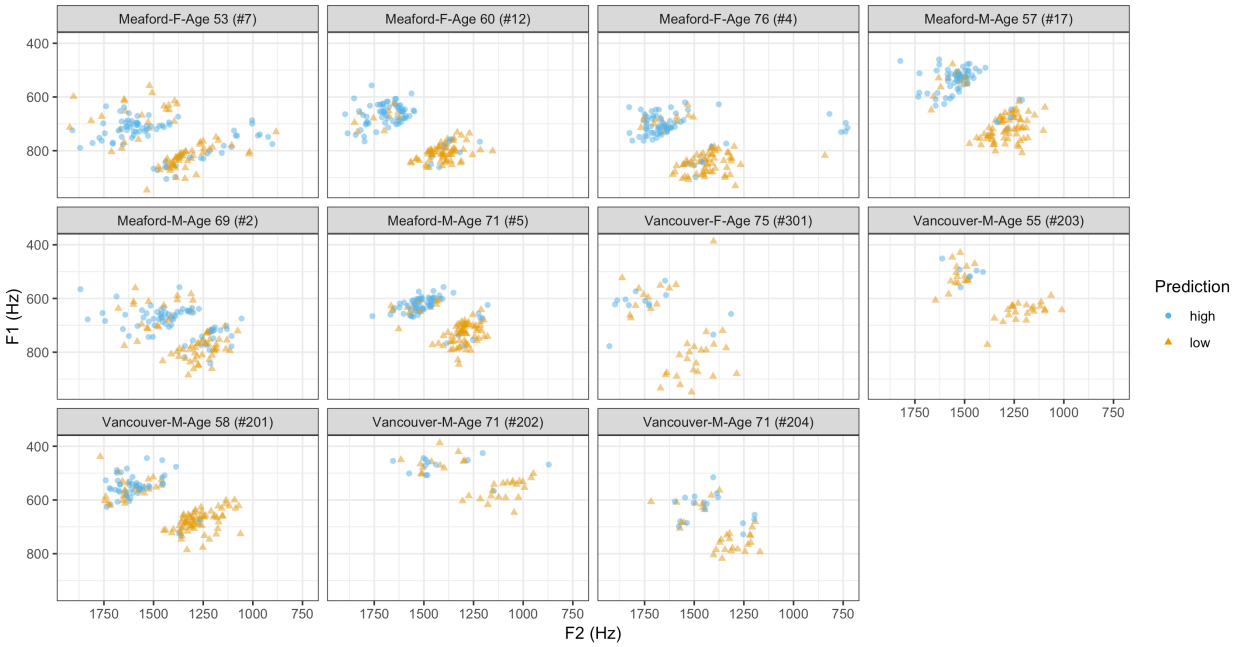

Figure 4: Individual vowel plots for six Meaford and five Vancouver speakers, upper and lower quartiles only

3.2.1 Correspondence between prediction and production - aggregates

As described above, Hall (2005)'s list of stimuli contained 299 target words, eight of which had more than one instantiation of /aɪ/, for a total of 307 target vowels. However, none of the participants in that study produced every stimulus without error. Likewise, none of the Vancouver speakers produced every one of the 96 target words correctly. Since mispronounced stimuli were discarded (see §2.4), the number of actual tokens analyzed varies from speaker to speaker; however, an average of 287 tokens per Meaford speaker (for a total of 1724) and an average of 95 tokens per Vancouver speaker (for a total of 662) were available for analysis. As explained above, those tokens realized in between the upper and lower quartiles (i.e. in the middle 50%) are not henceforth considered; the number of tokens analyzed is then 862 (Meaford) and 332 (Vancouver).

Across all Meaford speakers, 447 considered tokens were predicted to surface with the low variant. Of these, 346 (77.4%) did surface in the lower quartile, while 101 (22.6%) surfaced in the upper quartile. Of 415 considered tokens predicted to surface with the high variant, 330 (79.5%) did, while 85 (20.5%) were produced in the lower quartile.

Across Vancouver speakers, of 232 tokens predicted to surface with the low variant, 154 (66.4%) actually did, while 78 (33.6%) surfaced in the upper quartile. Meanwhile, 100 tokens were expected to surface with the high variant; 88 (88%) did in fact, while 12 (12%) surfaced with the low variant.

3.2.2 Correspondence between prediction and production - individuals

Within these aggregates, individual speakers performed variably. Although on average, Meaford speakers produced low-predicted vowels in the lower quartile 77.2%[1] of the time, individual scores ranged from as low as 67.1% (M7) to as high as 85.1% (M12). Likewise, high-predicted vowels surfaced in the upper quartile of Meaford speakers 79.9% of the time, on average. However, individual scores ranged even more, from 66.2% (M7) to 89.1% (M17). Notably, M7 produced the lowest percentage of prediction-aligned vowels in both low and high cases.

Among Vancouver speakers, low-predicted vowels surfaced in the lower quartile 64.2% of the time, on average. Individual score ranges from 59.0% (V203) to 72.3% (V201). High-predicted vowels surfaced in the upper quartile 87.5% of the time, on average, though individual scores ranged from 76.5% (V204) to 100.0% (V203). It is also noteworthy that the same speaker (V203) produced the lowest percentage of prediction-aligned low vowels and the highest percentage of prediction-aligned high vowels (see Appendix for individual surface distributions).

These results indicate that across all speakers, on average, the diphthong variants are produced as predicted roughly three-quarters of the time. Vancouver speakers show a much greater disparity in prediction-production correspondence between high- and low-predicted vowels, while Meaford speakers are relatively consistent. That is, Meaford speakers produced high- and low-predicted vowels as expected with nearly the same frequency (79.9% and 77.2%, respectively), while Vancouver speakers produced high-predicted vowels as expected much more reliably (87.5%) than they did low-predicted vowels (64.2%). Looked at another way, for Vancouver speakers, low-predicted vowels were much more likely to surface in either quartile than were high-predicted ones. Meaford speakers, on the other hand, were more or less just as likely to realize high vowels as expected as they were to realize low vowels as expected. Vancouver speakers seem then to be over-applying raising to more than a third of low-predicted vowels, while Meaford speakers do so roughly one-fifth of the time. However, for all speakers, a greater proportion of low-predicted vowels surfaced high than the reverse.

If the traditional account of Canadian Raising is correct, we would expect 100% of vowels to surface as predicted, with perhaps some allowance made for production error or token-level variation. Instead, roughly three-quarters of vowels across all speakers surfaced as predicted, with the remaining quarter surfacing in the opposite quartile. Naturally, it is unclear from this analysis what proportion of vowels would surface in their predicted categories, were those categories to include the 33rd, or even the 26th, percentiles. However, this does not diminish the fact that, as above, an average of 23% of vowels actually surfaced in the quartile opposite

[1] Note that the figures reported in this section represent averages across participants. Since the number of tokens analyzed for each participant was not constant (see § 3.2.1), these averages differ slightly from the aggregate percentages reported in § 3.2.1.

their predicted one. Whatever the explanation for this pattern may be, it is certainly problematic for a traditional account of Canadian Raising.

3.3 Composition of upper and lower quartiles

If the traditional formulation of Canadian Raising were to hold, and these variants are truly allophonic, we would expect that the upper quartile of productions would contain only high-predicted vowels, while the lower quartile would contain only low-predicted vowels. The above results indicate that this is not the case.

If, however, these variants were in fact fully contrastive, we might expect a distribution between quartiles corresponding to that in the set of stimuli. That is, if half of the dataset were predicted to surface with the high variant under an allophonic view, then half of both the upper and lower quartiles might be expected to consist of high-predicted vowels if the vowels have actually become fully unpredictable. The same should be expected of the low-predicted half of the dataset; the other half of each quartile would then consist of vowels predicted to surface low.

From the 862 Meaford tokens, the aggregate upper and lower quartiles contain 431 tokens each. Given the distribution of environments in Hall (2005)'s dataset, 207 (48.1%) of the tokens in each quartile should be high-predicted, assuming a fully random contrastive distribution. 224 (51.9%) of the tokens in each quartile, then, should be low-predicted. In reality, 330 (76.6%) of the tokens in the upper quartile are high-predicted, and 346 (80.3%) of the tokens in the lower quartile are low-predicted. Binomial tests reveal that each of these percentages is statistically significantly different from its predicted value (48.1% or 51.9%) ($p < 0.001$).

Environments in the Vancouver dataset were not as evenly distributed; however, their distribution arguably more closely aligns with that within the language. Given a corresponding distribution of the dataset across the two quartiles, of 166 tokens in each, 50 (30.1%) would be high-predicted and 116 (69.9%) would be low-predicted. However, 88 (53.0%) of the tokens in the upper quartile are predicted high, and 154 (92.8%) of the tokens in the lower quartile are predicted low. Again, binomial tests indicate that each of these percentages is statistically significantly different from its predicted value (30.1% or 69.9%) ($p < 0.001$).

Returning to the hypothetical "allophonic" distribution (i.e. 100% of the upper and lower quartiles are high- and low-predicted vowels, respectively), comparison of the actual percentages to 100% also shows a statistically significant difference in all cases ($p < 0.001$). The table below presents these quartile compositions. In other words, the actual distribution of predicted diphthongs in each quartile differs from what would be expected under either pure allophony or complete and unpredictable contrast.

% Low-Predicted in Lower Quartile		% High-Predicted in Upper Quartile	
Meaford	Vancouver	Meaford	Vancouver
80.3	92.8	76.6	53.0

Table 1: Proportion of prediction-matched tokens in upper and lower quartiles

How do the two populations compare to each other? Proportion tests show that they are significantly different in both the lower and upper quartiles ($p < 0.001$ for both), which indicates

that the patterns of Canadian Raising are different across the two populations. Nearly 20% of the Meaford speakers' low productions actually derive from vowels in raising environments, while almost 25% of their high productions are in non-raising contexts. For Vancouver speakers, on the other hand, low productions are somewhat more "appropriately" derived from non-raising contexts (only about 7% of low productions come from raising contexts), but an alarming situation exists in the raised productions, where nearly half are vowels in non-raising contexts.

While the two populations are different, they are similar in that high- and low-predicted vowels in both Meaford and Vancouver speakers clearly follow neither an allophonic nor a contrastive distribution. In both groups, Canadian Raising is not nearly as predictable as has been believed. Further, and most dramatically by the Vancouver speakers, it is occurring in more contexts than had been described previously, with less restriction in terms of phonological environments. What's more, this loosening of environmentally-specific conditioning is not sufficiently systematic to simply allow for revision of the rule. All of this is clearly problematic for the "received" account of Canadian Raising.

3.4 Inter-speaker conformity and variation

Hall (2005) reported that only 37 of 307 words (12%) showed agreement across the three speakers analyzed in that study, which is to say that all three speakers produced these words in the same quartile, be it upper or lower. Of these 37 words, 29 surfaced as predicted (see Hall 2005 for details of these findings). Across the six Meaford speakers analyzed in the present study, 25 words (out of 307, or about 8%) were produced in either the upper or lower quartile unanimously. Several of these words were the same as those agreed upon by Hall (2005)'s speakers, for example *dicey*, *exciting*, *five*, and *July*. Notably, all of these productions aligned with their predicted height. Further, an additional 36 words (or about 12%) were agreed upon by five of the six Meaford speakers. Of those, 32 were produced as predicted, while five (*digression*, *diverted*, *ice*, *gigantic* and *gigantically*) were produced with the raised variant [ʌɪ], contrary to prediction.

For their part, the five Vancouver speakers produced 12 unanimously agreed-upon words (of 96, or about 13%), 10 of which were realized with the predicted vowel height. The two words produced by all speakers with the unexpected variant [ʌɪ] were *dissected* and *gigantic*. It is noteworthy that *gigantic* was one of the words produced with the same vowel by nearly all Meaford speakers. In fact, 10 of 11 speakers across the two populations produced *gigantic* with the raised variant. It may also be noteworthy that *dissected*, while not exactly the same as the Meaford speakers' unexpectedly raised *diverted* and *digression*, does share the prosodic characteristics that would predict non-raising (in addition to the same preceding segment) as those two words, and was also produced unexpectedly with the raised variant by Vancouver speakers. Further, an additional 16 words (roughly 17%) were produced with the same variant by four of the five Vancouver speakers. Of these, 13 aligned with their predicted vowel height. The three that did not were *titanic*, *trifecta*, and *Nigerians*, all of which were produced with the raised variant [ʌɪ].

Interestingly, where consensus (or near-consensus) was reached by speakers, vowel realizations corresponded with prediction most of the time (92% for Meaford speakers, 82% for Vancouver speakers). When they didn't align with prediction, in every case but one (*ice* in five of six Meaford speakers), the vowels were raised despite being predicted low. This is perhaps not especially surprising, given the overall distribution of variants with respect to their predictions

(above). However, outside of these results, there is not enough consistency to make any real predictions for any particular word, either within or across the two groups of speakers.

3.5 Intra-speaker patterns and variation

Within participants, significant variation is evident as well. For example, one Meaford participant raised the target vowel in *hypodermic* but not in *hypothermia*, despite both words being the same with respect to syllable type, stress, following segment voicing, and even morphological status. Another Meaford participant produced *icily* with a raised vowel, but *ice* with a low one. The same participant also produced both *dichotomy* and *digression* with the high variant. Not only is the target vowel in these words in an open syllable, with primary stress falling on the following syllable, but in the case of *digression*, the following segment is voiced. Thus, in neither word would the vowel be expected to raise, least of all in the latter. This kind of pattern is seen in some other cases as well, where, within speakers, some phonologically or morphologically similar words seem to pattern together. For example, one Vancouver speaker raised both *bide* and *ride*, while a Meaford speaker produced both *bike* and *mike* with a low vowel. Again, however, there is not enough consistency in these aberrations, even within speakers, to make any generalizations. This is clearly perplexing for any notion of predictability in Canadian Raising, and confirms the similarly surprising results Hall (2005) found in her data.

4. Discussion

4.1 Individual and collective models of Canadian Raising

Hall (2005) suggests that some of the unexpected patterns found in her results may be explained by considering biphones in lexical neighbourhoods. She noticed that many instances of /sai/-, /dai/-, and /dʒai/-containing words seemed to pattern similarly, and posited that the raising/unraising behaviour of the most frequent member of each neighbourhood would pull the other members along, regardless of their individual characteristics. What makes this idea particularly interesting, and relevant to the current data, is that it *predicts* that individual speakers will have unique models of raising, based on their own lexical experience. Further, we might even expect some individuals within a group to overlap in some, but not all, instances. We might also expect different groups to behave differently. Indeed, some of these patterns are visible in the present data as well (see above, § 3.3). In particular, /dai/-initial words such as *digression*, *diverted*, and *dicey* and /dʒai/-initial words such as *gigantic* and *gigantically* do seem to pattern together for Meaford speakers. It should be noted here that the datasets for the Meaford and Vancouver experiments had some words in common, but not all, in fact not most. Thus it is with caution that any conclusions are drawn with respect to patterns of similarity between the two populations. That notwithstanding, however, recall that a number of words were included in the Vancouver list expressly for the purpose of this comparison (see § 2.2). In fact, nearly 30 words were common to both datasets, most of which were identical. Those that weren't entirely identical were identical with respect to context (for example *bike ~ biked*, *recital ~ recited*, *pricey ~ priceless*). It is notable, then, that of all the common or comparable words between the two datasets, only two (*dicey* and *gigantic*) were agreed upon by all speakers. This suggests that 1) individual speakers may have their own raising models, whether based on

phonological environment or some paradigm uniformity, or both; and 2) groups of individuals may have raising models with some common characteristics, distinct from those of other groups.

4.2 The case for contrast neutralization

The data presented to this point are clearly not supportive of any orthodox account of Canadian Raising. Its occurrence (or non-occurrence) has been shown to be largely unpredictable in the speakers from both Meaford and Vancouver. Given that word-level variation exists both between and within speakers, how are we to characterize these results? What does this mean for Canadian Raising?

It is true that some notable patterning was attested between speakers' productions of certain words, or morphophonologically similar words. However, whatever word-level similarities are present, notable though they may be, they are insufficient relative to the rest of the data to make any clear predictions or generalizations. One thing is clear, however; these Vancouver speakers are doing something different from their Meaford counterparts. Both groups are "misbehaving" with respect to rule-derived raising, but the Vancouver speakers are significantly more likely to over-apply raising, while rarely doing the opposite. We have seen that the distribution of vowels in both populations is hardly allophonic. However, it is still largely the case that vowels will raise predictably in raising contexts, though by no means is this universal. What is not predictable is what the vowels will do in low contexts. The suggestion here, then, is that what was perhaps allophony at one time is now becoming a contrast in the low environments. This contrast is then neutralized in raising contexts. This unidirectionality is particularly salient amongst the Vancouver speakers. Again, predictability (though not allophany) is preserved in raising environments, so it appears that a phonemic splitting, or contrast-in-progress may be underway.

If such a change is underway, it might not surprise us that it appears to be further advanced among the Vancouver speakers than the Meaford ones, for two reasons. The first is that the Vancouver speakers themselves are 14 years removed from the Meaford speakers, so that intervening time may have seen some of this change take place. Second, greater linguistic heterogeneity in a multi-cultural port city like Vancouver than in a rural, central Canadian town like Meaford may also contribute to such change, or at least to an increased rate of change if it is in fact occurring among both populations (and others). An examination of younger speakers in Meaford today might be instructive on this last point.

The case for a neutralization-in-progress would be strengthened by examination of productions from younger speakers. As it turns out, another group of Vancouver speakers did participate in this experiment. These were Vancouver-born, native English-speaking university students aged 18-28. A follow-up to this study, using data from these speakers might lend support to this hypothesis.

4.3 Future directions

Clearly, analysis of much more data from both Meaford and Vancouver, as well as other areas, needs to be done to establish any real patterns. As above, data from younger speakers in both cities should be looked at for evidence of a contrast-in-progress (or for evidence to the contrary). Recall also that, but for one speaker, only one of three blocks of Vancouver participants' data was analyzed, so another direction for future work would involve analysis of the remaining data.

A very preliminary and impressionistic examination of the one speaker's three sets of data indicate some intra-speaker variation within words themselves. Further evidence of this could have implications for the role of phonetic considerations in Canadian Raising.

Finally, the operationalization and measurement of raising in the current study is admittedly quite coarse. Consideration of factors beyond simple F1 values, such as formant trajectories, rates of formant change, and relative nucleus and offglide duration should be examined for more nuanced results.

5. Conclusion

The results of the present paper support the findings of recent studies (e.g. Rosenfelder 2007, Sadlier-Brown 2012, Swan 2017) that Canadian Raising is occurring in (at least parts of) British Columbia. However, the purpose of this study was to attempt to replicate the findings in Hall (2005) and to determine whether or not they might generalize to other speakers, both from Meaford and from Vancouver. These objectives were met; Hall (2005)'s rather surprising data from three Meaford speakers were confirmed in a larger subset of participants from that location, and in a subset of Vancouver speakers as well. Further, the pattern of raising found among these Vancouver speakers especially shows a clear tendency towards over-application of the process, potentially indicative of a contrast neutralization in raising environments.

This is not to contend that the environments thought to condition raising are of no predictive value at all; in fact, it is somewhat reassuring to see that raising contexts do in fact largely predict raised vowels. It may be that vowel raising is only one of several means at speakers' disposal of reliably differentiating vowels in raising and non-raising environments. If factors such as those listed above (§4.3) were to be considered in addition to vowel height, it may emerge that "raising" isn't necessarily raising at all, but a set of strategies involving some or all of these, from which speakers may choose, on a by-case basis, which to use to differentiate variants.

Such possibilities aside, however, the present results indicate that a traditional model of Canadian Raising is insufficient to account for its occurrence in both populations studied here. This is hardly the first report of unexpected realizations within what linguists thought was a stereotypical phonological rule. However, the results of this study should leave little doubt that Canadian Raising is far from the simple, predictable phenomenon it has long been considered to be.

References

Apple (1993). *Logic Pro X on the Mac App Store*. [Computer software]. Available at: https://itunes.apple.com/gb/app/logic-pro-x/id634148309?mt=12

Audacity Version 2.1.2 [Computer software]. (2016). Available: http://www.audacityteam.org/

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database.* Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In *Variation within Optimality Theory: Proceedings of the Stockholm workshop on 'variation within Optimality Theory'*, ed.Jennifer Spenader, Anders Eriksson and Östen Dahl, 25–36. Stockholm: Department of Linguistics, Stockholm University.

Boberg, Charles. 2008. Regional phonetic differentiation in Standard Canadian English. *Journal of English Linguistics* 36(2): 129–154.

Boersma, P. & Weenink, D. (2015*). Praat: doing phonetics by computer. Version 5.4.08.* Available from http://praat.org/

Britain, David. 1997. “Canadian raising” in the English Fens. *Language in Society* 26(1): 15–46. Cambridge University Press.
Chambers, J. K. 1973. Canadian raising. *The Canadian Journal of Linguistics / Revue canadienne de linguistique* 18(2): 113–135.
Chambers, J. K. 1989. Canadian raising: Blocking, fronting, etc. *American Speech* 64: 75–88.
Chambers, J. K. 2006. Canadian raising retrospect and prospect. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 51(2/3) : 105–118.
Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague: Mouton.
Dailey-O’Cain, Jennifer. 1997. Canadian raising in a midwestern U.S. city. *Language Variation and Change* 9: 107–120.
Fruehwald, Josef T. 2007. The spread of raising: Opacity, lexicalization, and diffustion. *CUREJ: College Undergraduate Research Electronic Journal*. University of Pennsylvania, http://repository.upenn.edu/curej/73
Gould, S. J. (1996). *The mismeasure of man* (Revised and expanded ed.). New York: Norton.
Hall, Kathleen C., Blake Allen, Michael Fry, Scott Mackie, and Michael McAuliffe. 2016. *Phonological CorpusTools*, Version 1.2. [Computer program]. Available from PCT GitHub page.
Hall, Kathleen C. 2005. Canadian raising revisited: Evidence for a gradient, lexicon-based approach. Paper presented at the Canadian English in the Global Context conference, University of Toronto, Toronto, ON.
Harris, Zellig S. 1960. *Structural linguistics*. Chicago: University of Chicago Press.
Hualde, José I., Tatiana Luchkina, and Christopher D. Eager. 2017. Canadian raising in Chicagoland: The production and perception of a marginal contrast. *Journal of Phonetics* 65: 15–44.
Idsardi, William. 2006. Canadian raising, opacity, and rephonemicization. *The Canadian Journal of Linguistics* 51(2/3): 119–126.
Jensen, John T. 2000. Against ambisyllabicity. *Phonology* 17: 187–235.
Joos, Martin. (1942). A phonological dilemma in Canadian English. *Language* 18: 141–144.
Kahn, Daniel. 1980. *Syllable-based generalizations in English phonology*. New York and London: Garland.
Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2003. Looking through opacity. *Theoretical Linguistics* 29: 123–129.
Milroy, James. 1996. Variation in /ai/ in Northern British English, with comments on Canadian raising. In *(N)WAVES and MEANS: A selection of papers from NWAVE 24*, 213–221. University of Pennsylvania working papers in linguistics.
Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1): 62–85.
Onosson, D. Sky. 2014. The prosodic structure of Canadian raising. In *Proceedings of the 2014 annual conference of the Canadian Linguistic Association* ed. Laura Teddiman.
Rosenfelder, Ingrid. 2007. Canadian raising in Victoria, B.C.: An acoustic analysis. *AAA: Arbeiten aus Anglistik und Amerikanistik* 32(2): 257–284.
Sadlier-Brown, Emily. 2012. Homogeneity and autonomy of Canadian raising. *World Englishes* 31(4): 534–548.
Swan, Julia T. 2017. Canadian raising on the rise in Vancouver? A phonetic comparison of Vancouver, BC and Seattle, WA. In *Proceedings of the 2017 annual conference of the Canadian Linguistic Association* ed. Andrew Alexander Monti.
Trudgill, Peter. 1986. *Dialects in contact.* Oxford: Basil Blackwell.
Vaden, K. I., Halpin, H. R., Hickok, G. S. (2009). *Irvine Phonotactic Online Dictionary, Version 2.0.* [Data file]. Available from http://www.iphod.com.
Vance, Timothy J. 1987. “Canadian raising” in some dialects of the northern United States. *American Speech* 62: 195–210.

Appendix

	Low				Mid				High			
	Stressed		Unstressed		Stressed		Unstressed		Stressed		Unstressed	
	Open	Closed	Open	Closed	Open	Closed	Open	Closed	Open	Closed	Open	Closed
__[+v]	transcriber chided livened misers	sidestep tithe liveliest ninths	Nigerians dilutes migrations hydroxy	oxides stockpiled coastlines minimized	visor priming undecided sided	bide limestone sidewalks chives	ideally digress hydraulics primeval	sundial waistline lullabies peacetime	rider hijacked climate siding	ride blindfold vine kindly	Siberia vibrations gigantic divert	paralyzed headline bedtime reptile
__[-v]	ripened sited pipers lysine	biked sniped typeset lifespans	microbial licentious pipette iconoclast	lookalike phenotypes bauxite urbanite	dicey biter recited hiker	ignites sufficed unites entice	trifecta hyperbole dissected vicarious	tailpipe poltergeist campsite graphite	writer sighting crisis icon	priceless write wife slightly	psychologist Titanic itinerary criteria	midnight sacrifice prototype afterlife

Table A1: Vancouver stimulus conditions and wordlist

Speaker	Predicted Low, Surfaced Low (%)	Predicted High, Surfaced High (%)
Meaford 2	68.1	68.7
Meaford 4	80.8	81.7
Meaford 5	80.0	87.5
Meaford 7	67.1	66.2
Meaford 12	85.1	86.1
Meaford 17	82.1	89.1
Meaford Average	**77.2**	**79.9**
Vancouver 201	72.3	92.0
Vancouver 202	62.5	84.6
Vancouver 203	59.0	100.0
Vancouver 204	64.5	76.5
Vancouver 301	62.9	84.6
Vancouver Average	**64.2**	**87.5**
All Speakers Average	**71.3**	**83.4**

Table A2: Surface distribution of high- and low-predicted tokens (by individual speaker)